ISSN : 2456-3307

IJSRCSEIT

A State Level Symposium & IT Meet - InnovIT'18

Organised by
PG and Reaseach Department of Computer science,
Shanmuga Industries Arts and Science College,
Tiruvannamalai, Tamil Nadu, India

UGC Approved Journal [Journal No : 64718]

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN COMPUTER SCIENCE, ENGINEERING AND INFORMATION TECHNOLOGY

Volume 4, Issue 3, January-February-2018

Email: editor@ijsrcseit.com

SHANMUGA INDUSTRIES ARTS AND SCIENCE COLLEGE (Co-Ed.,)
Certified under Section 2(f) & 12B of the UGC Act 1956
An ISO 9001: 2000 Certificated Institution
Permanently Affiliated to Thiruvalluvar University,
Vellore and Approved by the Government of Tamil Nadu & AICTE

A State Level Symposium & IT Meet - InnovIT'18

2nd FEB 2018

In Association with

International Journal of Scientific Research in Computer Science, Engineering and Information Technology

ISSN : 2456-3307

UGC Approved Journal [Journal No : 64718]

Volume 4, Issue 3, January-February-2018

International Peer Reviewed, Open Access Journal

Organised by:

PG and Reaseach Department of Computer science,
Shanmuga Industries Arts and Science College,
Tiruvannamalai, Tamil Nadu, India

Published By

Technoscience Academy

(The International Open Access Publisher)

Email: info@technoscienceacademy.com
Website: www.technoscienceacademy.com

About College

Tiruvannamalai is an ancient town exists since 8th century and it is one of the most ancient heritage sites of India and is a centre of the Saiva religion. The term "Annamalai" implies an inaccessible mountain and the word "Thiru" was prefixed to signify its greatness. These two terms coupled together to signify the name “Tiruvannamalai”. The temple is grand in conception and architecture and is rich in tradition, history and festivals. The main Deepam festival attracts devotees from far and wide throughout India. Tiruvannamalai, a complacent temple town which is rightly connected to the cities like Chennai, Pondicherry, Bangalore and Trichy. A galaxy of temples and ashrams situated in the quiet valleys of the town help people to undergo a Divine and Spiritual experience.

Shanmuga Industries Arts and Science College, popularly known as SIASC, is a Co-educational institution promoted by Shanmuga Industries Educational Trust, Tiruvannamalai. The objective of the Trust is to enable the college into an institution of excellence and to let the rural youth living in and around Tiruvannamalai to have easy access to higher education. The college is situated in Tiruvannamalai on the Tiruvannamalai-Manalurpet state highway. The premier institute of college education was established in the year1996. Since it’s founding, SIASC has distinguished itself by providing a higher level of culture, cultivating good discipline and finer value of life among students. The college covers a vast area of land comprising classrooms, laboratories, computing centers, auditorium, hostel, library etc. SIASC has facilities that are exceptional in every way. The environment- friendly green campus and calm atmosphere at SIASC helps an individual to discover himself and the contribution he can make to the human kind.

SIASC, is one among the leading institutions in the country to have been awarded with ISO 9001:2000 certificate in recognition of its quality standards. The facilities and infrastructure that the institution has, is much above the benchmark propounded by the University. The strength of the college can be attributed to the commitment of its management, its principal, its team of distinguished staffs, its creative and dedicated students. A reputation for excellence in higher education supported by a high caliber staff is reflected in the demand for entry to SIASC from achieving students. At Shanmuga, the staff members are not merely Lecturers, instead they are profound scholars and educators who will help an individual to grow towards his personal best during their three years stay at the campus. The essence of Education at SIASC is to ensure a smooth transition from Student to a Professional.

SHANMUGA

Shanmuga Industries Arts and Science College, popularly known as SIASC, is a Co-educational institution promoted by Shanmuga Industries Educational Trust, Tiruvannamalai.

CONTENTS

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3 | ISSN: 2456-3307

ACCUMNET : Optimal Routing Algorithms for Multi-Channel Multi-Hop Networks

Mrs. B.Arulmozhi[1], Ms .P.Agalya[2], Ms. A.Sivasankari[3]

[1]Head of the Department (BCA), Dept. of Computer Science and Applications, D.K.M College for Women (Autonomous), Vellore, Tamilnadu, India

[2]Research Scholar, Dept. of Computer Science and Applications, D.K.M College for Women (Autonomous), Vellore, Tamilnadu, India

[3]Head of the Department (CS), Dept. of Computer Science and Applications, D.K.M College for Women (Autonomous), Vellore, Tamilnadu, India

agalyapitchaimuthu@gmail.com[1]

ABSTRACT

In this paper, we tend to gift a routing and channel assignment protocol for multi-channel multi-hop wireless networks. We tend to contemplate a multi-hop network, wherever a mobile host might connect with associate access purpose victimization multi-hop wireless routes, via alternative mobile hosts or wireless routers. In addition, we tend to take into account a multi-channel network wherever multiple non-overlapping (orthogonal) channels area unit on the market, and every host or router will dynamically choose a channel to enhance performance. During this system, we tend to investigate the optimum routing downside in signal transmission from supply to destination for the multi-hop network. We tend to use the technique of rate less code that is employed to accumulate the information with every packet within the transmission. This could decrease the entire energy; scale back delays in transmission for transmission information from the supply to the destination. Proposed system allows vital performance through the shortest path routing victimization Floyd-Warshall algorithmic rule.

Keywords: Accumulative Multi-Hop, Energy Accumulation, Minimum Energy, Route discovery, Route selection, Route representation, Data forwarding, Route maintenance, Route energy efficiency

I. INTRODUCTION

In the wireless, network the information transmission between the supply and destination maintained by the cooperation between the two nodes. Within the tradition network, that information transmission between supply and destination achieved through the intermediate node that may receive the data from immediate nodes and transmits to next node. Typically, this drawback within the information transmission like delays in routing needs a lot of energy to transmit the information. Within the today's era of the network the relays idea wide used is relay channeling. Compared to ancient system during this, nodes use the data of all nodes rather than nearest one. This idea first projected by vander Meulen. The matter of routing communication networks, within which we tend to are instead interested here, is but removed from being understood nowadays. Within the simplest accumulative multi-hop network, one supply communicates to one destination aided by many relay nodes that may accumulate the received energy information from previous relay transmissions. In observe there are 2 main accumulation mechanisms at relays: energy and mutual information accumulation. Energy accumulation may be performed at the receiving nodes, e.g., through space-time continuum secret writing or repetition secret writing. Mutual-information accumulation may be realized victimization rate less codes e.g. fountain bird of prey codes. Accumulation mechanisms are thought of in current and next generation standards since they increase communication dependability and cut back energy consumption.

In this system, the relay channel considers one relay aided to data transmitted between supply and destination. This has robust management over the information transmission within the routing in smart rates. During this system self-addressed the matter of Accumulative multi-hop network routing within the communication between two nodes. The communication between two nodes through the only supply to single destination that is accumulated with relays gained from the immediate nodes. The buildup is completed by two ways that energy accumulation decoded packet in the end energy received from the supply node. Within the information transmission multi-hop information, we tend to principally specialize in decipher and forward strategy whereas transferring the data from single supply to destination. The mutual information accumulated till full message decoded. This will become totally alert to rate fewer codes like fountain bird of prey code. This will increase the dependability and reduces the energy demand within the transmission. Studied the matter of routing in multi-hop wireless network victimization the buildup of best mutual data with facilitates of distress optimality.

II. LITERATURE SURVEY

In the second section, we have a tendency to justify our work in short to line of methodology and summary of routing in multi-hop networks.

A. Background

Yaling principle Jun Wang introduced the necessity style and kinds of the routing protocols. There are 2 styles of routing protocols one is path calculation formula and packet forwarding theme. Path calculation formula is employed within the completely different network for path calculation. Knowledge is transmitted through a distinct path. For this they need used flooding-based route discovery, Dijkstra's formula and therefore the Bellman-Ford formula. These algorithms verified best for the wireless network. Within the alternative form of protocol supply routing and hop by hop routing forwarding theme is employed to send knowledge. supply routing theme knowledge is undemanding through the headers. Hop by Hop theme forwards the packet through a node by node to destination.

These are the theme to send the packets however whereas send forwarding and receiving the information node needs energy to remain active. whereas transmission packets nodes gain the energy from received overhead signals and assembling energy from re-transmission. This causes the low energy knowledge broadcast downside and needs additional power. To tackle this downside, one approach projected named as a cooperative strategy. This approach is employed whereas broadcasting, which may use native knowledge and loosely synchronized. By distinguishing the order of nodes and determinant the ability of every node.

Jiangzhuo subgenus Chen, Lujun Jia et.al addressed the matter of accumulative routing. Knowledge transfers ordinarily the most expensive activity of a wireless node in relationships of power consumption. Numerous ways are planned to shrink the energy expenses within the communication method. Address the matter of energy economical routing and identification of multipath routing supported numerous metrics. Cooperative relay theme is for a source to destination communication through the relays. Ancient energy consumption approach is often used with the fountain code employed in relays to accumulate energy. There is a drag of energy consumption and knowledge accumulation. Remaining work shows the review of routing ways and transmission within the network.

B. Review Of Package Recommendation Ways

João Luís Sobrinho [1] conferred the pure mathematics theory for the routing in wireless network investigated the pure mathematics for the dynamic routing. This used for shortest path routing to generalize the positive length cycles. Strengthen the convergence in routing inexplicit the properties referred to as monotonicity and isotonicity. First property converges in each network and second property converges in best path. Intra domain routing protocol wont to converge the short and wide methods in any network and repose domain to frame entree. The mathematical term projected for verification of routing policies.

Yaling principle and Jun Wang guided to style the routing metrics during a multihop network. First mentioned, that the characteristic is very important for the planning of metrics. A distinct network conjures up to structure the metrics and to induce numerous aspects of networks. It someday affects the operating to routing protocols. If these metrics not combined with correct protocols could cause the matter in routes and suboptimal methods. The author studied the importance of metric and protocol relationships and provided tips [2].

Ivana Maric and Roy D. Yates addressed matter of minimum energy broadcast problem. The nodes collect the energy whereas transmission the messages. They studied cooperative strategy for energy accumulation and chiefly targeted on the synchronized, low power network. That uses the native data to broadcast on the network. To beat lower energy downside projected the 2 approach 1st identification of nodes ordering within which message has got to be a pass. Alternative is finding of the ability thereto order. Among those second downside is resolved by victimization the applied math associate degreed used an formula for ordering nodes. Experimented it and therefore the result shows the higher performance [3].

Andreas F. Molisch and Neelesh B. Mehta et.al studied fountain code technique during a wireless network. N range of relays wont to transmit the data type supply to destination victimization fountain codes. The matter of ancient approaches is that it solely accumulates the energy during a cooperative manner whereas fountain codes are with efficiency accumulated the data. This reduces the specified energy to send the information from supply to destination. Whereas causing knowledge analyzed the behavior of supply node and relay after they begin and stops transmission the data and rewrite it. It the optimized and mutual data approach used for reducing energy consumption and time.

Zigui principle and Anders Høst-Madsen investigated energy economical cooperative multiple relay channel once carrier level synchronization not allowed and use of rewrite and forward approach. Showed rewrite and forward theme is economical logical thinking free victimization easy path and power allocation strategy.

III. ROUTING COMPONENT: AN EXHAUSTIVE VIEW

By breaking down the wireless routing protocol into smaller parts, we will analyze the weather that have to be compelled to be boxed-in in any wireless multi-hop routing protocol and show the interacting behavior between them. The behavior of these basic parts is customized to utterly completely different application profiles and needs, whereas keeping and maintaining the core helpful behavior and goals [1]. To satisfy network and application specific needs, extra parts are added to the routing protocol to control its behavior and maintain its performance procreate and specific by the appliance and network paradigm. Having the core parts, a routing protocol is solely extended to accommodate and support extra needs, services and choices by adding auxiliary parts

A. Route Discovery

Route discovery is that the initial stage of the perform of any wireless routing protocol. Route discovery is that the method of finding a route/set of potential routes between a supply and a meant destination. The method of finding a route may be classified into 3 categories: proactive, reactive or hybrid. Proactive route discovery, additionally called table-driven route discovery, depends on the utilization of up-to-date routing data regarding the complete network to search out a path from any supply to any destination within the network. This routing data is changed among nodes either sporadically or upon the prevalence of any modification within the topology. This data is unbroken at every node during a routing table. this sort of route discovery pre-determines routes between any 2 nodes regardless of the necessity for such routes. Once a node encompasses a packet to be sent, it ought not to look ahead to a route to be discovered. It consults its routing table, gets the up-to-date recorded route, then sends the packet while not acquisition a delay for the route to be discovered—the route is discovered a priori. There square measure 2 sub-categories underneath the proactive routing category: Distance Vector (DV) and Link State (LS). They dissent in however the topology data is unfolded. These techniques square measure borrowed from wired networks however they can be changed to handle the characteristics of MANETs.

(a) Distance Vector Proactive Routing

In DV route discovery, every node maintains a routing table wherever it stores information concerning all potential destinations, future node to succeed in that destination, and also the best glorious distance to succeed in the destination.1 These tables square measure updated by exchanging data with the neighbors. Every node sporadically sends a vector to its direct neighbors carrying the knowledge recorded within the routing table to take care of topology. The gap vector contains the destinations list and also the cost—the distance—to reach every destination.

(b) Link State Proactive Routing

Distance vector routing was utilized in ARPANET till 1979, once it absolutely was replaced by link state routing. the target of LS routing is to produce an alternate to DV that avoids routing loops and also the ensuing “count-to-infinity” downside. LS routing overcomes this by maintaining international configuration data at every node. In LS routing, every node sporadically sends data concerning the value to achieve every of its direct neighbors and it includes this data in what's called the link state packet. This link state packet is distributed to all or any the opposite nodes within the network by flooding. every node will constant link state flooding procedure and, eventually, every node can have link state packets from all alternative nodes, thus every node can have data concerning the whole topology and prices of all the links within the network. Then Dijkstra’s rule [3] is often run domestically to construct the shortest path to all or any attainable destinations. The results of this rule are often holding on within the routing tables for later use.

Reactive route discovery is additionally referred to as on-demand route discovery. Because the name implies, the route is discovered on demand. Once a supply includes a packet to be sent, it initiates a route discovery method to line up a path to the supposed destination. several approaches are often followed for path setup wherever the foremost common one has the supply node broadcast a route request packet carrying the destination address and inquiring for a route to it destination. once the route request reaches the destination or associate degree intermediate node that is aware of a route to it destination, a route reply packet is shipped back to the supply carrying details regarding the discovered route.

The class of **hybrid route discovery** is obtained by combining each the proactive and reactive techniques to form use of the benefits of each and mitigate their disadvantages. It tries to scale back the management overhead related to proactive route discovery and therefore the delay incurred within the reactive one.

B. Route Representation Data Forwarding

After choosing a route, it ought to be keeping to be followed for information transfer. We have a tendency to take into account each route illustration and information forwarding as one element as they're extremely integrated along and, in several protocols, they're done at the same time. Route illustration and information forwarding will follow one among 2 techniques: actual route and route steerage.

(a) Exact Route

In this technique, the sequence of intermediate nodes that a path ought to follow to succeed in a destination is delineated expressly. There are 2 approaches for mistreatment the precise route illustration and forwarding. These approaches are routing table and supply routing.

(b) Route Guidance

In route guidance-based protocols, the sequence of intermediate nodes isn't expressly delineated. The total path isn't determined before causing the packet by the supply, rather the trail is made on the fly (i.e., self-routing). Because the route isn't totally determined a priori, nodes cannot store data concerning the trail itself however they will store data concerning however future hop are chosen or data which will be used for choosing future hop. This can be what's known as route steerage.

C. Auxiliary Components

These parts don't seem to be essential for all routing protocols however they will be more to improve the performance of a protocol or to form it meet the wants and needs of a selected application or network paradigm. Samples of these parts are route maintenance, route energy potency, and route security.

D. Route Energy Potency

As a number of the wireless multi-hop networks are comprised of devices with restricted resources, e.g., device nodes in WSNs, such networks have energy potency as Routing for Wireless Multi-Hop Networks: Unifying options one amongst the main style issues that ought to be taken care of in any protocol designed for such networks as well as the routing ones. Routing protocols designed for such networks ought to embody mechanisms to conserve node energy to prolong the time period of the nodes and of the network as an entire. Samples of such techniques are information aggregation, use of meta-data, load reconciliation, restricted flooding, use of energy-aware metrics, use of a resource manager, and putt nodes into sleep mode.

E. Generic Routing Model

Each element is bestowed with its own numerous functionalities which will be accessible to the protocol designer to settle on from. The output and also the input of every element are shown to clarify the interactions between the varied parts.

The route discovery element has 5 options/functions for the designer to settle on from: (1) proactive with distance vector, (2) proactive with link state, (3) reactive with settled routing, (4) reactive with self-routing (which needs that every node discovers its neighbors; thus, it calls the neighbor discovery operate that feeds it with the neighbors list), and (5) hybrid discovery.

The route choice element has 3 functions for the protocol designer to settle on from: (1) source-based choice, (2) destination-based choice, and (3) intermediate-based choice. The selection of that operate to be used depends on the route discovery operate that has been chosen (e.g., the reactive self-routing discovery needs the utilization of intermediate-based route selection). Finally, the route illustration and knowledge-forwarding element has three Functions accessible for the designer's choice: (1) illustration and forwarding mistreatment actual route with routing tables, (2) illustration and forwarding mistreatment actual route with supply routing, and (3) illustration and forwarding mistreatment route steering. Again, the selection of the suitable operate strictly depends on the chosen discovery operate (e.g., the reactive self-routing discovery needs the utilization of route guidance). The subsequent pseudo-code shows the interaction and dependency of the route choice operate and also the route illustration and knowledge forwarding operate to be chosen and the already chosen discovery operates.

IV. CONCLUSION

We have studied the routing during a multi-hop network which will minimize the delay and energy consumption mistreatment mutual data. Component-based approach for breaking down a routing protocol into some core and auxiliary elements. we tend to bestowed the core elements that square measure thought-about a section of any wireless multi-hop routing protocol and are thought-about the common and unifying options of all wireless multi-hop routing schemes. The approaches like fountain code, rate less code that's used for routing purpose and metrics to formulate the network. Energy Associate in Nursing mutual data accumulation mistreatment relays is wont to notice an optimum path mistreatment print techniques and cut back communication delay.

V. REFERENCES

[1] J. L. Sobrinho, "An algebraic theory of dynamic network routing," IEEE/ACM Trans. Netw., vol. 13, no. 5, pp. 1160–1173, Oct. 2005.

[2] Y. Yang and J. Wang, "Design guidelines for routing metrics in multihop wireless networks," in Proc. IEEE INFOCOM, Apr. 2008, pp. 1615–1623.

[3] I. Maric and R. D. Yates, "Cooperative multihop broadcast for wireless networks," IEEE J. Sel. Areas Commun., vol. 22, no. 6, pp. 1080–1088, Aug. 2004.

[4] J. Chen, L. Jia, X. Liu, G. Noubir, and R. Sundaram, "Minimum energy accumulative routing in wireless networks," in Proc. IEEE INFOCOM, vol. 3. Mar. 2005, pp. 1875–1886.

[5] I. Molisch, N. Mehta, J. Yedidia, and J. Zhang, "Cooperative relay networks using fountain codes," in Proc. IEEE Global Commun. Conf. (GLOBECOM), Nov. 2006, pp. 1–6.

[6] J. Castura and Y. Mao, "Rateless coding over fading channels," IEEE Commun. Lett, vol. 10, no. 1, pp. 46–48, Jan. 2006.

[7] Z. Yang and A. Høst-Madsen, "Routing and power allocation in asynchronous Gaussian multiple-relay channels," EURASIP J. Wireless Commun. Netw., vol. 2006, no. 2, p. 35, 2006.

[8] R. Urgaonkar and M. J. Neely, "Optimal routing with mutual information accumulation in wireless networks," IEEE J. Sel. Areas Commun., vol. 30, no. 9, pp. 1730–1737, Oct. 2012.

[9] R. Yim, N. Mehta, A. F. Molisch, and J. Zhang, "Progressive accumulative routing in wireless networks," in Proc. IEEE Global

[10] Commun. Conf. (GLOBECOM), Nov. 2006, pp. 1–6. A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," IEEE Trans. Inf. Theory, vol. 57, no. 4, pp. 1872–1905, Apr. 2011

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3 | ISSN : 2456-3307

An Overview of Artificial Intelligence

A. Akilambigai[1], K.Vijayashanthi[2]
[12]Department of Computer Science, Kamban Arts and Science College for Women, Thiruvannamalai, Tamil Nadu, India

ABSTRACT

Artificial Intelligence (AI) is also Machine Intelligence (MI) is intelligence demonstrated by machine in contrast to the “Natural Intelligence” (NI) displayed by human and other animals. In computer science, all research is defined as the study of “Intelligent agent”: any device that perceives its environment and takes actions that maximize its change of successfully achieving its goal .The “Artificial Intelligence” is applied when a machine mimics “cognitive” function that human associate with other human minds such as “learning and problem solving”. AI is behavior of a machine, which, if performed by a human being, would be called intelligent. It makes machines smarter and more useful, and is less expensive than natural intelligence. Natural language processing (NLP) refers to artificial intelligence methods of communicating with a computer in a natural language like English. The main objective of a NLP program is to understand input and initiate action. Artificial intelligence involves two basic ideas. First, it involves studying the thought processes of human beings. Second, it deals with representing those processes via machines (like computers, robots, etc.).

Keywords: Artificial Intelligence, Machine Intelligence, Natural Language Processing

I. INTRODUCTION

Artificial intelligence is a branch of computer science that studies the computational requirements for tasks such as perception, reasoning and learning and develop systems to perform tasks.

Intelligence is the computational part of ability to achieve goals in the world. Varying kinds and degree of intelligence occur in people, many animals and some machine. It was founded as an academic discipline in 1956,and in the year since has experienced several waves of optimism, followed by disappointment and the loss of funding called as an AI winter.AI research has been divided into subfields that often fail to communicate with each other. These sub-fields are based on technical considerations such as particular goals (eg. ”robotics” or “machine learning”), the use of particular tools (“logic” or “neural networks”), or deep philosophical differences. Subfields have also been based on social factors (particular institutions or the work of particular researchers).

The traditional problem of AI research include reasoning, knowledge, planning, learning, natural language processing, perception and the ability to move and manipulate objects. General intelligence is among the field’s long term goals. The AI draws upon computer science, mathematics, psychology, linguistics, philosophy, neuroscience, artificial psychology and many others. It uses many tools including neural networks, search and mathematical optimization, method based on statistics, probability and economics. AI become an essential part of the technology industry, helping to solve many challenging problems in computer science

II. HISTORY OF ARTIFICIAL INTELLIGENCE

The birth of Artificial intelligence is at Dartmouth conference 1956. The proposal for the conference included this assertion: "every aspect of learning or any other feature of intelligence can be so precisely described that a machine can be made to simulate it.

The golden years started at 1956 and ended at 1974. There were many successful programs and new directions: Natural language understanding (first AI program to use a semantic net).Micro-world, Neural networks

The first AI winter at 1974 to 1880. AI was subject to critiques and financial setbacks. AI researchers had failed to appreciate the difficulty of the problems they faced. Their tremendous optimism had raised expectations impossibly high and when the promised results failed to materialize funding for AI disappeared.

Boom: In 1980s a form of AI program called "experts systems" was adapted by corporations around the world and knowledge became the focus of mainstream AI research, In those same years, the Japanese government aggressively funded AI with its fifth generation computer project.

The second AI winter at 1987 to 1993. The business community's fascination with AI rose and fell in the 80s in the classic pattern of an economic bubble. The collapse was in the perception of AI by government agencies and investors. The field continued to make advance despite the criticism.

According to Bloomberg's Jack clark,2015 was a landmark year for artificial intelligence, with the number of software projects that use AI within Google increased from a "sporadic usage" in 2012 to more than 2,700 projects. Clark also presents factual data indicating that error rates in image processing tasks have fallen significantly since 2011. He attributes this to an increase in affordable neural networks, due to rise in cloud computing infrastructure and to an increase in research tools and datasets.

III. TOOLS USED IN ARTIFICIAL INTELLIGENCE

Artificial intelligence has developed a large number of tools to solve the most difficult problems in computer science. A few of the most general of these methods are:

- ✓ Search and optimization
- ✓ Logic
- ✓ Probabilistic methods for uncertain reasoning

A. Search and Optimization

Many problems in AI can be solved in theory by intelligently searching through many possible solutions: Reasoning can be reduced to performing a search. Planning algorithm search through trees of goals and sub goals, attempting to find a path to a target goal, a process called means end analysis. Heuristics supply the program with a best guess for the path on which the solution lies. Heuristics limit the search for solutions into a smaller sample size.

A very different kind of search came to prominence in the 1990s, based on the mathematical theory of optimization. For many problems, it is possible to begin the search with some form of a guess and then refine the guess incrementally until no more refinements can be made. Evolutionary computation uses a form of optimization search. Example that may begin with a population of organisms and then allow them to mutate and recombine, selecting only the fittest to survive each generation. Forms of evolutionary computation include swarm intelligence algorithms such as ant colony or particle swarm optimization) and evolutionary algorithms such as genetic algorithms, gene expression programming and genetic programming.

B. Logic

Logic is used for knowledge representation and problems solving, but it can be applied to other problems as well. Example, the satplan algorithm uses logic for planning and inductive logic programming is a method for learning.

Several different forms of logic are used in AI research. Propositional or Sentential logic of statements, which can be true or false. First logic also allows the use of quantifiers and predicates and can express facts about objects, their properties and their relations with each other. Default logics, non-monotonic logics and circumscription are forms of logic designed to help with default reasoning and the qualification problem. Several extensions of logic have been designed to handle specific domains of knowledge, such as description logics; situation calculus, event calculus and fluent calculus (for representing events and time); casual calculus; brief calculus; and modal logics.

C. Probabilistic Methods For Uncertain Reasoning

AI researchers have devised a number of powerful tools to solve these problems using methods from probability theory and economics. Bayesian networks are a very general tools that can be used for a large number of problems. The key concept from the science of economics is “utility” a measure of how valuable something is to an intelligent agent. Precise mathematical tools have been developed that analyze how an agent can make choices and plan, using decision theory, decision analysis, and information value theory.

IV. APPROACHES

AI research into three approaches, which he calls computational psychology, computational philosophy, and computer science. Computational psychology is used to make computer programs that mimic human behaviour. Computational philosophy is used to develop an adaptive, free flowing computer mind. Implementing computer science serves the goal of creating computers that can perform tasks that only accomplish. Some of the approaches are,

- ✓ Cybernetics and brain simulation
- ✓ Cognitive simulation
- ✓ Knowledge based

A. Cybernetics And Brain Simulations

In the 1940s and 1950s, a number of researchers explored the connection between neurobiology, information theory, and cybernetics. Some of them built machines that are used electronic networks to exhibit rudimentary intelligence. Many of these researchers gathered for meetings of the teleological society at Princeton University and the ratio club in England. By 1960, this approach was largely abandoned, although elements of it would be revived in the 1980s

B. Cognitive Simulation

Economist Herbert Simon and Allen newell studied human problem solving skills, and attempted to formalize them, and their work laid the foundation of the field of artificial intelligence as well as cognitive science, operation research and management science. Their research team used the results of psychological experiments to develop program the simulated the techniques that people used to solve problems. This tradition, centered at Carnegie Mellon University would eventually culminate in the development of the soar architecture in the middle 1980s.

C. Knowledge Based

When computers with large memories became available around 1970, researchers from all three traditions began to build knowledge into AI applications. The “knowledge revolution” led to the development and deployment of expert systems. The first truly successful form of AI software. The knowledge revolution was also driven by the realization that enormous amounts of knowledge would be required by many simple AI applications.

V. APPLICATION USED FOR ARTIFICIAL INTELLIGENCE

AI is relevant to any intellectual task. Modern artificial intelligence techniques are pervasive and are too numerous to list here. Frequently, when a technique reaches mainstream use, it is no longer considered artificial intelligence; this phenomenon is described as the AI effect.

High profile examples of AI include autonomous vehicles such as drones and self-driving cars, medical diagnosis, creating art such as poetry, proving mathematical theorems, playing games such as chess or Go, search engines such as Google search, online assistants such as siri, image recognition in photographs, spam filtering, prediction of judicial decisions and targeting online advertisements.

- ✓ Competitions and prizes
- ✓ Healthcare
- ✓ Finance and economics
- ✓ Automotive
- ✓ Music

Finance and Economics

Bank use artificial intelligence to:

- ✓ Organize operations
- ✓ Invest in stocks
- ✓ Manage properties

Financial instutions have used AI to detect charges or claims.

Music

Composition, performance, music theory, sound processing are some of the major areas on which research in music and AI are focusing.

VI. PLATFORMS

A platform or computing platform is defined as "some sort of hardware architecture or software framework that allows software to run". As Rodney brooks pointed out many years ago, it is not just the artificial intelligence software that defines the AI features of the platform, but rather the actual platform itself that affects the AI that results, i.e., there needs to be work in AI problems on real world platforms rather than in isolation.

A wide variety of platform has allowed different aspects of AI to develop, ranging from expert systems such as Cyc to deep learning frameworks to robot platforms such as the roomba with the open interface. Recent advances in deep artificial neural networks are distributed computing have led to a proliferation of software libraries, including deep learning, tensor Flow, theano and torch.

VII. CONCLUSION

Although a conclusion may review the main points of the paper, do not replicate the abstract as the conclusion. A conclusion might elaborate on the importance of the work or suggest applications and extensions. Authors are strongly encouraged not to call out multiple figures or tables in the conclusion these should be referenced in the body of the paper.

VIII. ADVANTAGES

With AI, the changes of error are almost nil and greater precision and accuracy is achieved.AI finds application in space explorations. Intelligent robots can be used to explore space. They are machines and hence have the ability to endure the hostile environment of the interplanetary space.

AI has made daily life a lot easier like by use of applications or phones or computers that predict user actions and also make recommendations that suit user's choice e.g. application such as GPS, and maps application etc. Intelligent machines can replace human beings in many areas of work. Robots can do certain laborious tasks. Painstaking activities, which have long been carried out by human can be taken over by the robots.

IX. PROBLEMS OF ARTIFICIAL INTELLIGENCE

The overall research goal of artificial intelligence is to create technology that allows computer and machines to function in an intelligent manner. The general problem of simulating or creating intelligence has been broken down into sub problems. These consist of particular traits or capabilities that researcher except an intelligent system to display. The traits described below have received the most attention.

- ✓ Reasoning and problem solving
- ✓ Knowledge representation

A. Reasoning And Problem Solving

Human beings ordinarily use fast, intuitive judgements rather than step by step deduction that early AI research was able to model. AI progressed using "sub-symbolic" problem solving: embodied agent approaches emphasize the importance of sensorimotor skills to higher reasoning; neural net research attempts to stimulate the structure inside the brain rise to this skill; statistical approaches of AI mimic the human ability to guess

B. Knowledge Representation

Knowledge representation are central to AI research. Many of the problems machines are expected to solve will require extensive knowledge about the world. Among the things that AI needs to represents the object, properties, categories, and relations between objects, states, events and time, causes and effects , knowledge about knowledge and the domain. Knowledge representation is suitable for content based indexing and retrieval scene interpretation, clinical decision support, knowledge discovery via automated reasoning.

X. CONCLUSION

AI and the technology are one side of the life that always interest and surprise us with the new ideas, topics, innovations, products etc. AI is still not implemented as the films representing it; however, there are many important tries to reach the level and to compete in market, like sometimes the robots that they show in TV. AI is at the center of a new enterprise to build computational models of intelligence. The main assumption is that intelligence can be represented in terms of symbol structures and symbolic operations which can be programmed in a digital computer. Conventional digital computers may be capable of running such programs, or we may need to develop new machines that can support the complexity of human thought.

XI. REFERENCES

[1] Russell, Stuart J.; Norvig, Peter (2003).Artificial intelligence: A Modern Approach(2nd ed.). Upper Saddle River, New Jersey: Prentice Hall. ISBN 0 -13-790395-2

[2] Berglas, Antony(2008), Artificial intelligence will kill our grand children retrived 2008-06-13

[3] Russell, Stuart J.; Norvig, Peter (2009).Artificial intelligence: A Modern Approach(3rd ed.). Upper Saddle River, New Jersey: Prentice Hall. ISBN 0 -13-604259-7..

[4] Gelenter, david(2010), Dream logic, Internet and artificial thought retrieved 25 july 2010.

[5] Poole, David; Mackworth, Alan (2017). Artificial intelligence: Foundations of computational Agents (2nd ed.).Cambridge university press. ISBN 9781107195394.

[6] Neapolitan, Richard; jiang, xia (2018).Artificial intelligence: with an introduction to machine learning. Chapman & Hall/CRC .ISBN 978-1-13850-238-3

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3 | ISSN : 2456-3307

Distributed Intrusion Detection Using Mobile Agents

F. Asmathunnissa[1], A.Lavanya[2]
[1]M.Phil (Research scholar), Kamban College of Arts And Science For Women, Thiruvannmalai, Tamil Nadu, India

ABSTRACT

DIDMA (Distributed Intrusion Detection using Mobile Agents) is a novel architecture in the field of IDS (Intrusion Detection Systems), utilizing an agent-based approach in order to realize a distributed framework. The novelty in this architecture is the employment of mobile agents as its auditing components. This novel approach overcomes certain problems associated with traditional designs in IDS. In particular, problematic areas such as high-speed networks, not visible traffic, and fail-open architecture have been successfully managed. Moreover, the fault tolerant decentralized design of DIDMA clearly demonstrated resilience against active attacks.

Keywords: Intrusion Detection, Network Security, Distributed Networking, Mobile Agents

I. INTRODUCTION

The security of a local networking environment, or a single computer system, is a very important aspect of its infrastructure. A properly secured system is able to guarantee confidentiality, integrity, and availability of its resources and services. Thus, the ability to protect a system from outside interference is of the highest importance. However, individuals can specifically attack flaws in computer systems. These attacks result in rendering the system vulnerable, and compromising its entire security scheme. One aspect of the general security scheme is to detect if and when an attack against protected resources is attempted. The topic of network security that deals with this field is IDS [1]. The intrusion detection technology, employed in order to monitor resources distributed among several nodes in a local network, adopts a distributed approach to its design. However, contemporary IDS still suffer from many problems, like the inability to handle large amounts of network traffic, not visible network traffic, and a central monitor that provides a single point of failure, among others. The DIDMA prototype proposes a novel architecture that approaches the aforementioned problems by employing a distributed solution using mobile agents [2]. The components of the system are realized as software agents that have certain properties, and provide several advantages over traditional intrusion detection approaches. These software agents are its auditor components, and are implemented by using the Grasshopper platform [2]. Handling the information collected by an intrusion detection system requires extensive cross referencing in order to identify distributed attacks. Therefore, there is an obvious need for an automated distributed approach to the design procedure of detecting security violations. The rest of this paper is organized as follows. The second section presents the current framework of IDS and the associated problems, along with a brief introduction to the mobile agents technology. Moreover, related work is presented in the second section. The third section analyzes the DIDMA architecture. The results of the proposed design and possible future work are presented in the conclusion.

II. CURRENT FRAMEWORK

A. IDS and Existing Problems

Intrusion detection is a type of network security that aims to detect, identify, and isolate attempts of intrusion, or unauthorized usage of computer and network resources [1]. There are some problems that infest the traditional design approaches. IDS are not able to reliably intercept and examine the *high-speed traffic* of contemporary networking environments. The speed of networks increases at such a rate that IDS require many resources and computational power. Although the development of the IDS technology will continue to improve, likewise the speed of networks will continue to grow. Another problem closely related to the previous one is *not visible traffic*. The currently used high speed networks employ switched topologies in order to increase their speed. Although this approach has many benefits for the performance of the network, it interferes with the operation of IDS, because it distributes the traffic in ways that make it invisible to certain parts of the network. Furthermore, the employment of Virtual Private Networks (VPN) that encrypt the exchanged network traffic also creates monitoring problems for IDS. Consequently, a system deployed on such networks is unable to perform correctly. Nearly every intrusion detection system has the problem of *fail-open architecture*. If an attacker finds a way to disable the system, usually through the use of denial of service type attacks, then the previously monitored network becomes completely unprotected. Moreover, the system that failed provides absolutely no notification about its inability to continue providing detection services [3]. The *inability to detect certain types of attacks* is another common problem with IDS. Detection approaches are based on fixed, constant models that are matched against known attack patterns. Research on the field of detection approaches has proved that attacks which achieve the same results, but are slightly different than the ones expected by the system, are not detected [3]. Finally, one of the most important problems of IDS is their *high-rate susceptibility to false positives*. A false positive can be defined as the situation of threats that appear to be real to the system, but in reality they are just normal data transactions. This vulnerability of IDS is responsible for many problems, as an attacker can create many false positives, and when the system is reconfigured to ignore such patterns, he can launch the real attack undetected. Moreover, if the system is configured to respond aggressively to detected attacks, then a reaction to a false positive situation results to interference with legitimate operations.

B. Mobile Agents Technology

Agents are defined as software programs that are situated in an execution environment and are characterized by the property of autonomy [2]. The property of mobility constitutes one of the most useful characteristics of agents, though not necessarily of intelligent ones, since it can prove to be advantageous in the networked environments that exist today. Moreover, mobile agents (MA) offer a cleanly designed solution to the problem of distribution of a specific task over several entities that work together to achieve a common objective. The mobile agents' technology provides a development environment that can be utilized by designers in order to produce robust distributed applications that are able to operate efficiently in the contemporary wide area networks that are characterized by their dynamic nature. An investigation of the mobile agents technology reveals many advantages over the established distributed networking models.

C. Related Work

The problem of designing an IDS that follows the distributed networking paradigm has been researched in the past, and several prototype systems have been implemented. However, the technology of MA has not been very widely utilized in the intrusion detection field, and generally in the field of network security. The reason is basically the fact that this is a rather new approach to distributed computing and has not yet found many applications. The Autonomous Agents for Intrusion Detection (AAFID) project has approached the problem of intrusion detection with an agent-based solution. The project defines autonomous agents as independent software entities that are responsible for a specific monitoring task at a single host [5]. The proposed design of the system follows the distributed paradigm since it is based on multiple independent entities that work collectively and share information. The collected findings are reported to a central monitor that processes the data and is able to detect intrusions and intrusion attempts [5]. Although the system utilizes a distributed philosophy and design, it does not employ mobile agents. The agents of the system do not have the ability to migrate to other systems, but are stationed to a single host. Furthermore, the central monitor is a single point of failure, and can be attacked by a malicious party in order to render the system unusable. Another project that aimed to decentralize intrusion detection is the Distributed Intrusion Detection System (DIDS). The system employed several sensors in a restricted networking environment that monitored host and traffic events. Although the system distributed the monitoring to several different places in the network, there existed a single centralized director that received information from the sensors and reported intrusions [6]. Moreover, the developed system employed no agent-based technology or methodology. Thus, it was unable to receive several benefits of such an approach, like a true decentralized design that can robustly withstand an attack, and successfully overcome the problem of fail-open architecture.

A distributed administration system, named Sun Enterprise Network Security Service (SENSS), was recently released by Sun Microsystems. The system performs distributed administration functions in a local network environment by employing Java-based mobile agents. These agents are launched from a central location and they migrate to hosts in the network performing predefined administrative tasks [7]. It must be noted that the system was not developed to perform intrusion detection functions, thus it has certain limitations when it is employed to perform security-auditing procedures. The mobile agents it utilizes have no autonomy and they are able to travel only to a single host, not having the ability to migrate again. This fact proves to be a shortcoming since non-autonomous agents depend on the parent process, thus their functionality is reduced in the case of an unanticipated problem. The design process of the DIDMA system proposes a framework upon which further functionality can be added and implemented.

III. SYSTEM DESIGN

This approach has the benefit of abstracting and breaking down the problem into several parts that can be managed with greater ease. Thus, a simple blueprint of the design of the system can prove to be very helpful in expanding its basic functionality. Figure 1: Simple representation of the DIDMA system the system, at its most basic level, defines two different kinds of agents. The first are stationary agents that are situated at specific hosts in the local area network that must be monitored for intrusion attempts. The system refers to these entities as sensor agents (SA). The SA are responsible for collecting information about the specific host on which they are stationed. The data are collected from the log files of the system, as well as from operational statistics and events directly from the system. The collected data are processed by the SA and rendered into a common predefined format. It must be noted that the SA are not responsible for classifying events as intrusion attempts, nor are they able to alert the system of such an event. The party responsible for these activities is the second kind of agent defined by the architecture. The second kind of agents that exist in the proposed system are multi-hop mobile agents. These agents have the ability to transport themselves between the hosts that are under the surveillance of the intrusion detection system. They are responsible for inspecting and evaluating the information collected by the sensor agents. Depending on their predefined sensitivity, they can classify whether an event has been an intrusion attempt. When a MA classifies an event as an attack, it has the ability to alert the system. The proposed system is also consisting of a configuration console that provides a graphical user interface in order to interact with the party that has the responsibility of administrating the monitored network. The auditor agent configuration console (AAconf) provides the functionality of configuring certain run-time variables of the auditor agents and launching them to the network environment.

A. System Components

a. Sensor Agents

The SA is the sensor component of the developed intrusion detection system. The SA is a stationary agent that is bound to a specific host that is under the surveillance of the intrusion detection system. The SA is accompanied by a file, the attack patterns signature file that contains patterns of known attacks. The SA against the log files of the hosting system periodically checks these signatures. Any matching entries in the log files are extracted by the SA, formatted according to predefined format, and placed in the audit file. Therefore, the SA must be aware of the log files of the hosting system and the format utilized by those. The SA is the only component of the system that is platform dependent. If a new operating system needs to be supported and the corresponding host placed under the monitoring of the system, a SA that specifically understands the log files generated by this operating system should be designed. Moreover, this SA should be accompanied by an attack pattern signature file that contains known attacks related to the hosting operating system. Another detection approach that can be accommodated by the sensor agent is the incorporation of legacy, non-agent oriented, intrusion detection software [8]. A network-based intrusion detection system, like Snort [9], can be wrapped by an agent interface and monitor network activity for known attack signatures. The SA wrapper performs the translation between external agent requests to the legacy code, and between the legacy system's requests to the agent communication protocol [8].

b. Auditor Agents

The auditor agent subsystem is in essence the auditor component of DIDMA. The AA component is realized through a mobile agent that visits each host that is protected by the system and checks whether there has been a security violation that needs to be reported. The mobile AA knows from the moment that is launched the list of Internet Protocol (IP) addresses that correspond to the hosts that must be audited. The choice of the host that is visited next is not random, but the agent implements the oldest-node algorithm, that specifies that the agent should visit the host it last visited longest ago (or never visited, or does not remember visiting)[10]. When the mobile auditor agent visits a host, it accesses the audit file that has stored the events that should be evaluated. The agent evaluates these events, and according to its specified sensitivity it generates an alert or not. The alert methods that the AA supports are three. A generated alert can be sent via an e-mail message by utilizing the predefined destination e-mail address and a Simple Mail Transfer Protocol (SMTP) host. Another way of alerting is via a Short Message Service (SMS) text that is sent to a predefined mobile terminal number through an SMS center number. Finally, the AA can alert by sending a Transmission Control Protocol (TCP) message, which contains the text of the alert, to a predefined host. The format of the message follows the Intrusion Alert Protocol (IAP) which is an application level protocol for exchanging intrusion alert information. The protocol provides transport and security characteristics that are required in order to exchange alert data over insecure channels [11].

c. Auditor Agent Configuration Console

The auditor agent configuration console component is the subsystem that is mainly responsible for the interaction with the user, and the configuration, creation, and launching of the mobile auditor agents. It can be installed at any host of the internal local area network, and after the creation of the mobile auditor agents it can be safely ignored, even completely uninstalled. The user installs this component, and through the provided functionality the mobile auditor agents are configured. After the definition of the configuration parameters, the mobile agents are created and launched in the local area network. Since the AA conf subsystem plays no other role, nor it implements any other required operations, the system has absolutely no dependence to it. Therefore, the AA conf subsystem can be deactivated, or even uninstalled, without interrupting the operation of DIDMA. The absence of a central management component confers a distributed nature to the developed intrusion detection system. Thus, the system has no single point of failure which could be attacked by an outsider and disable it.

B. The Audit File

The communication between the stationary sensor agent of a specific monitored host and the visiting mobile auditor agent is realized through the audit file. The file should follow a predefined format in order achieve interoperability with other systems. Since the developed intrusion detection system employs agent-oriented software engineering approaches, the format of the audit file should be compliant to both intrusion detection and software agent standards. This fact guarantees interoperability with other intrusion detection systems that could be employed in a local area network, and also with other agents that belong to other development platforms and could utilize the resources of the present intrusion detection system in order to satisfy their own design requirements. Therefore, it gains all the advantages of the XML meta-language, like filtering and aggregation, and becomes easily extensible in order to be utilized by third party applications. Furthermore, the Knowledge Query and Manipulation Language (KQML) defines a standardized language for exchanging information and knowledge between agents. The main focus of KQML is to describe an extensible set of performatives that can be utilized in order to define the permissible operations that agents may attempt on the knowledge stores of each other [13]. Moreover, KQML provides a basic design architecture for knowledge sharing through agents. These communication facilitators coordinate the interaction and the exchange of knowledge and information between software agents [13]. The proposed format of the audit file follows a format that embeds the IDMEF in the KQML general directives. This is possible since KQML places no restriction in regard to the language of its messages, therefore IDMEF can be utilized.

C. System Implementation

A prototype of the DIDMA architecture has been implemented by utilizing the Java programming language, and the Grasshopper mobile agent platform [2, 14]. The Java programming language has been selected as the implementation language for its platform independency, and its strong security framework. Security features of Java, such as the automatic bounds checking, and the code verification signing, form a safety net for critical applications.

The Grasshopper mobile agent platform was chosen since it is readily and freely available for non-commercial endeavours. Grasshopper provides multi-protocol support, a naming service for migrating agents, advanced security features based on strong cryptographic algorithms, and a rich, fully documented interface for the creation of mobile, or other, software agents.

IV. EVALUATION

The approach of employing a distributed design based on MA for the infrastructure of an intrusion detection system, offers certain advantages over the traditional monolithic systems. These advantages form the motivation of DIDMA, the proposed architecture. The most important of these benefits addresses the problem of network traffic that is not visible to a traditional intrusion detection system, due to switched network environments. However, an intrusion detection system that employs the proposed design for the auditing process is able to monitor every host in the local network, despite the existing topology, since the mobile agents are free to travel to all nodes. Mobile agents can successfully resolve the inability of IDS to monitor high-speed networks. This central point of control is in essence a single point of failure, and constitutes the most probable target for an outside attacker. A distributed approach to the design of an intrusion detection system avoids this threat, since there are many mobile agents in the network that can alert interested parties if an attack on a node is detected. Although the proposed system defines a configuration management console, this is strictly for demonstration purposes. After launching the agents in the network, the configuration console can be completely disabled and the system will continue to operate normally. Furthermore, the proposed architecture can overcome the problem of false positives with its modular design. The mobile agents perform the auditing process on every host of the network and are able to cross-reference their findings. Thus, distributed attack patterns can be detected, and legal transactions can be confirmed. Moreover, the auditor agents can be configured to certain degrees of sensitivity dynamically, therefore the system is able to reduce the rate of false positives even further.

V. CONCLUSIONS AND FUTURE WORK

In this paper, we have presented DIDMA, a novel distributed architecture, upon which a prototype has been implemented. This implementation offers certain benefits to the intrusion detection methodology, and complements the traditional IDS technology. The detailed investigation that has been performed confirmed advantages of DIDMA over traditional IDS design approaches. The developed prototype was evaluated based on operational criteria, like the response time between an attempted intrusion and the actual alert, and the computational load on the hosting resources. Moreover, the fault tolerant architecture of DIDMA proved to be beneficial on active attacks against the system itself. The DIDMA architecture does not yet address the aspect of communication between the auditor mobile agents. Future work will be focused on the communication mechanisms between the auditor agents, in order to be able to recognize attack trails that spread over many monitored systems by cross-referencing their findings. Another area of possible further work is the realization of a minimal mobile agent platform that provides just the functionality and security mechanisms required by the DIDMA system. The advantages of implementing a specialized platform would be the low overhead of the system, and a distributed framework under the complete control of the intrusion detection system.

VI. REFERENCES

[1] J. McHugh, A. Christie, J. Allen, Defending Yourself: The Role of Intrusion Detection Systems, IEEE Software Magazine, Vol.17, No.5, 2000.

[2] W. R. Cockayne, M. Zyda, Mobile Agents, Prentice Hall, 1998.

[3] S. Northcutt, Network Intrusion Detection: An Analysts' Handbook, Second Edition, New Riders Publishing, 1999.

[4] J. J. Ordille, When Agents Roam, Who Can You Trust?, Bell Labs Computing Science Research Center

[5] J.S Balasubramaniyan, J. O. Garcia-Fernandez,D. Isacoff, E. Spafford, D. Zamboni, An Architecture for Intrusion Detection usingAutonomous Agents, CERIAS Technical Report 98/05, Purdue University, 1998.

[6] B. Mukherjee, T. L. Heberlein, K. N. Levitt, Network Intrusion Detection, IEEE Network Magazine, Vol.8, No.3, 1994.

[7] Sun Microsystems Inc., Sun EnterpriseNetwork Security Service

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3 | ISSN: 2456-3307

Neural Networks

Andrews Bernard K

PG Research, Department of Computer Science, Shanmuga Industries Arts & Science College Thiruvannamalai, Tamilnadu, India

ABSTRACT

Artificial neural networks have emerged from the studies of how brain performs. The human brain consists of many millions of individual processing elements, called neurons that are highly interconnected. Information from the outputs of the neurons, in the form of electric pulses is received by the cells at connections called synapses. The synapses connect to the cell inputs, or dendrites and the single output the neuron appears at the axon. An electric pulse is sent down the axon when the total input stimuli for all of the dendrites exceed a certain threshold. Artificial neural networks are made up of simplified individual models of the biological neuron that are connected together to form a network. Information is stored in the network in the form of weights or different connections strengths associated with synapses in the artificial neuron models.

Keywords: Human Brain, Neural Networks, ANN

I. INTRODUCTION

The importance of electricity in our day to day life has reached such a stage that it is very important to protect the power system equipments from damage and to ensure maximum continuity of supply. But there are power system blackouts by which the continuous power supply is being interrupted. What is more important in the case of a blackout is the rapidity with which the service is restored. Now- a -days power system blackouts are rare. But whenever they occur, the effect on commerce, industry and everyday life of general population can be quite severe. In order to reduce the social and economic cost of power system blackouts, many of the electric utility companies have pre-established guidelines and operating procedures to restore the power system. They contain sequential restoration steps that an operator should follow in order to restore the power system. They are based on certain assumptions whichmay not be present in the actual case. This reduces the success rates of these procedures.

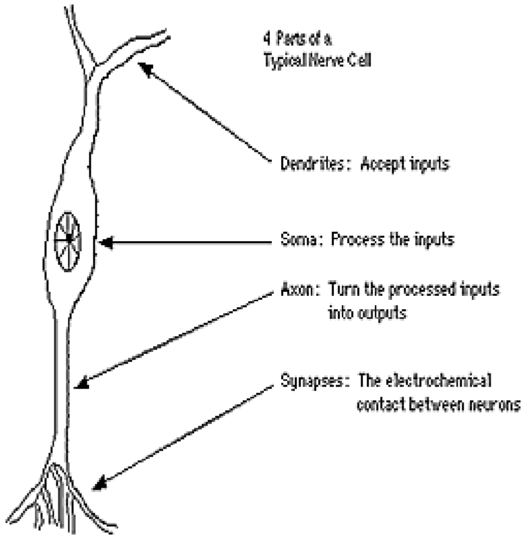
II. WHAT ARE ANNs?

Artificial Neural Network (ANN) is a system loosely modeled on human brain. It tries to obtain a performance similar to that of human's performance while solving problems. As a computational system it is made up of a large number of simple and highly interconnected processing elements which process information by its dynamic state response to external inputs. Computational elements in ANN are non-linear and so the results come out through non-linearity can be more accurate than other methods. These non-linear computational elements will be working in unison to solve specific problems. ANN is configured for specific applications such as data classification or pattern recognition through a learning process. Learning involves adjustment of synaptic connections that exist between neurons. ANN can be simulated within specialized hardware or sophisticated software. ANNs are implemented as software packages in computer or being used to incorporate Artificial Intelligence in control systems.

III. BIOLOGICAL NEURON

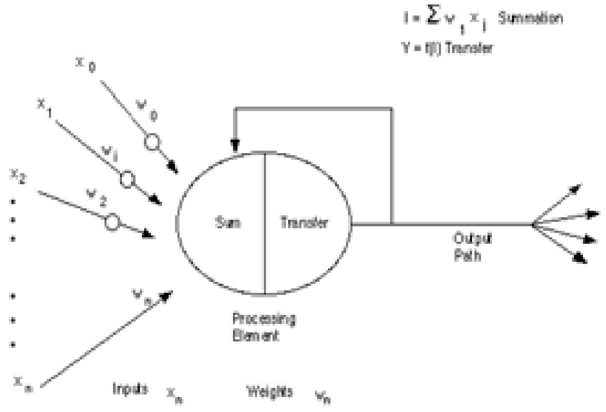
The most basic element of the human brain is a specific type of cell, which provides us with the abilities to remember, think, and apply previous experiences to our every action. These cells are known as neurons, each of these neurons can connect with up to 200000 other neurons. The power of brain comes from the numbers of these basic components and the multiple connections between them.

All natural neurons have four basic components, which are dendrites, soma, axon and synapses. Basically, a biological neuron receives inputs from other sources, combines them in some way, performs a generally non-linear operation on the result, and then output the final result. The figure below shows a simplified biological neuron and the relationship of its four components.



IV. ARTIFICIAL NEURON

The basic unit of neural networks, the artificial neurons, simulates the four basic functions of natural neurons. Artificial neurons are much simpler than the biological neurons. The figure below shows the basic structure of an artificial neuron.



Note that various inputs to the network are represented by the mathematical symbol, x(n). Each of these inputs are multiplied by a connection weight, these weights are represented by w(n). In the simplest case, these products are simply summed, fed through a transfer function to generate a result, and then output. Even though all artificial neural networks are constructed from this basic building blocks the fundamentals may vary in these building blocks and there are differences.

V. NEURAL NETWORKS

Artificial neural networks emerged from the studies of how brain performs. The human brain consists of many million of individual processing elements called neurons that are highly interconnected.

ANNs are made up of simplified individual models of the biological neurons that are connected together to form a network. Information is stored in the network in the form of weights or different connection strengths associated with the synapses in the artificial neuron models.

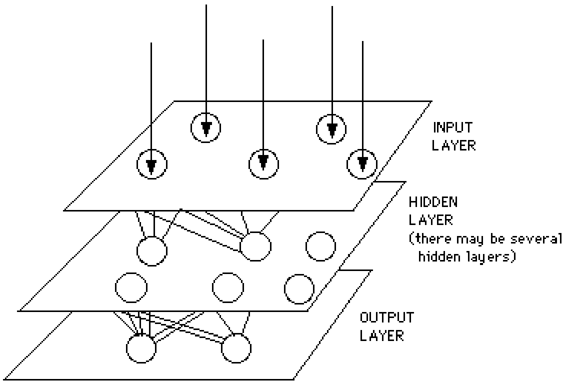
Many different types of neural networks are available and multilayered neural network are the most popular which are extremely successful in pattern reorganization problems. An artificial neuron is shown in the figure. Each neuron input is weighted by Changing the weights of an element will alter the behavior of the whole network. The output y is obtained summing the weighted inputs and passing the result through a non-linear activation function.

VI. PROCEDURE FOR ANN SYSTEM DESIGN

In realistic application the design of ANNs is complex, usually an iterative and interactive task. The developer must go through a period of trial and error in the design decisions before coming up with a satisfactory design. The design issues in neural network are complex and are the major concerns of system developers.

Designing of a neural network consists of:

- Arranging neurons in various layers.
- Deciding the type of connection among neurons of different layers, as well as among the neurons within a layer.
- Deciding the way neurons receive input and produces output.
- Determining the strength of connection that exists within the network by allowing the neurons learn the appropriate values of connection weights by using a training data set.



As the figure above shows, the neurons are grouped into layers. The input layer consists of neurons that receive input from external environment. The output layer consists of neurons that communicate the output of the system to the user or external environment. There are usually a number of hidden layers between these two layers. The figure above shows a simple structure with only one hidden layer.

When the input layer receives the input, its neurons produces output, which become input to the other layers of the system. The process continues until certain condition is satisfied or until the output layer is invoked and fire their output to the external environment.

VII. LEARNING TECHNIQUES

Learning rules are algorithm for slowly alerting the connections weighs to achieve a desirable goal such a minimization of an error function. The generalized step for any neural network leaning algorithm is follows are the commonly used learning algorithm for neural networks.

- Multi-Layer Neural Net (MLNN)
- Error Back Propagation (EBB)
- Radial Basis Functions (RBF
- Reinforcement Learning
- Temporal Deference Learning
- Adaptive Resonance Theory (ART)
- Genetic Algorithm

Selection of a particular learning algorithm depends on the network and network topology. As MLNN with EBP is most extensively used and widely accepted network for process application, namely for identification and control of the process.

VIII. FEATURES OF ANNs

ANNS have several attractive features:

Their ability to represent non-linear relations makes them well suited for non-linear modeling in control systems.

- Adaptation and learning in uncertain system through off line and on line weight adaptation.
- Parallel processing architecture allows fast processing for large-scale dynamic system.
- Neural network can handle large number of inputs and can have many outputs.
- ANNs can store knowledge in a distributed fashion and consequently have a high fault tolerance.

An ANN can been seen as a union of simple processing units, based on neurons that are linked to each other through connections similar to synapses. These connections contain the "knowledge" of the network and the pattern of connectivity express the objects represented in the network. The knowledge of the network is acquired through a learning process where the connections between processing elements is varied through weight changes.

Learning rules are algorithms for slowly alerting the connection weights to achieve a desired goal such as minimization of an error function. Learning algorithms used to train ANNs can be supervised or unsupervised. In supervised learning algorithms, input/output pairs are furnished and the connection weights are adjusted with respect to the error between the desired and obtained output. In unsupervised learning algorithms, the ANN will map an input set in a state space by automatically changing its weight connections. Supervised learning algorithms arecommonly used in engineering processes because they can guarantee the output.

In this power system restoration scheme, a multilayered perceptron(MLP) was used and trained with a supervised learning algorithm called back-propagation. A MLP consists of several layers of processing units that compute a nonlinear function of the internal product of the weighted input patterns. These types of network can deal with nonlinear relations between the variables; however, the existence of more than one layer makes the weight adjustment process for problem solution difficult.

IX. ANN BASED CONTROL CONFIGURATION

- Direct Inverse Control
- Direct Adaptive Control
- Indirect Adaptive Control
- Internal Model Control
- Model Reference Adaptive Control

X. ADAPTIVE CONTROL

The neural network approximates a wide variety of nonlinear control laws by adjusting the weights in training to achieve the desired approximate accuracy. One possible MRAC structure based on neural network is shown:

- In this configuration, control systems attempted to make the plant output YP (t) to follow the reference model output asymptotically. The error signal
- Used to train the neural network controller is the difference between the model and the plant outputs, principally; this network works like the direct adaptive neural control system.

XI. CONCLUSION

PSR has become a field of growing interest. Several techniques based on artificial intelligence have been proposed to improve power system restoration. These techniques propose the use of the computer as an operator aid instead of the use of predefined operating procedures for restoration. The stressful condition following a blackout and the pressure for achieving a restoration plan in minimum time can lead to misjudgment by system operator. This paper proposes the use of ANN for service restoration plan, since it has generalization capability and high processing speed. The large number of possible faulty conditions and the need to provide a restoration plan in minimum time are arguments in favor of this technique.

XII. REFERENCES

[1] IEEE Transactions On Power Delivery, Vol. 18, No. 4, October 2003

[2] "Neural Networks" – Control Systems Engineering (Thrid Edition) By I.J.Nagrath&M.Gopal

[3] L.M. Waghmare, Dr. Vinod Kumar & Dr. Saxena, "Electrical India" Januvary 1998.

[4] Http://Www.Electricalindia.Com

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018IJSRCSEIT| Volume 4 | Issue 3 | ISSN:2456-3307

High Performance Cloud Computing in Big Data

Bharathidasan K[1], UdhayaKumar U[2]

[1]Department of Computer Science, Shanmuga Industries Arts and Science College, Tiruvannamalai, Tamilnadu, India
[2]Assistant Professor, Department of Computer Science, Shanmuga Industries Arts and Science College, Tiruvannamalai, Tamilnadu,

ABSTRACT

Cloud computing is the product of the traditional computer technology and network technology development integration. Such as grid computing, distributed computing, parallel computing, utility computing, network storage, virtualization, and load balancing. To perform integrating multiple relative low-cost computing entities into one perfect system with powerful computing ability via the network and with the help of the SaaS, PaaS, IaaS, MSP and other advanced business models distributing this powerful computing ability to the hands of the end user. Cloud computing system mainly uses MapReduce model. The core design idea of MapReduce is to divide and conquer the problem and calculate on data rather than push data to calculate which effectively avoids many communication costs generated during data transmission.

Keywords: Big Data, Cloud computing, High-performance, HPC.

I. INTRODUCTION

With the advent of the digital age, the amount of data generated, stored and shared has been on the rise. From data warehouses, web pages and blogs to audio/video streams, all of these are sources of massive amounts of data. The result of this proliferation is the generation of massive amounts of pervasive and complex data, which needs to create efficiently, stored, shared and analysed to extract useful information [1, 15, 16].

Since innovations in data architecture are on our doorstep, the 'big data' paradigm refers to very large and complex data sets (i.e., petabytes and Exabyte's of data) that traditional data processing systems are inadequate to capture, store and analyse, to seek to glean intelligence from data and translate it into competitive advantage. As a result, Big data needs more computing power and storage provided by cloud computing platforms.

II. BIG DATA MANAGEMENT

The architecture of Big Data must synchronize with the support infrastructure of the organization. To date, all of the data used by organizations are stagnant. Data increasingly sourced from various fields that are disorganized and messy, such as information from machines or sensors and abundant sources of public and private data [2]. Previously, most companies were unable to either capture or store these data, and available tools could not manage the data in a reasonable amount of time. However, the new Big Data technology improves performance, facilitates innovation in the products and services of business models Big Data technology aims to minimize hardware and processing costs and to verify the value of Big Data before committing significant company resources [3]. Correctly, managed Big Data are accessible, reliable, secure, and manageable. Hence, Big Data applications can apply in various complex scientific disciplines (either single or interdisciplinary), including atmospheric science, astronomy, medicine, biology, genomics, and biogeochemistry[16]. In the following section, we briefly discuss data management tools and propose a new data life cycle that uses the technologies and terminologies of Big Data.

A. Management Tools

With the evolution of computing technology, large volumes can manage without requiring supercomputers and high cost. Many tools and techniques are available for data management, including Google BigTable, Simple DB, Not Only SQL (NoSQL), Data Stream Management System (DSMS), MemcacheDB, and Voldemort [3]. However, companies must develop specialized tools and technologies that can store, access, and analyze large amounts of data in near-real time because Big Data differs from the traditional data and cannot store in a single machine. Furthermore, Big Data lacks the structure of traditional data [4]. For Big Data, some of the most commonly used tools and techniques are Hadoop, MapReduce, and Big Table. These innovations have redefined data management because they efficiently process large amounts of data efficiently, cost-effectively, and promptly. The following section describes Hadoop and MapReduce in further detail, as well as the variousprojects/frameworks that are related to and suitable for the management and analysis of Big Data [5, 6].

B. Hadoop

Hadoop written in Java and is a top-level Apache project that started in 2006. It emphasizes discovery from the perspective of scalability and analysis to realize near-impossible feats. Doug Cutting developed Hadoop as a collection of open-source projects on which the Google MapReduce programming environment could apply in a distributed system. Presently, used on large amounts of data with Hadoop; enterprises can harness data that was previously difficult to manage and analyse. Hadoopused by approximately 63% of organizations to manage a vast number of unstructured logs and events (Sys.con Media, 2011).

C. HDFS

This paradigm applied when the amount of data is too much for a single machine. HDFS is more complicated than other file systems given the complexities and uncertainties of networks [22]. The cluster contains two types of nodes. The first node is a name-node that acts as a master node. The second node type is a data node that acts as a slave node. This type of node comes in multiples. Aside from these two types of nodes, HDFS can also have secondary name-node. HDFS stores files in blocks, the default block size of which is 64 MB. All HDFS files are replicated in multiples to facilitate the parallel processing of large amounts of data.

III. PROCESSING OF BIG DATA IN CLOUD COMPUTING

Cloud computing as an essential application environment for big data has attracted tremendous attention from the research community. Remarkable progress of Big data networking has also reported in this area. In this section, we introduce Big data research issues and solutions related to Cloud Computing [18, 20]. Individually, we are interested in the following topics: opportunities and challenges of Big data networking in Cloud Computing, cloud resource management of big data, and performance optimization of big data in Cloud Computing.

Data is the central element of communication and collaboration on the Internet and all the applications that built on this platform. The immense popularity of data-intensive applications like Facebook, LinkedIn, Twitter, Amazon, eBay, and Google+ contributes to increasing requirement of storage and processing of data in the cloud environment [20].

Therefore, they require high-performance processors to do the job. The cloud provides an excellent platform for big data storage, processing, and analysis, addressing two of the primary requirements of big data analytics, high storage, and high-performance computing [21].

The cloud-computing environment offers development, installation, and implementation of software and data applications 'as a service.' Three multi-layered infrastructures namely, platform as a service (PaaS), software as a service (SaaS), and infrastructure as a service (IaaS) exist. Infrastructure-as-a-service is a model that provides computing and storage resources as a service. On the other hand, in case of PaaS and SaaS, the cloud services provide software platform or software itself as a service to its clients [19].

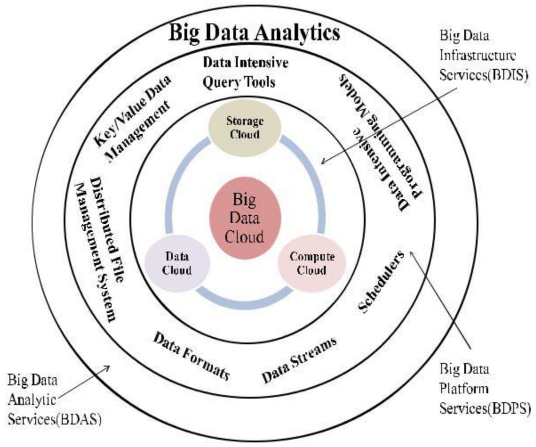

Big Data Analytics
Data Intensive Query Tools
Key/Value Data Management
Data Intensive Programming Models
Distributed File Management System
Schedulers
Data Formats
Data Streams
Storage Cloud
Big Data Cloud
Data Cloud
Compute Cloud
Big Data Infrastructure Services(BDIS)
Big Data Analytic Services(BDAS)
Big Data Platform Services(BDPS)

Figure 1. Big Data Cloud Computing

Since innovations in data architecture are on our doorstep, the 'big data' paradigm refers to very large and complex data sets (i.e., petabytes and exabytes of data) that traditional data processing systems are inadequate to capture, store and analyze, to seek to glean intelligence from data and translate it into competitive advantage [16]. As a result, Big data needs more computing power and storage provided by cloud computing platforms. In this context, cloud providers, such as IBM, Google, Amazon, and Microsoft, provide network-accessible storage priced by the gigabyte-month and computing cycles priced by the CPU-hour.

Although big data is still in the preliminary stages, comprehensive surveys exist in the literature [1, 9–11, 20]. This survey article aims at providing a holistic perspective on big data and big data-as-a-service (BDaaS) concepts to the research community active on big data-related themes, including a critical revision of the current state-of-the-art techniques, definition, and extensive researches issues. Following this introductory section, Sect. 2 presents related work approaches in the literature, including the architecture and possible impact areas. Section 3 demonstrates the business value and long-term benefits of adopting big data-as-a-service business [18].

Another significant challenge is the delivery of Big data capabilities through the cloud. The adoption of Big data-as-a-service (BDaaS) business models enables the efficient storage and management of massive datasets and data processing from an outside provider, as well as the exploitation of a full range of analytics capabilities (i.e., data and predictive analytics or business intelligence provided as service-based applications in the cloud).

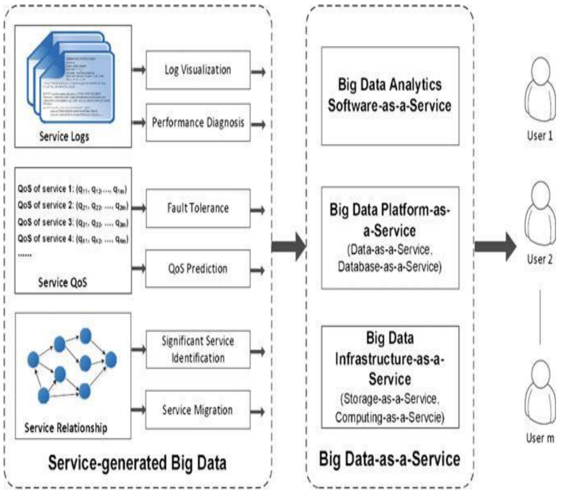

Service Logs
Log Visualization
Performance Diagnosis
QoS of service 1: (q11, q12, ..., q1m)
QoS of service 2: (q21, q22, ..., q2m)
QoS of service 3: (q31, q32, ..., q3m)
QoS of service 4: (q41, q42, ..., q4m)
......
Service QoS
Fault Tolerance
QoS Prediction
Service Relationship
Significant Service Identification
Service Migration
Service-generated Big Data
Big Data Analytics Software-as-a-Service
Big Data Platform-as-a-Service (Data-as-a-Service, Database-as-a-Service)
Big Data Infrastructure-as-a-Service (Storage-as-a-Service, Computing-as-a-Servcie)
Big Data-as-a-Service
User 1
User 2
User m

Figure2. Service-Generated Big Data and Big Data-As-A-Service

IV. PERFORMANCE OPTIMIZATION

Performance optimization is yet another classic and essential topic in cloud computing because appropriate optimization techniques will provide better application experiences with comparable or even less system resource consumption, compared to non-optimized cases.

A dataflow-based performance analysis tool for big data cloud, i.e., Hitune, presented in [Dai11]. Hitune shown to be useful in assisting users doing Hadoop performance analysis and system parameter tuning. Limitations of existing approaches, such as Hadoop logs and metrics compared and discussed. A few interesting case studies on Big data processing in cloud computing environment depicted [17, 18] . Efforts of the Fujitsu laboratory based on the data store and complex event processing, as well as workflow description in distributed data processing.

A recent online cost-minimization algorithm depicted in [Zhang10]. The proposed work specifically focused on real-time cost minimizations for uploading massive and dynamic data onto the cloud. The two online algorithms have achieved competitive cost reduction ratios [19]. However, the proposed methods evaluated on a limited scale. The proposed algorithms need further evaluated at more extensive and more competitive scales, e.g., data streaming applications with larger topologies.

V. PERFORMANCE OPTIMIZATION

In scientific applications, data commonly represented by a multi-dimensional array-based data model. For instance, the widely used Community Earth System Model (CESM) software package consists of four separate modules simultaneously simulating the earth atmosphere, ocean, land surface, and sea-ice, and each module uses the multi-dimensional arrays data model [9]. A typical example is a 3-dimensional temperature data with longitude, latitude, and time dimensions. It is often needed to compute the moving average, median, lowest and highest temperature with specified conditions such as areas and periods. Such computed results will further correlated with the computed results from other parameters, such as the humidity and wind velocity, to predict weather conditions [10].

The current way of conducting such processing is to read the required data (e.g., a sub-array of the affected area) from storage servers to compute nodes, perform computations on desired data with specified conditions, such as those data shown in a shaded area, and then write the output back to storage [7,8]. For CESM, an experimental test shows that the data access and movement time for the calculation of the moving average, median, lowest and highest degrees can occupy 88.2%, 95.4%, 96.6%, and 96.6% of the total execution time on a cluster, where 128GB of data retrieved to 272 nodes for processing.

CESM has data retrieval and processing phases and computing and simulation phases, as many scientific Big data applications do. The basic idea of the new decoupled HPC system architecture is to change the conventional architecture to handle these two phases differently on different nodes. Such architecture decouples nodes into compute nodes and data processing nodes [14, 15]. These nodes are mapped with computation-intensive operations and data-intensive operations respectively. Computation-intensive operations executed on massive compute nodes. Data-intensive operations executed on dedicated data processing nodes. In other words, the decoupled architecture reshapes the current pattern of retrieve - compute - store cycles into retrieving (generate) - reduce - compute - reduce - store cycles as shown in Figure 1, where the reduce phases are designed to conduct offloaded data-intensive operations and reduce data size before moving data across the network. This retrieval, reduce, compute, and store phases can be pipelined to overlap the I/O, communication, and computation times. From one point of view, the decoupled architecture is an enhanced framework of MapReduce [10], where one node with its local storage does not conduct the reduction, but a set of (data) nodes and the global storage so that that parallel computing features can maintain. From another point of view, the data nodes are the data-access accelerators, to speed up data accesses and reduce data size before sending data across the network.

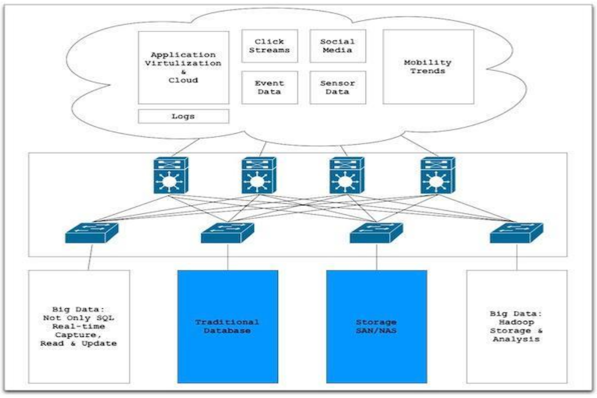

Application Virtulization & Cloud
Click Streams
Social Media
Mobility Trends
Event Data
Sensor Data
Logs
Big Data: Not Only SQL Real-time Capture, Read & Update
Traditional Database
Storage SAN/NAS
Big Data: Hadoop Storage & Analysis

Figure 3. General Framework of Big Data Networking.

VI. CHALLENGES

While the rise of big data yields enormous opportunities for individuals, organizations and the society, it also raises significant privacy and ethical issues [16, 17]. These issues are factors that may lead to situations in which the underlying analytic models and infrastructures are likely to impact privacy negatively from both a legal and an ethical perspective and hence represent possible obstacles for the big data's potential to be fully realized.

Big data analytics essentially requires very high computing capabilities to drive data into meaningful insights. High-performance data analytics, HPDA, seeks to widen the HPC and Big data analytics application domains by augmenting with other related technologies [17,19]. However, varied and complex requirements of big data analytics pose many challenges at micro as well as macro level. At the micro level, there are unusual and specific issues about statistical modelling of big data. At the macro level, big data analytics challenged by the complexities of working computational prototypes.

- **Data storage and management:** Since big data are dependent on extensive storage capacity and data volumes grow exponentially, the current data management systems cannot satisfy the needs of big data due to limited storage capacity. Also, the existing algorithms are not able to store data efficiently because of the heterogeneity of big data.

- **Data Transmission and Curation:** Since network bandwidth capacity is the major drawback in the cloud, data transmission is a challenge to overcome, especially when the volume of at enormous huge. For managing large-scale and structured datasets, data warehouses and data marts useful good approaches. Data warehouses are relational database systems that enable the data storage, analysis, and reporting, while the data marts are based on data warehouses and facilitate the analysis of them. In this context, NoSQL databases introduced as a potential technology for large and distributed data management and database design. The significant advantage of NoSQL databases is the schema-free orientation, which enables the quick modification of the structure of data and avoids rewriting the tables.

- **Data processing and analysis:** Query response time is a significant issue in big data, more time needed when traversing data in a database and performing real-time analytics. A flexible and reconfigured grid along with the big data pre-processing enhancement and consolidation of application and data-parallelization schemes more efficient active approaches for extracting more meaningful knowledge from the given data sets.

- **Data privacy and security:** Since the host of data or other critical operations performed by third party services or infrastructures, and security issues witnessed concerning big data storage and processing. The current technologies used in data security are mainly static data-oriented, although big data entails the dynamic change of current and additional data or variations in attributes. Privacy-preserving data mining without exposing sensitive personal information is another challenging field to investigate.

VII. CONCLUSION

The information-driven economy relies on the actionable insights extracted from data analytics. The era of data revolution has paved a way to the need of convergence of paradigms like High-Performance Computing and Big Data Analytics. The amalgamation of these paradigms is a herculean task involving various aspects of data management and computing efficiency. HPC with Big data has given rise to the evolution of the data storage technologies and computing models. The transformation of traditional analytical paradigms to cater to the requirement of the intense data applications and High-Performance Computing is the need of the hour.

The convergence of the paradigms “High-Performance Computing” and “Big Data Analytics” can lead to a sustainable solution for the data-driven applications. The continuous flow of “real” data which is the predominant type of data seen in data-intense applications needs to be handled by a different architectural platform termed as “Real Time Analytical Framework.” The computational requirements of these newer models are different from the traditional models, and hence the evolution of the models becomes the critical challenge of High-Performance Data Analytics.

VIII. REFERENCES

[1] Big Data: The next frontier for innovation, competition and productivity. James Maniyka, Executive summary ,McKinsey Global Institute ,May 2011, http://www.mckinsey.com/mgi/publication/big.data/MGI _big_data_exec_summary.pdf>.

[2] http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf [Accessed on 2nd January 2015]

[3] Beyond the hype: Big data concepts, methods, and analytics, Amir Gandomi , MurtazaHaide, International Journal of Information Management 35 (2015) 137–144

[4] https://analyticsacademy.withgoogle.com/course01/asses/pdf/DigitalAnalyticsFundamentals-Lesson2.1TheimportanceofdigitalanalyticsText.pdf [Acesses on 27th December 2014]

[5] Big Data Meets High Performance Computing Intel® Enterprise Edition for Lustre* software and Hadoop combine to bring big data analytics to high performance computing configurations.http://www.intel.com/content/dam/www/public/us/en/do cuments/white-papers/big-data-meets-high-performance-computing-white-paper.pdf

[6] Agrawal, D., Das, S., El Abbadi, A.: Big data and cloud computing: current state and futureopportunities. In: Proceedings of the 14th International Conference on Extending Database Technology (EDBT/ICDT’11), pp. 530–533 (2011)Amazon Web Services, Inc.: Elastic Compute Cloud (EC2). http://aws.amazon.com/ec2 (2015). Accessed 18 Oct 2015

[7] Assunção, M.D., Calheiros, R.N., Bianchi, S., Netto, M.A.S., Buyya, R.: Big data computing and clouds: trends and future directions. J. Parallel Distrib. Comput. 79–80, 3–15 (2015)

[8] Batalla, J.M., Kantor, M., Mavromoustakis, C.X., Skourletopoulos, G., Mastorakis, G.: A novel methodology for efficient throughput evaluation in virtualized routers. In: Proceedings of the IEEE International Conference on Communications (ICC 2015)—Communications Software, Services andMultimedia Applications Symposium (CSSMA), London, UK, pp. 6899–6905 (2015)

[9] Zheng, Z., Zhu, J., Lyu, M.R.: Service-generated big data and big data-as-a-service: an overview. Proceedings of the 2013 IEEE International Congress on Big Data (BigData Congress), pp. 403–410. Santa Clara, California (2013)

[10] Zhang, Linquan, et al. "Moving Big Data to The Cloud: An Online Cost Minimizing Approach." IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS 31.12 (2013): 1.http://i.cs.hku.hk/~fcmlau/papers/info13-lq-m.pdf

[11]Hashem, I.A.T., Yaqoob, I., Anuar, N.B., Mokhtar, S., Gani, A., Khan, S.U.: The rise of “big data” on cloud computing: review and open research issues. Inf. Syst. 47, 98–115 (2015)

[12]IBM Corporation: IBM big data & analytics hub: the four V’s of big data. http://www.ibmbigdatahub.com/infographic/four-vs-big-data (2014). Accessed 18 Oct 2015

[13]IBM Corporation: IBM social media analytics software as a service. http://www-03.ibm.com/software/products/en/social-media-analytics-saas (2015a). Accessed 18 October 2015.

[14]D. Ranjith, J. M. Balajee and C. Kumar,” Trust computation methods in mobile ADHOC network
using glomosim: A Review” International Journal of
Scientific Research and Modern Education, Vol. I, Issue I, pp. 777-780, Nov.2016.

[15]Janarthanan Y, Balajee J.M, and SrinivasaRaghava S.
"Content based video retrieval and analysis using image processing: A review."International Journal of
Pharmacy and Technology 8, no.4 (2016): 5042-5048.

[16]Jeyakumar, Balajee, MA SaleemDurai, and Daphne Lopez. "Case Studies in Amalgamation of Deep Learning and Big Data." In HCI Challenges and Privacy Preservation in Big Data Security, pp. 159-174. IGI Global, 2018.

[17]Kamalakannan, S. "G., Balajee, J., Srinivasa Raghavan.,“Superior content-based video retrieval system according to query image”." International Journal of Applied Engineering Research 10, no. 3 (2015): 7951-7957.

[18]Priya, V., Subha, S., &Balamurugan, B. (2017). Analysis of performance measures to handle medical E-commerce shopping cart abandonment in cloud. Informatics in Medicine Unlocked.

[19]Rangith. D Lakshmi narayanan. J and Balajee. J,” A study of behavior on information system in a universitycampus by analysis of people mobility” Internationaljournal of research in computer application& management, Vol. 6, Issue 7, pp. 29-31, Jul.2016.

[20]Ranjith, D., J. Balajee, and C. Kumar. "In premises of cloud computing and models." International Journal of Pharmacy and Technology 8, no. 3 (2016): 4685-4695.

[21]Sethumadahavi R Balajee J “Big Data Deep Learning in Healthcare for Electronic Health Records,” International Scientific Research Organization Journal, vol. 2, Issue 2, pp. 31–35, Jul. 2017.

[22]Ushapreethi P, BalajeeJeyakumar and BalaKrishnan P, Action Recongnition in Video Survillance Using Hipi and Map Reducing Model, International Journal of Mechanical Engineering and Technology 8(11), 2017,pp. 368–375.

Safe and Secure Data Transfer in Mobile AD-HOC Networks using Multilevel Encryption Techniques

Mr. P. Daniel Sundarraj[1],Dr. K. Arulanandam[2]

[1]Department of Computer Science and Application, K.M.G. College of Arts and Science, Gudiyattam,Tamil Nadu, India

[2]Department of Computer Application, Government Thirumagal Mills College, Gudiyattam,Tamil Nadu, India

ABSTRACT

At the time of sending any secret information from a source node to a destination node over a wireless network, it is very critical to transmit it in a safe and secure manner. A set of wireless nodes constructs an Ad Hoc network and this network does not have any central control or centralized administration. In self-mode, wireless Ad-hoc networks are organized and configured. All nodes in this network are set up by using a wireless transmitter and a wireless receiver. The wireless Ad Hoc network transmits data with other nodes within its communication range only. Using a common physical media, the data are transmitted between one node and another node in this network. Every node sends and receives signals using the same frequency band and it follows the same hopping method during data transmission. If the destination node is not inside the transmission range, the source node will use the other nodes to transmit the messages hop by hop. In order to send a message from one node to another node that is out of its frequency range, it needs the help of other nodes in the network for the data transfer. This technique is technically known as multi-hop communication. In this network, each and every node acts both as a host and as a router at the same time. Wireless Mobile Ad Hoc networks are usually attacked the sources such as intruders, hackers and other physical attacks. Constructing and configuring the safest and secure wireless ad-hoc network is very difficult for the reasons such as: the poor quality of communication paths and communication nodes, low quality infrastructure, frequently updating topologies and technologies. Due to these main factors, the wireless communication path or channel can be easily accessed by all the network users and the attackers and it makes the network operations very insecure and unsafe. Any user can easily break the network system and its operations by not following any specific protocol. Hence, a safe and secure protocol or an algorithm is to be developed for the safest data transfer. Also, there is another issue and it is the complexity of finding the routing mechanism to transfer our data from one node to another node in a safe way. In this paper, we are suggesting a multi-level encryption technique which can send the data over a wireless network in a safe and secure way.

Keywords: Ad-hoc Network, Encryption, Decryption, Routing, Multi Hopping, Cryptography, Cipher Text

I. INTRODUCTION

A. Ad Hoc Network: Characteristics

1. It does not have any fixed architecture.
2. It has a dynamic topology.
3. It is a Multi-hopping Network.
4. Scalability: It may have thousands of nodes.
5. Security: It is limited

B. Issues in Ad Hoc Networks

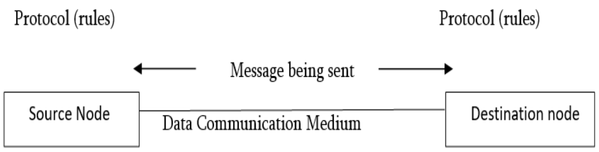
- Medium Access: Distributed, no time synchronization
- Routing: Route Acquisition Delay, Quick Reconfiguration, Loop Free
- Multicasting: The way of communication occurs between the nodes
- Transport Layer: Frequent path breaks
- Self-Organization: Neighbor discovery, Link failures
- Security: Jamming, Hackers and Intruders
- Energy Control: Transmission Power, Battery Monitoring, Processing Power.

C. Data Communication

Sharing our information with other people through a communication path or media is called data communication. This process can be local or remote. The local communication is done face to face. On the other hand, the data transfer in remote communication occurs over distance. The nodes or workstations in a network system are very useful devices to exchange information over a network. In the network system, each and every computer is known as a client machine or a node or workstation and it keeps a server machine to store the information in a central place. On request, the nodes will receive the required information from the server.

D. Computer Network

A group of computers form a computer network and the nodes in it are interconnected by using communication channels. Following are the important components of a network system.



While transmitting an secret message from a sender node to a receiver node over a network, the message should be protected from the unauthorized users. Hence proper techniques and methods are required to protect our secret message that we send from the source.

E. Cryptography

We can convert an intelligible message into an unintelligible message by using cryptographic techniques. This unintelligent format of message cannot be read by others.

In Cryptography, the following technical terms are used.

Plaintext: It is the message to be sent from a source node to a destination node in the network.

Ciphertext: After the message is converted into unintelligible format, this encrypted form of message is called Ciphertext.

Cipher: An algorithm which is used to convert an intelligible message into an unintelligible message is technically called as Cipher.

Key:A secret key is used by an algorithm (Cipher) for safe data transfer and this key is only known to the sender or receiver.

Encipher (Encoding): This is the process of converting a plaintext into a cipher text by using the cipher (algorithm) and the secret key is known as Encipher (Encoding)

Decipher (Decoding): This is the process of reconverting the cipher text into its original format (plaintext format) and it is technically called as Decipher (Decoding).

Cryptanalysis: The study of methods and principles to transform a cipher text (unintelligible form of message) into a plaintext (intelligible form of message) without using a secret key is called as Cryptanalysis or Code Breaking.

Cryptology: It combines both the cryptography and cryptanalysis.

II. OUR BASIC AND PROPOSED IDEA FOR THE IMPLEMENTATION OF SAFE AND SECURE DATA TRANSFER

1) To exchange our information between a source node and a destination node, the asymmetric cryptography method is used (public key and private key cryptography). A modified RSA algorithm is preferred here. The public key is known to all and the private key is kept secret in this technique.

2) The encryption is implemented in two stages by using a modified RSA algorithm which ensures security in data transferring.

III. RSA Algorithm in Cryptography (Modified Algorithm)

A. Steps to be used for converting data from plain text to cipher text format

1. The client node sends its public key to the server node and requests for its response.
2. The server encrypts the data asked by the client using the technique in RSA algorithm.
3. The Client node receives the requested data in its original form after decrypting the cipher text.

B. The Technique

1) The public key combines two numbers. In which, one number is the multiplication of two large prime numbers.
2) The private key is created with the help of the same two prime numbers.

C. Illustrating an example showing the Cryptographic Technique

Creating the Public Key

- Assume two prime numbers. (**M = 31 and N = 37**).
- The first part of the Public key will be now **n = M*N = 1147**.
- Calculate the exponent **e** with the following conditions :

e should be an integer value

e should not be a factor of n

Also, 1 < e <Φ(n)

Assume that e is equal to 3.

- Finally the Public Key is made of n and e

Generating Private Key

- Compute Φ(n) :

Φ(n) = (M-1)(N-1)

Hence, Φ(n) = 1086

- Compute to create the Private Key, **p** :
- **s = (r*Φ(n) + 1) / e** for some integer r
- For example if r = 2, the value of s is 724

Finally, we get the calculated Public Key is (n = 1147 and e = 3) and our Private Key is (s = 724)

D. Encryption

Suppose the message to be encrypted is **"BC"**:

- Convert the above characters into its sequence numbers (B = 2 and C = 3)
- Now the Encrypted Data will be **g = 23^{e} mod n**.
- Hence our Encrypted Data will be 697

E. Decryption

We need to decrypt **697** again into its original form:

- **The Calculation for Decrypting the Data is = g^{s} mod n**.
- Hence our Encrypted Data will be 23.

Finally we receive our original message "BC", since 2 = B and 3 = C.

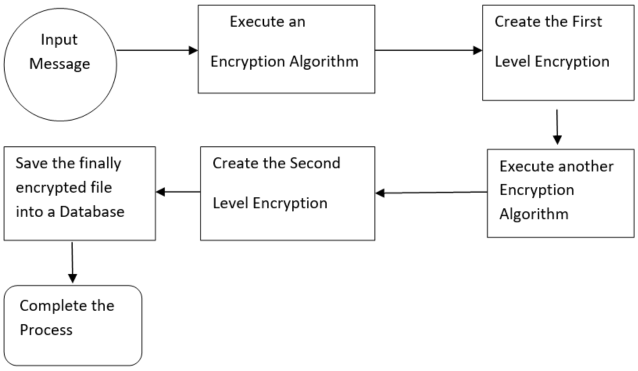
IV. FOLLOWING IS THE IMPLEMENTATION OF RSA ALGORITHM IN C PROGRAMMING LANGUAGE

```
// Our C program to implement the RSA algorithm

#include<stdio.h>

#include<math.h>

// Finding the gcd of i and j

intgcd(inti, int j)

{

    int t;

    while (1)

    {

        t = i%j;

        if (t == 0)

          return j;

        i = j;

        j = t;

    }

}

 // RSA algorithm

int main()

{

    // Let us take any 2 random prime numbers

    double p = 3;

    double q = 7;

     // Computing the first part of public key:

    double n = p*q;

     // Computing the other part of public key.

    // e – Encryption

    double e = 2;

    double hi = (p-1)*(q-1);

    while (e < hi)

    {

        // e is co-prime to hi which is smaller than hi.

        if (gcd(e, hi)==1)

            break;

        else

            e++;

    }

    // Creating the Private Key (d - Decryption)

    // selecting d such a way that it should satisfy

    // Calculating the d*e = 1 + k * totient

    int k = 2;  // A constant value

    double d = (1 + (k*hi))/e;

    // The Message to be encrypted is

    double mmesg = 20;

     printf("Our Message data is: = %lf", mmesg);

    // Calculating the message to be encrypted is : s = (mmesg ^ e) % n

    double s = pow(mmesg, e);

    s = fmod(s, n);

    printf("\The encrypted data is: = %lf", s);
    // Calculating the message to be decrypted is :
f = (s ^ d) % n

   double f = pow(s, d);

   f = fmod(f, n);

   printf("\nThe Original Message in Plain Text
Format is: = %lf", f);

   return 0;

}
```

V. MULTI-LEVEL ENCRYPTION

Multi-Level encryption is a method or technique which encrypts the message more times in different levels.

(The inverse algorithms are to be used for decrypt the message and converting it into original message)


Input Message
Execute an Encryption Algorithm
Create the First Level Encryption
Execute another Encryption Algorithm
Create the Second Level Encryption
Save the finally encrypted file into a Database
Complete the Process

VI. CONCLUSION

After we encrypt the intelligible plain text message in multiple levels as stated above, the final encrypted message is decrypted by using the inverse algorithms in order to get the original message. If the message is transmitted like this, the hackers will find it very difficult in accessing our data and damage it and this type of technique is more secure and safe data transmission as well. While following this technique, the time synchronization process should be there to synchronize our data being sent and we need to also ensure that we get back the data in original format without any acquisition delay. For this time synchronization process, we need to create proper encryption and decryption algorithms so that it should not take much time for encryption and decryption process.

VII. REFERENCES

[1] Kejun Liu, Jing Deng, Member, IEEE, Pramod K. Varshney, Fellow, IEEE, and KashyapBalakrishnan, Member, IEEE

[2] L. Buttyan and J.-P. Hubaux, “Stimulating Cooperation in Self-Organizing Mobile Ad Hoc Networks,” ACM/Kluwer Mobile Networks and Applications, vol. 8, no. 5, 2003.

[3] K. Balakrishnan is with the Security Services Group, Deloitte and Touche LLP, 1750 Tysons Boulevard, Suite 800, McLean, VA 22102. E-mail: kbalakrishnan@deloitte.com.

[4] S. Marti, T. Giuli, K. Lai, and M. Baker, “Mitigating Routing Misbehavior in Mobile Ad Hoc Networks,” Proc. MobiCom, Aug. 2000.

[5] V.-N. Padmanabhan and D.-R. Simon, “Secure Traceroute to Detect Faulty or Malicious Routing,” SIGCOMM Computer Comm. Rev., vol. 33, no. 1, Jan. 2003.

[6] Srinivasan, P. Nuggehalli, C.F. Chiasserini, and R.R. Rao, “Cooperation in Wireless Ad Hoc Networks,” Proc. INFOCOM, Mar.-Apr. 2003.

[7] D. Johnson, D. Maltz, Y.C. Hu, and J. Jetcheva, “The Dynamic Source Routing Protocolfor Mobile Ad Hoc Networks (DSR),” Internet draft, Feb. 2002.

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3| ISSN: 2456-3307

Innovating iTwin Technology Advances an Overview

V.Sujitha[1], M.Vijayakumar[2]

[1]Assistant Professor, Department of Computer Application, Sun Arts and Science College, Tamilnadu, India
[2]Assistant Professor, Department of Computer Application, ArunaVidya Arts and Science College, Tamilnadu, India

ABSTRACT

USB flash drive is a data or information storage device that is used to storing the data. Cloud storage is mainly used to store the data on the middle server. USB stands for Universal Serial Bus, which is a kind of computer port, whichcan be used to attach equipment to a computer. An example of a USB is the interface used to upload pictures from a digital camera to a computer. In this research Paper, it is anticipated that a USB flash drive is a information storage device that abide of flash memory with an USB. A Cloud Storage wherever the data is store in the cloud that is distantly present and can be accessed as and when needed. But a problem of this USB is its tiny size due to which it can easily be misplaced. This is a difficulty if the data present in it is sensitive. In this Cloud Storage the data be capable of misused if the username and password of an account to contact the storage is hacked by someone. There are many drawback like some degree of storage, security, back up, temporary files. iTwin is a 'limitless' USB mechanism that permit users to access, change & exchange all their files & media whichever two online computers anywhere in the world. The simply limit is the size of your computer's hard drive. iTwin is a innovative unfolding that permits remote file access without the any problems of the cloud and USB flash drives. It's can be access as a USB with no exact installation. iTwin used by cryptographic protocols and algorithms. USB contains attributes like two directional file access, no temporary files.

Keywords:AES -Advanced Encryption Standard, PC -Personal Computer, USB -Universal Serial Bus, VPN -Virtual Private Network

I. INTRODUCTION

Accessing the data while you are far away from your PC, there is one option that is to construct use of high capacity USB storage device. But what if you told that, you could privacy access your PC or any other computer while you are travelling on the road? iTwin is a device that connect to your PC and it allows you to access files and devices that are linked to your home network remotely.

It looks similar as USB device and it is designed by union of two halves. One half is associated to your house or office PC and you have to carry other half always with you. The part that you bring with you is used as a key for obtain the connection to your PC when you are far away from your PC. When you connect one more part of the iTwin device to your laptop when you are on the road, the device makes a VPN is Virtual Private Network to your house or to the PC that you use in your office. iTwin is completely new file sharing and remote access device created by a company named as iTwin. It is very related like two ends of a cable, but is does not need the cable. It is simpler to use than a flash drive. It is just a plug and play device. With iTwin, it is possible to connect any two online computers located anywhere in the world. iTwin was invented by an

Indian named Lux Anantharaman. He has completed a Bachelors degree in Electrical and Electronic Engineering. He studied in IIT in Chennai and he completed a Masters degree from IISc in Bangalore. Lux was completing a part-time MBA at NUS Business School in Singapore, but he put studies on hold because of the potential of iTwin.

II. WORKING OF iTWIN

The iTwin Connect is a device similar to a USB flash drive but, is designed with two USB connections. The iTwin device is very dense and it establishes a secure connection between two computers or a secure connection between one computer and the iTwin server. When you connect the iTwin Connect device to the main computer in your home or at office, the software is automatically installed configure the computer for remote connection.

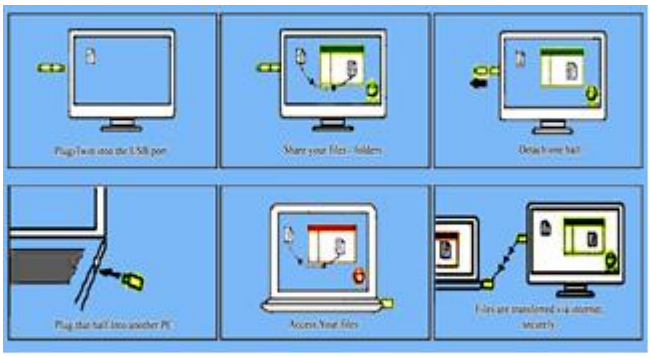

Plug iTwin into the USB port
Share your files - folders
Detach one half
Plug that half into another PC
Access Your files
Files are transferred via internet securely

Figure 1. Working of iTwin

When the device is disconnected from the main computer, you have to separate the two parts of the USB ports, which are separated in two separate USB devices. The two separate devices are very dense at less than two inches. Small size makes it convenient and easy to carry with you all the time. When you connect the second half to your laptop while travelling on the road, it will routinely install itself without any user interference. In addition, you can set up a special password that disables the device if you are going to lose it. This ensures you can lock down your files to avoid access by an illegal user.

Figure 2.iTwin Features

One of the best features of iTwin device is the ability to securely access your data by establishing a personal Virtual Private Network which is capable of managing several tasks. You can able to access all your files and data on your home or office PC by using Windows Explorer. Otherwise you can access definite files that are stored in desktop applications. As well as it is possible to access files from your isolated computer and save that files to the device that you are currently using, till both the devices are enabled with iTwin Connect device.

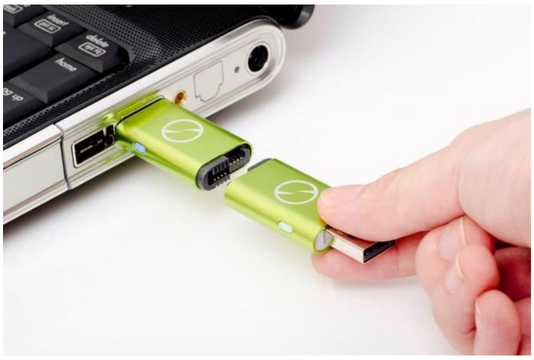


Figure 3. SplitingiTwin

III. ADVANCED FEATURES OF iTWIN

A. Remote Desktop

This feature enables you to observe the desktop of the main computer and allows you to manage the device. This is very useful medium for managing your computer from a remote position and it can also be used to offer tech support to somebody who experiencing computer problems without actually being seated in front of the PC.

By making use of isolated Desktop you can also start Windows Remote Desktop. This can be done with a single click which provides you the access to multiple different applications as well as the data enclosed in them on the remote PC.

B. Teleport Me

The Teleport Me feature is the secret browsing tool that enables you to surf the Internet lacking any limitations. Any websites that you open or any information that you transmit passes through the protected Virtual Private Network channel and uses the similar Internet connection that your house or workplace PC uses. If the main workstation cannot be left operating for some reason, Teleport Me is intended to connect to the iTwin dedicated network services to provide you a private connection. The company maintains dedicated servers all over the Europe, Asia Pacific and the United States.

The Teleport Me feature is a helpful tool, mainly if you are browsing the Internet on hotel or other types of public wireless networks where security is an important issue. Teleport Me takes care that your private information is protected from snooping eyes. It also ensures that no one is tracking your browsing movement. There are no browsing limits which denote that you can browse your social media accounts, can watch programs, and connect in any other activity you usually do online.

C. Office and Home Network Access

You can access approximately everything that is connected to your office or home network using the iTwin connect device. This includes devices such as drives or network applications on an office network, or devices like media servers, cameras, and televisions on your home server. For accessing the network it does not require any set of connections and arrangement. It automatically makes a secure VPN to your network.

D. Secure AES 256-bit Encryption

Hardware grade security is provided in iTwin. AES (Advanced Encryption Standard) 256-bit encryption is a security technology adopted by the US government to defend top secret classified information. When every part of the iTwinConnect device is paired with one another, a unique encryption key is generated for each session to make sure all information is protected prior to being transmitted over the Internet.

You can also configure the two-factor authentication integrated with the iTwin Connect device. This enables you to setup an optional second password on the one half of the USB device which you bring with you. If you come about to lose this half of iTwin device, without using the second password it cannot be accessed.

E. No subscription or Contract promise

While using iTwin Connect, there is no subscription necessity or contract commitment, you just have to pay a one-time cost of $130 for lifetime access to the iTwin device. You can also own the device and be able to use it in several ways you like. You do not have to be a mobile expert. You can also use iTwin Connect for personal use as well as to help family members that frequently call you for help with a computer problem. If they have one part of the iTwin device, you can access and managed the Desktop and repair the problem.

IV. ITWIN BENEFITS AND ITS USAGE

Most of the mobile professionals and individuals that want to access their files and information in spite of where they are, select cloud services for backing up and storing important documents. A cloud service is suitable and enables you to access your files from some device with an Internet connection. Many cloud service providers deploy security technologies to guarantee their customers that documents are securely transmitted and stored. On the other hand, not anything is one hundred percent perfect that means a device like iTwin Connect can help you cover up all of your bases in the event of data break or loss.

iTwin Connect device makes sure that your files reside private and protected. as you own the device, it is just functional when attached to the computer; it uses two-factor authentication and military grade security, as well as performs functions that we have discussed in this article. Even if you leave the main computer powered up so you can connect to it from any place, your data and records stay protected.

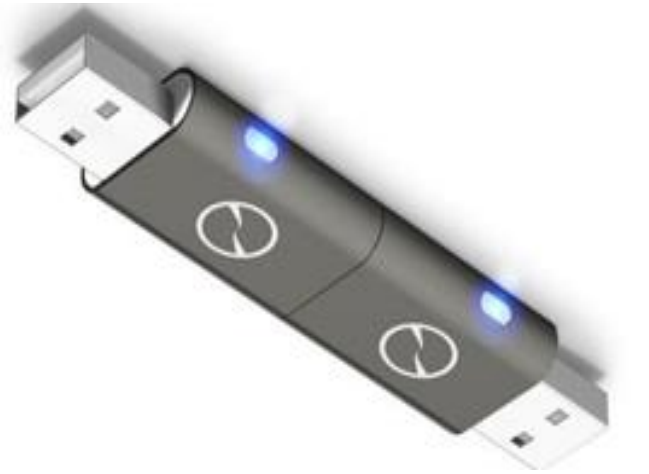


Figure 4.iTwin Device

iTwin Connect device also provides suitable browsing in any case of where you are positioned across the world. 'Teleport Me' feature enables you to select your continent even if you are travelling across the road. For example, if you live in the United States and you are travelling out of the country, you can select your continent as the US and still have the benefit of the programs you watch in the US. This benefit may also work vice versa. If you are travelling in the US and enjoy shows in UK, you can set browsing tool to the UK and access each and every one of the shows you enjoy while travelling. if you are sitting in an airfield waiting for your flight, you can access your desired shows and news sites from any continent. This is a benefit because your device and its IP address otherwise would not allow you to access programming outside of a specific geographic region.

iTwin Connect enables you to choose from locations in the United States, Europe or the Asia Pacific. This means you can grab up on your beloved shows and news while sitting in an airport waiting for your flight since iTwin directs the traffic through their dedicated servers in these locations.

V. PRIVACY OF ITWIN

A. Hardware Grade Security

When two parts of iTwin connect are attached together and inserted into a computer, a arbitrary 256-bit cryptographic key is generated on-board the iTwin device. This cryptographic key is shared among the two halves of the iTwin device using the particular iTwin connector. The cryptographic key never leaves the device. All data and information transferred by means of the two halves of iTwinis encrypted using this cryptographic key. The user can 'join up' the device to generate the keys anytime and any number of times. The keys are saved only inside the device and not known to any other entity.

B. Two Factor Authentication

iTwin device provides 2-factor authentication for advanced security. Access to your data is provided based on two different factors: Something you have – your physical iTwin device.

Something you be familiar with –a password which you have to enter before the iTwin device can be used. Setting an iTwin device password is optional however we suggest that you set the optional password throughout device initialization to protect your data and network in the event that you lose the device. The password set for your iTwin is stored only on the two halves of your iTwin and nowhere else.

C. Remote Disable

In the one half of the iTwin is lost; connection between the two halves of the device can be disconnected using the Remote Disable Feature. This is done by entering a unique disable code in the iTwin Disable Web Centre. The connection between the two halves of iTwinwill be disabled within 90 seconds and after that, it is impossible to gain access to your data via the lost device even if somebody finds it.

Figure 5. Example of iTwin

D. End –to-End Encryption

The shared cryptographic key stored in the two parts of the iTwin device are used to produce session keys which protect all information transmitted over the Internet using industrial strength AES-256 bit encryption.

E. Twin Trust Authentication

Every iTwin half has given a unique device ID and an linked device authentication key, adapted during manufacturing. Every iTwin device also carries certified public certificates of Twin Trust servers, inserted during manufacturing. Before allowing any data transfer, every iTwinis authenticated by iTwin's Twin Trust servers. After authenticating with Twin Trust, two halves of iTwin commonly authenticate each other using their previously shared AES 256 bit crypto key. All communication between iTwin and Twin Trust server is secured using HTTPS protocol.

VI. ITWIN ADVANTAGES

- One-time straight payment gives your lifetime access of the device.
- The capability to disable the device remotely if it is lost or stolen.
- There are no restrictions in terms of file size or type.
- It has secure military grade AES encryption ensures secure file and data transport.
- Personal VPN protects you on hotel as well as public Wi-Fi networks.
- Two-factor authentication provides additional security.
- Access to additional features such as your home or office network as well as the devices connected to it.
- Stretchy and safe browsing allows you to maintain your usual browsing activities while you are travelling.
- iTwin Connect is well-suited with both Windows and Mac devices.

VII. iTWIN DISADVANTAGES

- iTwin connect device is 3.5 inches long in total that means when the parts are separated they are very easy to misplace.
- When the iTwin Connect USB dongle is plugged in it can slow down network performance.
- You have to set up files to share in advance to access them.
- Require support for mobile devices.

VIII. CONCLUSION

Without a hesitation, iTwin Connect represents a unique solution for providing secure access to your files and information from a remote location, credit to the AES 256-bit encryption technology. The iTwin connect device performs like to Peer to Peer access excluding the data is only being shared between you and your main computer. If you are travelling with a Windows device and your main computer is a Mac, you can still access your files since iTwin Connect will work with both systems.

The iTwin is a unbelievable solution for the house user who desires to access and Change their files remotely and securely. The iTwin bypasses the virtual world of cloud services to turn your physical storage into its own networking solution. The iTwinis simple to use and inexpensively solves the Drop box limitations. For peoples who want to maintain files up-to-date among two computers the iTwin is for them.

IX. REFERENCES

[1] http://www.seminarsonly.com/computer%20science/itwin-seminar-report-ppt-pdf.php

[2] https://www.ijarcce.com/upload/2016/april-16/IJARCCE%2076.pdf

[3] data.conferenceworld.in/ICRTESM3/P1048-1054.pdf

[4] http://www.ijirse.com/wp-content/upload/2016/02/530V.pdf

[5] http://searchsecurity.techtarget.com/definition/Advanced-Encryption-Standard

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3| ISSN: 2456-3307

Intrusion Detection System Using Raspberry Pi Honeypot in Network Security

M.Devi Priya[1], A.Lavanya[2]

[1]M.Phil (Research scholar), KambanCollegeof Arts And Science For Women, Thiruvannmalai, Tamil Nadu, India

[2]Head of Department, Department Of Computer Science, Kamban CollegeOf Arts And Science For Women,Thiruvannmalai, Tamil Nadu, India

ABSTRACT

In the ever-changing world of global data communication, inexpensive Internet connection and fast-paced software development, security has become more and more of an issue in this world. Security is the basic requirement in today's world as any type of interaction and storage of data on the internet is becoming unassertive. Protecting the information access and data integrity are the basic security characteristics of computer security. A decoy based technology; Honeypot along with a Raspberry Pi makes network security simple,cost effective and easy for implementation. This paper is devoted to implementing the Raspberry Pi based Honeypot in a network that will attract attackers by simulating vulnerabilities and poor security too. Honeypot will record all the attackers' activities and after data, analysis not only displays the type of attack done but also allow improvements in the security of the network.

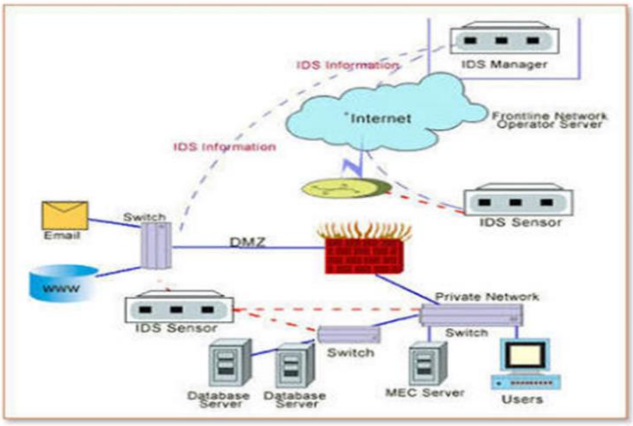
Keywords : Communication, Security, Information, Honeypot, Raspberry Pi, Data Analysis, Data Security, Network

I. INTRODUCTION

Information is one of the strategic resource, every organizations spends a significant amount of budget on managing of information resources. Computer security has several security related objectives among them the three fundamental objective are: Secrecy (to protect information), Incorruptibility, to protect accuracy of information; to ensure information delivery. It is necessary to put high priority to system security, minimize loopholes and secure the computer system against intrusion. Today's standard of security implements a configured firewall with an intrusion detection system. If an intruder is able to acquire the weakness in the network by scanning the host network, he can easily penetrate into the system and can obtain valuable data. If an intruder is masking his identity for a firewall-enabled service, intrusion detection systems cannot minimize the damages. Most of the security approaches now a day's focus on defense rather than aggressive form of a security. One of the aggressive for of defense mechanism uses Honeypots. It also acts as a Booby trap equipment, which are configured as a system weakness to attract intruders and gather all the information to eliminate future attacks thus, eliminating security loopholes, these are known as Honeypots. For example, honeypots like Honeyd1 are already being used to detect attackers and protect information. This architecture puts forth a simple, cost effective and an autonomous deployment in any environment. Subsequent chapters contain a description of the security system using Intrusion Detection System in combination with Raspberry Pi Honeypot.

II. INTRUSION DETECTION SYSTEM

IDS is a security application for computers and networks that gather and analyze information by scanning all the inbound and outbound network activities and identifies suspicious patterns that may indicate a network or system attack from someone attempting to break into or compromise a system.



III. TOOLS FOR DETECTING INTRUSIONS

Snort is a versatile and an open source tool used for intrusion detection. It is a network intrusion detection system (NIDS), a packet sniffer that captures and scans the network traffic in real time, examining each packet closely to detect an intrusion. Snort is based on libpcap (for library packet capture) one of the tool used in TCP/IP traffic sniffers and analysers. Snort also combines abnormal behaviour detection signatures and different methods of protocol detection. Observing vindictive exercises in PC frameworks is perplexing and costly. Using a Raspberry Pi in a network makes the network administrators work less complex and easy to implement. Described form of protection provides use of advanced security method called Honeypot along with a Raspberry Pi.

IV. HONEYPOT

Honeypot systems setup to gather information regarding an attacker or intruder into your system. Honeypots are an addition to your traditional internet security systems; they are an addition to your network security systems. Honeypots can be setup inside or outside of a firewall design or any strategic location within a network. In a sense, they are variants of standard Intrusion Detection Systems (IDS) but with more of a focus on information gathering and deception. Honeypots are deployed on an unused IP address, which is monitored by the administrator. This decoy system is waiting for attackers to start an interaction with the system. Any type of interaction with the honeypot is considered as suspicious. The main goal of this system is to gather as much data as possible in a manner that will protect the system and network from future attacks and thus remove any computer as well as network security loopholes.

A. Honeypot Types

i) Purpose of Honeypot

These are specific to the area of deployment.

Research Honeypot

Research honeypots are difficult to deploy and maintain. Their sole purpose is to extract information about intruders, attackers their methods and tools.

Production Honeypot

Production honeypot these are designed for directly enhancing system protection. They provide real time security by slowing down an attack on real system targets.

ii) LEVEL OF INTERACTION

Honeypots are categorized into three types depending upon the level of interaction.

1. Low-interaction Honeypot

Low-interaction Honeypot does not contain a real-time system. They are used for gathering information and low interaction honeypots can’t be used to utilize the full potential of a honeypot. These type of honeypots are easy to deploy and maintain.Honeyd10 is one ofthe low interaction Honeypot.

2.Medium-Interaction Honeypot

These type of Honeypots give an illusion of a false operating system with which the attack can communicate. Thus capturing all the attackers’ activities. Honey trap is a type of medium action Honeypot.

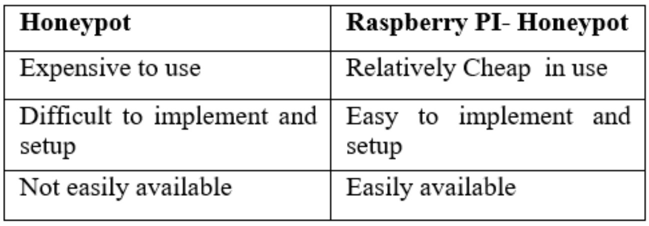
3. High-Level Of Interaction Honeypot

These are the most advanced honeypots, which are complex and difficult to setup. These type of honeypots have their own OS. Then the risk of deploying is high. Honey net is an example of this type of honeypot. It is a combination of decoys all working as one with different interaction level.

iii) Hybrid Honeypot

Monitoring malicious activities in computer systems is very complex and expensive. Using a Raspberry Pi in a network makes the network administrators work less complex and easy to implement. Described form of protection provides use of advanced security method called Honeypot along with a Raspberry Pi.

Difference between Honeypot and Raspberry Pi Honeypot



Honeypot	Raspberry PI- Honeypot
Expensive to use	Relatively Cheap in use
Difficult to implement and setup	Easy to implement and setup
Not easily available	Easily available

B. Raspberry Pi-Honeypot Advantages and Disadvantages

I) Using The Raspberry Pi-Honeypot Has Some Significant Advantages:

1. Cost Effective- As Raspberry Pi are very cheap and easily available, also setting up a Raspberry Pi is very easy.Hence setting up a Raspberry Pi-Honeypot in a network becomes easy.
2. Simple – Honeypots do not require any complex operation or algorithm for deployment. They are flexible.
3. Record new tactics – they capture all interaction with the intruder and discover new tactics.
4. Data – They produce high quality data.

ii)Honeypot Technology Also Has Its Drawbacks

1. Gain control-attacker can gain control of a honeypot and retrieve all the information.
2. Divulge identity – An experienced attacker can detect presence of incorrectly configured system acting as a decoy.

V. RASPBERRY PI

The Raspberry Pi is a low cost, credit card sized computer that plugs into a computer monitor or TV, and uses a standard keyboard and mouse. The Raspberry Pi has the ability to interact with the outside world; it plugs into a computer monitor or TV and uses a standard keyboard and mouse. It uses programming language like Scratch and Python. Low power consumption with headless setup. It can simply turn into a powerful Honeypot or attack detector.

VI. USAGE OF RASPBERRY PI-HONEYPOT WITH AN INTRUSION DETECTION SYSTEM

Proposed architecture deals with implementing a Raspberry Pi-Honeypot with Snort IDS. Thus a solution to minimize failures in detection process and collection of important data based on honeypot consists of combining security tools: Snort IDS, Modern Honeypot Network. This detection mechanism based on Raspberry Pi-Honeypot is implemented as a client server architecture. It has a central main sever interacting with multiple clients in the network. Client work station serve to capture suspicious activities or directly record the malicious code which is then sent to server for processing. Server analyses received data decides to issue or not to issue a security warning and display cumulative information through a web interface.

A. Server Architecture

Due to centralization of collected data the server is connected to multiple clients and is set to receive all incoming messages which are stored in knowledge database. The proposed server architecture consists of:

1. Modern Honeypot Network (MHN): You can observe and control the honeypot from a central location.
2. Verification Process: Receive the amount of data from client and integrates diversified data formats.

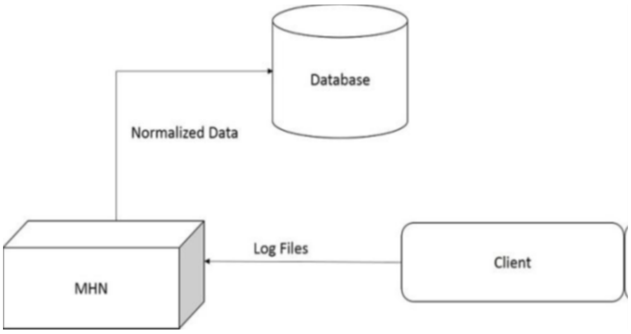


Figure 1. Server Side Architecture

B. Client Architecture

This architecture consists of Raspberry Pi-Honeypot which captures all the attackers' activities .The data is delivered to the server for further analysis and updating network security.

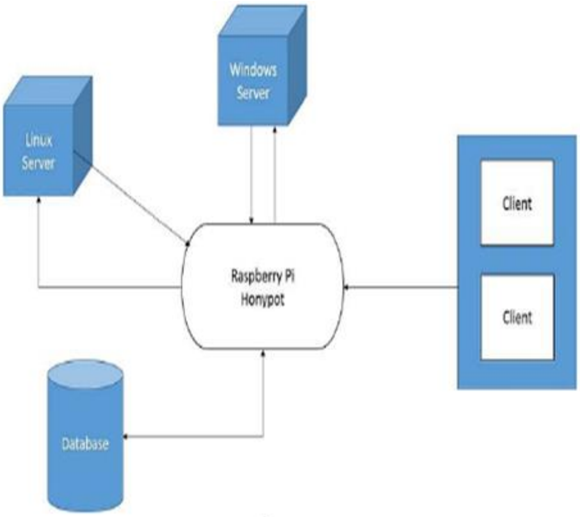


Figure 2. Client Side Architecture

Client architecture consists of:

1. Kippo: it is a SSH Honeypot tool written in python that will log brute force attacks and shell interaction performed by the attacker.
2. Dionaea: will capture the patter malware by simulating basic system services and vulnerabilities.
3. Glastopf: it is a web application Honeypot, it gathers data by emulating thousands of vulnerabilities. Unlike many other honeypots, Glastopf focuses on replying the correct response to the attacker exploiting the targeted Web application, and not the specific vulnerability.
4. Snort: Intrusion detection system that monitor has and filter packets during detecting intrusion.

VII. RASPBERRY PI-HONEYPOT

This proposed Honeypot is developed as a separate device (Raspberry Pi) physically present in the network. It will be deployed with Dionaea or Glastopf or Kippo which will collect all the data and send it to the server. Raspberry Pi-Honeypots can merge in any environment making them more difficult to identify and reveal. Deployment of multiple Raspberry Pi-Honeypots are easy and affordable.

VIII. CONCLUSION

The usage of Raspberry Pi-Honeypot as a decoy in the network represents a simple and an efficient solution for enhancing network security using raspberry pi and open source tools. Deployment and management of raspberry pi as a honeypot is cost effective and also provides easy integration. The support of this work is to introduce a new and cost effective mechanism for network security. This mechanism combines the security tools in order to minimize the disadvantages and maximize the security capabilities in the process of securing the network.

IX. REFERENCES

[1] LiberiosVokorokos, Peter Fanfara, JánRadusovský and Peter Poór,Sophisticated Honeypot Mechanism - the Autonomous Hybrid Solution for Enhancing Computer System Security, SAMI 2013 IEEE 11th International Symposium on Applied Machine Intelligence and Informatics, January 31 - February 2, 2013, Herl'any, Slovakia.

[2] R. Chandran, S. Pakala, Simulating Network with Honeyd, Technical Paper, Paladion Networks, December 2003.

[3] Article Title: http://www.snort.org

[4] https://www.zeltser.com/mpdernhoneynetworkexperiments/

[5] Article Title: http://bob.k6rtm.net/kippo.html

[6] L. Spitzner, Honeypots: Tracking Hackers, Boston, USA: Addison- Weasley, Parson Education, ISBN 0-321-10895-7, 2003.

[7] L. Spitzner, The value of Honeypots, Part One: Definitions and value of Honeypots, Security Focus, 2001.

[8] S.Karthik, B.Samudrala, A.t.Yang,Design of Network Security Projects using Honeypots Journal of Computer Sciences in Colleges, 2004.

[9] http:/www.raspberrypi.org/help/what-is-a-raspberry-pi/.

[10] E. Dankova et al.,An Anomaly-Based Intrusion Detection System, Electrical Engineering and Informatics 2,Kosice,ISBN 978-80-553-0611-7,2011.

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3| ISSN : 2456-3307

Network Component Development for Xml Migration

Ms. A.Sivasankari[1], Ms. D.Janani[2], Ms. G.Arunkumari[3]

[1]Head of the Department (CS), Dept. of Computer Science and Applications, D.K.M College for Women (Autonomous), Vellore, Tamilnadu, India

[2]Research Scholar, Dept. of Computer Science and Applications, D.K.M College for Women (Autonomous), Vellore, Tamilnadu, India

[3]Assistant Professor, Dept. of Computer Science and Applications, D.K.M College for Women (Autonomous), Vellore, Tamilnadu, India

ABSTRACT

This Network Component Development for XML Migration is developed for Company. In the cutting edge Internet, age a large portion of the business is being directed through the Internet. As a result, space comes at a premium, more space that is needed the more it costs. In the existing scenario a client wants to store data in the web server has to rent space in the database server of the deploying agency as well as paying for the web space in the web server. This system is not economical and furthermore the client has lesser control over the data as the Database Administrator belongs to the ISP. To counter these disadvantages the proposed system implements the changes as follows, when a user submits data to the deploying agency instead of storing it in a separate database server the data is stored in a flat file format like XML in the web server space itself, thereby eliminating the cost of renting a database server completely. The client can then login to the web server and downloads the necessary XML files that are stored in the local database. This provides an added advantage in that the client can have better control over the data as they can have their own Database Administrator.

Keywords: Internet Service Provider (ISP), Data Base Administrator (DBA), Extensible Markup Language (XML), Electronic Data Interchange (EDI)

I. INTRODUCTION

The system will produce data as per XML format. This will be a web application hosted on standard web server. The network component checks the validity using a Login. The Username and password are stored in a separate database; the given password is checked for authentication against the user name. If there is a match, the user is allowed to enter in to the session, failing which he is denied permission.

If by chance a phonetic error occurs, the user has to re-login to establish his identity. New users need not undergo this password check, the very first time they try this. However user names and password are registered, the very first time itself and any future login needs a valid username and password. Once logged in, then they both can use EDI standard to grab the data residing in the web application or server using socket programming.

Here the information is saved in XML files which will be residing in the application tier. XML supports an elegant way of storing the data in a tag format as per the specification. This completely avoids the need for having the database in server. The networking component developed will download the XML files over the internet, parse the data available in XML files and migrate the data to the database, which will be available locally.

For the past couple of years Internet applications have relied on the databases installed in the server for storage of information. An Internet application, which requires persistent data, uses a database provided by the hosting agency (ISP).

Usually clients who use the Internet solutions will have Database servers running in their local Intranet. In other words, if a client wants to store some data in the deploying agency he has to buy space with the database server of the deploying agency. It will be very efficient if the company can use the existing databases, since consolidation of data can be made quickly. Also in the administration point of view, it is very hard for the company to maintain the data.

The Internet plays a major role in achieving the centralization of data and applications. This provides a mechanism for accessing the same resource from different locations all over the world. The major drawbacks of existing database oriented Internet applications have to be eliminated by some alternative solutions.

II. NEED FOR THE SYSTEM

In the existing system, database has to be present in the middle-tier where the Internet application is deployed.

The existing system has the following disadvantages:

a. For the user this is very costly as deploying agency will be billing a huge amount for the usage of database.

b. Moreover user is not having full control over the database and it is not possible to do DBA level administration. So data consolidation (Data warehousing) is not possible.

Today's advancement in the field of data formatting with a clear separation between presentation and data has made the IT industrialists to rethink on the existing solutions and proposing a new system wherein actual data is saved in the applications tier in the file formats like XML. By using EDI standards this data can be transferred to any other location.

III. PROPOSED MODEL

With the rapid development of the Internet services there is a far reaching need for a low cost and efficient data storage mechanism that can support quick data transmission and management. This solution provides a better utilization of network bandwidth and also reduces the traffic by avoiding congestion. Here the information is saved in XML files which will be residing in a tag format as per the specifications. This is very efficient since storing the data into a flat file like XML file will be very quick. This completely avoids the need for having a database in the server. The data will be downloaded as XML file and the file will be parsed for the data and is then migrated to the local database.

Summarizing, the proposed system aims to achieve the following goals:

a. Eliminating the need for buying a separate database server space from the deploying agency thereby reducing the cost.

Have a better control over the information by transfer of data and subsequent migration of data to the local database.

IV. PROBLEM FORMULATION

This enables persistent data storage for an internet application in the local machine. The utility when implemented for intranet can also relieve the overhead of maintaining a database in an application server which is costly and difficult to manage over the network. At the same time, the utility can be used as a means of having a local database thereby avoiding the overhead of maintaining a remote DBA. So, a network component is developed for downloading the XML document that contains data from the server, then to extract actual data from document using XML Parser and Transfer the data from XML file to RDBMS table residing on the local client machine.

When a user submits data to the deploying agency instead of storing it in a separate database server the data is stored in a flat file format like XML in the web server space itself, thereby eliminating the cost of renting a database server completely. The client can then login to the web server and downloads the necessary XML files that are stored in the local database. This provides an added advantage in that the client can have better control over the data as they can have their own Database Administrator.

V. USER DESIGN AND DEVELOPMENT STAGES

The following are the various design and development stages:

a. User authentication
b. Building the XML data with DTD
c. Initializing & Establishing connection with server
d. Data transfer from server to local machine
e. XML data parsing
f. Migration of the parsed data to a local database
g. Connection termination
h. Maintenance

A. User Login

The network component checks the validity using a login form. The username and password are stored in a separate database. The given password is checked for authentication against the username. If he is an authorized user he is allowed to enter a session. Once logged in, they both can use EDI standard to grab the data residing in the web application or server using Socket programming.

B. Data Grabbing As Per EDI Specification

Java socket Programming allow the utility to transfer the data from a remote application server to the local machine. Data in the server will be in the form of XML files which will be validated as per the DTD for the application which resides in server.

C. XML Parsing Using SAX Parser

Once the data is grabbed into the local machine it has to be parsed to extract the data inside.

D. JDBC Connectivity for migration of the data to RDBMS

The extracted data has to be ported to a database. This will be along using JDBC. This Software is a start to the next generation of software that separates the data and presentation.

VI. IMPLEMENTATION

A. Steps to Store the User Data

This describes the series of steps to store the user data into the web server in XML format.

a. User data to web page
b. Converted to XML file in specific format
c. Stored in web server

B. Steps Performed By the Administrator

This describes the series of steps performed by the Administrator

a. User authenticated to web server
b. Receives list of files stored in the server
c. Download the necessary files
d. Parse with XML parser
e. Stored in local database

C. Transferring data and conversion to XML

a. User enter data to web browser
b. Web browser transfer data into ISP server

c. ISP server converts data into XML files
d. XML files store company's web space

D. Connection establishment by client

a. Company's DBA request connection ISP server
b. ISP server establish connection

E. Database Creation

a. Company's DBA generate database tables

F. File download and database update

a. Company's DBA downloads XML files

XML files update query into local database

VII.CONCLUSION

This application "Network Component development for XML migration" will produce data as per XML format. This will be a web application hosted on standard web server. The network component checks the validity using a Login. Here information is stored in XML files which will be residing in the application tier. XML supports an elegant way of storing the data in a tag format as per the specification. The system is very efficient since storing the data into a flat file like XML file will be very quick. This completely avoids the need for having the database in server. The networking component developed will download the XML files over the Internet, parse the data available in XML files and migrate the data to the database which will be available locally.

Low-cost and efficient data storage mechanism that is implemented in this system can support quick data transmission and management. This solution provides a better utilization of network bandwidth and also reduces the traffic.

Future enhancements for this XML component can also be created for other tasks by which the performance and can be better utilized and by using this, the maintenance can be done efficiently.

VIII.REFERENCES

[1] Gomez P. and Zadrozny P, "PROFESSIONAL Java 2 Enterprise Edition", Tata McGraw Hill Publishers, 2000.

[2] Schildt, H. "JAVA2-The Complete Reference", V Edition, Tata McGraw Hill Publishers, 2002.

[3] Weaver. B, "Beginning Java XML", Tata McGraw Hill Publishers, 1999.

[4] HTML Black Book by Holzner, HTML 4 for Dummies Quick Reference by Eric Ray

[5] Instant XML Programmer's Reference by Trevor Jenkins

[6] Software Engineering by Roger Pressman

[7] Software Engineering, A practitioner's Approach by Roger Pressman, McGraw Hill International Edition, 6th Edition

[8] Software Engineering by Somerville, Pearson Education, 7th edition

[9] Software Engineering by K.K.Aggarwal&Yogesh Singh, New Age International Publishers

[10] Software Engineering, An Engineering Approach by James F.Peters, WitoldPedrycz, John Wiley

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3| ISSN : 2456-3307

Variety of Cloud Computing and Authentication with Artificial Intelligence

V. Jeevitha[1], G. Pavithra[2], R. Valarmathi[3]
[123]Research Scholar, Department of Computer Science, Shanmuga Industries Arts and Science College, Thiruvannamalai, Tamil Nadu, India

ABSTRACT

Cloud made it easy for an organization to increase its capability without actually adding new infrastructure, new software or updating existing technology; as it is Internet based system for providing services to the end users on pay per usage basis. Cloud computing reduces cost of computation & storage to a large extend and also improves productivity. From few days cloud has grown from a promising business application to fastest growing IT industries. Cloud offers services such as storage, computation etc for different types of markets such as health care, net banking, several government organizations and other financial applications. Now many popular educational institutes and enterprises are also getting their applications and data shifted to the cloud. And discussing the Authentication of cloud and Artificial Intelligence integrated. In the last sections major benefits and downsides of cloud computing has been discussed.

Keywords : Cloud Computing, Security, Services, transparency, Private, Public, Community, Hybrid, Multiple Cloud, Cloud Authentication, Artificial Intelligence, Cloud Benefits

I. INTRODUCTION

A Cloud is a type of distributed and parallel system that consists of a collaboration of inter-connected and virtualized computers that are dynamically presented and provisioned as one or more unified computing resource(s) based on service-level agreements established through the negotiation between the consumers and the cloud service provider.”

Saving your document to the Internet rather than saving it to computer memory is cloud computing. This will facilitate access to it from anywhere and through any device connected to the internet. A paradigm in which information is stored permanently and also replicated in servers on the Internet by the expert cloud providers and cached temporarily on clients that include entertainment centers, desktops, table computers, wall computers, notebooks, handhelds etc. In short it is model for convenient on-demand network access to sharable & configurable computing resources, such as information, services, applications, storage, and networks that can be easily released and provisioned with minimal service provider’s interaction. Figure 1 shows how cloud computing can be used to access applications and data from any of the network devices.

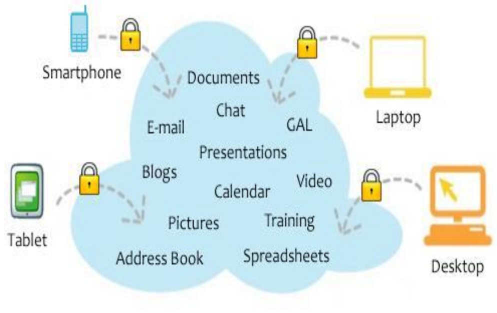


Figure 1: Applications & data access from cloud through any network device.

A. Characteristics of Cloud Computing

Cloud computing implies four main characteristics as follows:

- The end user has “no-need-to-know” about the internal details of the cloud infrastructure. The application itself interfaces with it through the API (Applications Programming Interface).
- The cloud provides “elasticity and flexibility” to the users to scale up and scale down in utilizing resources of all kinds (server capacity, databases, storage, load balancing etc.) according to their requirements.

The cloud offers “Anywhere and always on” type of network based on the computing and the “pay as much as used and needed” type of utility computing to its customers.

II. CLASSIFICATION OF CLOUD COMPUTING

The cloud customers can access data, applications, software, servers and heterogeneous platforms.

A. Private Cloud:

This type of cloud is rented and owned by an organization. The organization uses cloud resources for its private use only. These types of special clouds are personally built by an enterprise for serving their critical business processing needs.

B. Public Cloud:

In this type of cloud all the resources are owned by cloud provider and they sell the resources to public on demand. End users can rent required resources and pay as per usage. Google, Amazon, Salesforce, Rackspace and Microsoft are some main examples of public clouds.

C. Community Cloud:

It is another type of Private cloud. But here cloud resources are shared among the members of a closed community having same resource requirements and interest. The Media Cloud is the example of community cloud setup by Siemens IT Solutions and Services. This type of community cloud may be operated by collaborate efforts of all or by a third party alone.

D. Hybrid Cloud:

It is the collaboration of two or more above mentioned cloud infrastructures (private, community, or public). The sole purpose of hybrid cloud is to provide extra services and resources to end users to serve their high demands.

Advantages:

- Security , Cost efficiency
- Scalability, Flexibility
- Preservation of investments

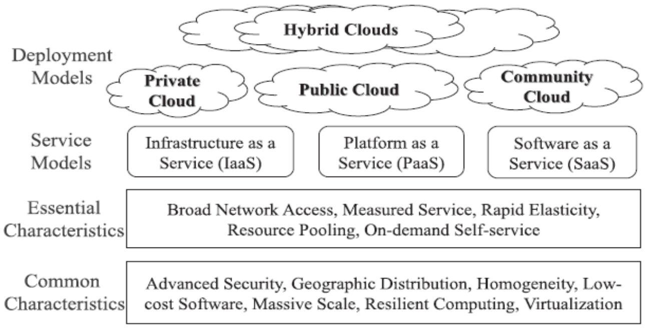

Hybrid Clouds
Deployment Models
Private Cloud
Public Cloud
Community Cloud
Service Models
Infrastructure as a Service (IaaS)
Platform as a Service (PaaS)
Software as a Service (SaaS)
Essential Characteristics
Broad Network Access, Measured Service, Rapid Elasticity, Resource Pooling, On-demand Self-service
Common Characteristics
Advanced Security, Geographic Distribution, Homogeneity, Low-cost Software, Massive Scale, Resilient Computing, Virtualization

Figure 2: The cloud definition framework by NIST

E. Multi- Cloud

“Multi-cloud” describes an environment that relies on multiple clouds such **as OpenStack®, Microsoft® Azure® or AWS.**

For instance, may be running a workload that requires large pools of storage and networking resources on a private cloud, such as OpenStack.

At the same time, you may have a workload that needs to scale up or down quickly on a public cloud, such as Microsoft Azure or AWS. Each workload is running on the ideal cloud, but now you have multiple clouds to manage. With that in mind, let’s look at why CIOs are pursuing multi-cloud strategies, often in concert with their hybrid cloud approach.

For many, it's about more fully realizing the powerful potential of cloud and giving IT teams increased flexibility with and control over their workloads and data.

i) Types of Multi Cloud Web Services

- OpenStack
- Microsoft Azure
- Amazon
- VM ware

"Multi-cloud strategy allows an organization to meet specific workload or application requirements – both technically and commercially – by consuming cloud services from several cloud providers, "Not every department, team, business function, or application or workload will have similar requirements in terms of performance, privacy, security, or geographic reach for their cloud. Being able to use multiple cloud providers that meet their various application and data needs is critical as cloud computing has become more mature and mainstream."

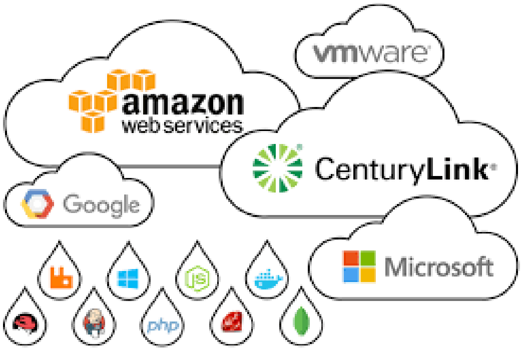


Figure: 3 Types of Web Services

"Business units may begin using a cloud provider for a particular project, then IT will need to fold use of that provider into an overall cloud plan."

ii) Work with Multi-cloud

IT may see geographic benefits to using multiple providers, to address app latency concerns, for example. But another reality is that some business units may begin using a cloud provider for a particular project, then IT will need to fold use of that provider into an overall cloud plan.

Additionally, vendor lock-in concerns and possible cloud provider outages are two issues that pop up frequently when IT leaders advocate for multi-cloud strategy.

"Multi-cloud strategy can be an enabler for preventing vendor lock-in, a means to avoid single points of failure and downtime, or simply a mechanism to consume unique innovations from several providers.

A multi-cloud strategy – which is always means "multi-vendor," too – as a way of mitigating vendor lock-in risks. But that's actually a secondary benefit. The real advantage driving multi-cloud strategies is greater flexibility and agility to adapt to the breakneck pace of modern business.

"If your business needs change, your cloud can change with them" with a multi-cloud strategy, It's not just a business enablement strategy, either. It's also an IT-forward strategy. "Technology and cloud are changing so rapidly, and [they are changing] a lot. If you are less locked down, you will be able to grow with technology. You will be able to grow with the cloud. You will have a lot more options and flexibility. It's a really good business case."

Figure 4: Multiple Cloud

F. Cloud Authentication

Cloud computing is the new way to interact with device, software, data and processes. Needed true things across old and new computing paradigms is "AUTHENTICATION".

Authentication forms the basis for Security in Cloud Computing Network. Private, Public, Hybrid clouds are adding yet another layer of Complexity. When user need a resource to continuum their work to complete in cloud, Sometimes it may cause difficulty due to unauthorized access by Theft or Hacker for a same Resource. Authentication can control all the IT Resources Who can access and When they need to Access .Authentication for a user can provided by 2 Ways.

1) Using Local Credential
2) Using Active Directory Credential

When a User Log in to machine, if a machine is not joined with AD then Username and Password can be Validated by Local Credential. When a User Log in to a machine, if a machine is joined with AD then Username and Password can be Validated only when they match information which is stored in Database by Active Directory Credential.

Whereas AD is a Active Directory is a Directory Service that Microsoft developed for Windows Domain network. AD service consists of Mulitple Directory Services. The Best known Active Directory Domain Services commonly abbreviated as ADDS or simply AD.

i) Two Factor Authentication

Two Factor Authentication is Processed by Two things
1) Something you know (username & password)
2) Something you have(Authentication)

Two Factor Authentication Technology helps to protect user to Login securely in corporate Environment. Two Factor Authentication is also known as Two Factor Security or Two Step Verification. It reduce number of Incidents which is Processed by Unauthorized user.

In Organization /Company, the User can store personal details and sensitive financial Information in the System that can be secured only Two Factor Authentication otherwise it can be easily hacked by unauthorized user. Two Factor Authentication can enable for Gmail, FaceBook, Apple, Twitter, Outlook, Yahoo Accounts.

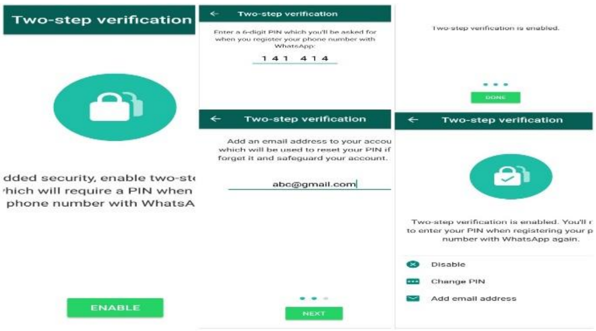
In FaceBook we Can use Two Factor Authentication without using phone

- Backup Phone
- Backup Code
- Register Your Device

ii) Example for two factor authentication

Whatsapp is also One of the Best example in Two Factor Authentication. In Whatsapp when we enable the option of "Two-step Verification" we can secure our data from unauthorized person who known your number,(i.e) unauthorized Person can enter your Mobile Number in their Mobile Whatsapp and they can restore your messages easily.

If we create a Pin Number using Two-Step Verification in Whatsapp setting Nobody can missuse our data.(i.e)After we use this authentication method,If unauthorized Person can enter your Mobile Number in their Mobile Whatsapp and then second stage of creating whatsapp account will ask you a PassCode, that PassCode can get through Only phone call who is using that mobile number[Authorized Person who is Using that number] so they cannot restore/Backup your messages.



dded security, enable two-st
hich will require a PIN when
phone number with WhatsA

Enter a 6-digit PIN which you'll be asked for when you register your phone number with WhatsApp:

1 4 1 4 1 4

Two-step verification is enabled.

Add an email address to your accou which will be used to reset your PIN if forget it and safeguard your account.

abc@gmail.com

Two-step verification is enabled. You'll r to enter your PIN when registering your p number with WhatsApp again.

Figure 5: Two Factor Authentication

G. Artificial Intelligence

Artificial Intelligence is like an iceberg, there is a lot hidden than what is visible. The true potential of AI is yet to come out. The way AI and cloud computing is changing the landscape of corporate world; it is believed to be the future of technology. Artificial Intelligence has the potential to further streamline the immense capabilities of cloud computing. Artificial Intelligence equips cloud computing with tremendous power. It enables machines to learn, think, act, and react like human beings. AI helps machines to analyze and learn from the historical data, identify patterns and make real-time decisions. This will lead to process automation which will eradicate the possibility of human errors.

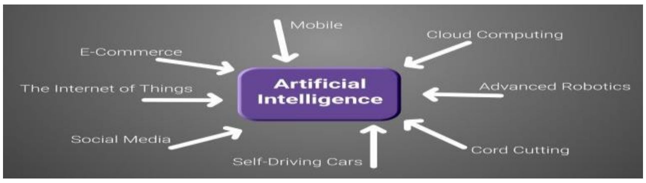


Figure 6: Artificial Intelligence with cloud

The combination of cloud computing and AI has bought a major change in the world of information technology and various other industries and it is seen as the way forward. It has the potential to change the way data used to get stored and processed across various geographies.

The combination of cloud computing and AI also presents a unique opportunity for cloud and artificial intelligence professionals to explore the endless possibilities for future. Looking at the current trend in the growth of cloud and AI, one thing is for sure that there is going to be a tremendous demand of trained professionals in these fields. It is going to be an amazing opportunity for IT professionals who have just started their careers, if they want to make a career in the technology for the future. They can easily get trained and certified on cloud computing and artificial intelligence.

Figure 7: Future Technology of AI

H. Artificial Intelligence in Future Technology

The cloud technology can help AI's by providing the required information for the learning processes while the AI can help cloud by providing information that can offer more data. AI is capable of streamlining the immense capacities of the cloud. It equips cloud technology with enormous powers. It enables the machines to act, react, think and learn in the manner human beings do. AI assists different machines in learning and analyzing the historical data, making decisions and identifying the patterns. Such a process helps in eradicating the chances of human errors. Therefore, AI enhances the process of decision making of various organizations.

Cloud technology is spread among a number of servers in various languages with huge data storage and across various geographies. Organizations can make use of this data to make up intelligent and automated solutions for customers and clients. Cloud computing is getting more powerful with AI as its applications are extended across multiple diversified sectors in the economy. Thus, even organizations can make use of AI cloud computing to attain long-term goals for their businesses.

III. CLOUD COMPUTING SERVICE MODELS

All the cloud resources are provided as services to the end users. The service models of cloud computing are mainly Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS).

A. Software as a Service (SaaS):

All the applications running on the cloud are provided as the services to the end users. This eliminates the software up-gradation and software licensing investments for the clients. On the other hand cost of the cloud is rather low. Cloud also delivers business applications such as accounting, enterprise resource planning (ERP) and customer relationship management (CRM). The SaaS cloud's example includes Salesforce CRM and Google Apps.

B. Platform as a Service (PaaS): This service allows user to built applications using several software tools along with programming languages (e.g. Java, .Net, Python) and also deploy user's applications onto the cloud infrastructure. The user needs not to manage the cloud infrastructure, operating system and other requirements for them. The PaaS cloud's example includes Google App Engine and Microsoft Windows Azure

C. Infrastructure as a Service (IaaS):

By this user can use storage, network, servers, processing and other resources on rental basis. The user can run and deploy the applications and guest OS. The user does not control or manage the infrastructure but has control over applications, OS, storage etc. The PaaS cloud's example is Amazon EC2.

IV. HOW CLOUD COMPUTING WORKS

As an organization recruits new employees they also need to purchase computers, software or software licenses for them. They also need to check whether current software license allows another user. But with cloud they only need to install an application for each new hiring. By this application workers can log into the cloud; as the cloud is hosting all the relevant programs for their jobs. Cloud is owned by another company called the cloud service provider. Cloud provides the shifting of workload from the user's computer to a remote application and also reduces software & hardware demands on user's side. The user only needs to run the system's interface software of cloud. The cloud system includes several storage servers and a master control server. Using cloud information is stored at a remotely located database owned by a third party (i.e. cloud provider) instead of your computer's hard drive. The internet serves as a medium between the user and the cloud.

A. General Cloud Computing Architecture:

Cloud provider needs to maintain quality parameters as negotiated in SLA. Many critical QoS parameters are considered for a service request, such as cost, time, trust/security and reliability.

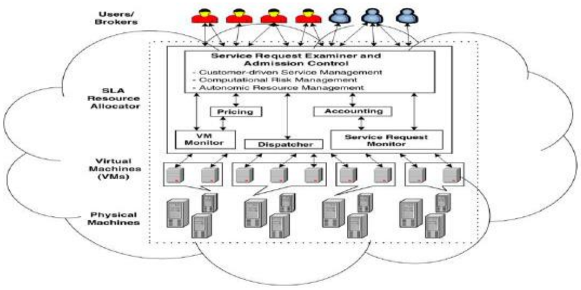

Users/ Brokers
SLA Resource Allocator
Service Request Examiner and Admission Control
- Customer-driven Service Management
- Computational Risk Management
- Autonomic Resource Management
Pricing
Accounting
VM Monitor
Dispatcher
Service Request Monitor
Virtual Machines (VMs)
Physical Machines

Figure 8: A typical cloud computing system architecture

B. Requirements for Cloud Computing Implementation

Cloud delivers services in an on-demand environment. Several applications supported by the cloud must be secure, fast and always available. For this, they need to build a dynamic and intelligent cloud infrastructure with four core properties in mind.

- Transparency
- Scalability
- Intelligent Monitoring
- Security

V. BENEFITS OF CLOUD

- Anytime & Anywhere access
- Transferring the Risk
- Online Editing
- Online collaboration
- Location and Device independence
- Increased pace of innovation & Environmentally Friendly
- Recovery & Backups

VI. DOWNSIDES OF CLOUD

- Service Availability
- Data mobility and ownership

- Privacy
- No direct control
- Security issues
- Cost
- Identity Management
- Inflexibility

VII. CONCLUSION

Cloud offers resources sharing in a cost effective and independent way. Through cloud providers are sharing their resources and capabilities with external users on rental basis. Surely, many organizations are benefitting from Cloud computing, as cloud provides facility to run OS for several servers on Virtual machine. Apart from multinational organizations several small enterprises and educational institute are also using cloud services. There are several risks involved in the cloud. The problems discussed in this paper have made adaptation of hybrid and public cloud difficult. Multi Hybrid Cloud, Authenticate the cloud by two factor method and using cloud in future by artificial Intelligence of all cloud computing major problems are findout and clearing the ideas.

ACKNOWLEDGEMENT

We wish to express our deep sense of indebtedness and sincerest gratitude to all the experts and Mr. U. Udhayakumar for his valuable guidance and appreciation. We also wish to thank our Parents for their encouragement at every step of progress.

VIII. REFERENCES

[1] R. Buyya, et al., Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility, Future Generation Computer Systems (2009), doi:10.1016/j.future.2008.12.001

[2] National Institute of Standards and Technology. The NIST definition of cloud computing; 2011. <http://www.nist.gov/itl/cloud/upload/cloud-defv15.pdf> [retrieved 14.04.11].

[3] Baek Sung-Jin, Park Sun-Mi, Yang Su-Hyun, Song Eun-Ha, Jeong Young-Sik, Efficient server virtualization using grid service infrastructure, J Inform Process Syst 2010;6(4):553–62.

[4] Rong C et al, Beyond lightning: A survey on security challenges in cloud computing, Elsevier, Comput Electr Eng (2012), http://dx.doi.org/10.1016/j.compeleceng.2012.04.015

[5] Donlin Chen, Mingming Ma, and Qiuyun Lv, A Federation Model for Education under Hybrid Cloud Computing, Elsevier, 2nd International Conference on Future Computers in Education, Vols.23-24, PP 340-343, 2012

[6] Aida Ghazizadeh, Cloud computing benefits and architecture in e-learning, Seventh IEEE International Conference on Wireless, Mobile and Ubiquitous Technology in Education, pp. 199-201, 2012.

[7] https://en.wikipedia.org/wiki/Cloud_computing

[8] I. Foster, Y. Zhao, I. Raicu, and S. Lu, Cloud Computing and Grid Computing 360-Degree Compared, IEEE Grid Computing Environments Workshop, pp. 1-10, 2008.

[9] LIN Yu- hua , Practice and Innovation of an experimental teaching mode in universities under environment of cloud computing, Research and exploration in laboratory, vol. 30, pp. 271-274, 2011.

[10] Siemens IT Solutions and Services. Community clouds: supporting business ecosystemswith cloud computing; 2011. <http://www.it-solutions.siemens.com/b2b/it/en/global/Documents/-Publications/Community-Clouds-Whitepaper_PDF_e.pdf> [retrieved 18.04.11].

[11] Mell Peter, Grance Tim. Effectively and securely using the cloud computing paradigm; 2011. <http://csrc.nist.gov/groups/SNS/cloud-

computing/cloudcomputing-v26.ppt>
[retrieved 18.04.11].

[12] Salesforce. Salesforce CRM applications and software solutions. <http://www.salesforce.com/eu/crm/products.jsp>.

[13] Google, Google Apps. <http://www.google.com/apps/>.

[14] Google. Google App Engine. <http://code.google.com/appengine/>.

[15] Microsoft. Microsoft Windows Azure. <http://www.microsoft.com/windowsazure/>.

[16] Amazon. Amazon Elastic Compute Cloud (EC2). <http://aws.amazon.com/ec2/>.

[17] http://www.howstuffworks.com/cloud-computing/cloud-computing.htm

[18] https://devcentral.f5.com/blogs/us/4-things-you-need-in-a-cloud-computing-infrastructure

[19] http://knowcloudcomputing.blogspot.in/2012/02/grid-and-cloud-computing-technically.html

[20] http://www.javacodegeeks.com/2013/04/advantages-and-disadvantages-of-cloud-computing-cloud-computing-pros-and-cons.html

[21] http://www.thebeckon.com/pros-and-cons-of-cloud-computing/

[22] Mathias Mujinga, Baldreck Chipangura, Cloud computing concerns in developing economies, 9th Australian Information Security Management Conference, Edith Cowan University, Perth Western,Australia, 5th -7th December, 2011

[23] http://venturebeat.com/2012/01/16/the-downside-of-cloud-computing-4-reasons-to-think-twice/

[24] http://sbinfocanada.about.com/od/itmanagement/a/Cloud-Computing-Disadvantages.htm

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3 | ISSN : 2456-3307

Cryptography in Cloud Computing: A Basic Approach to confirm Security in Cloud

E. Manigandan[1], **Prof. C. Kalaiarasi**[2]
[1]Research Scholar, Government Arts College, Thiruvannamalai, Tamil Nadu, India
[2]Assistant Professor, Government Arts College, Thiruvannamalai, Tamil Nadu, India

ABSTRACT

Cloud computing is associate degree Internet-based computing model that provides many resources through Cloud Service suppliers (CSP) to Cloud Users (CU) on demand basis while not shopping for the underlying infrastructure and follows pay-per-use basis. It supports virtualization of physical resources so as to enhance potency and accomplishment of multiple tasks at identical time. Cloud Computing atmosphere (CCE) provides many readying models to represent many classes of cloud owned by organization or institutes. However, CCE offer resources to Cloud Users through many services like PaaS, SaaS, IaaS. Cloud Computing may be a notion supported the construct of summation physical resources associate decreed displaying them as an unacknowledged resource. it's a model for manufacturing resources, for searching for applications, and for manifesto-independent user access to services. Cloud will are available differing kinds, and therefore the services and therefore the applications that probably run on clouds might or might not be provided by a cloud service supplier. There are 2 distinctive cluster of models specifically readying models and repair models. Service models consists of IaaS, SaaS, PaaS. The readying or readying model consists of Public Cloud, non-public Cloud, Hybrid Cloud, Community Cloud .Cloud Computing has scores of distinct properties that create it vital. Privacy looks to be associate degree distinctive concern in cloud .Various sorts of service models beneath cloud computing facilitate varied levels of privacy services. We will get the minimum security in IaaS (Infrastructure as a Service) and most with a SaaS supplier. During this paper, we will focus upon the reviewing and understanding cloud security problems by proposing crypto algorithms and effective measures thus on make sure the knowledge security in cloud. Beside this, we will elucidate a little a lot of concerning some security aspects of cryptography by displaying some privacy problems with current cloud computing surroundings.

Keywords: Cloud Computing, Cryptography, Security Issues Privacy, Security Algorithms, Encryption, Decryption

I. INTRODUCTION

Cloud computing is one among the favoured topics of the present world. net has started driving of these new technologies. net was designed first of all to be robust, however not utterly safe. Distributed applications like these is way vulnerable to attacks. Cloud Computing has all the feebleness related to these net utilization and therefore the further threats arise from the combined, Virtualized and decentralised resources. There are several knowledge privacy issues in cloud computing. Incorrect revelation of a knowledge employed in businesses in cloud to third parties is one among the foremost problems that has been found. Encryption ought to be properly used and therefore the crypto algorithms embody AES, RSA, DES and three DES .In this paper, we tend to describe concerning victimisation crypto algorithms therefore on increase security concern. Cloud Security is ensured by knowledge integrity, Secured knowledge transfer and by Cryptography. There are types of crypto graphical algorithms, which may be enforced therefore on guarantee security within the cloud. The two forms of algorithms are symmetric and Asymmetric encoding key algorithms. symmetric contains algorithms like DES, AES, three DES and Blowfish formula. Asymmetric contains algorithms like RSA, Differ-dramatist Key Exchange. Symmetric key and Asymmetric key algorithms is employed to cipher and decipher the information in cloud.

II. CONNECTED WORKS

a. Within the paper [1] the authors alter the matter of security information of knowledge throughout data transmission. the most issue to worry regarding this paper is that the secret writing of knowledge so confidentiality and privacy are often simply achieved. The algorithmic program used here is Rijndael secret writing algorithmic program at the side of EAP-CHAP.

b. This paper [2] presents a protocol or set of directions that uses the services of a 3rd party auditor or checker not solely to verify and attest the integrity of knowledge hold on at remote servers however additionally in retrieving and obtaining the info back as shortly as attainable in intact type. The most advantage of this theme is that the use of digital signature to assure the integrity of native knowledge. However, the general method is sort of problematic and sophisticated because the keys and knowledge also are encrypted and decrypted severally.

III. CRYPTOGRAPHY: SECURITY PRINCIPLES & ALGORITHMS

Cryptography will facilitate break of day integration of Cloud Computing by increased range of privacy connected corporations. the first level of privacy wherever cryptography will facilitate Cloud computing is safe and secure storage. Cryptography is that the science of storing messages firmly by changing the information into forms that isn't decipherable. In today’s world cryptography is taken into account as a group of three algorithms. These algorithms area unit Symmetric-key algorithms, Asymmetric-key algorithms and Hashing. In Cloud computing, the most issues area unit associated with drawback in information security, backup information, network traffic, file storage system, and security of host, and cryptography alone will solve these problems to extents. For a secure and secure communication between the guest domain and therefore the host domain, or from hosts to management systems, coding technologies, like Secure hypertext transfer protocol, encrypted VPNs, TLS, Secure Shell, and then on ought to be used. Coding can facilitate United States forestall such exploits like man-in-the-middle, spoofed attacks, and session hijacking. Cloud computing provides purchasers with a computing facilities or infrastructure on prime of that they'll store information and run applications. whereas the benefits of cloud computing area unit pretty clear, it introduces new security challenges as cloud operators area unit purported to manipulate information for purchasers while not essentially being totally trustworthy . we area unit going to be making an attempt to style crypto graphical primitives and protocols that are tailored to the setting of cloud computing, trying to strike a balance between security, potency and practicality. Cloud information storage enhances the danger of outflow of knowledge and doesn't offer access to unauthorized users. Cloud information management can't be totally trust worthy by information house owners. Cloud information method and computation might expose the privacy of users, owning the information or connected entities to parities that doesn't have unauthorized access. For overcoming the higher than issues, cryptography has been wide applied to make sure information security, privacy and trust in cloud computing.

A. Symmetric key algorithms

Symmetric uses single key that works for each encoding and decoding. The isobilateral systems offer a two channel system to their users. It ensures authentication and authorization. Symmetric-key algorithms are those algorithms that uses just one and solely key for each. The key's unbroken as secret. isobilateral algorithms have the advantage of not taking in an excessive amount of computation power and it works with terribly high speed in encoding. Symmetric-key algorithms are divided into two types: Block cipher and Stream cipher. In bock beer cipher input is taken as a block of plaintext of fastened size counting on the kind of symmetric encoding algorithmic program, key of fastened size is applied on to dam of plain text so the output block of identical size because the block of plaintext is obtained. just in case of stream cipher one bit is encrypted at a selected time. Some widespread Symmetric-key algorithms utilized in cloud computing includes: encryption normal (DES), Triple-DES, and Advanced encoding normal (AES).

a) Advanced encoding normal (AES)

In cryptography, the Advanced encoding normal [3] is kind of symmetric-key encoding algorithmic program. Every of the ciphers incorporates a 128-bit block size and having key sizes of 128, 192 and 256 bits, severally. AES algorithmic program assures that the hash code is encrypted in an exceedingly secure manner. AES incorporates a block size of 128 bits. Its algorithmic program is as follows: Key growth, Initial spherical - spherical Keys are other. Rounds, Sub Bytes a non-uniform substitution step wherever every computer memory unit is substituted with another in keeping with a table. Rows are shifted a transposition step wherever every row of the state is shifted cyclically a precise range of steps. Columns are mixed a intermixture operation that operates on the columns of the state, combining the four bytes in every column eight. Add spherical Key each computer memory unit of that exact state is combined with the spherical key; every spherical key's derived from the given cipher key employing a key schedule. Final spherical, Sub Bytes, Shift Rows, Add spherical Key. The DES algorithmic program was finally tame 1998 employing a system that prices concerning $250,000.Triple DES clothed to be too slow for potency because the DES algorithmic program was developed for mid-1970’s hardware and didn't manufacture economical and effective computer code. Triple DES has thrice as several rounds as DES and is correspondingly slower.

b) Encryption normal (DES)

The info encoding normal (DES) may be a block cipher and comes below isobilateral key cryptography. found in Jan 1977 by the National Institute of Standards and Technology, named as authority. At the encoding web site, DES merely takes a 64-bit plaintext and creates a 64-bit cipher text, at the decoding method, it takes a 64-bit cipher text and creates a 64-bit plaintext, and same fifty six bit cipher key's used for each encoding and decoding. The encoding method is formed victimisation two permutations (P-boxes), that we tend to decision initial and final permutation, and sixteen Fiestel rounds. every spherical uses a special kind of 48-bit spherical key that is generated from the cipher key in keeping with a predefined algorithmic program.

c) Blowfish algorithmic

Program Blowfish conjointly comes below isobilateral block cipher which will be used as a substitute for DES. It takes a variable-length key, ranging from thirty two bits to 448 bits, creating it significantly higher for each domestic and marketable use. Blowfish was designed in 1993 by Bruce Schneider as a free, quick substitute to existing encoding algorithms. Since then it's been verified significantly, and it's bit by bit gaining quality as a powerful encoding algorithmic program. Blowfish is non-proprietary and license-free, and is out there free for all uses.

B. Asymmetric Key Algorithms

It is comparatively a brand new idea not like cruciform cryptosystem. completely different keys area unit used for secret writing and decoding. this is often a property that set this theme completely different than cruciform secret writing theme. every receiver possesses a decoding key of its own, typically cited as his personal key. Receiver must generate associate secret writing key, cited as his public key. Generally, this sort of cryptosystem involves trustworthy third party that formally declares that a specific public key belongs to a particular person or entity solely.

a) RSA Cryptosystem

This cryptosystem is one the initial systems and oldest of uneven cryptosystem. It remains most used and used cryptosystem even currently. The system was fabricated by 3 students named West Chadic Rivest, Adi Shamir, and Len International Journal of applied science and Computing, Adelman and thus, it's termed as RSA cryptosystem. This rule is employed for public-key cryptography and not personal key cryptogram. It's the primary and still most ordinarily used uneven rule. It involves 2 keys specifically a public key and a non-public key. The general public secret's used for encrypting messages and is understood to everybody. Messages encrypted with the utilization of public key may be decrypted solely by victimization the personal key. during this verification method, the server implements public key authentication by sign language a singular message with its personal key, that is named as digital signature. The signature is then came to the consumer. Then it verifies victimization the server's noted public key.

b) Diffie-Hellman Key Exchange

Whitfield Diffie and Martin playwright introduced a key exchange protocol with the assistance of the separate power downside in 1976. during this key exchange protocol sender and receiver can manage to line up a secret key to their cruciform key system, victimization associate unsafe channel. to line up a key Alice chooses a random whole number aE[1;n] computes ga, equally Bob computes gb for random bE[1;n] and sends it to Alice. the key secret's chat, that Alice computes by computing (gb)a and Bob by computing (ga)b. The necessary ideas on that the protection of the Diffie-Hellman Protocols defend upon DDH, DHP, DLP like etc,.

C. Hashing Algorithms

a) MD5- (Message-Digest formula 5)

A wide used hash perform formula in cryptography with a 128-bit hash price and possesses a variable length message into a fixed-length output of 128 bits. initial the input message is divides up into lump of 512- bit blocks then the message is cushiony so its total length is portable by 512. The sender of the info uses the general public key to code the message and therefore the receiver uses its personal key to decode the message.

IV. SECURITY PROBLEMS FACED BY CLOUD COMPUTING

When it involves privacy and security, cloud is greatly plagued by the threat of that. The folks like the vendors should make certain that the folks victimization cloud doesn't face any downside like information loss or thievery of information. There is an opportunity wherever a malicious user or hacker will get into the cloud by impersonating a legitimate user, there by poignant the completely complete cloud so poignant many of us who area unit victimization the infected or affected cloud. a number of the matter that is visage by the Cloud computing are:

i. Information thievery

ii. Integrity of information

iii. Privacy issues

iv. Loss of information

v. Infected Applications

vi. Precise location of information

vii. Seller level Security

viii. User level Security

The current generation of cloud computing facilities does not offer any privacy against un trusted cloud operators and thence they're not alleged to store vital data like medical records, money records or high impact business information. To handle this we have a tendency to area unit following varied analysis comes that vary from theory to follow. The most use of coding is to produce privacy through abstraction of all helpful data concerning the plaintext. Coding modifies information useless within the sense that one does not get to access it. We are going to be creating algorithms for cryptosystems, which will facilitate to perform a spread of computations on encrypted information, ranging from traditional purpose of computation to the special purpose computations so as to eradicate this downside. analysis on homomorphic cryptography includes work on fully-homomorphic coding, searchable coding, structured coding, useful coding.

a. Proofs of storage

A client can verify whether the cloud operator has tampered with its data using proof of storage. Particularly, this is done without the client storing a copy of the data and without it having to store back any of the data. In fact, the work for the client is negligible no matter how large the data is.

b. Secure Storage system

We have a tendency to try to style cloud storage systems that give privacy, security, integrity of consumer information against Associate in nursing malicious cloud supplier. Systems can give privacy with none loss of potency and higher functioning can got to be taken care of by creating use of latest cryptology encoding techniques like homomorphic encoding, searchable encoding, verifiable computation and proofs of storage and lots of others.

V. CONCLUSION

Cloud computing is growing as a replacement issue and it's the new trend so and lots of the organizations and large corporations are moving toward the cloud however insulation behind owing to some security issues. Cloud security is associate degree final idea which is able to crush the drawbacks the acceptance of the cloud by the large MNCs, corporations and organizations. There are lots of security algorithms which can be enforced to the cloud. DES, Triple-DES, AES, and Blowfish etc. are some symmetrical algorithms. DES and AES are principally used symmetrical algorithms as they're comparatively safer. DES is sort of straightforward to implement than AES. RSA and Diffie-Hellman Key Exchange is that the uneven algorithmic program. RSA and Diffie-Hellman Key Exchange is employed to get encoding keys for symmetrical algorithms in cloud. However the protection algorithms which permit linear looking on decrypted information are needed for cloud computing, which is able to watch out concerning the security of the information. There's an outsized scope of improvement during this field of analysis. We will use cryptography in varied places so as security in cloud. as an example, Cryptography are often used for maintaining cloud information access management, cloud information trust management, verifiable computing, cloud information authorization and authentication and secure information storage. Aside from of these, Lattice based mostly Cryptography and ID based Cryptography are the two vital sectors that is making certain cloud information security in gift world. Still there's lots of analysis to be tired this field.

VI. REFERENCES

[1] Sanjoli Singla, Jasmeet Singh,"Cloud computing security using encryption technique", IJARCET, vol.2, ISSUE 7.

[2] R. Bala Chandar, M. S. Kavitha , K. Seenivasan," A proficient model for high end security in cloud computing", International Journal of Emerging Research in Management &Technology, Vol.5, Issue 10.

[3] Bokefode Jayant.D, Ubale Swapnaja A, Pingale Subhash V., Karane Kailash J. , Apate Sulabha S. ,"Developing Secure Cloud Storage System by Applying AES and RSA Cryptography Algorithms with Role bases Access Control Model", International Journal of Computer Applications, Volume 118-No.12, May2015

[4] Karun Handa, Uma Singh," Data Security in Cloud Computing using Encryption and Steganography", International Journal of Computer Science and Mobile Computing", Vol.4 Issue.5, May-2015, pg.786-791

[5] M.Vijayapriya,"security algorithm in cloud computing: overview", International Journal of Computer Science & Engineering Technology (IJCSET), Vol.4, ISSN: 2229-3345.

[6] Rashmi Nigoti, Manoj Jhuria, Dr. Shailendra Singh, "A survey of Cryptographic algorithms for cloud computing", International Journal of Emerging Technologies in Computational and Applied Sciences, March 2013, ISSN (online)-2279-0055.

[7] Douglas R. Stinson," Cryptography: Theory& Practice", Chapman and Hall Publications.

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3| ISSN : 2456-3307

Computer Wireless Networking and Communication

R. Deeparani[1], S. Surya[1]
[1]Research Scholar, Department of Computer Science, Shanmuga Industries Arts and Science College, Thiruvannamalai, Tamil Nadu, India

ABSTRACT

Wireless communications and networking technologies have drastically changed the way we live. An explosion of innovation over the past two decades has resulted in wireless networking capabilities that have fundamentally changed the way we create, share, and use information. Combined with advances in computing and networking technology, the wireless Internet ushered into reality the information age predicted long ago. This information era has undeniable effects on global socioeconomic and cultural conditions. These effects have had a profound impact on the operations of governments and military forces. Timely and reliable access to information is key to the success of virtually all government and military functions. A review of how wireless networks can be used in education and training is then given and it is demonstrated that the education field has benefited from the growth of wireless technology and the cost effectiveness of this technology.

Keywords: Bluetooth, Wi-Fi, WiMAX, and Cellular Networks

I. INTRODUCTION

The invention of the computer and the subsequent creation of communication networks can be hailed at the most significant accomplishment of the 21st century. This invention has transformed the way in which communication and information processing takes place. The network functionality of computer systems has been exploited by the government, businesses, and individual with immense benefits being reaped by all. The two major types of networks in existence are the fixed connection (which makes use of cables) and wireless networks (which use waves to transmit data). The backbone of the vast communication network is made up of fixed connections which mostly utilize fiber optics as well as Ethernet. Even so, wireless networks have gained increased popularity in the course of the past decade. Malone (2004) reveals that as of the year 2000, wireless networks were limited in existence due to the prohibitive cost of wireless devices such as integrated routers and access points and laptops. The hardware cost has significantly decreased making wireless networks affordable to many individuals and organization. In addition to this, technological advances have increased the capacity and efficiency of wireless networks, which have made them favourably, compare with wired networks. This paper will set out to discuss wireless
networking with particular focus on the types of wireless technologies commonly employed and the security measures used to protect wireless technology.

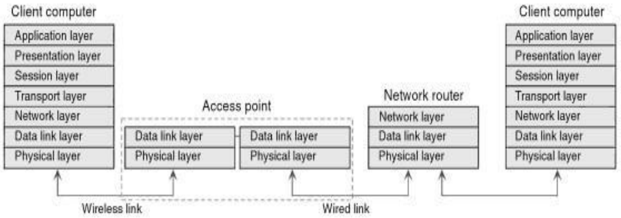
II. COMPUTER NETWORKS: AN OVERVIEW

Computer networks are made up of interconnected computing devices which communicate with each other and these networks are categorized by their sizes. The smallest is the Personal Area Networks (PANs) which extend to a few meters and connect adjacent devices together. Wireless PANs make use of technologies such as Bluetooth to replace cabling as data is moved from device to device. Local Area Networks (LANS) extend from a few hundred meters to a few kilometers and they were designed to cover buildings which are close together or large facilities. Wireless LANs are implemented in facilities such as campuses and busy business locations. Metropolitan Area Networks (MANs) connect different buildings and facilities within a city. These networks mostly make use of wired connections with fiber optic transmissions providing the fastest speeds. The biggest networks are Wide Area Networks (WANs) which connect cities and countries together and they typically make use of fiber-optic cables, which operate at speeds of up to 40Gbps.

III. WHAT IS WIRELESS NETWORKING?

Wireless networking refers to the "utilization of cross-vendor industry standards, such as IEEE 802.11, where nodes communicate without needing to be wired" (Mamoukaris& Economides 2003, p.1). The infrastructure of wireless networks makes use of standard protocols that are oriented according to the demands of the network. This makes the capacity as well as the quality of services of wireless networks vary based on the devices. Wireless networks are typically expected to deal with devices that are made from various manufactures. The networks are therefore supposed to be able to support different hardware technologies, architectures, and transport protocols and also control the flow of traffic within the network.

All wireless networks make use of waves in the electromagnetic spectrum range. For example, Wireless local-area networks (Wireless LANs) make use of high frequency electromagnetic waves to transmit data. Modulation and demodulation of the radio waves used to transmit data occurs at the transmitter and receiver respectively. They operate in the industry, scientific, and medical (ISM) radio bands and unlicensed-national information infrastructure (U-NII) bands (Zheng 2009). The networks are often connected to routers in order for them to access the internet. Reynolds (2003) declares that Wi-Fi has the potential to let anyone with a computing device to connect to the internet at impressive speeds without the need Wireless networks also use the Open System Interconnect (OSI) reference model in the transmission of data. The manner in which this reference model applies to wireless networks is similar to wired networks with some differences in the data link layer where wireless networks coordinate access by data to a common air medium and also deal with errors which occur due to the inherent nature of the wireless medium. At the Physical layer, the data is transmitted in the form of radio waves.



IV. WHY WE NEED TO BUILD A WIRELESS NETWORKING

In most cases, wireless networks are also connected to the internet. A router, which is a device that enables a single internet connection to be shared by many computing devices on the same network, is applicable in such a scenario. The range personal networking devices that can access the wireless networks is great and it includes; laptops, personal digital assistants, tablet PCs, and pocket PCs. All the devices accessing the network need to be equipped with an operating system that allows for communication across a wireless network.

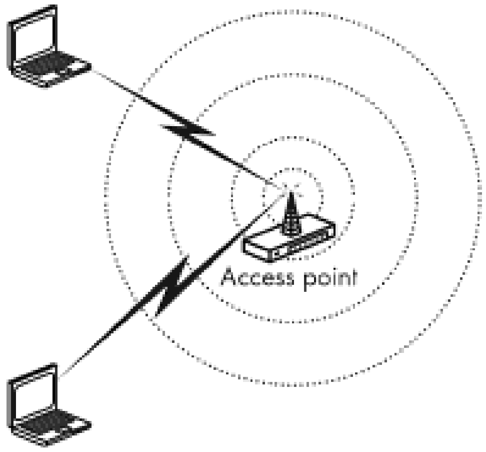
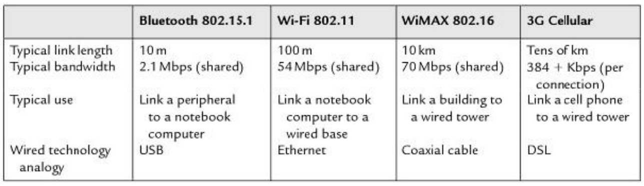


Figure 2. Access Point

V. WIRELESS TECHNOLOGIES

There are a myriad of wireless technologies and they differ in the amount of bandwidth they provide as well as the distance over which the nodes in the network can communicate. Zheng (2009) observes that wireless technologies also differ in the part of the electromagnetic spectrum that they use and the amount of power consumed. To provide physical connectivity, wireless network devices must operate in the same part of the radio spectrum and two wireless cards therefore need to be configured to use the same protocol on the same channel in order for communication to occur. There are four prominent wireless technologies which are; Bluetooth,Wi-Fi, WiMAX and 3G cellular wireless.



	Bluetooth 802.15.1	Wi-Fi 802.11	WiMAX 802.16	3G Cellular
Typical link length	10 m	100 m	10 km	Tens of km
Typical bandwidth	2.1 Mbps (shared)	54 Mbps (shared)	70 Mbps (shared)	384 + Kbps (per connection)
Typical use	Link a peripheral to a notebook computer	Link a notebook computer to a wired base	Link a building to a wired tower	Link a cell phone to a wired tower
Wired technology analogy	USB	Ethernet	Coaxial cable	DSL

Table 1: Popular Wireless Technologies

A.Bluetooth

Bluetooth (IEEE 802.15.1) is the technology that is employed to undertake short-range communicationbetween notebook computers, PDAs, mobile phones and other personal computing devices. The technology is more convenient than connecting devices with a wire to communicate. Bluetooth operates in a license free band at 2.45GHz, the communication range is about 10m, and due to this short range, the technology is sometimes categorized as a personal area network (PAN) (Zheng 2009). A major consideration with Bluetooth technology is power usage and typically, the technology provides speeds of up to 2.1Mbps with low power consumption.

B.Wi-Fi

Wi-Fi stands for wireless fidelity technology and the term is commonly used to describe a wireless local area network based on the IEEE 802.11 series of standards. The IEEE 802.11 standards resolve compatibility issues between manufacturers of wireless networking equipment by specifying an "over the air" interface consisting of "radio frequency technology to transmit and receive data between a wireless client and a base station as well as among wireless clients communicating directly with each other" Wi-Fi describes a family of radio protocols which include 802.11a, 802.11b, and 802.11g. 802.11b is the most popular wireless networking protocol in use and it uses a modulation called Direct Sequence Spread Spectrum in a portion of the ISM band from 2.412 to 2.484GHz (Zheng 2009).

C. Wi-MAX

A popular form of broadband wireless access for fast local connection to the network is WiMAX. WiMAX is the abbreviation for Worldwide Interoperability for Microwave Access and it was standardized as IEEE 802.16 (Zheng 2009). WiMAX technology has a typical range of 1-6 miles but the technology can span a maximum of 30miles which has made the technology classified as a MAN. This specification has gained great success in the provision of internet access and broadband services through wireless communication systems. WiMAX has a high capacity which makes it efficient in data transmission with speeds of up to 70Mbps being provided to a single subscriber station. The original

WiMAX physical layer protocol is designed to propagate signals at a frequency of 10-66 GHz and the technology is able to provide both line of sight coverage and optimal non line of sight coverage as well.

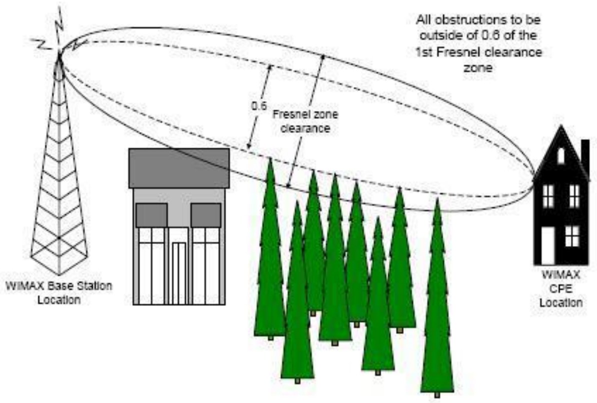

All obstructions to be outside of 0.6 of the 1st Fresnel clearance zone
0.6
Fresnel zone clearance
WiMAX Base Station Location
WiMAX CPE Location

Figure 3. LOS Signal Transmission

The components of a WiMAX include; a Base Station, Subscriber Station, Mobile Subscriber and a Relay Station. The Base station connects and manages access by the devices in the network. This component is made up of multiple antennas pointed in different directions and transceivers which are necessary for the wireless data network communication. A subscriber station is a fixed wireless node which communicates with the base station and forms a link between networks. A mobile subscriber is a wireless node that receives or transmits data through the Base Station while the relay station is a Subscriber Station whose purpose is to retransmit traffic to the relay stations or subscriber stations. A significant merit of WiMAX is that it supports high mobility by user devices. A user can access the network so long as they do not exceed the threshold speed which is normally valued at 120km/H. This property of the technology allows for portability since the user can traverse a significant area which is covered by multiple base stations without having to interrupt their current session.

D.Cellular Networks

While mobile phones have gained overwhelming prominence in the past decades, mobile phone networks were introduced as far back as the early 1980s and this technology was able to provide access to the wired phone network to mobile user. The area of coverage by the cellular wireless network can range from a few hundred meters to a few kilometers in radius. In each cell, there is a base station which is connected to the wired network and which allows the mobile devices in the range to communicate with each other.

Figure 4. Cellular Transmission Towers.

VI. ADVANTAGES OF WIRELESS TECHNOLOGY

The main advantage of a wireless network over a wired one is that users can move around freely within the area of the network with their laptops, handheld devices etc and get an internet connection.

Users are also able to share files and other resources with other devices that are connected to the network without having to be cabled to a port.

Not having to lay lots of cables and put them through walls etc. can be a considerable advantage in terms of time and expense. It also makes it easier to add extra devices to the network, as no new cabling is needed.

If you are a business such as a café, having a wireless network that is accessible to customers can bring you extra business. Customers generally love wireless networks because they are convenient.

Wireless networks can sometimes handle a larger amount of users because they are not limited by a specific number of connection ports.

Instant transfer of information to social media is made much easier. For instance, taking a photograph and uploading it to Facebook can generally be done much quicker with wireless technology.

VII. DISADVANTAGES OF WIRELESS TECHNOLOGY

It can require extra costs and equipment to set up, although increasingly routers have built-in wireless capability, as do devices such as laptops, handheld devices, modern DVD players, and TVs.

Setting up a wireless network can sometimes be difficult for people who are not experienced with computers. (Although there are issues with setting up a wired network too, off course!)

File-sharing transfer speeds are normally slower with wireless networks than they are with cabled. The speeds can also vary considerably according to your location in relation to the network.

The general speed of a wireless connection is also usually much slower than a wired one. The connection also gets worse the farther you are from the router, which can be a problem in a large building or space.

Everyday household items and structures such as walls, ceilings, and furniture can obstruct wireless connections.

Wireless networks are generally less secure. There can also be problems with neighbors stealing bandwidth, if the network has not been set up to be password protected. Information is also less secure too and can be easier to hack into.

VIII. USING WIRELESS TECHNOLOGY INEDUCATION AND TRAINING

Wireless networks have had a profound impact in the area of schools where the exchange of data was previously unattainable due to the complications associated with wired networked.

The education field has benefited from the growth of wireless technology and the cost effectiveness of this technology. Before wireless networks were feasible, the education area suffered from the inherent setbacks of wired networks such as a lack of mobility, the complexity of deployment and difficulty in expanding the network.

The members of the educational institutes want to access the network for wide ranges of purposes and from various locations.

Wireless networks can be less expensive to implement in a school setting that wired networks are. For instance, establishing a wireless LAN in the school may only require the administration to provide the basic connectivity. The users will being their own laptops and therefore save the school money that would have been spent on buying computer hardware as well as Ethernet drops and power outlets. The students will then be able to access the network using their own personal computing devices without incurring additional costs to the schools.

Educational institutes which make use of centralized databases for educational material and information can benefit from wireless networks since the students are able to access the available resources at different areas in the school.

Training sessions may occur in places that are not equipped with wired networks. In such settings, implementing wired networks may be impractical and expensive.

Wireless networks can be quickly deployed for temporary use and then moved when the training is over.

For small training sessions, which have a small number of people, ad-hoc networks can be very useful since they do not require any additional infrastructure to set up. The various individuals in the networks can therefore share resources after configuring their devices to communicate in an ad-hoc manner.

This computer networks do not require the use of access point but rather allow the wireless devices, which are within range of each other to discover each other and proceed to communicate in a peer to- peer manner. Mamoukaris and Economides (2003) argue that implementation so of an ad-hoc wireless networks can help overcome some of the drawbackscaused by the changing educational environment. The networks provide the flexibility and dynamic interaction that is required to foster the success of group communication.

IX. CONCLUSION

In conclusion, wireless communications globally is something that people can expect as technology advances. Wireless communications has many benefits and can make the world a lot more efficient. From the discussions provided in this paper, it is clear that wireless network solutions are increasing in popularity as they become more affordable and are adopted by more people. This paper has elaborated how wireless networks provide freedom from place restriction, scalability and flexibility. The most popular technologies are; Bluetooth, Wi-Fi, WiMAX and Cellular networks. The paper has confirmed that the mobility of wireless networks is their most desirable characteristic. It has been noted that in spite of their merits, there are a few significant issues with wireless networks, which are primarily: quality assurance and security issues. Wireless links are noisier and less reliable than wired links due to the interference that occurs as the signals are transmitted. Engaging in site surveys before setting up a wireless network can help to mitigate this issue. Using strong encryption standards and can resolve the security issues inherent with wireless networks.

ACKNOWLEGEMENT

We wish to express our deep sense of indebtednessand sincerest gratitude effort to all the experts andMr. V.SAKTHIVEL for his valuable guidance and appreciation. We also wish to thank our parents for their encouragement at every step of progress.

X. REFERENCES

[1] Chenoweth, T Robert, M & Sharon, T 2010, “Wireless Insecurity: Examining User Security Behavior on Public Networks”, Communications of the ACM, 53(2): 134-138.

[2] Ganesh, R &Pahlavan, K 2000, Wireless Network Deployments, Springer, Boston.

[3] Jordan, R &Abdallah, C 2002, “Wireless communications and networking: an overview”, IEEE Antenna's and Propagation Magazine, 44 (1): 185-193.

[4] Kumar, A &Manjunath, K 2008, Wireless Networking, Morgan Kaufmann, Boston.

[5] Kumar, A 2010, “Evolution of Mobile Wireless Communication Networks: 1G to 4G”, International Journal of Electronics & Communication Technology, 1(1): 68-72.

[6] Malone S, 2004, Case Study: A Path towards a Secure, Multi-role Wireless LAN in a Higher Education Environment, SANS Institute, Massachusetts.

[7] Mamaukaris, K V and Economides, AA 2003, Wireless technology in educational systems. International PEG Conference, St. Petersburg.

[8] Reynolds, J 2003, Going Wi-Fi: A Practical Guide to Planning and Building an 802.11 Network, CMP, New York.

[9] Schmidt, A &Lian, S 2009, Security and Privacy in Mobile Information and Communication Systems, Springer, Boston.

[10] Singh, L 2009, Network Security and Management, PHI Learning Pvt. Ltd., New Delhi.

[11] Wi-Fi Alliance, 2004, WPA Deployment Guidelines for Public Access Wi-Fi Networks. Wi-Fi alliance, Massachusetts.

[12] Zheng, P 2009, Wireless Networking Complete, Morgan Kaufmann, Boston.

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3| ISSN: 2456-3307

Task Scheduling and Resource Allocation Using a Heuristic Approach In Cloud Computing

K. Durailingam[1], Dr.V.S. Prakash[1]
[1]Indian Arts and Science College, Kondam, Tiruvannamalai, Tamil Nadu, India

ABSTRACT

Cloud computing is needed by trendy technology. Task planning and resource allocation are vital aspects of cloud computing. This paper proposes a heuristic approach that mixes the changed analytic hierarchy method (MAHP), bandwidth aware divisible scheduling (BATS) + BAR optimization, longest expected processing time preemption (LEPT), and divide-and-conquer strategies to perform task planning and resource allocation. During this approach, every task is processed before its actual allocation to cloud resources using a MAHP process. The resources are allocated victimization the combined haywire + BAR optimization methodology, that considers the information measure and cargo of the cloud resources as constraints. Additionally, the planned system preempts resource intensive tasks exploitation LEPT preemption. The divide-and-conquer approach improves the planned system, as is established by experimentation through comparison with the existing bats and improved differential evolution algorithmic rule (IDEA) frameworks once turnaround and time interval square measure used as performance metrics.

Keywords: Cloud computing, Task planning, Heuristic, Resource management, Analytic hierarchy system, BATS, BAR

I. INTRODUCTION

Cloud computing is an accelerating technology within the field of distributed computing. Cloud computing will be utilized in applications that embody storing information, information analytics and IoT applications [1]. Cloud computing may be a technology that has modified ancient ways in which during which services square measure deployed by enterprises or people. It provides differing kinds of services to registered users as net services so the users don't need to invest in computing infrastructure. Cloud computing provides services like IaaS (Infrastructure as a Service), PaaS (Platform as a Service), and SaaS (Software as a Service) [2]. In every type of service, the users square measure expected to submit the requests to the service supplier through the medium of the internet. The service supplier is responsible for managing the resources to fulfill the requests generated by users. Service suppliers use programing algorithms to schedule the incoming request (tasks) and to manage their computing resources with efficiency. Task programing and resource management allow providers to maximize revenue and the utilization of resources up to their limits. In follow, in terms of the performance of cloud computing resources, the programing and allocation of resources are vital hurdles. For this reason, researchers are interested in studies of task programing in cloud computing. Task programing is that the method of arrangement incoming requests (tasks) in a very bound manner so the obtainable resources are properly utilized. as a result of cloud computing is that the technology that delivers services through the medium of the web, service users should submit their requests on-line. because every service features a variety of users, variety of requests (tasks) is also generated at a time. Systems that do not use scheduling might feature longer waiting periods for tasks what is more, some short-run tasks might terminate, as a result of the waiting amount. At the time of programing, the hardware must think about variety of constraints, as well as the character of the task, the dimensions of the task, the task execution time, the avail- ability of resources, the task queue, and the load on the resources. Task programing is one amongst the core problems in cloud computing. correct task programing might end in the economical utilization of resources. the key advantage of cloud computing is that it promotes correct utilization of resources [3]. Thus, task programing and resource allocation square measure 2 sides of a single coin. every affects the opposite. Currently, net users will access content anyplace and anytime, with no need to contemplate the hosting infrastructure. Such hosting infrastructure consists of various machines with numerous capabilities that square measure maintained and managed by the service supplier. Cloud computing enhances the capabilities of such infrastructure, which may access the web. Cloud service suppliers earn profits by providing services to cloud service users.

The cloud service user will use the complete stack of computing services, that ranges from hardware to applications. Services in cloud computing use a pay-as-you-go basis. The cloud service user will scale back or increase the obtainable resources, per the stress of the applications. this is often one the key benefits of cloud computing, however service users are also accountable for paying extra prices for this advantage. The cloud service user will rent the resources at any time and unharness them with no issue. The cloud service user has the liberty to use any service supported application would like. The liberty of service alternative for users has semiconductor diode to problems; that's future user request cannot be dead expected. Thus, task programing and re- supply allocation square measure mandatory components of cloud computing analysis. The potency of resource uses depends on the programing and cargo leveling methodologies, rather than the random allocation of resources. Cloud computing is wide used for finding complicated tasks (user requests). In finding complicated task problems, the employment of programing formula is suggested. Such programing algorithms leverage the resources. The projected system employs options of the Cybershake scientific workflow and the Epigenomics scientific progress, that are represented in Section computer file. The major contributions of this paper are summarized as follows.

1. The analytic hierarchy method is changed to rank scientific tasks.
2. To manage the resources given information measure constraints and therefore the load on the virtual machine, the projected system incorporates a version of the prevailing round the bend formula that has been changed by introducing BAR system optimization.
3. Bipartite graphs square measure utilized to map tasks to appropriate virtual machines once the condition is glad.
4. A pre-emption methodology offers US the standing of the virtual machine, and a changed divide-and-conquer methodology has been projected to mixture the results when tasks pre-emption.
5. The projected resolution is by experimentation-investigated victimization the Cloud Sim machine.

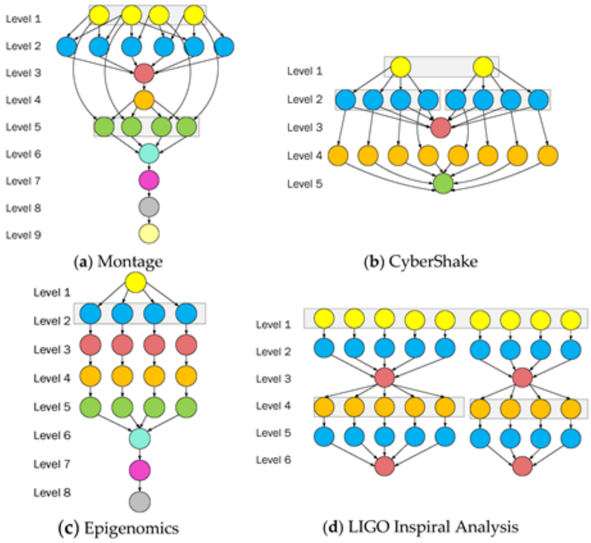
The remainder of the paper is organized as follows. Section “Introduction” provides Associate in Nursing introduction to cloud computing and its outstanding problems, particularly task programing and resource allocation. Section “Related work” focuses on connected studies that investigate task programing and resource allocation. Section “Input data” describes the computer file provided to the Cybershake scientific progress and the Epigenomics scientific workflow. Section "Proposed system" addresses the design of the projected system. Section "Proposed Methodology," explains the projected methodology. Section "Evaluation of the projected heuristic approach" focuses on evaluating the projected heuristic approach. Section "Results and discussion" describes the results and discusses the projected system compared with the prevailing round the bend and plan algorithms. Finally, concluding remarks and future directions are conferred in Section "Conclusion".

II. RELATED WORK

This section provides a short review of task programing and resource allocation ways. Several researchers have projected solutions to beat the matter of programing and resource allocation. However, additional improvements will still be created. Tsai et al. [4] projected a multiobject approach that employs the improved differential evolution algorithmic rule. This existing technique provides a value and time model for cloud computing. However, variations within the tasks aren't thought-about during this approach. Magukuri et al. [5] projected a load balancing and programing algorithmic rule that doesn't contemplate job sizes. The authors considered the refresh times of the server in fulfilling requests. Cheng et al. [6] introduced the programing of tasks supported a vacation queuing model. This technique doesn't show the correct utilization of resources. Lin et al. [7] projected the programing of tasks whereas considering information measure as a resource. A nonlinear programming model has been shaped to portion resources to tasks. Ergu et al. [8] proposed AHP ranking-based task programing. Zhu et al. [9] introduced rolling-horizon programing architecture to schedule real-time tasks. Authors have illustrated the link between task programing and energy conservation by resource allocation. Lin et al. [10] projected programing for parallel workloads. Authors have used the FCFS approach to order jobs once resources are avail in a position. The projected system doesn't specialize in aborting the roles and starvation. Ghanbari et al. [11] projected a priority-based job programing algorithmic rule to be used in cloud computing. Multi criteria selections and multiple attributes square measure thought-about. Polverini et al. [12] introduced the optimized value of energy and queuing delay constraints. Alejandra et al. [13] projected the employment of meta-heuristic optimization and particle swarm optimization to reduce execution prices through programing. Keshk et al. [14] projected the employment of changed ant colony optimization in load equalization. This technique improves the makespan of a job. This technique doesn't contemplate the supply of resources or the load of tasks. Shamsollah et al. [15] projected a system supported a multi-criteria algorithmic rule for programing server load. Shamsollah et al. [16] pro- posed a system supported priority for acting separable load programing that employs analytical hierarchy method. Gougarzi et al. [17] projected a resource allocation drawback that aims to reduce the full energy value of cloud computing systems while meeting the required client-level slas in a very probabilistic sense. Here, authors have applied a reverse approach that applies a penalty if the client doesn't meet the SLA agreements. Some authors have enforced a heuristic algorithmic rule to solve task programing and resource allocation drawback de- scribed higher than. Radojevic et al. [18] introduced central load equalization call model to be used in cloud environments; this model automates the programing method and reduces the role of human administrators. However, this model is deficient in decisive the capabilities of nodes and, configuration details, and therefore the complete sys- tem has no backup, therefore leading to one purpose of failure. Additionally, Ghanbari et al. [12] and Goswami et al. [14] specialise in programing tasks whereas considering various constraints. This state-of the art motivates the authors of this study to conduct additional analysis on task programing and resource allocation.

III. INPUT WORK

Cybershake scientific workflow Cloud computing is that the service supplier paradigm within which users submit requests for execution. Thus, the responsibility of the cloud service supplier is to schedule various requests and manage resources with efficiency. To the most effective of the authors' data, most existing work involves scheduling tasks once they enter a task queue. However, the particular procedure of planning tasks and resource management begins with however the service supplier addresses incoming tasks. The pro- posed system uses Cybershake scientific advancement knowledge as in place tasks [12]. Fig. 1 shows a visualization of the Cybershake scientific advancement, that is employed by the Southern CA Earthquake Center (SCEC) to characterize earthquake haz- wet lung using the Probabilistic Unstable Hazard Analysis (PSHA) technique. It additionally generates inexperienced strain tensors (GSTs). Table one shows the Cybershake seismogram synthesis tasks with their sizes and execution times. The Cybershake could be a collection of assorted node knowledge that area unit out there for study [14]. The Cybershake scientific work- flow sample tasks area unit out there with task size 30,50,100 and 1000. From a machine purpose of read, the seismogram synthesis tasks area unit quite stringent. The Cyber- shake spends lots of time on seismogram synthesis throughout its execution. These kinds of tasks additionally need large amount of computational resources, like central processing unit time, and memory.



(a) Montage

(b) CyberShake

(c) Epigenomics

(d) LIGO Inspiral Analysis

Figure 1. Cybershake scientific workflow

Cybershake scientific workflow has been divided into 5 steps.

1. Extract GST - This step of the workflow extracts the GST (Green strain tensor) data for processing.
2. Seismogram synthesis – These tasks are the most computationally intensive. Most of the time spent in running the Cybershake algorithm is employed on this step.
3. ZipSeis–This step aggregates the processed data.
4. PeakValCalcOkaya – The highest-strength values of each seismogram are calculated in this step.
5. ZipPSA-This step aggregates the processed data.

Table 1 Cybershake seismogram synthesis tasks

Tasks	Size	of
tasks	Time	
Task 3	62,69,51,663	39.06
Task 5	69,47,76,323	38.49
Task 7	58,57,63,637	36.27
Task 9	53,68,97,326	32.29
Task 11	67,05,35,542	62.25
Task 14	40,67,28,38,798	96.91
Task 16	45,23,96,996	45.60
Task 18	50,27,64,231	28.67
Task 20	62,41,88,532	24.56
Task 22	42,65,77,006	31.05
Task 24	51,58,32,878	54.87
Task 26	68,14,99,417	23.99
Task 28	44,14,51,516	26.46

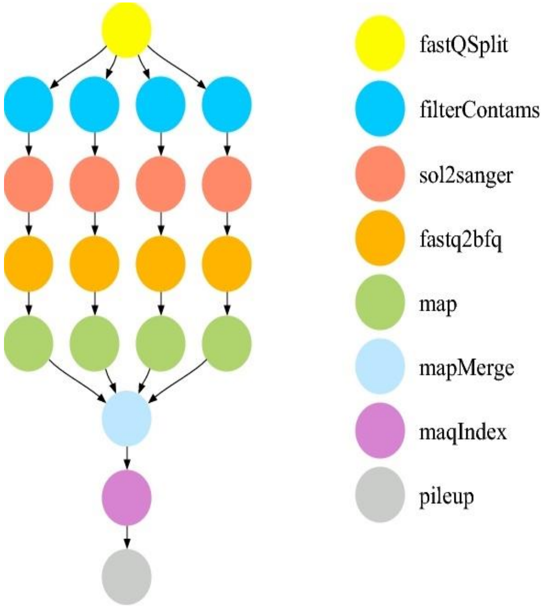


Figure 2. Epigenomics Scientific Workflow

IV. EPIGENOMICS SCIENTIFIC WORK FLOW

Figure a pair of shows the Epigenomics scientific workflow [18] that is used to change the method of ordering sequencing. This operation is related to resource- intensive tasks. The generated information area unit born-again into files and forwarded to magazine system. This method additionally involves several operations, andthese operations area unit time over whelming.

V. PROPOSED SYSTEM

Figure three shows the design of the planned system. In apply, varied varieties and sizes of tasks gain the cloud information centers for execution. The planned system takes the $64000 tasks as Associate in Nursing input, as delineate in Section 3. In general, scientific tasks represent collections of various types and sizes. To manage the tasks that get a cloud information center, the planned system uses the analytic hierarchy method (AHP). The first aim of this planned system is to manage incoming tasks. Therefore, the planned system uses the AHP method-field of study to assign a rank to every task supported its length and run time. The procedure for ranking the tasks for scientific workflows is delineate in section 5.

1. As before long because the tasks area unit assigned individual rankings, they are collected and organized into task queues. The tasks within the task queue area unit strictly organized following the AHP ranking. Thus, the primary stage of the planned system is completed. Next, within the second stage, the planned system additionally addresses the computing resources of cloud information centers, like central processing unit, memory and information measure mistreatment the planned BATS+BAR optimized allocation methodology. This technique works as follows. It takes the task to be dead from the task' queue. The assignment of resources and tasks follows the allocation combining weight. 4. A detailed clarification is givenin section five.2. This stage is that the second a part of the procedure during which the allocations of resources are distributed using BATS+BAR. Within the next half, the planned system uses a preemption methodology, i.e., the preemption technique. LEPT incessantly checks the load of the Virtual machine. If it's exceeded theproposed system then uses a virtual machine standing table to confirm this standing of alternative Virtual machines (vms). During this regard, if this virtual machine is overladen et al. Area unit idle, then such vmsarea unit set. Once this identification, the proposed system uses a divide-and-conquer methodology that breaks up the task and distributes it to alternative virtual machines, as described intimately in section five.3. During this means, the propose system has overcome the restrictions of buggy in terms of the allocation of resources based on CPU, memory and information measure. If anyone resource (CPU, memory, bandwidth) is not available in sufficient amounts, then the tasks must wait. In addition, existing systems do not consider preemption, and the inputs to existing systems are tasks of the same size. Fig.4 presents a flow chart that represents the proposed heuristic approach.

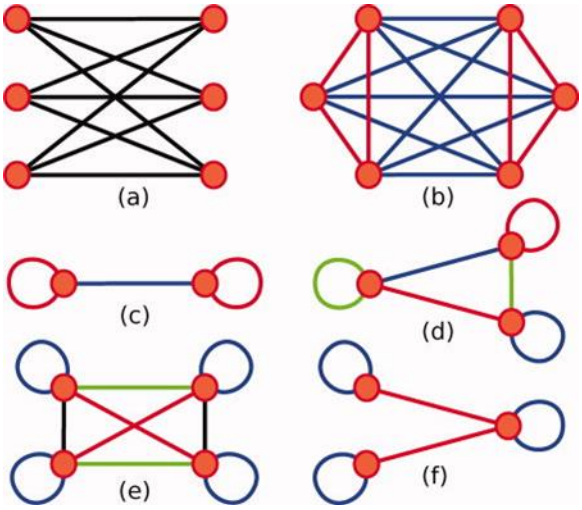


Figure 4. Complete Bipartite Graph

VI. PROPOSED METHODOLOGY

Here, we offer a detailed explanation of the proposed system to beat the programing challenge.

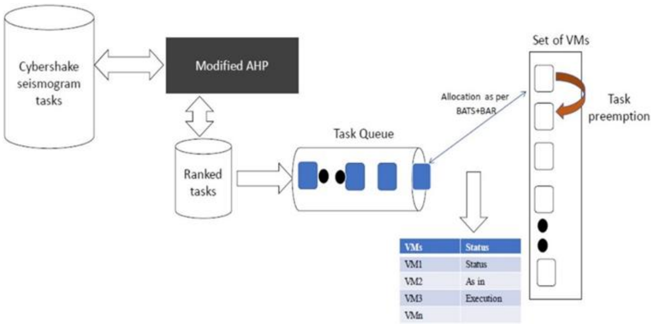


Figure 3. Proposed System Architecture

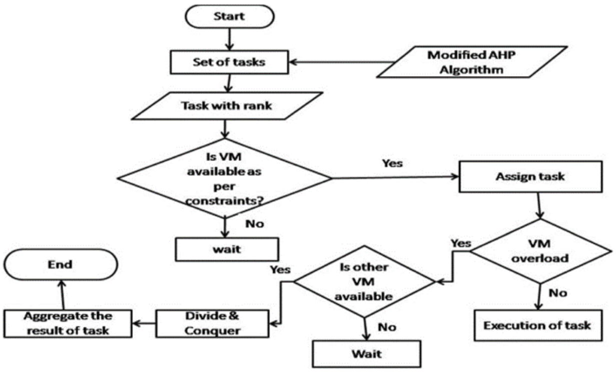
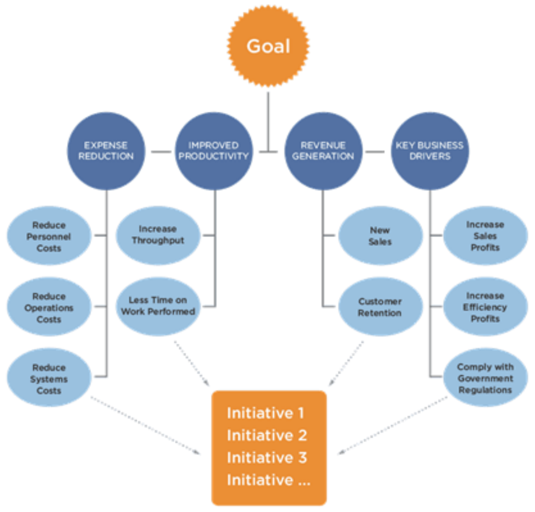


Figure 4. Proposed System Flowchart

VII. ANALYTIC HIERARCHY PROCESS

The analytic hierarchy process [18] is intended to solve complicated issues with multiple criteria. The proposed system uses this procedure in cloud computing environments to rank the incoming tasks in a sure manner. The projected system uses scientific work flow tasks, like those of Cybershake and Epigenomics, for experiments because such need long execution times. Initially, the advancement is Split into 5 stages, which are introduced in the input information section. Before continuing with the planned system, the AHP methodology is applied for the Overall Cybershake advancement.The Cybershake workflow is management flow Dependent; so, the second stage can execute only once the execution of the primary stage.

To evaluate preferences, the projected system uses the Saaty preference table, that is given in Table two with its numerical ratings. To push understanding whereas accounting for house limitations, the projected system divides every calculation table into 2 elements. The primary half extends from Task three to Task , whereas the opposite half shows the calculations from Task twenty to Task . Here, the projected system considers two important criteria that are concerned in scientific tasks; task length and task run time. The comparison numerical ratings are given in Table 2, that is thought because the Saaty preference table. Before the actual calculation is begun, the projected system assigns preference values to the tasks. Here, the preferences associated with the tasks are supported their lengths and also the execution times of the various tasks. The projected system slightly modifies the Saaty table preferences as a result of, as tasks with totally different ranks are on a server, the ranks of resulting tasks amendment, and new rankings should be calculated.



THE ANALYTIC HIERARCHY PROCESS (AHP) MODEL

The pro- display system calculates such rankings of tasks. Tables 3 and 4 show the assignment of Saaty preferences in line with examination the sizes and runtimes of tasks. Within thebottom row, the add of every column is noted. Tables 5 and 6 show the multiplication of the Saaty preference values by the results organized within the bottom rows of Tables 3 and 4 and so gift the results of adding every column at rock bottom. Tables 2 and 4 show the normalized values of Tables 5 and 6, which appear earlier in the manuscript. These tables include average at the bottom. The results show that the summation of each column is equal to 1.

Table 2 Numerical saaty preferences

Numerical rating	Judgment preference
9	Extremely preferred
8	Very strongly to extremely preferred
7	Very strongly preferred to preferred
6	Strongly to very strongly
5	Strongly preferred
4	Moderately to strongly preferred
3	Moderately preferred
2	Equally to moderately preferred
1	Equally preferred

VIII. BATS+ BAR SYSTEM

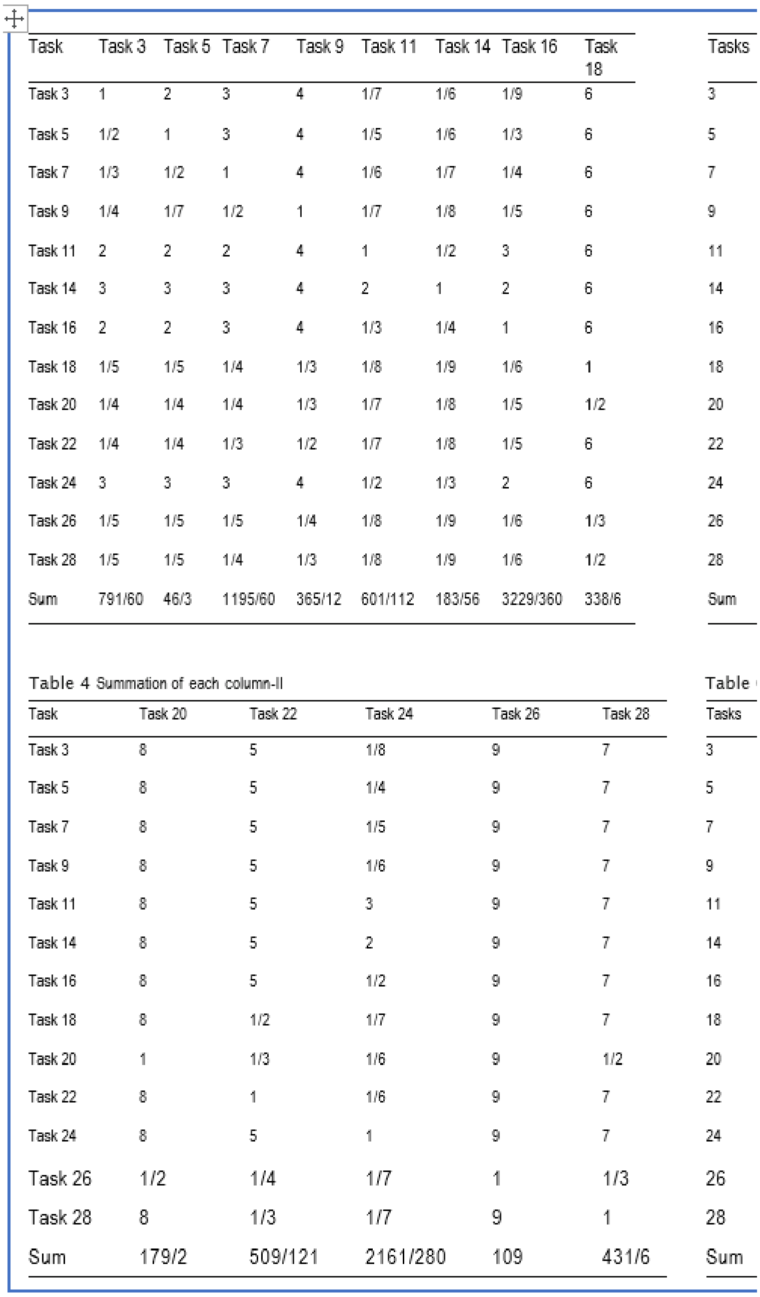
The planned system has two aspects that involve planning tasks and managing resources. Here, we improve upon the round the bend algorithmic rule, that was originally planned by Weiwei Lin [7]. freelance tasks of equal size ar considered within the design of this technique. How- ever, in allocating resources, the system doesn't consider the load on virtual machines as a result of the waiting amount for the tasks is long. In alternative cases, one virtual machine is busy whereas it executes a task, whereas others ar occupied and waiting for jobs. The bar systems (BSs) algorithmic rule was planned by Acebo and Rosa (2008) [14]. The social behavior of bartenders is that the basis of BS systems. Swarm intelligence has else an optimisation side to BS. In a bar, bartenders should act in a very extremely dynamic, asynchronous and time-critical setting, and no obvious greedy strategy (such as serving the most effective client initial, serving the closest client initial or serving the first-arriving client first) offers smart results. Thus, multi-agent systems offer a decent framework at intervals that to handle the challenge of developing a brand new category of adaptational and sturdy systems. In general, the crucial step within the SB algorithmic rule is that the selection of the task that the agent should execute within the next time step. In BSs, agents acting as bartenders, operate at the same time in an environment during a exceedingly in a very synchronous manner; that's, they execute tasks by deciding that drinks to pour. once Associate in Nursing initial section, the “bartenders” build their selections according to totally different problem-dependent properties (e.g. weight, speed, location, time interval, most load, etc.), instead of constructing selections willy-nilly. Over time, if an agent is unable to adapt the environment to the preconditions of the task (such because the value for the agent to execute the task within the current state of the environment) or if it's unable to hold the task out by itself, it'll be eliminated. To over- return this behavior, we propose modifying nuts by adding a BAR system.

The procedure is as follows:

- Aggregate all of the task information that is ordered by rank.

- Virtual machine (server) information is collected. This information includes the initial load on the virtual machine, its bandwidth and the time required to process the tasks on the server.

- A bipartite graph is generated with the number of tasks. The ranking prioritiescan be used to construct a graph, by which each task is allocated to a virtual machine.



Task	Task 3	Task 5	Task 7	Task 9	Task 11	Task 14	Task 16	Task 18	Tasks
Task 3	1	2	3	4	1/7	1/6	1/9	6	3
Task 5	1/2	1	3	4	1/5	1/6	1/3	6	5
Task 7	1/3	1/2	1	4	1/6	1/7	1/4	6	7
Task 9	1/4	1/7	1/2	1	1/7	1/8	1/5	6	9
Task 11	2	2	2	4	1	1/2	3	6	11
Task 14	3	3	3	4	2	1	2	6	14
Task 16	2	2	3	4	1/3	1/4	1	6	16
Task 18	1/5	1/5	1/4	1/3	1/8	1/9	1/6	1	18
Task 20	1/4	1/4	1/4	1/3	1/7	1/8	1/5	1/2	20
Task 22	1/4	1/4	1/3	1/2	1/7	1/8	1/5	6	22
Task 24	3	3	3	4	1/2	1/3	2	6	24
Task 26	1/5	1/5	1/5	1/4	1/8	1/9	1/6	1/3	26
Task 28	1/5	1/5	1/4	1/3	1/8	1/9	1/6	1/2	28
Sum	791/60	46/3	1195/60	365/12	601/112	183/56	3229/360	338/6	Sum

Table 4 Summation of each column-II

Table

Task	Task 20	Task 22	Task 24	Task 26	Task 28	Tasks
Task 3	8	5	1/8	9	7	3
Task 5	8	5	1/4	9	7	5
Task 7	8	5	1/5	9	7	7
Task 9	8	5	1/6	9	7	9
Task 11	8	5	3	9	7	11
Task 14	8	5	2	9	7	14
Task 16	8	5	1/2	9	7	16
Task 18	8	1/2	1/7	9	7	18
Task 20	1	1/3	1/6	9	1/2	20
Task 22	8	1	1/6	9	7	22
Task 24	8	5	1	9	7	24
Task 26	1/2	1/4	1/7	1	1/3	26
Task 28	8	1/3	1/7	9	1	28
Sum	179/2	509/121	2161/280	109	431/6	Sum

The Load on the virtual machine(S) is calculated as,

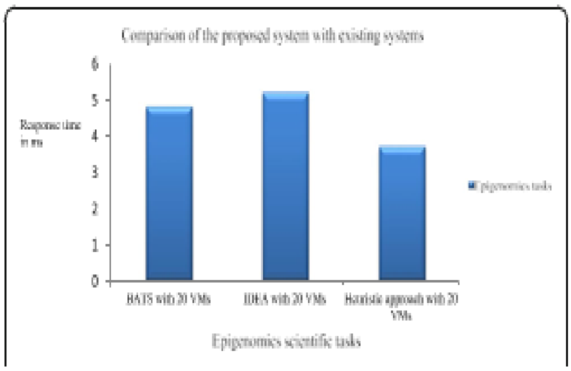
L^{ini} ¼ L_s^{ini} js⊂S 1

The bandwidth is calculated as,

DB_w¼bˈi<¼bi 2

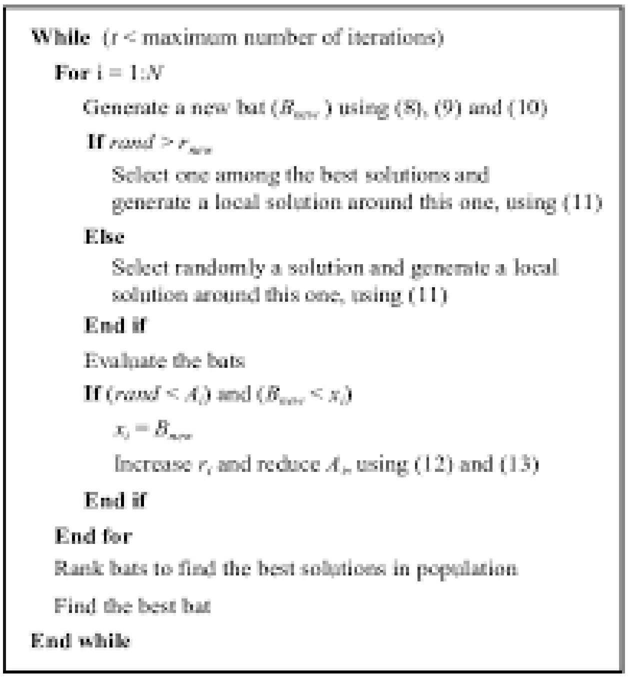
Thetotaltimetakentoprocessthetasksiscalculatedas,

L_s^{fin}ðαÞ¼L_sðαÞ 3



CPU utilizing Bar Diagram

Where, (α) = any task.



```
While  ($r$ < maximum number of iterations)
   For i = 1:$N$
      Generate a new bat ($B_{new}$ ) using (8), (9) and (10)
      If $rand > r_{new}$
         Select one among the best solutions and
         generate a local solution around this one, using (11)
      Else
         Select randomly a solution and generate a local
         solution around this one, using (11)
      End if
      Evaluate the bats
      If ($rand < A_i$) and ($B_{new} < x_i$)
         $x_i = B_{new}$
         Increase $r_i$ and reduce $A_i$, using (12) and (13)
      End if
   End for
   Rank bats to find the best solutions in population
   Find the best bat
End while
```

Bipartite graph

A bipartite graph is produced based on the following conditions:

1. A bipartite graph is constructed as-, G = (T_nU S, E) in which ‘T_n’ represents the number of tasks, ‘S’represents the servers, and ‘E ⊆TX S’ that is, the set of edges that are present between the task and the server. An edge represents the tasks ‘Ti ⊂Tn’, which are present on virtual machine‘s⊆S’. 2.Agraph isconstructed using bipartite graph with the number of tasks.
2. Balancetheconstructedgraphwithconstraints includingthelocalcost,theinitialloadandthe bandwidth.
3. Based on the local cost and the initial load we computethetotalloadonthevirtualmachine.
4. Next,weapplytheconditionrepresentedbyEq.If thisconditionissatisfied,thenweallocatethetasks tothatparticularvirtualmachine.Ifthisconditionis not satisfied by that virtual machine, then we move on the next server and check thiscondition.
5. After allocating the tasks, the constructed bipartite is updated if any task remain to be processed. It is the bipartite graph of the set of virtual machines and set ofresources.

IX. EVALUATION OF TIME INTERVAL

As a second performance metric, we take into account the time interval of the algorithm to incoming tasks. The time interval is actually the time during that the request is actually considered. In different words, we are able to say that the time interval is directly addicted to the supply of resources. The availability of resources depends abreast of the programming of tasks. If the programming of tasks is performedproperly, then the resources can naturally be free early or earlier of deadlines, the response times can be less in such cases.

By, scrutiny the response times obtained for our proposed heuristic approach with those obtained using the existing bats and idea frameworks, we are able to see that our system’s response time is almost 500th less. The latent period comparisons for Cybershake and Epigenomics are conferred in Figs. and nine respectively. The comparison is additionally shown in tabular we have a tendency to take into account two parameters the latent period and turnaround compare the proposed heuristic approach with the prevailing nutty and plan frameworks. Because we have a tendency to be evaluating these frame- works in a very cloud computing surroundings, the response time is generally less effective.

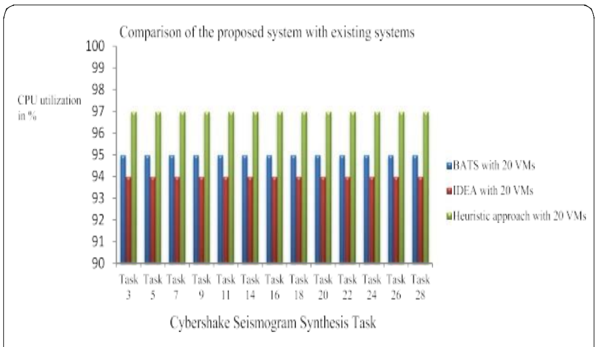
On the other hand, we have a tendency to additionally evaluate our planned heuristic approach to determine its resource performance compare it to those of the prevailing nutty and plan frameworks.

X. EVALUATION OF CPU UTILIZATION

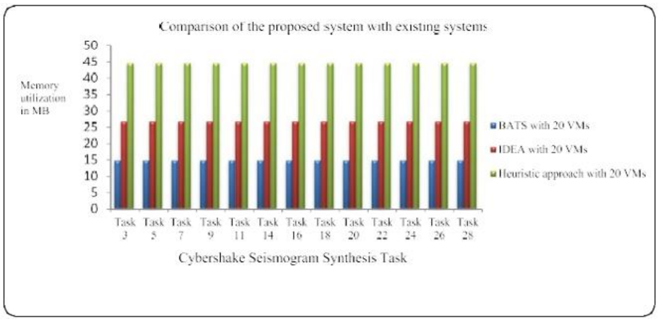
Key comparison of resource utilization between the planned heuristic approach and existing nutty and plan frameworks. The proper utilization of resources produces profits for cloud computing service suppliers. The experimental results shows that the planned heuristic approach utilised the CPU resource more efficiently than the prevailing nutty framework.

XI. EVALUATION OF MEMORY UTILIZATION

The second key comparison of resource utilization between the planned heuristic approach and therefore the existing nutty and idea frameworks. The experimental results shows that the proposed heuristic approach utilizes memory re- sources more efficiently than the prevailing idea and bats frameworks.



CPU utilization by Cybershake tasks



XII. CONCLUSION

In this study, we proposed heuristic formula that performs task scheduling and allocates resources efficiently in cloud computing environments. We use real Cybershake and Epigenomics scientific workflows as input tasks for the system. After we compare our projected heuristic approach with the existing fruity and plan frameworks with respect to turnaround time and interval, we find that our approach provides improved results. On the opposite hand, from the point of view of resupply utilization, the projected heuristic approach efficiently allocates resources with high utility. We tend to obtain the utmost utilization result for computing resources like computer hardware, memory and bandwidth. Most existing systems consider only 2 resources, computer hardware and memory, in evaluating their performance the projected system adds bandwidth as a resource. Future work can specialize in simpler scheduling algorithms in which turn- around time and response time will be improved.

XIII. REFERENCES

[1] Mezmaz M, Melab N, Kessaci Y, Lee YC, Talbi E-G, Zomaya AY, Tuyttens D (2011) A parallel bi-objective hybrid meta heuristic for energy-aware scheduling for cloud computing systems. J Parallel Distributed Computing 71(11):1497–1508

[2] Armbrust M, Fox A, Griffith R, Joseph AD, Katz R, Konwinski A, Lee G, Patterson D, Rabkin A, Stoica I et al (2010) A view of cloud computing. Commun ACM 53(4):50–58

[3] Tsai J-T, Fang J-C, Chou J-H (2013) Optimized task scheduling and resource allocation on cloud computing environment using improved differential evolution algorithm. ComputOper Res 40(12):3045–3055

[4] Cheng C, Li J, Wang Y (2015) An energy-saving task scheduling strategy based on vacation queuing theory in cloud computing. Tsinghua SciTechnol 20(1):28–39

[5] Ergu D, Kou G, Peng Y, Shi Y, Shi Y (2013) The analytic hierarchy process: task scheduling and resource allocation in cloud computing environment. The Journal of Supercomputing. 64(3):835-848

[6] Zhu X, Yang LT, Chen H, Wang J, Yin S, Liu X (2014) Real-time tasks oriented energy-aware scheduling in virtualized clouds. IEEE Transactions on Cloud Computing 2(2):168–180

[7] Handfield R, Walton SV, Sroufe R, Melnyk SA (2002) Applying environmental criteria to supplier assessment: a study in the application of the analytical hierarchy process. Eur J Oper Res 141(1):70–87

[8] Calheiros RN, Ranjan R, Beloglazov A, De Rose CA, Buyya R (2011) Cloudsim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. Software: Practice and Experience

[9] Maguluri ST, Srikant R (2014) Scheduling jobs with unknown duration in clouds. IEEE/ACM Trans Netw (TON) 22(6):1938–

[10] Lin W, Liang C, Wang JZ, Buyya R (2014) Bandwidth-aware divisible task scheduling for cloud computing. Software: Practice and Experience 44(2):163–174

[11] Shamsollah G, Othman M (2012) Priority based job scheduling algorithm in cloud computing. Procedia Engineering 50:778–785

[12] Polverini M, Cianfrani A, Ren S, Vasilakos AV (2014) Thermal aware scheduling of batch jobs in geographically distributed data centers. IEEE Transactions on Cloud Computing 2(1):71–84

[13] Keshk AE, El-Sisi AB, Tawfeek MA (2014) Cloud task scheduling for load balancing based on intelligent strategy. Int J IntellSystAppl 6(5):25

[14] Rodriguez MA, Buyya R (2014) Deadline based resource provisioningand scheduling algorithm for scientific workows on clouds. IEEE Transactions on Cloud Computing 2(2):222–235

[15] Liu X, Zha Y, Yin Q, Peng Y, Qin L (2015) Scheduling parallel jobs with tentative runs and consolidation in the cloud. J SystSoftw 104:141–151

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3| ISSN: 2456-3307

iPath : Path Inference in Wireless Sensor Networks

P.Florance Rincy[1], Mrs. K.Sumalatha[2]

[1]M.Phil Research Scholar,Dept. Of Computer Science,Kamban College of Arts & Science for Women
Tiruvannamalai, Tamil Nadu, India

[2]Head ofthe Department, Dept. Of Computer Science,Kamban College of Arts & Science for Women
Tiruvannamalai, Tamil Nadu, India

ABSTRACT

Recent wireless sensor networks (WSNs) are becoming increasingly complex with the growing network scale and the dynamic nature of wireless communications. Many measurement and diagnostic approaches depend on per-packet routing paths for accurate and fine-grained analysis of the complex network behaviors. In this paper, we propose iPath, a novel path inference approach to reconstructing the per-packet routing paths in dynamic and large-scale networks. The basic idea of iPath is to exploit high path similarity to iteratively infer long paths from short ones. iPath starts with an initial known set of paths and performs path inference iteratively. iPath includes a novel design of a lightweight hash function for verification of the inferred paths. In order to further improve the inference capability as well as the execution efficiency, iPath includes a fast bootstrapping algorithmto reconstruct the initial set of paths. We also implement iPath and evaluate its performance using traces from large-scale WSN deployments as well as extensive simulations. Results show that iPath achieves much higher reconstruction ratios under different network settings compared to other state-of-the-art approaches.

Keywords: Measurement, Path Reconstruction, WirelessSensor Networks

I. INTRODUCTION

Wireless sensor networks (WSNs) can be applied in many application scenarios, e.g., structural protection [1], ecosystem management [2], and urban CO monitoring [3]. In a typical WSN, a number of self-organized sensor nodes report the sensing data periodically to a central sink via multihop wireless.Recent years have witnessed a rapid growth of sensor network scale. Some sensor networks include hundreds even thousandsof sensor nodes [2], [3]. These networks often employ dynamic routing protocols [4]–[6] to achieve fast adaptation to the dynamic wireless channel conditions. The growing network scale and the dynamic nature of wireless channel make WSNs become increasingly complex and hard to manage. Reconstructing the routing path of each received packet at the sink side is an effective way to understand the network's complex internal behaviors [7], [8].

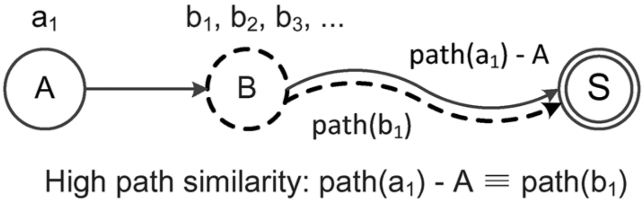


Fig.1. Example to illustrate the basic idea of iPath.

With the routing path of each packet, many measurement and diagnostic approaches [9]–[13] are able to conduct effective management and protocol optimizations for deployed WSNs consisting of a large number of unattended sensor nodes. For example, PAD [10] depends on the routing path information to build a Bayesian network for inferring the root causes of abnormal phenomena. Path information is also important for a network manager to effectively manage a sensor network. For example, given the per-packet path information, a network manager can easily find out the nodes with a lot of packets forwarded by them, i.e., network hop spots.

The contributions of this work are the following.

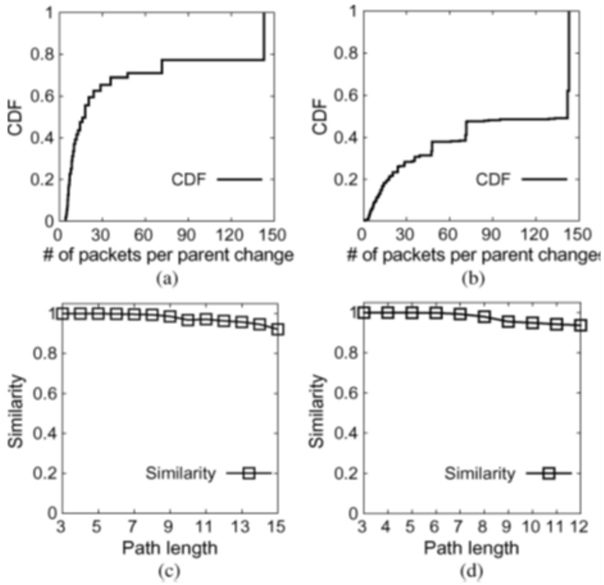
We observe high path similarity in a real-world sensor network. Based on this observation, we propose an iterative boosting algorithm for efficient path inference.We propose a lightweight hash function for efficient verification within iPath. We further propose a fast bootstrapping algorithm to improve the inference capability as well as its execution efficiency.

We propose an analytical model to calculate the successful reconstruction probability in various network conditions such as network scale, routing dynamics, packet losses, and node density.

II. MEASUREMENT STUDY

In order to quantify the path similarity in real-world deployment, we conduct a measurement study on two deployed networks—citysee[3] and greenorbs [2]. The cityseeproject is deployed in an urban area for measuring carbon emission. All nodes are organized in four subnets. Each subnet has one sink node, and sink nodes communicate to the base station through 802.11 wireless links. We collect traces from one sink of a subnet with 297 nodes. The greenorbs project includes 383 nodes in a forest area for measuring the carbon absorbance. These two networks use the Collection Tree Protocol [4] as its routing protocol. In order to reduce the energy consumption and prolong the network lifetime, all nodes except the sink node. Work at low-power listening states. The wakeup interval of the low power setting is 512 ms.Each node reports data packets to a sink with a period of 10 min. Each data packet carries the routing path information directly for offline analysis. We first look at the routing dynamics of the networks.

We measure a quantity that is defined to be the average number of periods (i.e., local packets) between two parent changes by a node. It is simply the inverse of the number of parent changes per period at a node. A smaller means more frequent parent changes. Fig. 2(a) and (b) shows the cumulative distribution function (CDF) of for all nodes in the two networks. We can see that these two network have different degrees of routing dynamics. On average, there is a parent change every 46.9 periods in cityseeand 89.1 periods in greenorbs. As a comparison, the MNT paper [8] reports a parent change every 88.2-793.3 periods of the networks tested, which have less frequent parent changes. We see that cityseeand greenorbs have high routing dynamics, making per-packet path inference necessary for reasoning about complex routing behaviors. On the other hand, we observe high path similarity in the networks, i.e., it is highly probable that a packet from node and one of the packets from's parent will follow the same path startingfrom's parent toward the sink. To quantitatively measure path similarity, wedefine sim (len) such that among all packets with path length len, there are sim(len)ratio of packets that follow the same path as at least one(len-1) hop packet. Fig. 2(c) and (d) shows the sim(len) values with varying len . We see that the values of sim (len) are close to 1, indicating that a high path similarity in both the cityseenetwork and greenorbs network. Note that the paths shown in these two figures include more than 99% of the total packet paths in these two traces. Therefore, the path similarity observation is not biased. The above results show that although there are severe routing dynamics, the path similarity can still be very high. This key observation gives us important implications for efficient path inference: If a similar short path is known, it can be used toreconstruct a long path efficiently.



III. NETWORK MODEL

In this section, we summarize the assumptions made and data fields in each packet. We assume a multihop WSN with a number of sensor nodes. Each node generates and forwards data packets to a single sink. In multisink scenarios, there exist multiple routing topologies.The path reconstruction can be accomplished separately based on the packets collected at each sink.In each packet,there are several data fields related to iPath.

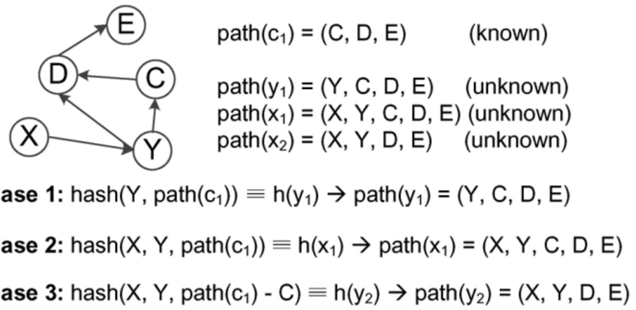
We summarize them as follows.

- The first two hops of the routing path, origin o(k) and parent p(k).Including the parent information in each packet is common best practice in many real applications for different purposes like network topology generation orpassive neighbor discovery [8], [22].
- The path length (k). It is included in the packet header in many protocols like CTP [4]. With the path length, iPath is able to filter out many irrelevant packets during the iterative boosting (Section V-A).
- A hash value h(k) of packet 's routing path. It can makethe sink be able to verify whether a short path and a long path are similar. The hash value is calculated on the nodes along the routing path by the PSP-Hashing (Section V-B).

IV. IPATH DESIGN

A. Iterative Boosting

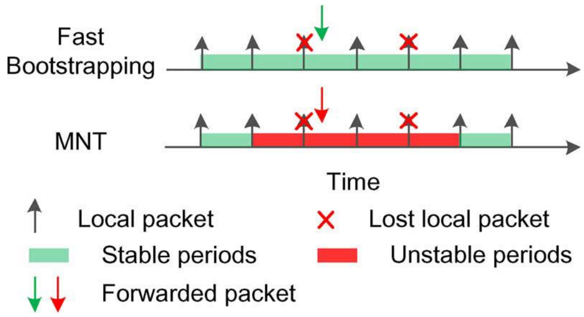
iPath reconstructs unknown long paths from known short paths iteratively. By comparing the *recorded hash value* and the *calculated hash value*, the sink can verify whether a long path and a short path share the same path after the short path's original node. When the sink finds a match, the long path can be reconstructed by combining its original node and the shortpath.



E
D
C
X
Y
path(c₁) = (C, D, E) (known)
path(y₁) = (Y, C, D, E) (unknown)
path(x₁) = (X, Y, C, D, E) (unknown)
path(x₂) = (X, Y, D, E) (unknown)
Case 1: hash(Y, path(c₁)) ≡ h(y₁) → path(y₁) = (Y, C, D, E)
Case 2: hash(X, Y, path(c₁)) ≡ h(x₁) → path(x₁) = (X, Y, C, D, E)
Case 3: hash(X, Y, path(c₁) - C) ≡ h(y₂) → path(y₂) = (X, Y, D, E)

B. Fast Bootstrapping

The iterative boosting algorithm needs an initial set of reconstructed paths. In addition to the one/two-hop paths, the fast bootstrapping algorithm further provides more initial reconstructed paths for the iterative boosting algorithm. These initial reconstructed paths reduce the number of iterations needed and speed up the iterative boosting algorithm. The fast bootstrapping algorithm needs two additional data fields in each packet, parent change counter ***Pc (K)*** and global packet generation time. The parent change counter records the



accumulated number of parent changes, and the global packet generation time can be estimated by attaching an accumulated delay in each packet [12]. For packet, there are an upper bound and a lower bound of the difference between the estimated packet generation time and the real value . The basic idea is to reconstruct a packet's path by the help of the local packets at each hop.

V. ANALYSIS

In order to quantify the reconstruction performance of ipathand two related approaches, we analyze these approaches by a novel analytical model. Here, the performance means the probability of a successful reconstruction, which is the most important metric. We use the following definitions for analysis.

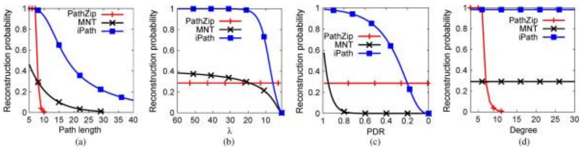
- Local packet generation period. Ipathdoes not require all nodes have the same local packet generation period. In order to simplify the presentation, we assume all nodes have the same packet generation period in this analysis section.
- Routing dynamics, which is the number of parent changes in a single period. On average, there is one parent change every local packets. We call these consecutive periods as one *cycle* for analysis.
- Packet delivery ratio PDR of packet. It can be calculated as the product of the packet reception ratios (PRR) along the routing path of packet.The average node degree.

A. Performance of ipath

The fast bootstrapping algorithm reconstructs an initial set of paths for the iterative boosting algorithm. Therefore, we first analyze the performance of the fast bootstrapping algorithm.

Performance of Iterative Boosting:

The iterative boosting algorithm reconstructs long paths based on short paths. Specifically, a path with length can be reconstructed by another path with length or (three cases in Section V-A). That path can be reconstructed by another even shorter path .A path can be successfully reconstructed only when the path helping to reconstruct it can also be successfully reconstructed. In other words, a



B. Methodology

Ipathis implemented in tinyos2.1. In the trace-driven study, we use traces collected from the citysee[3] project and the greenorbs project [2]. As mentioned in Section III, cityseeis a large-scale deployed network in an urban area for monitoring the carbon emission. Greenorbs is a large-scale sensor network for forest monitoring. A customized Collection Tree Protocol [4] is used as the routing protocol in these two projects. The cityseeand greenorbs traces include the first 10 hops in each packet for further offline analysis. Therefore, in the trace-driven study, we can use the collected routing information to reproduce the local operations on each node for each approach. Take pathzipas an example, we calculate the hash value according to the path included in each received packet at the sink side. Then, we run path zip'salgorithm to reconstruct paths and compare them to the collected ones to calculate the error ratio.

VI. CONCLUSION

In this paper, we propose iPath, a novel path inference approach to reconstructing the routing path for each received packet. iPath exploits the path similarity and uses the iterative boosting algorithm to reconstruct the routing path effectively. Furthermore, the fast bootstrapping algorithm provides an initial set of paths for the iterative algorithm. We formally analyze the reconstruction performance of iPath as well astwo related approaches. The analysis results show that iPath achieves higher reconstruction ratio when the network setting varies. We also implement iPath and evaluate its performance by a trace-driven study and extensive simulations.

Compared to states of the art, iPath achieves much higher reconstruction ratio under different network settings.

VII. REFERENCES

[1] M. Ceriotti et al., "Monitoring heritage buildings with wireless sensor networks: The Torre Aquila deployment," in Proc. IPSN, 2009, pp. 277–288.

[2] L. Mo et al., "Canopy closure estimates with GreenOrbs: Sustainable sensing in the forest," in Proc. SenSys, 2009, pp. 99–112.

[3] X. Mao et al., "CitySee: Urban CO2 monitoring with sensors," in Proc. IEEE INFOCOM, 2012, pp. 1611–1619.

[4] O. Gnawali, R. Fonseca, K. Jamieson, D. Moss, and P. Levis, "Collection tree protocol," in Proc. SenSys, 2009, pp. 1–14.

[5] S. J. D. Couto, D. Aguayo, J. Bicket, and R. Morris, "A highthroughput path metric for multi-hop wireless routing," in Proc. MobiCom, 2003, pp. 134–146.

[6] Z. Li, M. Li, J. Wang, and Z. Cao, "Ubiquitous data collection for mobile users in wireless sensor networks," in Proc. IEEE INFOCOM, 2011, pp. 2246–2254.

[7] X. Lu, D. Dong, Y. Liu, X. Liao, and L. Shanshan, "PathZip: Packet path tracing in wireless sensor networks," in Proc. IEEE MASS, 2012, pp. 380–388.

[8] M. Keller, J. Beutel, and L. Thiele, "How was your journey? Uncovering routing dynamics in deployed sensor networks with multi-hop network tomography," in Proc. SenSys, 2012, pp. 15–28.

[9] Y. Yang, Y. Xu, X. Li, and C. Chen, "A loss inference algorithm for wireless sensor networks to improve data reliability of digital ecosystems.," IEEE Trans. Ind. Electron., vol. 58, no. 6, pp. 2126–2137, Jun. 2011.

[10] Y. Liu, K. Liu, and M. Li, "Passive diagnosis for wireless sensor networks," IEEE/ACM Trans. Netw., vol. 18, no. 4, pp. 1132–1144, Aug. 2010.

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3| ISSN: 2456-3307

Comparative Study on Various Text Mining Algorithms in Data Mining

M.Prakash[1], A.Jesudasan[2]

[1]Assistant Professor, Department of Computer Science, Shanmuga Industries Arts and Science College, Tamilnadu, India

[2]Department of Computer Science, Shanmuga Industries Arts and Science College, Tamilnadu, India

ABSTRACT

This paper describes about text mining from the source of data mining. Data mining is nothing but an extraction of hidden knowledge from the huge database. There are lot of domains in data mining as text mining, image mining, sequential pattern mining, web mining and so. Here text mining can be used for extracting the information of the text using various algorithms using data mining software called WEKA. The data sets are taken from the UCI repository for performing the text mining techniques.

Keywords: Text mining, Data mining, WEKA, UCI repository, Algorithms

I. INTRODUCTION

Data mining is a technique which can be used for extracting the hidden knowledge from the huge database. The data mining can be classified into various domains named as text mining, image mining, sequential pattern mining, and web mining and so. Now, we are going to discuss about the text mining, how the information can be extracted from the database of text mining. The text mining has various fields like information retrieval, document similarity, information extraction, clustering, classification and so. Searching the similar document has an important role in text mining and document management. Classification is one of the main tasks in document similarity. It is used to classify the documents based on their category. Text mining also referred as text data mining which is similar to data analytics. Text mining is the process of deriving the highly valuable information from the text. Text mining can involve the process of structuring the input text, deriving the patterns within the structure of the data, and finally evaluation and interpretation of the output can execute. Text mining tasks includes the following methods as text categorization, text clustering, concept extraction, sentiment analysis, document summarization, and entity relational modeling. The main goal of text mining is to turn the text into the data using the application called Natural Language Processing (NLP). The term text analytics is modified as text mining by the author named Ronen Feldman. The term text analytics is also describes the application of text mining to respond the business demerits independently with queries and analyze the fielded numerical data.

II. TEXT MINING

Text mining or knowledge discovery from text (KDT) deals with the machine supported analysis of text. The text mining is a method of extracting the text and retrieves the highly valuable information's. It uses methods from information retrieval, information extraction and natural language processing (NLP) and also connects them with the algorithms and methods of Knowledge discovery of data, data mining, machine learning and statistics. This method of text mining process cannot do any mining process without the help of any algorithms. In this paper, there are three Meta classification algorithms have been used for text mining in a comparative manner. Finally resulted which algorithm will produce the high accuracy in execution of the information retrieved.

The three Meta classification algorithms are named below:

1. Attribute selected classifier.
2. Filtered classifier.
3. Logit Boost

The above mentioned three algorithms are used for mining process in the text. These algorithms are used for classifying the computer files based on their extension. For example, .docx, .pdf, .xls, .ppt, and so. The performance of Meta algorithms are analyzed by applying the performance factor such as classification accuracy and error rate. The current research in the area of text mining tackles the problems like text representation, classification, clustering or searching the hidden patterns. It is used to describe the application of data mining techniques to automated discovery of useful or interesting knowledge from unstructured or semi-structured text. The procedure of synthesizing the information by analysing the relations, the patterns, and the procedures among textual data semi-structured or unstructured text.

III. APPLICATIONS

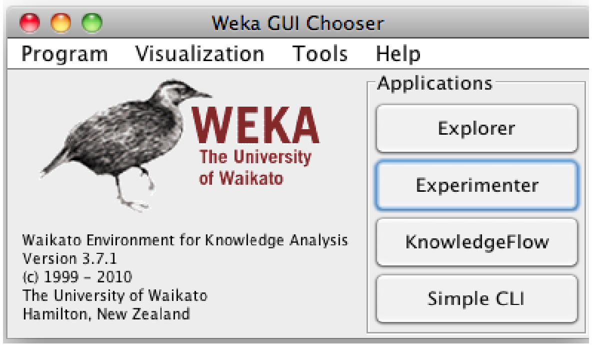
- Enterprise Business Intelligence,
- Data Mining Competitive Intelligence,
- E-Discovery,
- National Security,
- Intelligence Scientific Discovery,
- Records Management,
- Search Or Information Access And
- Social Media Monitoring.

IV. WEKA TOOL

WEKA is a machine learning software written in java and developed by the University of Waikato, New Zealand. It is free open source software licensed under the GNU general public License. The term Waikato Environment for Knowledge Analysis is shortly called as WEKA. Weka contains a collection of visualization tools and algorithms for data analysis and predictive modeling, together with graphical user interfaces for easy access to these functions. It can be used in many different application areas, in particular for educational purposes and research. Advantages of Weka are mentioned below:

- Free availability under the GNU General Public License.
- Portability, because it is implemented in the Java programming language and it runs on any modern computing platform.
- A comprehensive collection of data preprocessing and modeling techniques.
- Ease of use due to its graphical user interfaces.

Weka's main user interface is the Explorer, but essentially the same functionality can be accessed through the component-based Knowledge Flow interface and from the command line. There is also the Experimenter, which allows the systematic comparison of the predictive performance of Weka's machine learning algorithms on a collection of datasets. There are several features in explorer provide access to the main components such as pre-process, classify, cluster, visualize panel, select attribute.



Waikato Environment for Knowledge Analysis
Version 3.7.1
(c) 1999 - 2010
The University of Waikato
Hamilton, New Zealand

V. METHODOLOGY

Text classification is one of the important research issues in the field of text mining where the documents are classified with supervised knowledge. The main objective of this process is to find the best classification algorithm among:

- Attribute Selected Classifier,
- Filtered Classifier and
- Logit Boost.

The methodology of this paper work is as collection of the data set from the UCI repository, and implement that data into the Meta classification algorithms, and checks about the performance factor of these algorithms and finally producing the best algorithm, which can produce the best result by having high accuracy of data and low error rate in execution.

Here, the computer files from the system hard disk can be taken as dataset for text mining process. By taken these dataset, it will be implemented into the above mentioned three Meta classification algorithms and proceed for the mining process. After the implementation of data set into the algorithms, the performance factors can be calculated by producing the classification accuracy and error rate. From this methodology, we are finalizing the best algorithm by monitoring which will produce the best result.

VI. DATASET

A dataset can be collected from the computer systems, which are stored in the hard disk. The dataset can contains minimum of 9000 instances and four attributes namely file name, file size, extension and file path. Weka data mining tool is used for analyzing the performance factor of the classification algorithms.

VII. CLASSIFICATION ALGORITHMS

Classification is an important data mining technique with broad applications. It is used to classify each item in a set of data into one of predefined set of classes or groups. Classification algorithm plays an important role in document classification. There are various Meta classification algorithms such as Attribute Selected Classifier, Bagging, Decorate, Vote, Filtered Classifier, Logit Boost, END, Rotation Forest, and so on. Now, we have analyzed three Classification Meta Algorithms. The algorithms are namely Attribute Selected Classifier, Filtered Classifier and Logit Boost.

A. Attribute Selected Classifier

Dimensionality of training and test data is reduced by attribute selection before being passed on to a classifier. Some of the important options in attribute selected classifier as Classifier, Debug, Evaluator, and Search. Here, the base classifiers are used and from debug method it set to true and the classifier may output additional information to the console. Evaluator set the attribute evaluator to use and is used during the attribute selection phase before the classifier is invoked. Set the search method for using during the attribute selection phase before the classifier is invoked.

B. Filtered Classifier

From the filtered classifier class is used for running an arbitrary classifier on data that has been passed through an arbitrary filter. Similar to classifier, the structure of the filter is based exclusively on the training data and test instances will be processed by the filter without changing their structure. Some of the important options in Filtered classifier are Classifier, Debug, and Filter.

C. Logit Boost

Logit Boost algorithm is an extension of Ada boost algorithm. It replaces the exponential loss of Ada boost algorithm to conditional Bernoulli likelihood loss. This Class is used for performing additive logistic regression. This class performs classification using a regression scheme as the base learner, and can handle multiclass problems.

VIII. EXPERIMENTAL RESULT ACCURACY AND ERROR RATE

There are various measures used for classification accuracy such as true positive rate, precision, F Measure, ROC Area, and kappa Statistics. The TP Rate is the ratio of play cases predicted correctly cases to the total of positive cases. F Measure is a way of combining recall and precision scores into a single measure of performance. Precision is the proportion of relevant documents in the results returned. ROC Area is a traditional to plot the same information in a normalized form with 1-false negative rate plotted against the false positive rate.

Table I
Accuracy Measures of Classification Algorithms

Parameter	Attribute selected classifier	Filtered classifier	Logit boost
Correctly classified instances	95.44	97.12	97.91
Incorrectly classified instances	4.56	2.88	2.09
TP rate	95.40	97.10	97.90
Precision	95.30	95.40	98.10
F measure	94.90	96.10	97.70
ROC area	99.00	99.80	99.90
Kappa statistics	94.25	96.37	97.37

From the above mentioned table, the accuracy measures of classification algorithms the values executed for Logit Boost algorithm produces a high performance factor in accuracy by comparing with other two algorithms.

IX. ERROR RATES

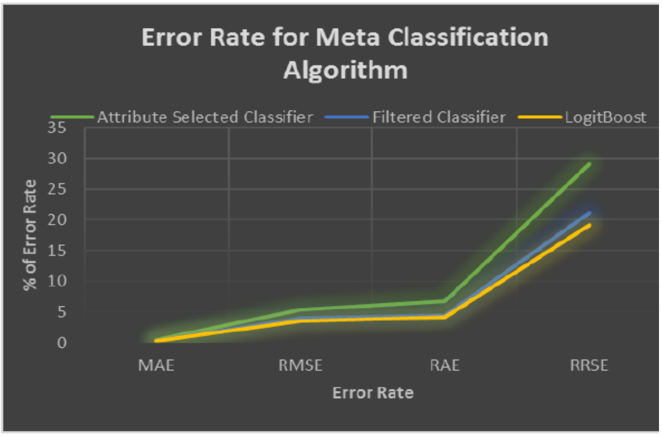
There are some types of errors as mentioned below,

- Mean absolute error (M.A.E),
- Root mean square error (R.M.S.E),
- Relative absolute error(R.A.E) and
- Root relative squared error (R.R.S.R).

Table IIError rate

Algorithm	MAE	RMSE	RAE	RRSE
Attribute selected classifier	0.46	5.41	6.69	29.11
Filtered classifier	0.31	3.94	4.45	21.19
Logit boost	0.29	3.56	4.18	19.13

The mean absolute error (MAE) is defined as the quantity used to measure how close predictions or forecasts are to the eventual outcomes. The root mean square error (RMSE) is defined as frequently used measure of the differences between values predicted by a model or an estimator and the values actually observed. Relative error is a measure of the uncertainty of measurement compared to the size of the measurement. The root relative squared error is defined as a relative to what it would have been if a simple predictor had been used.



From the above table and figure we conclude that the Logit Boost algorithm produces the low error rate than the attribute selected classifier and filtered classifier.

X. CONCLUSION

Data mining can be defined as the extraction of useful knowledge from large data repositories. Text mining is a technique which extracts information from both structured and unstructured data and also finding patterns which is novel and not known earlier. Here, the classification Meta algorithms are used for classifying computer files which are stored in the computer. The Classification Meta algorithms include three techniques namely Attribute Selected Classifier, Filtered Classifier and Logit Boost. By analyzing the experimental results it is observed that the Logit Boost classification technique has yields better result than other techniques.

XI. REFERENCES

[1] Abdullah Wahbeh H, Mohammed Al-Kabi., “Comparative Assessment of the Performance of Three WEKA Text Classifiers Applied to Arabic Text”, Vol. 21, No. 1, pp. 15- 28, 2012.

[2] Abdullah Wahbeh H, Qasem Al-Radaideh A, Mohammed Al-Kabi N, and Emad Al-ShawakfaM., “A Comparison Study between Data Mining Tools over some Classification Methods”.

[3] Artur Ferreira., “Survey on Boosting Algorithms for Supervised and Semi-supervised Learning”.

[4] Christophe Giraud-Carrier., “Meta learning - A Tutorial”.

[5] ChristophGoller, Joachim Löning., Thilo Will, Werner Wolff., “Automatic Document Classification: A thorough Evaluation of various Methods”.

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3| ISSN : 2456-3307

Cheating Mechanism in Visual Cryptography a Novel Method with Efficient Halftoned Image with SBR Technique By Sealing The Algorithm

Ms. R. Nandhini[1], Mrs. S. Shanthi[2], Ms. A. Sivasankari[3]

[1]Research Scholar, Dept of Computer Science and Applications, D.K.M. College for Women (Autonomous), Vellore, Tamilnadu, India

[2]Assistant Professor, Dept of Computer Science and Applications, D.K.M. College for Women (Autonomous), Vellore, Tamilnadu, India

[3]Head of the Department (CS), Dept of Computer Science and Applications, D.K.M. College for Women (Autonomous), Vellore, Tamilnadu, India

ABSTRACT

The cheating is main issue while transferring the secret image against the owner. To avoid this Visual Cryptography novel method technique of cryptography is used. In this we divide secret images into multiple shares and are distributed to various entities. To get back the secret image we need all the shares. Using different operations we can reconstructed by superimposing these share. The pixel expansion and noise at output is a major drawback in Traditional. The major problem is cheating between shareholders cheating owners and between them. To avoid these limitation sealing algorithm with two application of visual cryptography (VC) are MIVC and EVC is used. The image will be changed into halftoned representation before sealing the algorithm. For this we are using SBR technique. The output overcomes common limitation likes pixel expansion and clarity of image along with this the system also provides a cheating prevention mechanism. Two secret image can be send at the same time by connecting them into halftonedrepresentations which are partitioned as totally three shares.

Keywords: Cheating, Shares, Halftoned Image, Pixel Expansion, Sealing Algorithm, Visual Cryptography.

I. INTRODUCTION

Visual Cryptography technique allows the encryption of an image without having any complex computations or cryptography knowledge. It is simply secrete sharing technique. In this secrete image are distributed to different entries. This technique initially proposed by Naor and Shamir in the year 1994. In this we need two transparent images. One image contains random pixels and another contains secret information. To retrieval the information, we need all the shares layers are to be overlapped. Because decryption is done only when all shares are overlap together, we get original image.

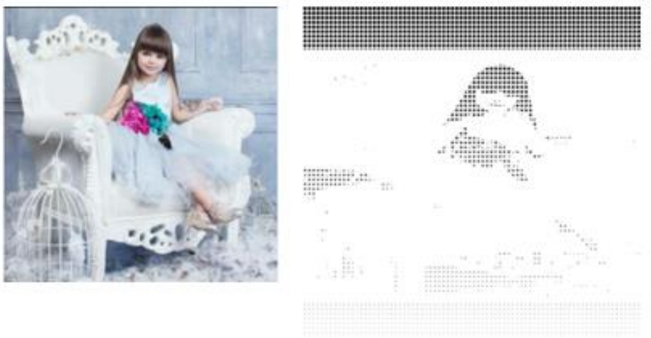


Figure 1. Basic Model of Visual Cryptography

In Visual Cryptography Biometric is an important security application. It has a Facial, Fingerprints, Signature image are used as secret key. To kept these images secret we distribute these images and shared. After different entities release the shares it overlapped to get back the secrete image. To solve the common and traditional drawbacks in Visual cryptography we use novel method. This includes image clarity, pixel expansion, cheating between shareholders and image content owner. Here two secret image are dived into three shares. Therefore three cover image are sealed into these shares for high security protection. For increase the security we can also add the pin number or password can be entered before the share creation and it should be identified correctly while overlapping of shares. Then only secret image is viewed to the user.

II. OBJECTIVE

In this the research is to test secure visual cryptography with halftoned images, do not require more pixels in the shared images. Better Image clarity after overlap secret image.

The overlap secret image was proposed with

To get secret image as better clarity and no loss of pixels.

As in the original image pixels is equal to the pixel in after overlapped secret image.

Apply cheating mechanism method used for secured transmission of secret image preventing from cheating.

III. PROPOSED METHOD

Novel Algorithm:

Step1: Read the input image.

Step 2: It is decomposed into three-share image based on RGB.

Step 3: The halftone technique is applied to these images to binary images.

Step 4: Read the share images or key images.

Step 5: Perform XOR operation between binary images obtained in halftone process.

Step 6: Repeat this process to every binary images.

Step 7: Now to overlap to get the original image will be back.

For receiver side process

Step 1: Read the encrypted image.

Step 2: Split the image.

Step 3: For each image perform these: Split the image into RBG.

Step 4: De-embedded the share image and each binary secrete image.

Step 5: Inverse the halftoning is performed to each image.

Perform Halftoning image process for each Pixel:

Halftoning: It is the process of converting images with greater amplitude resolution to lesser amplitude resolution.

Input: The matrix D cxd and a pixel with gray level g in input image I

Process [I]

For i=0 to c-1 do

If g=Dij then print black pixelat position (i,j)

Else print while pixel at position (i,j)

Output:halftoned image contain in the position of(i,j)

SBR Technique

Performs fully automatic mono-model registration of both images.

Step 1: Select image near the center of the dataset as a template image.

Step 2: Compute information and edge direction for each pixel.

Step 3: For each target image landmark pixel is assigned automatically.

Step 4: Compute edge information for landmark pixel

Step 5: Search matching pixel for landmark pixel in the template using kernel equation.

Embedded EVCs

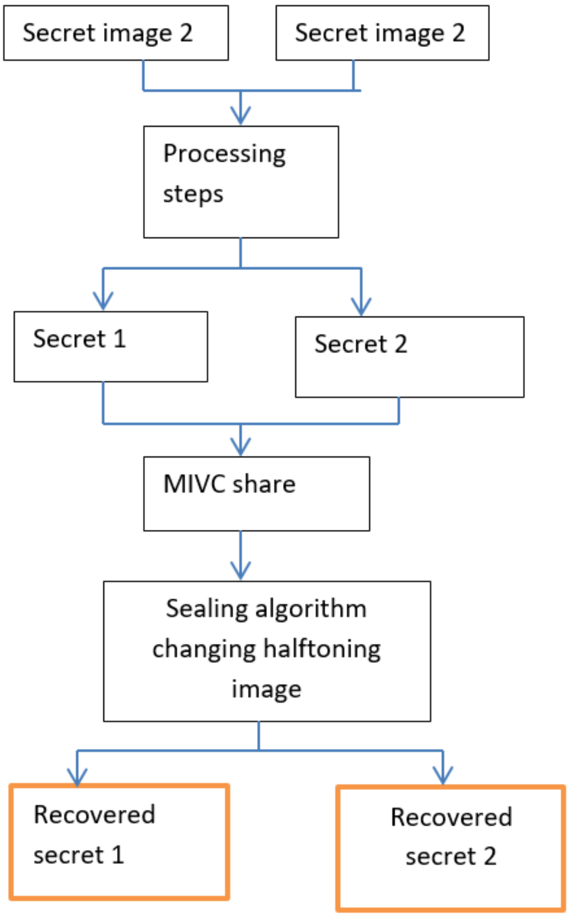
It contain two main steps
Step 1: n covering share will be generated.
Step 2: By embedding the corresponding vcs into the n shares.
First to generate the n covering shares for an access take gray-scale original share images and output n binary meaningful share.
Second use corresponding vcs to encode a secret image and then embedded shares that were generated.
Third decrypt the embedded shares them again to get original image.

Process

Input: n covering shares of (c0,c1) with pixel expansion and secret image I1
Step 1: Divide the covering shares into blocks that contain sub pixels
Step 2: Choose m embedding position in each block in n covering shares.
Step 3: For each black and white pixel in I1 randomly choose a share matrix M belongs to c1
Step 4: Embedded the m sub pixel of each row and column of the share matrix M into the m embedded position.
Output: n embedded share are e0,e1,e2…en-1

MIVC Share

Step 1: Two secrete image is masked.
Step 2: Rotate the pixel in the binary image of share 1 divided into two shares.
Step 3: Secret image became visible by superimposing the share 1 and share 2.
Step 4: It will combined with same share 3.



IV. CONCLUSION

For security issues while sharing secret images or keys is an important part. VCs is an efficient secret sharing technique where secret image into number of encrypted images. This work proposed sealing algorithm of novel method where secret image is divided into three shares. Share 1 and 2 generate secret image 1. Secret image 2 is obtained by rotating pixels of binary image in share 1 and combined the same with share 3. It prevents cheating between shareholders and owners. It will overcome loss of image clarity as well as pixel expansion as in the same ratio in the original image.

V. REFERENCES

[1] NazaninAskari, Howard M. Heys, Member, IEEE, and Cecilia R. Moloney, Member , IEEE "Novel Visual Cryptography Schemes without Pixel Expansion for halftone images" IEEE trans. 2014.

[2] Xiatian Wu and Wei Sun, "Extended Capabilities for XOR based Visual Cryptography" in IEEE 2013.

[3] Young-Chang Hou, Shih-Chieh Wei and Chia-Yin Lin "Random-Grid –Based Visual Cryptography" Schemes in the year of2013 in IEEE,2013.

[4] Pallavi V. Chavan and Dr. Mohammad Atique, "Design of Hierarchical Visual Cryptography" in IEEE, 2012.

[5] Z.Zhou, G. R. Arce, and G. D. Crescenzo, "Halftone Visual Cryptography" IEEE Trans. Image Process., vol 15, no 8, pp. 24412453, Aug 2006.

[6] A. Ross and A. A. Othman, "Visual cryptography for biometric privacy" IEEE Trans Inf Forensics Security, vol. 6, no 1, pp.7081, Mar. 2011.

An Efficient Management for Map Reduce Using Partition and Aggregation in Software Application

P. Ramya[1],Dr A. Saravanan[2]

[1]Research Scholar,Department of CSC, Sun arts and science college, Thiruvannamalai, Tamil Nadu, India
[2]Professor, Department of CSC, Sun arts and science college, Thiruvannamalai, Tamil Nadu, India

ABSTRACT

In this paper, we study to reduce data traffic and to avoid duplication using a MapReduce technique and data partition scheme. Aggregator problem and large-scale optimization problem of duplication and data traffic were made in online or offline. In these problem we use map reduce technique to clustering the data and use k-nearest algorithm is used to reduce the time and cluster the nearest data to avoid conflict. Then we also use ProMiSH based on random projections and hashing for partition process to avoid data traffic in the search engine. In this, we also using k-nearest algorithm for aggregation to clustering the nearest neighbor data and using ProMiSH based on random projections and hashing for partition process to avoid data traffic in the search engine. Then these we use algorithms to decrease the traffic and conflict while processing the data and improve the speed to access the data faster.

Keywords: MapReduce, NSK, K-nearest, ProMiSH

I. INTRODUCTION

We use the technology of MapReduce concept of clustering to reduce the data traffic and to aggregate the data by using nearest neighbour of k-nearest algorithm then we using the ProMiSH based on random projections and hashing for partition process to avoid data traffic in the search engine. Then these we use algorithms to decrease the traffic and conflict while processing the data and improve the speed to access the data faster.

II. PROBLEM DESCRIPTION

Problem is while we searching a data in any applications or search engine it has having data traffic and unrelated data will be shown. It will create a conflict to the user and also these data traffic will be increase the time of the user. In these problem we use map reduce technique to clustering the data and use k-nearest algorithm is used to reduce the time and cluster the nearest data to avoid conflict. Then we also use ProMiSH based on random projections and hashing for partition process to avoid data traffic in the search engine.

III. EXISTING SYSTEM

In this paper, we study nearest keyword set (referred to as NKS) queries on text-rich multi-dimensional datasets. NKS query is two-dimensional data access and it process the top-k clusters these algorithmsis used to increase size or dimensionality in datasets. In three-based indexes, suggest possible solutions to NKS queries on multidimensional dataset. Then NKS is the user-defined keywords.

IV. DISADVANTAGES OF THE EXISTING SYSTEM

- NKS queries are useful for graph pattern search, where the graphs are in a high dimensional space.
- Nearest neighbor queries usually require coordinate information for queries, which makes it difficult to develop an efficient method to solve NKS queries by existing techniques for nearest neighbor search.

V. PROPOSED SYSTEM

A. Design Considerations

To reduce network traffic within a MapReduce job, we have to consider aggregate data with similar keys before sending them to remote reduce tasks. Even though we have a similar function, called combiner, which has been already adopted by ProMiSH, it operates immediately after a map task for its generated data, failing to exploit the data aggregation opportunities among multiple tasks on different machines. Objective is to minimize the total network traffic by Data partition and aggregation for a MapReduce job.

In this paper, we propose ProMiSH (short for Projection and Multi-Scale Hashing) to enable fast processing for NKS queries. In particular, we develop an exact ProMiSH (referred to as ProMiSH-E) that always retrieves the optimal top-k results, and an approximate ProMiSH (referred to as ProMiSH-A) that is more efficient in terms of time and space, and is able to obtain near-optimal results in practice.

ProMiSH-E uses a set of hash tables and inverted indexes to perform a localized search. Based on this index, we developed ProMiSH-E that finds an optimal subset of points and ProMiSH-A which searches near-optimal results with better efficiency. ProMiSH is faster than state-of-the-art tree-based techniques, with multiple orders of magnitude performance improvement.

B. System Architecture

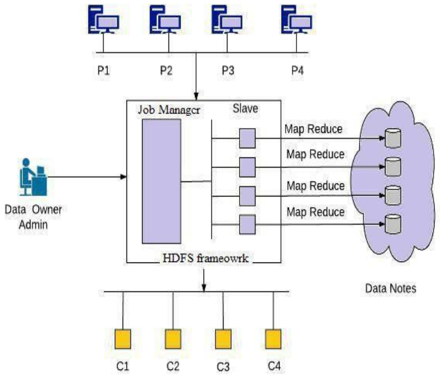


Figure 1. Architecture of Proposed System

The incoming applications from data generators is received by the Job Manager where it is partitioned and Map/Reduce Tasks are carried out. The data is portioned and stored on the nodes by using load balancing techniques to minimize traffic. The Clients ask queries to the system.

Advantages:

1. The performance of ProMiSH on both real and synthetic datasets.
2. We develop efficient search algorithms that work with the multi-scale indexes for fast query processing.

VI. RESULT AND ANALYSIS

Through this project we plan to solve the below two problems:

- Data traffic: by designing a novel intermediate data partition scheme we aim to reduce network traffic cost. And solve the problem of data traffic in search engines and software
- Aggregator placement problem: A decomposition based distributed algorithm is proposed to with the large-scale optimization problem for an software application.

VII. CONCLUSION

In this paper, we proposed solutions to the problem of top-k nearest keyword set search in multi-dimensional datasets. We proposed a novel index called ProMiSH based on random projections and hashing. Based on this index, we developed ProMiSH-E that finds an optimal subset of points and ProMiSH-A that searches near-optimal results with better efficiency.

VIII. FUTURE WORK

The future scope of the works should be made on complex data partitioning in the database where more intelligent methods have to be employed. This includes analyzing computation cost, skew record etc. So that optimization of data partition is done in map reduce. If the cluster is dynamically growing, the index of the cluster also keep grows.

IX. REFERENCES

[1] C. Long, R. C.-W. Wong, K. Wang, and A. W.-C. Fu, “Collective spatialkeyword queries: a distance owner-driven approach,” in SIGMOD, 2013.

[2] D. Zhang, B. C. Ooi, and A. K. H. Tung, “Locating mapped resourcesin web 2.0,” in ICDE, 2010, pp. 521–532.

[3] V. Singh, S. Venkatesha, and A. K. Singh, “Geo-clustering of imageswith missing geotags,” in GRC, 2010, pp. 420–425.

[4] X. Cao, G. Cong, C. S. Jensen, and B. C. Ooi, “Collective spatialkeyword querying,” in SIGMOD, 2011.

[5] W. Li and C. X. Chen, “Efficient data modeling and querying systemfor multi-dimensional spatial data,” in GIS, 2008, pp. 58:1–58:4.

[6] V. Singh, A. Bhattacharya, and A. K. Singh, “Querying spatial patterns,”in EDBT, 2010, pp. 418–429.

[7] J. Bourgain, “On lipschitz embedding of finite metric spaces in Hilbert space,” Israel J. Math., vol. 52, pp. 46–52, 1985.

[8] H. He and A. K. Singh, “Graphrank: Statistical modeling and mining ofsignificantsubgraphs in the feature space,” in ICDM, 2006, pp. 885–890.

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3| ISSN: 2456-3307

Mobile Ad Hoc Networks (MANET)-An Overview

S.Revathi[1], **E.Bhuvaneswari**[2]
[1]Computer Science, Kamban college of Arts and Science for women/ Thiruvannamalai, TN, India
[2]Assistant professor of Computer Science, Kamban College of Arts and Science for women /Thiruvannamalai, TN, India

ABSTRACT

Mobile ad hoc networks (MANETs) are a subclass of wireless ad hoc networks having special characteristics of dynamic network topology and moving nodes. Mobile ad hoc networks (MANETs) are infrastructure-less self-configuring networks designed to support mobility. The main of this paper is to provide an overview of MANET including its concept, features, types, and its applications along with the routing protocols used for communication.

Keywords: Mobile ad hoc networks, Vehicular ad hoc networks, Smart phone ad hoc networks, Flying ad hoc networks

I. INTRODUCTION

Mobile computing is one of the most important technologies supporting pervasive computing. Advances in both hardware and software techniques have enabled the spread of mobile hosts and wireless networking to masses. Generally, there are two modes in which wireless mobile nodes can communicate,

1. **Infrastructured**: In this mode, all the communication among the mobile nodes goes through a base station. A Base station is also known as an access point. These base stations are connected to the fixed infrastructure or wired networks.

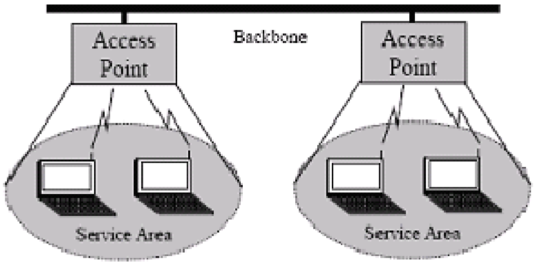


Figure 1. Infrastructure-Based Wireless Network

2. **Infrastructure less**: This mode of communication is known as a mobile ad hoc network. A MANET is a collection of wireless nodes that can dynamically form a network to exchange information without using any pre-existing fixed network infrastructure. This is a very important part of communication technology that supports truly pervasive computing because in many contexts information exchange between mobile units cannot rely on any fixed network infrastructure, but on the rapid configuration of wireless connection on-the-fly. A typical example of this mode of communication is people sitting in the conference room and exchanging data among them without any fixed infrastructure.

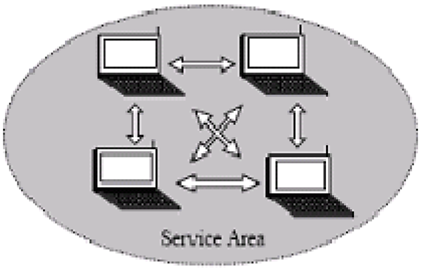


Figure 2. Infrastructure less (Ad Hoc) Wireless Network

II. CONCEPT OF MANET

A mobile ad-hoc network is a collection of wireless nodes that can dynamically be set up anywhere and anytime without using any pre-existing network infrastructure. It is an autonomous system in which mobile hosts connected by wireless links are free to move randomly. In MANET, nodes can be act as both host and routers.

III. FEATURES OF MANET

The features of MANET are given below:

1. Autonomous terminal: In MANET, each mobile host is an autonomous node, which may function as both a host and a router. In other words, besides the basic processing ability as a host, the mobile nodes can also perform switching functions as a router. So usually endpoints and switches are indistinguishable in MANET.

2. Distributed operation: There is no background network for the central control of the network operations, the control and management of the network is distributed among the terminals. The nodes involved in a MANET should collaborate amongst themselves and each node acts as a relay as needed, to implement functions e.g. security and routing.

3. Multi-hop routing: Basic types of ad hoc routing algorithms can be single hop and multi-hop. Single-hop MANET is simpler than multihop in terms of structure and implementation, with the cost of lesser functionality and applicability. When delivering data packets from a source to its destination out of the direct wireless transmission range, the packets should be forwarded via one or more intermediate nodes.

4. Dynamic network topology: Here the nodes are mobile, the network topology may change rapidly and unpredictably and the connectivity among the terminals may vary with time. MANET should adapt to the traffic and propagation conditions as well as the mobility patterns of the mobile network nodes. The mobile nodes in the network dynamically establish routing among themselves as they move about, forming their own network on the fly. Moreover, a user in the MANET may not only operate within the ad hoc network but may require access to a public fixed network.

5. Fluctuating link capacity: The nature of high bit-error rates of wireless connection might be more profound in a MANET. One end-to-end path can be shared by several sessions. The channel over which the terminals communicate is subject to noise, fading, and interference, and has less bandwidth than a wired network. In some scenarios, the path between any pair of users can traverse multiple wireless links and the link themselves can be heterogeneous. One effect of the relatively low to moderate capacities is that congestion is typically the norm rather than the exception i.e. aggregate application demand will likely approach or exceed network capacity frequently.

6. Energy-constrained operation: Some or all of the nodes in a MANET may rely on batteries or other means for their energy. Such devices need optimized algorithms and mechanisms that implement the computing and communicating functions.

7. Limited physical security: MANETs are generally more vulnerable to physical security threats than are

fixed cable networks. The increased possibility of eavesdropping, spoofing and denial-of-service attacks should be carefully considered.

IV. TYPES OF MANET

1. **Vehicular ad hoc networks** (VANETs) **are** used for communication between vehicles and roadside equipment. Intelligent vehicular ad hoc networks (In VANETs) are a kind of artificial intelligence that helps vehicles to behave in intelligent manners during vehicle-to-vehicle collisions, accidents.

2. **Smart phone ad hoc networks** (SPANs) leverage the existing hardware (primarily Bluetooth and Wi-Fi) in commercially available smart phones to create peer-to-peer networks without relying on cellular carrier networks, wireless access points, or traditional network infrastructure. SPANs differ from traditional hub and spoke networks,such as Wi-Fi Direct, in that they support multi-hop relays and there is no notion of a group leader so peers can join and leave at will without destroying the network.

3. **Internet-based mobile ad-hoc networks (iMANETs)** is a type of wireless ad hoc network that supports Internet protocols such as TCP/UDP and IP. The network uses a network-layer routing protocol to link mobile nodes and establish routes distributedly and automatically.

4. **Flying ad hoc networks (FANETs)** are composed of unmanned aerial vehicle, allowing great mobility and providing connectivity to remote areas.

V. APPLICATIONS OF MANET

Mobile ad-hoc networks are the only choice for mobility support where there is no infrastructure or it is too expensive. Some application areas of such use of MANET are given below:

1. **Instant Infrastructure:** Unplanned meetings, spontaneous interpersonal communications etc. Cannot rely on any infrastructure, therefore ad-hoc connectivity has to be set up.

2. **Disaster relief:** Disaster break infrastructure and emergency teams have to rely on the infrastructure they set up themselves. Therefore, ad-hoc networks can be the solution.

3. **Military Activities:** Many military activities are confidential and for security reasons it is good to use ad-hoc connectivity for communication.

Remote Area: In sparsely populated and hilly areas it is too expensive to set up an infrastructure. Depending on the communication pattern, ad-hoc networks can be the solution.

VI. CLASSIFICATION OF ROUTING PROTOCOLS IN MANET

Mobility of nodes and rapidly changing topology are such characteristics of the MANET network that make routing decisions more challenging. Several other factors such as power and storage constraints and security makes routing more challenging in VANET. Routing protocols can be classified on various basis such as on the topology of network for routing i.e. proactive and reactive routing protocols, on the basis of communication strategy used for delivery of information from source to destination i.e. unicast, multicast and broadcast. In this paper, classification is done using topology information. Topology-based routing protocol uses topology information which is stored in the routing table as a basis to forward packets from source node to the destination node. They are further divided into three groups as Proactive, Reactive and Hybrid Protocols.

The Classification done using Topology is represented below,

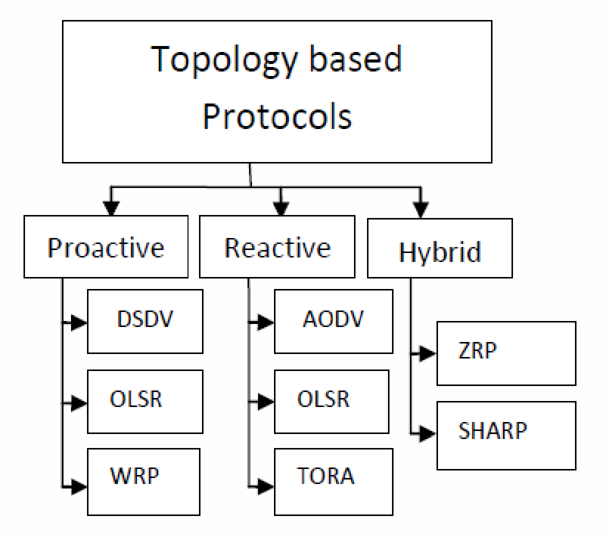


Figure 3. Classification of Routing Protocols of MANET

A. Proactive Routing Protocol

Proactive MANET protocols are table-driven and will actively determine the layout of the network. Through a regular exchange of network topology packets between the nodes of the network, a complete picture of the network is maintained at every single node. There is hence minimal delay in determining the route to be taken. This is especially important for time-critical traffic. However, a drawback to a proactive MANET of protocol is that the life span of a link is significantly short. This phenomenon is brought about by the increased mobility of the nodes, which will render the routing information in the table invalid quickly. When the routing information becomes invalid quickly, there are many short-lived routes that are being determined and not used before they turn void. Hence, another drawback resulting from the increased mobility is the amount of traffic overhead generated when evaluating these unnecessary routes. This is especially aggravated when the network size increases. The fraction of the total control traffic that consists of actual practical data is further decreased.

Example of Proactive Routing Protocol

- Destination-Sequenced Distance Vector (DSDV)
- Optimized Link State Routing (OLSR)
- Wireless Routing Protocol (WRP)

B. Reactive Routing Protocol

On-demand routing is a popular routing category for wireless ad hoc routing. It is a relatively new routing philosophy that provides a scalable solution to relatively large network topologies. The design follows the idea that each node tries to reduce routing overhead by only sending routing packets when communication is requested. Common for most on- demand routing protocols are the route discovery phase where packets are flooded into the network in search of an optimal path to the destination node in the network.There are numerous on-demand routing protocols, but only two of them are more significant. These are Ad Hoc On-Demand Distance Vector Routing (AODV) and Dynamic Source Routing (DSR). These two have been chosen because both have been extensively evaluated in the MANET literature and are being considered by the Internet Engineering Task Force (IETF) MANET Working Group as the leading candidates for standardization. Thus, reactive MANET protocols are most suited for networks with high node mobility or where the node transmits data infrequently.

Example of Reactive Routing Protocol

- Ad Hoc On-Demand Distance Vector (AODV)
- Dynamic Source Routing (DSR)
- Temporally Ordered Routing Algorithm (TORA)

C. Hybrid Routing Protocol

Since proactive and reactive routing protocols each work best in oppositely different scenarios, there is good reason to develop hybrid routing protocols, which use a mix of both proactive and reactive routing protocols. These hybrid protocols can be used to find a balance between the proactive and reactive protocols. The basic idea behind hybrid routing protocols is to use proactive routing mechanisms in some areas of the network at certain times and reactive routing for the rest of the network. The proactive operations are restricted to a small domain in order to reduce the control overheads and delays. The reactive routing protocols are used for locating nodes outside this domain, as this is more bandwidth-efficient in a constantly changing network.

Example of Hybrid Routing Protocol

- Zone Routing Protocol (ZRP)
- Sharp Hybrid Adaptive Routing Protocol (SHARP)

VII. CONCLUSION

In this paper, an overview on Mobile ad hoc networks (MANETs) is presented including concept of MANETs, its applications and features that distinguish it from other wireless networks. Due to these features, there is need of separate routing protocols for MANET. Classification of routing protocols for MANET has been done on the basis topology of the network i.e. proactive or table- driven and reactive or demand- driven. A summarized overview of routing protocols belonging to each type of classification has also been presented which is very useful when we get to know about an overview of MANET. We concluded that MANET routing protocols are designed based on the application area and environment and it is not possible to design a single protocol, which is suitable for all MANETs.

VIII. REFERENCES

[1] J. H. Schiller, mobile communications, Pearson Education, 2003.

[2] Referring www.google.com/manets Wikipedia.

[3] L.Abusalah, A.Khokhar - A survey of secure mobile ad hoc routing protocols, Tutorials

[4] Referring http:// Gupta - A literature Survey of Manet

[5] Referring www.google.com/types of manets wireless network.

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3| ISSN: 2456-3307

A Review on Compression and Encryption of Data for Secure Transmission

R. Saikumar[1], M.Sivaranjani[2]
[12]Department of Computer Science, Bharathiar University, Coimbatore, India

ABSTRACT

Data is the digital information to be stored in the computer in the form of text documents, audios, videos, or other types of data. Security is about, the protection of those assets stored in a computer. Compression algorithms can be used to reduce the redundancy of the data representation. Data compression is a very good technique which can be used to reduce the size of the data and storing the same amount of data comparatively smaller bits resulting in reducing the data storage space, resource usage. There are a lot of techniques have been implemented for the process of data compression which can be categorized as Lossy and Lossless data compression techniques. Data compression is an attractive approach to reduce the communication cost by effectively utilizing the available bandwidth in the data links. This data represents a variety of objects from the various multimedia data such as text, images, videos, sound clippings, computer programs, graphs, charts, maps, tables, etc. Over the last era, there has been recording explosion in the amount of digital data transmitted through Internet, representing text, images, video, sound, computer programs, etc. The researchers are developing the novel algorithms which can be used for data compression. It is also important to consider the security aspects of the data being transmitted while compressing it, as most of the data transmitted over the Internet is very much vulnerable to an aggregate of attacks. This presentation is focused on addressing this problem of lossless compression of multimedia files with an added security.

Keywords: Data compression, Digital Information, Security, Lossy and Lossless, Multimedia, Encryption/ Decryption

I. INTRODUCTION

Data Compression is one of the best technique for encoding the data so that it takes less storage space or less transmission time over the internet. Compression is possible because of, most of the real biosphere information's are jobless. Data Compression is a technique that can decrease the size of the data by removing unnecessary information from the file and duplicity of the file. Data compression is a skill of plummeting the number of bits needed to store or transmit the data over the network transmission. Two types of data compression techniques are there: Lossy and Lossless data compression.

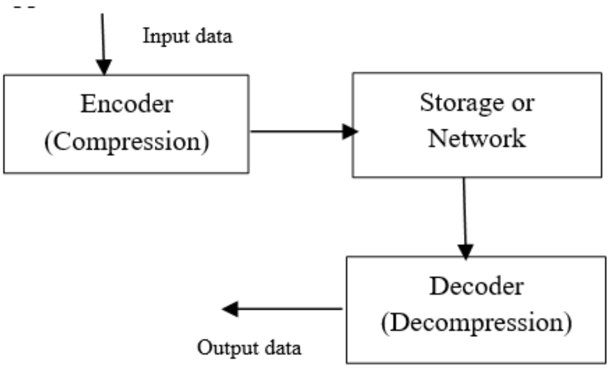


Figure 1. Diagrammatic Representation of Compression

A Lossy data compression method is one where the compressing data and then decompressing it retrieves the data that may well be different from the original. Lossy data compression is used frequently the Internet for the techniques of streaming media and telephony applications.

Lossy compression is similarly known as irreversible compression. It is a class of data encoding method, which can be used approximations and fractional data neglect to signify the content. These techniques are used to reduce the size of the data for a storage, handling the data, and transmitting the content. This data represents a variety of objects from the multimedia field such as text, images, video, sound, computer programs, graphs, charts, maps, tables, mathematical equations etc. Digital Libraries Initiative (DLI), have furnished several research projects whose goal is to collect, store, and organize an information in a digital form, and make it available for searching, retrieval, and processing via communication networks. The speed of the data transfer from the disk to memory is faster than the normal data. The security goals for the data are Confidential, Authentication, Integrity, and Non-repudiation. Information security is an emergent issue among IT organizations of all sizes. To grab this growing concern, more and more IT firms are moving towards cryptography to protect their valuable information. In addition to above concerns over securing the stored data, IT organizations are also facing challenges with ever-increasing costs of storage required to make sure that, there is enough storage capacity to meet the organization's present and future demands. It involves transforming data of a given format, called source message to data of a smaller sized format called code word. Data encryption is known for protecting an information from the spying. It transforms data of a given format, called plaintext, to another format, called ciphertext, using an encryption key. Now, the compression and encryption methods are done distinctly. Cryptography prior to the modern age was effectively synonymous with encryption, the conversion of information from a readable state to apparent nonsense. Cryptography's potential for use as a tool for espionage and sedition has led many governments to classify it as a weapon and to limit or even prohibit its use and export. Lossless compression is one of the data compression methods which can be used to allow the exceptional data to be perfectly restored from the compressed data. Lossless data compression is a method existed in many applications. For example, these methods can be used in the ZIP format and in the format of GNU tool GZIP. Some of the image files can use the formats like PNG or GIF, can be used only lossless compression, while other files like TIFF and MNG can be used either lossless or Lossy method. Maximum, lossless compression methods can ensure two things: the first step, generates a statistical model for the input data, and the second step, uses this model to map input data to bit sequences in such a way that "probable" data will produce shorter output than "improbable" data.

II. COMPRESSION

The importance of an information and the communication systems for society and the global economy is intensifying with the increasing value and quantity of data that is transmitted and stored on those systems. Here, those systems and data are also highly vulnerable to the variety of threats, such as unauthorized access, alteration of the data, and destruction of the data. Encryption and decryption is nothing but hiding the data and un-hiding the data from the unknown users due to some security reasons A cipher is used to accomplish the encryption and decryption. Merriam-Webster's Collegiate Dictionary defines cipher as ―a method of transforming a text in order to conceal its meaning. An information which is veiled is called plaintext; once it has been encrypted, it is called ciphertext. To hide any data two techniques are mainly used one is Cryptography other is Steganography. In this paper we use

Cryptography. Cryptography is nothing but a method of protecting the data, which provide the countless approaches for transforming the data into an unreadable form, so that Valid User can access an information at the destination.

1) Advantages of Compression:

- Less disk space
- Reading and writing faster
- Faster file transfer
- Variable dynamic range
- Byte order independent

2) Disadvantages of Compression:

- Effects of error in transmission
- Added complication
- Slower for sophisticated method
- Unknown byte/pixel resolution
- Decompress all data

Data compression is the best method for reducing the cost of communication by encrypting the data for secure transmission using bandwidth, which is available. Compression algorithms reduce the redundancy in data representation to decrease the storage required for that data. Over the last decade, there has been an unprecedented explosion in the amount of digital data transmitted via the Internet, representing text, images, video, sound, computer programs etc. Data compression implies sending or storing a smaller number of bits. Compression is the reduction in the size of data in order to save space or transmission time. Many methods are used for this purpose, in general, these methods can be divided into two broad categories: Lossy and Lossless methods. Lossy Compression generally used for compress an image. In this original data is not identical to compressed data that means there is some loss e.g. Block Truncation Coding,Transform Coding, etc... Lossless Compression used for compress any textual data.

III. CRYPTOGRAPHY METHOD

Computers are used all over the world by people for many purposes such as banking purpose, shopping, in the military, maintaining student records, etc. Privacy is a critical issue in many of these applications, how are we need to make sure that an unauthorized party cannot read or modify messages. Cryptography is a method for the transformation of the readable and understandable data into a form which cannot be understood by the third party in order to secure the data. The information that we need to hide, is called plaintext, It's the original text, It could be in a form of characters, numerical data, executable programs, pictures, or any other information, The plaintext, for example, sending a message through our mobile can be encrypted before sending to the destination and the data can be decrypted automatically with the use of key which can be known as source code, at the time of receiving the message at the receiver side. The data which will be transmitted after encryption is called cipher text, it's refers to the string of "meaningless" data, or unclear text that nobody must understand, except the recipients. It is the data that will be transmitted exactly through the network, many algorithms are used to transform plaintext into ciphertext. Cipher is one of the method, which can be used to transform plaintext to ciphertext, this method is called encryption, and in other words, mechanism of converting readable and understandable data into "meaning less" data called encryption. Computer security is a generic term defined as a collection of tools designed to protect any data from hackers, theft, corruption, or natural disaster while allowing these data to be available to the users at the same time. The example of these tools is the antivirus program. Network security refers to any activity can be designed to protect the usability, integrity, reliability, and safety of the data during their transmission over a network, Network security deals with the hardware and software. The activity can be done by one of the following methods as anti-virus and anti-spyware, firewall, Intrusion prevention systems, and Virtual

Private Networks. Internet Security is a protocol used to protect the data during their transmission over a collection of interconnected networks, while information security is about how to prevent attacks, and to detect attacks on information-based systems. Cryptography Goals Confidentiality, Authentication, Data Integrity, Non-Repudiation, Access Control. There are two types of cryptography methods as symmetric cryptosystem and asymmetric cryptosystem which can be figured below:

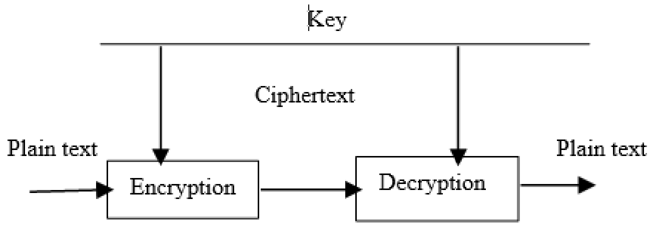


Figure 2. Symmetric Cryptosystem

In symmetric key cryptography is also known as private-key cryptography, a secret key may be held by one person or exchanged between the sender and the receiver of a message. If private key cryptography is used to send secret messages between two parties, both the sender and receiver must have a copy of the secret key.

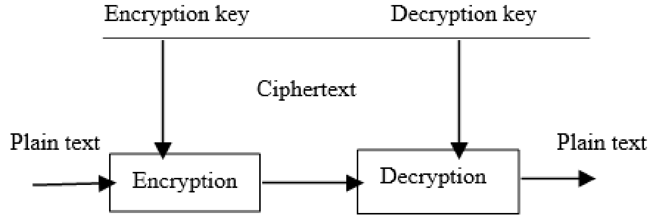


Figure 3. Asymmetric Cryptosystem

In the two-key system is also known as the public key system, one key encrypts the information and another, mathematically related key decrypts it. The computer sending an encrypted message uses a chosen private key that is never shared and so is known only to the sender. If a sending computer first encrypts the message with the intended receiver's public key and again with the sender's secret, private key, then the receiving computer may decrypt the message, first using its secret key and then the sender's public key. Using this public-key cryptographic method, the sender and receiver are able to authenticate one another as well as protect the secrecy of the message.

IV. DATA ENCRYPTION/ DECRYPTION

A. Encryption

A data encryption is a random string of bits created explicitly for scrambling and unscrambling data. Data encryption is designed with algorithms intended to ensure that every key is unpredictable and unique. Cryptography uses two types of keys: symmetric and asymmetric. Symmetric keys have the longest key, they can utilize a single key for both the encryption and decryption method of the cipher text. This type of key is called a secret key. Secret-key ciphers generally fall into one of two categories:stream ciphers or block ciphers. Data is encrypted with algorithm called encryption algorithm and encryption key. The process of encryption result in cypertext and it can be decrypted by the correct key at the receiver side. Most cryptographic processes use symmetric encryption to encrypt data transmissions but use asymmetric encryption to encrypt and exchange the secret key. Symmetric encryption is also known as the private key encryption, which can be used the same private key for both encryption and decryption methods. The risk in this system is that if either party loses the key or the key is intercepted, the system is broken and messages cannot be exchanged securely.

B. Decryption

One of the main reasons for implementing an encryption-decryption system is for data privacy. As information travels over the World Wide Web, it becomes subject to access from unauthorized individuals or organizations. Decryption is the process of taking encoded or encrypted text or other data and converting it back into text that you or the computer can read and understand. The term could be used to describe the method of un-encrypting the data manually or un-encrypting the data using the proper codes or keys. Encryption is the process of translating plain text data (plaintext) into something that appears to be random and meaningless (ciphertext). Decryption is the process of converting cipher text back to plaintext with the presence of proper key at the receiver end.

V. CONCLUSION

Cryptography is used to ensure that the contents of a message are confidentiality transmitted and would not be altered. Confidentiality means nobody can understand the received message except the one that has the decipher key, and "data cannot be changed" means the original information would not be changed or modified.

VI. REFERENCES

[1] Swarnalata, Bollavarapu, Ruchita Sharma "Data Security using Compression and Cryptography Techniques"

[2] Manoj Patil, Prof. VinaySahu "A Survey of Compression and Encryption Techniques for SMS"

[3] Bobby Jasuja, Abhishek Pandya "Crypto-Compression System: An Integrated Approach using Stream Cipher Cryptography and Entropy Encoding"

[4] M. Burrows and D. J. Wheeler. "A Block-sorting Lossless Data Compression Algorithm", SRC Research Report 124, Digital Systems Research Cente

[5] H. Kruse and A. Mukherjee. "Data Compression

[6] Using Text Encryption", Proc. Data CompressionConference, 1997, IEEE Computer Society Press,1997, p. 447.

[7] F. Awan, Nan Zhang N. Motegi, R.Iqbal, A. Mukherjee, LIPT: A reversible Lossless Text Transformation to Improve Compression Performance., Proceedings of Data Compression

[8] The conference, Snowbird, Utah, March 2001.

[9] Dr. V.K. Govindan, B.S. Shajee Mohan "An Intelligent Text Data Encryption and Compression for High Speed and Secure Data Transmission over the Internet" CSED, L.B.S.C.E., Kasaragod, Kerala.

[10] Robert Franceschini, Amar Mukherjee" Data Compression Using Encrypted Text" o-8186-7402-4196 $5.00 0 1996 IEEE Proceedings of ADL '96

[11] Amandeep Singh Sidhu, Er. MeenakshiGarg " An Advanced Text Encryption & Compression System Based on ASCII Values & Arithmetic Encoding to Improve Data Security" International Journal of Computer Science and Mobile Computing, IJCSMC, Vol. 3, Issue. 10, October 2014, pg.45 – 51

[12] V.K. Govindan and B.S. ShajeeMohan,"IDBE - An Intelligent Dictionary Based Encoding Algorithm for Text Data Compression for High-Speed Data Transmission over Internet

[13] P.G.Howard and J.C.Vitter, Fellow IEEE" Arithmetic Coding For Data Compression".

[14] S. Kaur and V.S.Verma," Design and Implementation of LZW Data Compression Algorithm", International Journal of Information Sciences and Techniques (IJIST) Vol.2, No.4, July 2012

[15] A.Mukherjee, R.Franceschini, "Data compression Using Encrypted Text"

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3| ISSN : 2456-3307

Cryptography Techniques

R.Sakthi Uma[1], Prof. R. Angelin Preethi[2]

[12]Department of Computer Science, Kamban College Of Arts and Science for Women, Tiruvannamalai, Tamil Nadu, India

ABSTRACT

Security is main issue of this generation of computing because many types of attacks are increasing day by day. Establishing a network is not a big issue for network administrators but protecting the entire network is a big issue. There are various methods and tools are available today for destroying the existing network. In this paper we mainly emphasize on the network security also we present some major issues that can affect our network.

Keywords: Security, Cryptography,

I. INTRODUCTION

For the first few decades of their existence computer networks were primarily used by university researchers for sending email, and by corporate employees for sharing printers. Under these conditions, security did not get a lot of attention. But now, as millions of ordinary citizens are using networks for banking, shopping, and filling their tax returns network security is looming on the horizon as a potentially massive problem.

Information systems have evolved in the last few decades from centralized and highly secure host-based systems to be decentralized. It is often said that in the enterprise model, "the network is the computer".Ultimately, the systems so large that they were hard to manage effectively.

To make things still worse, users of laptop and remote systems demanded connection into corporate offices from their homes, from hotel rooms, and from customer sites. Then the Internet becomes popular, and people inside the company wanted to connect out to it. To most administrators, the Internet is a nightmare that can potentially open the company's entire internal network to outsiders.

II. WHAT ARE THE THREATS?

Threats by floods and fires are easy to understand: the techniques for protecting against them area well known. But threats perpetrated by malicious users, disgruntled employees, and unknown hackers are a true nightmare. Every day some new technique for attacking systems is developed.

You may not know you are being attacked or have been attacked. No site is an exception. They break in using their computer and modem just for the fun or challenge. Professional hackers are quite busy as well.

A. Areas Of Security Weakness

The following list describes some of the weakest areas on company-wide networks:

- Well-known (and easily guessed) passwords, or leaked passwords, that compromise user logon and authentication

- Poorly implemented logon settings, user account rights, and file access permissions
- Disks and electronic mail that carry viruses
- Open doors into internal networks, created by users that access the Internet or by poorly implemented Internet firewalls
- Dial-up mobile and remote computers that have been stolen along with logon information

B. WHO ARE THE HACKERS?

You may not know any hackers personally; On the other hand, a hacker might be your next-door neighbor's son--someone with a computer and modem who is familiar with what you do, and who might guess your logon password because you use some derivative of your kids' names. Dangerous hackers are very knowledgeable about computers and security techniques, and they use sophisticated techniques to break into computer systems. Your competitor may hire such a hacker. If hackers cover their tracks, you might never know that they have stolen your customer mailing list or trade secrets.

Hackers often intend to make a profit or want to obtain free services. A phone hacker (or preacher) is intent on obtaining logon information to online services or on making long-distance phone calls through your phone system so that you pick up the charges. A hacker often uses information obtained during one break-in to access and break into another computer system.

C. The Internal Threat

A recent online survey by Network World magazine revealed that most security experts and readers felt that internal employees were the biggest threat to their information systems. Employees are familiar with the network, know which systems hold valuable information, and may have easy access to those systems through their own account or the account of another use. The American Society for Industrial Security estimates that 77 percent of information theft is perpetrated by insiders.

III. METHODS OF ATTACK

A. Phone Attacks

A preacher is a person who takes advantage of the telecommunications system to make free lone-distance telephone calls, listen to private conversations, access internal systems, or hack into other systems via the system broken into. Preachers are familiar with telephone switches, networks, and other equipment, and often have manuals from the manufacturers of telecom equipment that describe exactly how to operate and repair that equipment. Experienced preachers can manipulate telephone billing, access codes, and call routing.
Preachers can make free long-distance phone calls by gaining "dial-in / dial out" capabilities.

B. Hackers User Accounts and Passwords Attack

An attacker's first priority is to obtain user account names and passwords since this provides easy access to a system. Once inside, the hacker will find away to elevate his privileges. The attacker can obtain a list of user account names from a number of likely sources. Once a user account list is obtained, the hacker will try to determine which account will give the most access if broken into the pc support staff may inadvertently provide this information in the form of list of uses to contact in case of problems. Once a hacker obtains alginate user account name, cracking the password is the next step. Hackers take advantage of common passwords: if they know the user of an account, they may try various combinations of the user's kids and pets' names. Many people use the same password to log on to other systems, such as ATM machines. If a hacker obtains a user account name, but not a password, he can try brute force methods of breaking into the account. A program is set up to try thousands or millions of different passwords until the account opens. This method is ineffectively if logon restrictions that limit the number of attempted.

C. Electronic Eavesdropping and Cable Sniffing

A packet snifter is a device or software that can read transmitted packets. Packet sniffing is a passive eves dropping technique that is hard to detect. The packet-sniffing devices may be installed on internal or external networks. Although packet sniffing an internet transition line is not necessarily informative, sniffing a cable that runs into your facilities who are armed with packet snifters, or from hackers who have penetrated your building and planted listening devices.

D.Viruses and Trojan Horeses

Viruses are small programs that mimic the activities of real-life viruses. They get into computer systems by being copied from contained disks or downloaded from online services by unsuspecting users. Once a system is contaminated, the virus executes some immediate action, or waits until a specified time or for a specific command executed by the user. Viruses may display harmless messages or destroy the information stored on entire hard disks. A Trojan horse is similar to a virus, but contaminates a system by posing as some other type of program.

Viruses are created by authors who are fascinated by how quickly their virus may spread through computer systems. Terrorists and industrial spies create viruses that cause damage in order to seek revenge on an opponent or to viruses that cause damage in order to seek revenge on opponent or tom damage the operations of a competitor. Some viruses are intended targets.

E.Natural Threats

Obviously, not all threats to the integrity of your network come from people. Power surges, failing components, and other problems may bring down systems and cost your organization thousands or millions of dollars in down time. In some cases, continuous access to information is critical to the operation of the entire business.

IV. COUNTER MEASURES

A. Defining Security

Information security is the practice of protecting resources and data on computer systems and networks, including information on storage devices and in transmission. Make it your business to control and monitor the security of your systems and to implement security policies and procedures that people can follow.

- Identification and authentication
- Access control
- Accountability and auditing
- Accuracy
- Reliability
- Data exchange

B.Security Costs

Consider how much your organization can afford to spend on security. At the physical level, power surges, failing components and other problems may bring down systems and cost your organization thousands or millions of dollars in downtime. In some cases, continuous access to information is critical to the operation of the entire business.There are also direct costs, such as equipment costs, as well as administrative expenses. Beyond the dollar costs, there are expenses related to the inconvenience of the security system. It may simply take more time to get things done when complex procedures are in place to provide security

C. Protective Measures

There are a number of protective measures that help you "harden" your defenses. A few obvious steps are:

- Create security policies, plans, and job positions as appropriate.
- Set up a security-response team, experts who handle security problems.
- Perform background checks on personnel and keep tabs on employees who are disgruntled, who are working closely with other companies, and who are in the process of leaving the company.

- Classify your employees much the way the military classifies its personnel, giving some people higher clearance for access to sensitive information than others.

D. Backups

Backups are essential. If your systems are stolen, destroyed by fire, or corrupted by hackers, you'll need to go back to the last uncorrupted backup. The National Computer Security Association provides some interesting figures.The procedures you use to restore backups are critical in the case of virus attacks. Your backups may be corrupted, in which case you'll need to go back in the archive until you find a non-corrupted backup set. Back up as frequently as possible and place back up media into permanent archives as often as possible.

E.Encryption

You can use cryptographic techniques to protect files stored on disks and backups from prying eyes, or to conceal data transmissions and electronic mail. Encryption utilities scramble files and lock them with a password key. Using encryption may cause a drop in performance.

Encryption may give you the feeling that your files are private, when in fact someone might have cracked your encryption key and begun reading all your files. The stronger the encryption system, the better, but sure to implement additional security measures as appropriate. Also be aware that someone who gains access to your system might replace your encryption program with a Trojan horse version of the program that steals your password. Make sure the encryption software is protected and secure. Then take actions to monitor for possible virus infections.

F. Virus Protection

Viruses are a real threat to your network. They are easily contracted form unknown disks or by downloading files form online services, bulletin boards, and the Internet. Any of your network users can contract a virus at any time and spread it to the network. A virus is often hard to detect. It may wait on your system before it executes. Vigilant users or network administrators may detect unusual activity or notice an increase in the size of files (indicating potential infection).

Even after detecting and cleaning up a virus infection, there is still a good chance that the virus is lurking somewhere in your organization, ready to re-infect systems. It may even have infected the backup sets. You may need to implement a plan to detect and remove the virus throughout your organization. Check all workstations, disks, and other data sources for infections.

G. Advantages

These advantages can be lined up simply as

1. Protects personal data of clients on the network.
2. Protects information been shared between computers on the network.
3. Protects the physical computers from harm based from possible attacks on the network from the outside
4. Provides levels of access if the network has many computers attached so some computers may have more access to information than others. (Account system)
5. Private networks can be closed off from the internet making them protected from most outside attacks. Which makes them secure from Virus attacks.

V. CONCLUSION

As internet has become a huge part of our daily life, the need of network security has also increased exponentially from the last decade. As more and more users connect to the internet it attracts a lot of criminals. Today, everything is connected to internet from simple shopping to defense secrets as a result there is huge need of network security. Billions of dollars of transactions happens every hour over the internet, this need to be protected at all costs. Most of the attacks can be easily prevented, by following many simply methods as outlined in this paper. As new and more sophisticated attacks occur, researchers across the world find new methods to prevent them.

VI. REFERENCES

[1] B. Daya ,“Network Security: History, Importance, and Future ,”University of Florida Department of Electrical and Computer Engineering , 2013. http://web.mit.edu/~bdaya/www/Network%20Security.pdf

[2] Li CHEN,Web Security : Theory And Applications,School of Software,SunYat-sen University, China.

[3] J. E. Canavan, Fundamentals of Network Security, Artech House Telecommunications Library, 2000.

[4] A. R. F. Hamedani, “Network Security Issues, Tools for Testing,” School of Information Science, Halmstad University, 2010.

[5] S. A. Khayam, Recent Advances in Intrusion Detection, Proceedings of the 26th Annual Computer Security Applications Conference, Saint-Malo, France, pp. 224-243, 42, 2009

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018IJSRCSEIT| Volume 4 | Issue 3 | ISSN: 2456-3307

Big Data Analytics in Healthcare

Saranya K[1], UdhayaKumar U[2]
[1]Department of Computer Science, Shanmuga Industries Arts and Science College, Tiruvannamalai, Tamilnadu, India
[2]Assistant Professor, Department of Computer Science, Shanmuga Industries Arts and Science College, Tiruvannamalai, Tamilnadu, India

ABSTRACT

The rapidly expanding field of big data analytics has started to play a pivotal role in the evolution of healthcare practices and research. It has provided tools to accumulate, manage, analyse, and assimilate large volumes of disparate, structured, and unstructured data produced by current healthcare systems. Big data analytics recently applied towards aiding the process of care delivery and disease exploration. In this paper, we discuss some of these significant challenges with a focus on three upcoming and promising areas of medical research: image, signal, and genomics-based analytics.

Keywords: Big Data, Big Data Analytics, 3V's, Image Processing, Signal processing, Genomics.

I. INTRODUCTION

The concept of "big data" is not new; however, the way it is defined is continually changing. Various attempts at defining big data essentially characterize it as a collection of data elements whose size, speed, type, and/or complexity require one to seek, significant, and invent new hardware and software mechanisms in order to successfully store, analyse, and visualize the data. Healthcare is a prime example of how the three V's of data, velocity (speed of generation of data), variety, and volume is an innate aspect of the data it produces.

Historical approaches to medical research have focused on the investigation of disease states based on the changes in physiology in the form of a confined view of the specific singular modality of data Important physiological and pathophysiological phenomena are concurrently manifest as changes across multiple clinical streams. This results from strong coupling among different systems within the body (e.g., interactions between heart rate, respiration, and blood pressure) thereby producing potential markers for clinical assessment.

In this paper, three areas of big data analytics in medicine discussed. These three areas do not comprehensively reflect the application of big data analytics in medicine.

Image Processing: Medical images are an essential source of data frequently used for diagnosis, therapy assessment, and planning. Computed tomography (CT), magnetic resonance imaging (MRI), X-ray, molecular imaging, ultrasound, are some of the examples of imaging techniques.

Signal Processing: Similar to medical images, medical signals also pose volume and velocity obstacles especially during continuous, high-resolution acquisition and storage from a multitude of monitors connected to each patient. Currently, healthcare systems use numerous different and continuous monitoring devices that singular physiological waveform data or discretised vital information to provide alert mechanisms in case of open events.

Genomics: The cost to sequence the human genome (encompassing 30,000 to 35,000 genes) is rapidly decreasing with the development of high-throughput sequencing technologythe predictive, preventive, participatory, and personalized health, referred to as P4, medicine paradigm as well as an integrative personal omics profile. The P4 initiative is using a system approach for (i) analysing genome-scale datasets to determine disease states, (ii) moving towards blood-based diagnostic tools for continuous monitoring of a subject, (iii) exploring new approaches to drug targetdiscovery, developing tools to deal with big data challenges of capturing, validating, storing, mining, integrating, and finally (iv) modelling data for each.

II. MEDICAL IMAGE PROCESSING FROM BIG DATA POINT OF VIEW

Medical imaging provides essential information on anatomy and organ function in addition to detecting diseases states. Moreover, it is utilized for organ delineation, identifying

Tumors in lungs, spinal deformity diagnosis, artery stenosis detection, aneurysm detection, and so forth. In these applications, image processing techniques such as enhancement, segmentation, and denoising in addition to machine learning methods employed. As the size and dimensionality of data increase, understanding the dependencies among the data and designing efficient, accurate, and computationally efficient methods demand new computer-aided techniques and platforms. The rapid growth in the number of healthcare organizations as well as the number of patients has resulted in the excellent use of computer-aided medical diagnostics and decision support systems in clinical settings.

Many areas of health care such as diagnosis, prognosis, and screening can improve by utilizing computational intelligence. The integration of computer analysis with appropriate care has potential to help clinicians improve diagnostic accuracy. The integration of medical images with other typesofelectronic health record (EHR) data and genomic data can also improve the accuracy and reduce the time taken for diagnosis. In the following, data produced by imaging techniques reviewed and applications of medical imaging from a big data point of view are discussed.

A. Data Produced by Imaging Techniques

Medical imaging encompasses a broad spectrum of different image acquisition methodologies typically utilized for a variety of clinical applications. For example, visualizing blood vessel structure can be performed using magnetic resonance imaging (MRI), computed tomography (CT), ultrasound, and photoacoustic imaging [30]. From a data dimension point of view, medical images might have 2, 3, and four dimensions. Positron emission tomography (PET), CT, 3D ultrasound, and functional MRI (fMRI) considered as multidimensional medical data. Modern medical image technologies can produce higher solution images such as respiration-correlated or “four-dimensional” computed tomography (4D CT) Molecular imaging is a noninvasive technique of cellular and subcellular events which has the potential for clinical diagnosis of disease states such as cancer. However, to make it clinically applicable for patients, the interaction of radiology, nuclear medicine, and biology is crucial that could complicate its automated analysis. Microwave imaging is an emerging methodology that could create a map of electromagnetic wave scattering arising from the contrast in the dielectric properties of different tissues. It has both functional and physiological information encoded in the dielectric properties which can help differentiate and characterize different tissues and pathologies. However, microwaves have scattering behavior that makes retrieval of data a challenging task. Advanced

Multimodal Image-Guided Operating (AMIGO) suite designed which has angiographic X-ray system, MRI, 3D ultrasound, and PET/CT imaging in the operating room (OR). This system has been used for cancer therapy and showed the improvement in localization and targeting an individual's diseased tissue [1, 2].

B. Methods

The volume of medical images is growing exponentially. For instance, Image CLEF medical image dataset contained around 66,000 images between 2005 and 2007.

While just in the year of 2013 around 300,000 images were stored every day [3]. In addition to the growing volume of images, they differ in modality, resolution, dimension, and quality, which introduce new challenges such as data integration and mining especially if multiple datasets are involved. Compared to the volume of research that exists on single modal medical image analysis, there is the considerably lesser number of research initiatives on multimodal image analysis.

C. Analytical Methods

The goal of medical image analytics is to improve the interpretability of depicted contents [4].Many methods and frameworks developed for medical image processing. However, these methods are not necessarily applicable for big data applications. One of the frameworks developed for analyzing and transformation of huge datasets is Hadoop that employs Map Reduce Map Reduce is a programming paradigm that provides scalability across many servers in a Hadoop cluster with a broad variety of real-world applications. However, it does not perform well with input output intensive tasks Map Reduce framework.It has been used in [5] to increase the speed of three large-scale medical image processing use-cases, (i) finding optimal parameters for lung texture classification by employing a well-known machine learning method, support vector machines (SVM), (ii) content-based medical image indexing, and (iii) wavelet analysis for robust texture classification.

D. Collecting, Sharing, and Compressing Methods

In addition to developing analytical methods, efforts have made for collecting, compressing, sharing, and anonymizing medical data. One example is IDASH (integrating data for analysis, anonymization, and sharing) which is a center for biomedical computing [6]. It focuses on algorithms and tools for sharing data in a privacy-preserving manner. The goal of iDASH is to bring together a multi-institutional team of quantitative scientists to develop algorithms and tools, services, and a biomedical cyberinfrastructure to be used by biomedical and behavioral researchers [7]. Another example of a similar approach is Health-e-Child consortium of 14 academic, industry, and clinical partners with the aim of developing an integrated healthcare platform for European pediatrics [8].

There are some limitations in implementing the application-specific compression methods on both general purpose processors and parallel processors such as graphics processing units (GPUs) as these algorithms need highly variable control and the complex bit well suited to GPUs and pipeline architectures. To overcome this limitation, an FPGA implementation proposed for LZ-factorization which decreases the computational burden of the compression algorithm [9]. Lossy image compression has been introduced in [10] that reshapes the image in such a way that if the image is uniformly sampled, sharp features have a higher sampling density than the rough ones. This method is claimed to be applicable for big data compression. However, for medical applications lossy methods are not applicable in most cases as fidelity is essential and information must preserve. These techniques are among a few techniques that have been either designed as prototypes or developed with limited applications. Developing methods for processing/analysing a broad range and large volume of data with acceptable accuracy and speed is still critical. In Table 1, we summarize the challenges facing medical image processing. When dealing with big data, these challenges seemed to be more severe and on the other hand analytical methods could benefit the big data to handle them.

III. MEDICAL SIGNAL ANALYTICS

Streaming data analytics in healthcare can be defined as a systematic use of continuous waveform (the signal varying against time) and related medical record information developedthrough applied analytical disciplines (e.g., statistical, quantitative, contextual, cognitive, and predictive) to drive decision making for patient care. The analytics workflow of real-time streaming waveforms in clinical settings can be broadly described.

Firstly, a platform for streaming data acquisition and ingestion is required, which has the bandwidth to handle multiple waveforms at different fidelities. Integrating these dynamic waveform data with static data from the EHR is a crucial component to provide situational and contextual awareness for the analytics engine. Enriching the data consumed by analytics not only makesthe system more robust but also helps balance the sensitivity and specificity of the predictive analytics. The specifics of the signal processing will largely depend on the type of disease cohort under investigation. A variety of signal processing mechanisms can be utilized to extract a multitude of target features which are then consumed by a trained machine learning model to produce actionable insight.

These actionable insights could either be diagnostic, predictive, or prescriptive. These insights could further be designed to trigger other mechanisms such as alarms and notification to physicians.

A. Data Acquisition

Historically streaming data from connected physiological signal acquisition devices rarely stored. Even if the option to store this data were available, thelength of these data captures was typically short and downloaded only using proprietary software and data formats provided by the device manufacturers. Although most major medical device manufacturers are now taking steps to provide interfaces to access live streaming data from their devices, such data in motion very quickly poses archetypal big datachallenges. The fact that there are also governance challenges such as lack of data protocols, lack of data standards, and data privacy issues are adding to this.

B. Data Storage and Retrieval

With massive volumes of streaming data and other patient information that can gather from clinical settings, sophisticated storage mechanisms of such data are imperative. Since storing and retrieving can be computational and time expensive, it is critical to have a storage infrastructure that facilitates rapid data pull and commits based on analytic demands. With its capability to store and compute large volumes of data, usage of systems such as Hadoop, Map Reduce, and Mongo DB [11, 12] is becoming much more familiar with the healthcare research communities. Mongo DB is a free cross-platform document-oriented database

C. Data Aggregation

Integration of disparate sources of data, developing consistency within the data, standardization of data from similar sources, and improving the confidencein the data especially towards utilizing automated analytics are among challenges facing data aggregation in healthcare systems. Analysis of continuous data heavily utilizes the information in time domain. However, static data does not always provide accurate time context and, hence, when combining the waveform data with static electronic health record data, thetemporal nature of the time context during integration can also add significantly to the challenges. There are considerable efforts in compiling waveforms and other associated electronic medical information into one cohesive database that are made publicly available to researchers worldwide [13]. For example, MIMIC II [14] and some other datasets included in Physionet [15] provide waveforms and other clinical data from a wide variety of actual patient cohorts.

D. Signal Analytics Using Big Data

Research in signal processing for developing big data-based clinical decision support systems (CDSSs) is getting more prevalent. In fact, organizations such as the Institution of Medicine have long advocated use of health information technology including CDSS to improve care quality CDSSs provide medical practitioners with knowledge and patient-specific information, intelligently filtered and presented at appropriate times, to improve the delivery of care. A study presented by Lee and Mark uses the MIMIC II database to prompt therapeutic intervention to hypotensive episodes using cardiac and blood pressure time series data another study shows the use of physiological waveform data along with clinical data from the MIMIC II database for finding similarities among patients within the selected cohorts.

As complex physiological monitoring devices are getting smaller, cheaper, and more portable, personal monitoring devices are being used outside of clinical environments by both patients and enthusiasts alike. However, similar to clinical applications, combining information simultaneously collected from multiple portable devices can become challenging.Pantelopoulos and Bourbakis discussed the research and development of wearable biosensor systems and identified the advantages and shortcomings in this area of study. Similarly, portable and connected electrocardiogram, blood pressure and body weight devices used to set up a network-based study of telemedicine. The variety of fixed as well as mobile sensors available for data mining in the healthcare sector and how such data can be leveraged for developing patient care technologies are surveyed.

IV. BIG DATA APPLICATIONS IN GENOMICS

The advent of high-throughput sequencing methods has enabled researchers to study genetic markers over a wide range of population improve efficiency by more than five orders of magnitude since sequencing of the human genome was completed and associate genetic causes of the phenotype in disease states. Genome-wide analysis utilizing microarrays has been successful in analyzing traits across a population and contributed successfully to treatments of complex diseases such as Crohn's disease and age-related muscular degeneration. Analytics of high-throughput sequencing techniques in genomics is an inherently big data problem as the human genome consists of 30,000 to 35,000 genes Initiatives are currently being pursued over the timescale of years to integrate clinical data from the genomic level to the physiological level of a human being. These initiatives will help in delivering personalized care to each patient.

A. Pathway Analysis

Resources for inferring functional effects for "-omics" big data are largely based on statistical associations between observed gene expression changes andpredicted functional effects. Experiment and analytical practices lead to error as well as batch effects Interpretation of functional effects has to incorporate continuous increases in available genomic data and corresponding annotation of genes.There are variety of tools, but no "gold standard" for functional pathway analysis of high-throughput genome-scale data. Three generations of methods used for pathway analysis are described as follows. The first generation encompasses overrepresentation analysis approaches that determine the fraction of genes in a particular pathway found among the genes, which are differentially expressed. Examples of the first generation tools are Onto-Express Go Miner and Clue Go. The second generation includes functional class scoring approaches, which incorporate expression levelchanges in individual genes as well as functionally similar genes.

B. Reconstruction of Regulatory Networks

Pathway analysis approaches do not attempt to make sense of high-throughput big data in biology as arising from the integrated operation of a dynamical system. There are multiple approaches to analyzing genome-scale data using a dynamical system framework [30, 31]. Due to the breadth of the field, in this section, we mainly focus on techniques to infer network models from big biological data. Reconstruction of metabolic networks has advanced in last two decades. One objective is to develop an understanding of organism-specific metabolism through reconstruction of metabolic networks by integrating genomics, transcriptomics, and proteomics high-throughput sequencing techniques [33]. Constraint-based methods are widely applied to probe the genotype-phenotype relationship and attempt to overcome the limited availability of kinetic constants Reconstruction of gene regulatory networks from gene expression data is another well-developed field.

Network inference methods can be split into five categories based on the underlying model in each case: regression, mutual information, correlation, Boolean regulatory networks, and other techniques [16, 35]. Over 30 inference techniques were assessed after DREAM5 challenge in 2010 [17, 34]. Boolean regulatory networks [18] are a particular case of discrete dynamical models where the state of a node or a set of nodes exists in a binary state. The actual state of each node or set of nodes is determined by using Boolean operations on the state of other nodes in the network [19]. Boolean networks are beneficial when an amount of quantitative data is small [10, 20] but yield the high number of false positives (i.e., when a given condition is satisfied while actually, that is not the case) that may be reduced by using prior knowledge [21, 22]. Another bottleneck is that Boolean networks are prohibitively expensive when the number of nodes in the network is large. This is due to the number of global states rising exponentially in the number of entities [23].A method to overcome this bottleneck is to use clustering to break down the problem size. For example, Martin et al. [24] broke down a 34,000-probe microarray gene expression dataset into 23 sets of metagenes using clustering techniques. This Boolean model successfully captured the network dynamics for two different immunology microarray datasets. The dynamics of gene regulatory network can be captured using ordinary differential equations (ODEs) [12-15]. This approach has been applied to determine regulatory network for yeast [25]. The study successfully captured the regulatory network which has been characterized using experiments by molecular biologists. Reconstruction of a gene regulatory network on a genome-scale system as a dynamical model is computationally intensive [26]. A parallelizable dynamical ODE model has been developed to address this bottleneck [27]. It reduces the computational time to O(N2) from the time taken in other approaches which are O(N3) or O(N2 log N) [28].Determining connections in the regulatory network for a problem of the size of the human genome, consisting of 30,000 to 35,000 genes [16, 17], will require exploring close to a billion possible connections. The dynamical ODE model has been applied to reconstruct the cardiogenic gene regulatory network of the mammalian heart [29].

V. CONCLUSION

Big data analytics which leverages legions of disparate, structured and unstructured data sources is going to play a vital role in how healthcare practiced in the future. One can already see a spectrum of analytics being utilized, aiding in the decision-making and performance of healthcare personnel and patients. Here we focused on three areas of interest: medical image analysis, physiological signal processing, and genomic data processing.

The acquisition, formation/reconstruction, enhancement, transmission, and compression. New technological advances have although there are some genuine challenges for signal processing of physiological data to deal with, given the current state of data competency and no standardized structure, there are opportunities in each step of the process towards providing systemic improvements within the healthcare research and practice communities. Apart from the apparent need for further research in the area of data wrangling, aggregating, and harmonizing continuous and discrete medical data formats, there is also an equal need for developing novel signal processing techniques specialized towards physiological signals. Research about mining for biomarkers and hidden patterns within bio signals tounderstand and predict disease cases has shown potential in providing actionable information. However, there are opportunities for developing algorithms to address data filtering, interpolation, transformation, feature extraction, feature selection, and so forth. Furthermore, with the notoriety and improvement of machine learning algorithms, there are opportunities in improving and developing robust CDSS for clinical prediction, prescription, and diagnostics.

Integration of physiological data and high-throughput "- omics" techniques to deliver clinical recommendations is the grand challenge for systems biologists. Although associatingFunctional effects with changes in gene expression have progressed, the continuous increase in available genomic data and its corresponding impact of annotation of genes and errors from experiment and analytical practices make analyzing functional effect from high-throughput sequencing techniques a challenging task. Reconstruction of networks on the genome-scale is an ill-posed problem. Robust applications developed for reconstruction of metabolic networks and gene regulatory networks. Limited availability of kinetic constants is a bottleneck, and hence various models attempt to overcome this limitation. There is an incomplete understanding of this large-scale problem as gene regulation, an effect of different network architectures, and evolutionary effects on these networks are still being analysed [135]. To address these concerns, the combination of the careful design of experiments and model development for reconstruction of networks will help in saving time and resources spent in building the understanding of regulation in genome-scale networks. The opportunity of addressing the grand challenge requires close cooperation among experimentalists, computational scientists, and clinicians.

VI.REFERENCES

[1] A. McAfee, E. Brynjolfsson, T. H. Davenport, D. J. Patil, and D. Barton, "Big data: the management revolution," *Harvard Business Review*, vol. 90, no. 10, pp. 60–68, 2012.

[2] C. Lynch, "Big data: how do your data grow?" *Nature*, vol. 455,no. 7209, pp. 28–29, 2008.

[3] A. Jacobs, "The pathologies of big data," *Communications of the ACM*, vol. 52, no. 8, pp.

[4] P. Zikopoulos, C. Eaton, D. deRoos, T. Deutsch, and G. Lapis, *Understanding Big Data: Analytics for Enterprise Class Hadoopand Streaming Data*, McGraw-Hill Osborne Media.

[5] J.Manyika,M. Chui, B. Brown et al., Big Data:The Next Frontierfor Innovation, Competition, and Productivity,McKinsey Global Institute, 2011.

[6] J. J. Borckardt, M. R. Nash, M. D. Murphy, M. Moore, D. Shaw, and P. O'Neil, "Clinical practice as natural laboratory for psychotherapy research: a guide to case-based time-series analysis," *The American Psychologist*, vol. 63, no. 2, pp. 77–95, 2008.

[7] L. A. Celi, R.G.Mark, D. J. Stone, and R.A.Montgomery, "'Big data' in the intensive care unit: closing the data loop,"

AmericanJournal of Respiratory andCriticalCareMedicine, vol. 187, no. 11,pp. 1157–1160, 2013.

[8] F. Ritter, T. Boskamp, A. Homeyer et al., "Medical image analysis," *IEEE Pulse*, vol. 2, no. 6, pp. 60–70, 2011.

[9] J. A. Seibert, "Modalities and data acquisition," in *PracticalImaging Informatics*, pp. 49–66, Springer, New York, NY, USA, 2010.

[10] B. J. Drew, P. Harris, J. K. Z`egre-Hemsey et al., "Insights into the problem of alarm fatigue with physiologic monitor devices: a comprehensive observational study of consecutive intensive care unit patients," *PLoSONE*, vol. 9, no. 10, Article IDe110274, 2014. BioMed Research International

[11] K. C. GrahamandM. Cvach, "Monitor alarmfatigue: standardizing use of physiological monitors and decreasing nuisance alarms,"*The American Journal of Critical Care*, vol. 19, no. 1, pp. 28–34, 2010.

[12] M. Cvach, "Monitor alarm fatigue: an integrative review," *Biomedical Instrumentation &Technology*, vol. 46, no. 4, pp. 268–277, 2012.

[13] J. M. Rothschild, C. P. Landrigan, J. W. Cronin et al., "The Critical Care Safety Study: the incidence and nature of adverse events and serious medical errors in intensive care," *CriticalCare Medicine*, vol. 33, no. 8, pp. 1694–1700, 2005.

[14] P. Carayon and A. P. G¨urses, "A human factors engineering conceptual framework of nursing workload and patient safety in intensive care units," *Intensive and Critical Care Nursing*, vol.21, no. 5, pp. 284–301, 2005.

[15] P.Carayon, "Human factors of complex sociotechnical systems," *Applied Ergonomics*, vol. 37, no. 4, pp. 525–535, 2006.

[16] E. S. Lander, L. M. Linton, B. Birren et al., "Initial sequencing and analysis of the human genome," *Nature*, vol. 409, no. 6822, pp. 860–921, 2001.

[17] R. Drmanac, A. B. Sparks, M. J. Callow et al., "Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays," *Science*, vol. 327, no. 5961, pp. 78–81, 2010.

[18] T. Caulfield, J. Evans, A.McGuire et al., "Reflections on the cost of 'Low-Cost' whole genome sequencing: framing the health policy debate," *PLoS Biology*, vol. 11, no. 11, Article ID e1001699,2013.

[19] F. E. Dewey, M. E. Grove, C. Pan et al., "Clinical interpretation and implications of whole-genome sequencing," *JAMA*, vol. 311, no. 10, pp. 1035–1045, 2014.

[20] L. Hood and S. H. Friend, "Predictive, personalized, preventive, participatory (P4) cancer medicine," *Nature Reviews ClinicalOncology*, vol. 8, no. 3, pp. 184–187, 2011.

[21] L. Hood and M. Flores, "A personal view on systems medicine and the emergence of proactive P4 medicine: predictive, preventive, personalized and participatory," *New Biotechnology*, vol. 29, no. 6, pp. 613–624, 2012.

[22] L. Hood and N. D. Price, "Demystifying disease, democratizing health care," *Science Translational Medicine*, vol. 6, no. 225, Article ID 225ed5, 2014.

[23] R.Chen, G. I. Mias, J. Li-Pook-Than et al., "Personal omics profiling reveals dynamic molecular and medical phenotypes," *Cell*, vol. 148, no. 6, pp. 1293–1307, 2012.

[24] G. H. Fernald, E. Capriotti, R. Daneshjou, K. J. Karczewski, and R. B. Altman, "Bioinformatics challenges for personalized medicine," *Bioinformatics*, vol. 27, no. 13, Article ID btr295, pp.1741–1748, 2011.

[25] P. Khatri, M. Sirota, and A. J. Butte, "Ten years of pathway analysis: current approaches and outstanding challenges," *PLoSComputational Biology*, vol. 8, no. 2, Article ID e1002375, 2012.

[26] J. Oyelade, J. Soyemi, I. Isewon, and O. Obembe, "Bioinformatics, healthcare informatics and analytics: an imperative for improved healthcare system," *International Journal of*

AppliedInformation Systems, vol. 8, no. 5, pp. 1–6, 2015.

[27] T. G. Kannampallil, A. Franklin, T. Cohen, and T. G. Buchman, “Sub-optimal patterns of information use: a rational analysis of information seeking behavior in critical care,” in *CognitiveInformatics in Health and Biomedicine*, pp. 389–408, Springer, London, UK, 2014.

[28] H. Elshazly, A. T. Azar, A. El-korany, and A. E. Hassanien, “Hybrid system for lymphatic diseases diagnosis,” in *Proceedingsof the International Conference on Advances in Computing,Communications and Informatics (ICACCI ’13)*, pp. 343–347, IEEE, Mysore, India, August 2013.

[29] G.Dougherty, *Digital Image Processing forMedical Applications*, Cambridge University Press, 2009.

[30] R. C. Gessner, C. B. Frederick, F. S. Foster, and P. A. Dayton, “Acoustic angiography: a new imaging modality for assessing microvasculature architecture,” *International Journal of BiomedicalImaging*, vol. 2013,Article ID 936593, 9 pages, 2013.

[31] Janarthanan Y, Balajee J.M, and SrinivasaRaghavaS."Content based video retrieval and analysis usingimage processing: A review."International Journal ofPharmacy and Technology 8, no.4 (2016): 5042-5048.

[32] Jeyakumar, Balajee, MA SaleemDurai, and DaphneLopez. "Case Studies in Amalgamation of DeepLearning and Big Data." In HCI Challenges andPrivacy Preservation in Big Data Security, pp. 159-174. IGI Global, 2018.

[33] Kamalakannan, S. "G., Balajee, J., Srinivasa Raghavan., “Superior content-based video retrievalsystem according to query image”." InternationalJournal of Applied Engineering Research 10, no. 3(2015): 7951-7957.

[34] Sethumadahavi R Balajee J “Big Data Deep Learningin Healthcare for Electronic Health Records,”International Scientific Research OrganizationJournal, vol. 2, Issue 2, pp. 31–35, Jul. 2017.

[35] Ushapreethi P, BalajeeJeyakumar and BalaKrishnanP, Action Recongnition in Video SurvillanceUsingHipi and Map Reducing Model, International Journalof Mechanical Engineering and Technology 8(11),2017,pp. 368–375.

Spreading of Hearsay over the Networks

Ms. E. Rathidevi[1], Mrs. S. Shanthi[2], Ms. R. Nandhini[3]

[13]Research Scholar, Dept. of Computer Science and Applications, D.K.M. College for Women (Autonomous), Vellore, Tamilnadu, India

[2]Assistant Professor, Dept. of Computer Science and Applications, D.K.M. College for Women (Autonomous), Vellore, Tamilnadu, India

ABSTRACT

Hearsay spreading is one of the essential instruments for data scattering in networks. Randomized gossip spreading is a class of basic randomized conveyed calculations, all expanding on the worldview that hubs of a system contact irregular neighbors to trade information. Initially, a solitary hub is aware of gossip. In each prevailing round, each hub picks an arbitrary neighbor, and the two hubs share the gossip in the event that one of them is now mindful of it. These bits of gossip can harm organization, change voting conduct or spread buildup about an item before its launch. We demonstrate that, in these systems: (a) The standard PUSH– PULL procedure conveys the message to all hubs within O (log2 n) rounds with high likelihood; (b) without anyone else's input, PUSH and PULL require polynomial numerous rounds. We additionally research the non concurrent rendition of the push-pull convention, where the hubs don't work in rounds, however trade data as per a Poisson process with rate 1.

Keywords: Hearsay, Spreading, Push, Pull, Hub, Trade data, Gossip

I. INTRODUCTION

The previous decade has seen an emotional increment in the use of online social networking (OSN) and microblogging administrations like Facebook, Google+, Twitter, and so on. Aside from their value in helping people stay in contact, they are progressively being utilized for dispersing data about occasions happening continuously. Randomized gossip spreading is a class of basic randomized disseminated calculations, all expanding on the worldview that hubs of a system contact irregular neighbors to trade data.

Gossip spreading is one of the fundamental instruments for data scattering in systems. In this paper we investigate the execution of gossip spreading in the Preferential Attachment display [1]. We demonstrate that, while neither PUSH nor PULL independent from anyone else ensure quick data scattering, with PUSH- - PULL the data achieves all hubs in the system inside O(log2 n) rounds with high likelihood, n being the quantity of hubs in the system.

A standout amongst the most considered inquiries concerning gossip spreading is the accompanying: How many number of rounds will it take for to disperse the data to all hubs in the graph, expecting a most pessimistic scenario source?

We think about this inquiry for the particular connection model and demonstrate the accompanying:

– Paying little mind to the beginning hub, the PUSH system requires, with Ω (1) likelihood, polynomially numerous rounds;

– There are beginning hubs to such an extent that the PULL system requires, with Ω (1) likelihood, polynomially numerous rounds;

– Paying little mind to the beginning hub, the PUSH- - PULL system requires, with likelihood 1 o(1), O(log2 n) numerous rounds.

Regardless of being exceptionally basic conventions, they turned out to be extremely effective both in theoretical investigations and in practical applications .In the present work, we address the major inquiry concerning whether and how the structure of specific models for genuine systems impacts the spread of data. Assume that we are given a graph whose hubs speak to singular substances, and each edge remains for some sort of cooperation between them. At first, there is a solitary hub that is aware of gossip. The convention at that point continues in rounds. In each such round, each hub picks an irregular neighbor and the two hubs share the gossip, if no less than one of them knows about it.

As to execution of the push-pull convention on the special connection demonstrate, where β= 3, demonstrated that it spreads the data to all hubs of the traditional particular connection irregular graph in n(log n) rounds. The point of this work is to utilize an exploratory examination so as to (a) better understand the execution of randomized gossip spreading conventions on special connection systems; over the long haul, this may help in the outline of proficient correspondence systems; and (b) better understand the upside of furnishing hubs with a little measure of memory, which is utilized to abstain from reaching a consistent number of past contactees.

II. SPREADING OF HEARSAY

It is normal to ask whether a high edge extension or conductance1 infer that gossip spreading is quick. The graph in Fig. 1 has high edge development however gossip spreading takes straightly numerous rounds. The diagram comprises of √n many independent sets, every one of size √n. These autonomous sets are organized in a cycle. Two neighboring free sets shape a total bipartite diagram. The focal hub is associated with one vertex in every autonomous set. The graph likewise has a high edge development yet PUSH- - PULL requires polynomially numerous rounds regardless of the distance across being steady. Mihail et al. [2] ponder the edge extension and the conductance of diagrams that are fundamentally the same as PA graphs. We should allude to these as ''nearly'' PA diagrams. They demonstrate that the edge development and conductance are steady in these diagrams, when m ≥ 2.

Ensuing to our work in this paper it was appeared in [3,4] that if a diagram has high conductance at that point gossip spreading is quick. Specifically, if a graph has a conductance ϕ then PUSH- - PULL achieves each hub inside numerous rounds with high likelihood, paying little respect to the source hub. Note that while nearly PA diagrams are known to have steady conductance, the same isn't known for PA graphs.

Previous to [3,4], it was realized that the high conductance infers that non-uniform gossip spreading succeeds. By nonuniform we imply that, for each requested match of neighbors i and j, hub I will choose j with likelihood pij for the gossip spreading advance (by and large, pij ≠ pji). Boyd et al. [5] consider the ''averaging'' issue on general diagrams, which is firmly identified with the

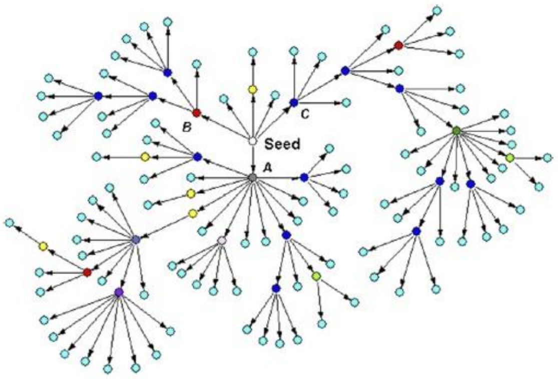


Figure 1. Rumor spreading in spite of a high edge expansion.

Joining of PUSH- - PULL. A culmination of their fundamental outcomes is that, if the pij are appropriately picked, non-uniform PUSH- - PULL gossip spreading prevails inside O(log n) adjusts in nearly PA diagrams. They additionally demonstrate that this dissemination can be discovered proficiently utilizing neighborhood calculations in these diagrams, however their strategy requires Ω(log n) steps. While the commitment of [5] is imperative, this culmination is in our setting to some degree trifling. That such a likelihood circulation exists is direct. As a result of their high conductance, nearly PA diagrams have distance across O(log n). Along these lines, in a synchronous system, it is conceivable to choose a pioneer in O(log n) numerous rounds and set up a BFS tree beginning from it. By relegating likelihood 1 to the edge between a hub and its parent one has the coveted non-uniform likelihood conveyance. Accordingly, from the perspective of this paper the presence of a non-uniform issue is somewhat uninteresting. Boyd et al. [5] likewise demonstrate that these circulations can be discovered effectively utilizing neighborhood calculations, however their strategy requires Ω(log n) numerous means. The nearby calculations of every hub, at each progression, incorporate a telecom of a few esteems to all neighbors. Nearby communicating, utilized for width (that is, O(log n)) numerous rounds, is an inconsequential data spread procedure.

Additionally, Mosk-Aoyama and Shah [6] consider the issue of figuring detachable capacities. Specifically, they consider the uniform gossip spreading issue on diagrams weighted by a high-conductance doubly stochastic network "that doles out equivalent likelihood to every one of the neighbors of any hub" (that is, if pij is the likelihood that hub i starts an association with hub j in the nonexclusive round t, at that point ∀ij∈ E(G) pij = pji = Δ–1, where Δ is the most extreme degree in the graph). Their work infers that if the conductance of a graph is Ω(1) at that point gossip spreading closes in O(δlog n) numerous rounds—this, while being a decent headed for steady degree diagrams, is polynomially expansive for PA graphs.

III. GOSSIP CONTROLLING

Anti-gossip can likewise be spread from individual to individual not at all like antibodies for infections which must be managed to people. In some sense this makes our concern more tractable yet in addition it implies it must be contemplated in view of various finishes. In this paper we consider a suite of techniques. Our key knowledge in concentrate hostile to bits of gossip in a decentralized setting is this: The propagation of the anti-gossip does not depend principally on the definitiveness of the source that issues the anti-goosip yet on the trust clients put in their companions in the interpersonal organization.

The principal procedure we ponder, the Slow Start Model, models a circumstance where a nearby expert may find gossip n days after it begins and choose to spread a hostile to gossip.

In the second model, called the Bonfire Model, we expect that the informal community contains an arrangement of cautious specialists, signals that are watchful for the spread of gossipy tidbits. Once a reference point gets gossip it quickly begins spreading hostile to bits of gossip to battle the gossip. This technique relates to a semi-brought together situation where coalitions of specialists may proactively choose to seed the system with watchful clients who can both distinguish bits of gossip and react to them.

IV. ANTI-GOSSIP MODEL

A. Slow Start Model

Here we demonstrate the circumstance that an expert with constrained purview distinguishes the spread of gossip and afterward battles it by beginning a free course from an arbitrarily chose contaminated hub. We fight that there will dependably be a period slack between the beginning of gossip and its location. This time slack is refered to as postpone time and is represented by n. The procedure begins from a solitary tainted hub Vi, n time units after the gossip begun, Vi spreads the counter gossip messages to its neighbors Ni. Every hub w 2 Ni acknowledges the counter gossip with likelihood q. Through the course of our tests we have taken q to be 0.05.

B. Bonfire Model

Between the time an expert identifies the spread of gossip and chooses how to battle it, the gossip keeps spreading apace. With a specific end goal to proactively battle bits of gossip, specialists may implant operators in the system that are equipped for identifying the spread of gossip and are approved to begin spreading against bits of gossip when they recognize the spread of gossip. We call these specialists reference points. In this paper, the reference point hub utilizes an indistinguishable system from the Delayed begin model to spread hostile to gossip, i.e., it spreads the counter gossip to every one of its neighbors with some likelihood. In our analyses the guide hubs are chosen aimlessly. Be that as it may, in genuine systems, the hubs can be chosen in light of different qualities like network, expert, trust and so on. In addition, signal hubs can likewise be point particular. For instance, one hub may go about as a guide for innovation based gossipy tidbits however not for amusement based bits of gossip. This determination will include theme and expertize mining from the system. The Bonfire demonstrate with one signal is equivalent to the Delayed begin show. In the Delayed begin demonstrate, the beginning time of the counter gossip process is settled however here it relies on the time when the guide is enacted.

V. CONCLUSION

We have indicated how quick the PUSH- - PULL system disperses some data all through the hubs of a PA diagram, also, how moderate the PUSH, PULL systems acquire a similar outcome. We trust that our outcomes may offer a few bits of knowledge into genuine gossip spreading among people. To be specific, it appears conceivable that in an informal community there exists a ''center'' of individuals that won't not be VIPs, but all things considered can achieve a dominant part of their group in a couple of steps.

VI. REFERENCES

[1] B. Bollobás, O. Riordan, J. Spencer, G. Tusnády, The degree sequence of a scale-free random graph process, Random Structures & Algorithms 18 (3) (2001) 279–290.

[2] M. Mihail, C.H. Papadimitriou, A. Saberi, On certain connectivity properties of the internet topology, in: Proceedings of the 44th Symposium on Foundations of Computer Science, FOCS 2003, pp. 28–35.

[3] FlavioChierichetti, Silvio Lattanzi, Alessandro Panconesi, Almost tight bounds for rumour spreading with conductance, in: Proceedings of the 42nd ACM Symposium on Theory of Computing, STOC 2010.

[4] FlavioChierichetti, Silvio Lattanzi, Alessandro Panconesi, Rumour spreading and graph conductance, in: Proceedings of the 21st ACM–SIAM Symposium on Discrete Algorithms, SODA 2010.

[5] S.P. Boyd, A. Ghosh, B. Prabhakar, D. Shah, Gossip algorithms: design, analysis and applications, in: Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM 2005, pp. 1653–1664.

[6] D. Mosk-Aoyama, D. Shah, Fast distributed algorithms for computing separable functions, IEEE Transactions on Information Theory 54 (7) (2008) 2997–3007.

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3| ISSN : 2456-3307

Attribute-Based Encryption with Equality Test in Cloud Computing Using Key-Policy

Mr.R. Venkatesan[1], Dr. M. Geetha[2]

[1]Research Scholar, Indian Arts and Science College, Kondam,Tiruvannamalai, Tamil Nadu, India

[2]Indian Arts and Science College, Kondam,Tiruvannamalai, Tamil Nadu, India

ABSTRACT

The privacy of users should be thought of because the utmost priority in distributed networks. To protect the identities of users, attribute-based encoding (ABE) was presented by Sahai et al. ABE has been wide utilized in several situations, significantly in cloud computing. During this paper, public key encoding with equality check is concatenated with key-policy ABE (KP-ABE) to present KP-ABE with equality test (KP-ABEwET). The projected theme not solely offers ne-grained authorization of cipher texts however additionally protects the identities of users. In contrast to ABE with keyword search, KP-ABEwET will take a look at whether or not the cipher texts encrypted by completely different public keys contain constant data. Moreover, the authorization process of the conferred theme is additional edible than that of Ma et al.'s scheme. Moreover, the projected scheme achieves one-way against chosen-cipher text attack supported the additive Dife Hellman (BDH) assumption. Additionally, a brand new procedure drawback referred to as the twin-decision BDH downside (tDBDH) is proposed during this paper. tDBDH is established to be as laborious because the decisional BDH downside. Finally, for the rest time, the protection model of authorization is provided, and also the security of authorization supported the tDBDH assumption is proved within the random oracle model.

Keywords: Cloud service, attribute-based encryption, public key encryption, equality test, keyword Search

I. INTRODUCTION

In the current network era, cloud service suppliers provide in - nite space for storing and computing power for users to manage their information. To fancy these services, people and organizations store their non-public information on cloud servers. However, within the case of security breaches, users' non-public information hold on within the cloud is not any longer safe. once users source their information to cloud servers, they expect complete privacy of their information hold on within the cloud. Protective the privacy and information of users has remained a awfully crucial drawback for cloud servers. To avoid any inconvenience, users store their non-public information in encrypted kind. For ne-grained sharing of encrypted information, Sahai and Waters conferred attribute-based cryptography (ABE) [2]. ABE may be a public key cryptosystem variant that enables users to access secret information supported their attributes. This cryptosystem enriches the property of the cryptography policy and therefore the description of users' rights and it changes from a one-one to one-many situation throughout the encryption and decoding phases. Moreover, it hides the identities of the users in acceptable terms. During a resultant work, Goyal et al. projected key-policy attribute-based cryptography (KP-ABE) in 2006 [18]. The underlying cryptonyms-tem combines the key key and therefore the access structure. Bettencourt et al. projected cipher text-policy attribute-based cryptography (CP-ABE) in 2007 which mixes the cipher text and therefore the access structure. Thereafter, various cryptographers conferred several analyses works supported ABE shortly once its conceptualization, ABE reached prime importance in our existence (for example, in tv payment systems, personal health record sys-teams and then on). Moreover, ABE is additionally being wide incorporate-rated in cloud computing. However, if one needs to check plaintexts adore 2 cipher texts, the key should be wont to decipher the 2 cipher texts. To overcome this drawback, Yang et al. conferred a replacement cryptosystem referred to as public key cryptography with equality take a look at (PKEwET) in 2010. His planned system will take a look at whether or not 2 cipher texts contain constant plaintexts with-out secret writing. However, this theme permits anyone to perform such a check. to beat this defect, Tang created some enhancements to the theme (e.g., PKEET with ne-grained authorization (FGwPKEET), all-or-nothing PKEET (AoNwPKEET) [28] associate degree an extension of FG-PKEwET). In 2015, Ma et al. projected a replacement primitive referred to as PKEwET supporting edible authorization (PKEwET-FA). There area unit four forms of edible authorizations in their theme. To change the certificate management of PKEwET, Ma combined the ideas of PKEwET and identity-based cryptography to gift identity-based cryptography with equality check (IBEET). Recently, in 2017, Wu et al. improved Ma et al.'s theme by reducing the machine time value. To offer additional ne-grained authorization, we have a tendency to propose a replacement primitive known as key-policy attribute-based encoding with equality check (KP-ABEwET). we tend to mix the ideas of PKEwET and KP-ABE. As conferred in suppose that there area unit four users. S and S0 area unit the sets of attributes for encoding, and T and T0 check with the access structures utilized by the coding secret key. S00 denotes the set of attributes of the tester, and T0A is that the access structure used for the authorization of the attribute set of SA0. T0B is that the access structure used for the authorization of the attribute set of SB0. We tend to describe the underlying situation as follows: User one will store his personal information within the cloud and might decode the cipher texts that area unit encrypted by a group of attributes S with T(S) D one. User a pair of will store his personal information within the cloud, however he cannot decode the cipher texts that area unit encrypted by a group of attributes S with T(S) 6D1. User three has the attribute S00, wherever T0A(S00) D one and T0B(S00) D one, and he will perform the check over 2 completely different cipher texts encrypted by attribute SA0 and attribute SB0. User four doesn't have the attribute S00 satisfying T0A(S00) D one and T0B(S00) D one, and he cannot perform the check over 2 completely different cipher texts encrypted by attribute SA0 and attribute SB0.

A. Contribution

This paper presents a replacement primitive known as key-policy attribute-based encoding with equality take a look at (KP-ABEwET). Our objective is to realize a ne-grained authorization of cipher texts. the most technologies in our theme embrace key-policy attribute-based encoding (KP-ABE) [18] and public key encoding with equality check (PKEwET) the most contributions will be summarized as follows:

1) First, we tend to style a replacement theme by combining KP-ABE with PKEwET. Compared with the present PKEwET schemes, our projected theme supports activity the ne-grained take a look at of cipher texts and changes from one-one to one-many for users within the testing algorithmic rule.
2) Our theme will be viewed as associate degree extension of attribute-based encoding with keyword search (ABEwKS). at the side of different aspects, the planned theme permits testing whether or not the cipher texts contain

identical data that square measure encrypted by completely different public keys.

3) The projected theme achieves unidirectional against chosen-cipher text attack (OW-CCA) supported the additive Dif e-Hellman (BDH) assumption within the random oracle model.
4) A new process drawback known as the twin-decision additive Dif e-Hellman drawback (tDBDH) is additionally conferred and is established to be as laborious because the DBDH drawback.
5) We give the protection model of authorization and prove the protection of authorization supported the tDBDH assumption within the random oracle model. To the most effective of our data, this work is that the rst to prove the protection of authorization in such a way.

B. Related Work

Deterministic encoding, planned by Bellare et al. [8], is another primitive that supports the equality take a look at on cipher-texts. This primitive was completely studied in several subsequent works [1], [7] however all of them square measure settled algorithms. Conversely, PKEwET could be a probabilistic algorithmic rule that supports the equality take a look at on cipher texts.

PKEwET may be viewed as associate extension of public key encoding with keyword search (PEKS). The construct of PEKS was projected by Boneh et al. [4]. It will perform keyword searches over cipher texts while not decrypting them. Later, many modi male erectile dysfunction schemes of PEKS were projected [6], [9], [11], [12]. to resolve the matter of access management in a very multi-user setting, PEKS was combined with ABE for achieving the applied perspective in cloud computing. In [5], [10], [13], [15], [17], the authors com-binned PKES with KP-ABE. In another works, including [3], [14], [16], the authors combined PKES with CP-ABE whereas incorporating the access structure with the cipher text of the keyword search. Though the results were slightly completely different, none of the works conferred a mechanism to see whether or not 2 {different totally different completely different} cipher texts encrypted by different public keys contain a similar data. to beat this limitation, we tend to gift a good KP-ABEwET mechanism.

C. Organization

The remainder of this paper is organized as follows. In Section two, we have a tendency to introduce connected preliminaries. Section three describes the system and also the security model. Our theme is conferred in Section four. Section five provides the protection proof of our theme and of authorization. In Section half-dozen, the performance evaluations area unit cheese y mentioned. Finally, Section seven presents the final remarks.

II. PRELIMINARIES

In this half, we tend to introduce some basic data, as well as cryptographically assumptions, Shamir's secret sharing theme and access tree, that's utilized during this paper

A. Cryptographic Assumptions

The following section presents the Diamond State nations of linear maps and also the drawback formulation.

De nation 1: linear Maps: Let G1 and G2 be multiplicative teams of prime order letter, e V G1 G1 ! G2 be a linear map, and g be a generator of G1. linear maps West African ll the subsequent conditions:

(1) Bilinearity: 8g1; g2 a pair of G1 and 8a; b a pair of Zq, we've got e(ga1; gb2) D e(g1; g2)ab.

(2) Non-degeneracy: e(g; g) 6D1.

(3) Computability: 8g1; g2 a pair of G1, we are able to cipher e(g1; g2).

De nation 2: linear Dif e-Hellman (BDH) problem: Let G1 and G2 be increasing teams of prime order letter, e V G1 G1 ! G2 be a linear map, and g be a generator of G1. The BDH drawback is that given a 4-tuple (g; ga; gb; gc), the aim is to cipher e(g; g) abc, wherever a; b; c a pair of Zq.

De nation 3: Diamond Statecisional linear Dif e-Hellman (DBDH) problem: Let G1 and G2 be increasing teams of prime order letter, e V G1 G1 ! G2 be a linear map, and *g* be a generator ofG$_1$. The DBDH problem is to distinguishbetween the distributions of 5-tuples

(g; g^a; g^b; g^c; e(g; g)^{abc}) and (g; g^a; g^b; g^c; e(g; g)^d), where a; b; c; d 2 Z_q.

De nation 4: Twin-Decision Bilinear Dif e-Hellman(tDBDH) problem: Let G$_1$ and G$_2$ be multiplicative groups of prime order *q*, *e* V G$_1$ G$_1$! G$_2$ be a bilinear map, and *g* be a generator of G$_1$. The tDBDH problem is to In general, the tDBDH problem appears to be weaker than the DBDH problem. However, this problem is in fact as hard as the DBDH problem. (The tDBDH problem is different from the twin bilinear Dif e-Hellman inversion problem that proposed by Chen et al.)

Theorem 1: *The tDBDH problem is as hard as the DBDH problem. Proof:* It is quite clear that tDBDH DBDH. Next, wepresent the proof of DBDH tDBDH.

To prove DBDH tDBDH, we suppose that there is an algorithm A that can solve the tDBDH problem in polynomial time. We construct an algorithm B as follows. B takes a 4-tuple (g^a; g^b; g^c; e(g; g)^d) as input, and its objective is to determine whether e(g; g)^d D e(g; g)^{abc} holds.

B chooses a random range x and constructs a 7-tuple (ga; gb; gc; e(g; g)d ; gbx ; gcx ; e(g; g)dx2). Then, it calls thealgorithm A. The rule A checks whether or not e(g; g)d D e(g; g)abc and e(g; g)dx2 D e(g; g)abcx2 hold.

If A outputs affirmative, then it implies that e(g; g)d D e(g; g) abc and e(g; g)dx2 D e(g; g)abcx2 . Apparently, it's doubly con-riming that the input could be a affirmative DBDH instance. Thus, B replies "yes".

If A outputs no, then it implies that either e(g; g)d 6D e(g; g)abc or e(g; g)dx2 6De(g; g)abcx2 . no matter that is true, will quickly deduce that the input could be a no DBDH instance. Thus, B replies "no".

B. SHAMIR'S SECRET SHARING SCHEME

Shamir's (t; n)-threshold secret sharing theme is predicated on the Lagrange interpolation polynomial. an in depth introduction is delineated as follows:
Given t distinct points (xi; f (xi)), wherever f (x) may be a polynomial of degree but t, f (x) is set as follows: Shamir's theme is Delaware need for a secret s a pair of Zp by setting a0 D s and selecting a1; a2; ; at one a pair of Zq. For all one xi q; one i n, the trustworthy party computes f (xi), wherever f (x) D noble metal one a xk . The shares (x ; f (x)) ar distributed to n distinct

parties. Since the key may be a constant term s D a0 D f (0), the key will be recovered from any t shares (xi; f (xi)) as follows:

C. ACCESS TREE

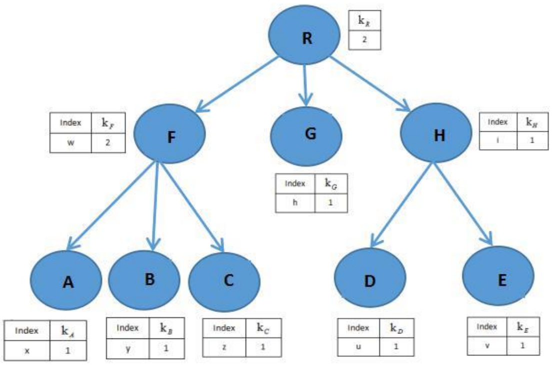


Figure 1. Access Tree

We suppose that T is AN access tree composed of leaf nodes and non-leaf nodes (e.g., Fig. 2). every leaf node represents AN attribute, and every non-leaf node represents a logic element. every logic element is drawn by its youngsters and also the threshold price. Let numb be the quantity of kids of a node x and kx be the brink price of the node x; we've got zero kx numb . Then, every leaf node includes a threshold price kx D one.

We suppose that the kids of each node do have orders from one to num. Next, we tend to First State ne some new functions. The perform parent(x) represents the parent of node x. The function att(x) is First State need as AN attribute related to the leaf node. The perform index(x) returns the quantity related to node x.

Let r be the basis of AN access tree T, expressed as Tr . Lone-Star State refers to the sub tree of T unmoving at node x. Lone-Star State (S) D one means the set of attributes S sates atomic number 99 the tree Lone-Star State . Here, we tend to use a algorithmic rule to reckon Lone-Star State (S).

If x may be a non-leaf node, we tend to reckon Tx0 (S) for all children x0 of x. If a minimum of youngsters come one, then Lone-Star State (S) returns one.

If x may be a leaf node, then Lone-Star State (S) returns one if att(x) a pair of S.

III. PROBLEM FORMULATION

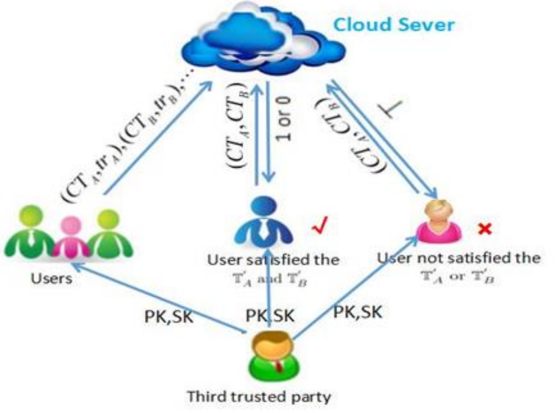

Cloud Sever
$(CT_A, tr_A),(CT_B, tr_B),\dots$
(CT_A, CT_B)
1 or 0
(CT_A, CT_B)
$\perp$
✓
✗
Users
User satisfied the $\mathbb{T}'_A$ and $\mathbb{T}'_B$
User not satisfied the $\mathbb{T}'_A$ or $\mathbb{T}'_B$
PK,SK
PK,SK
PK,SK
Third trusted party

FIGURE 2. System model for KP-ABEwET.

A. System Model

Fig. 3 illustrates the system model of KP-ABEwET. The sys-tem has 3 taking part entities: the cloud server, the users and a sure third party. The trusty third party generates public key pk and personal key sk for users. The users code and send their non-public information to the cloud server. If a user needs the cloud server to check the cipher text, then the cloud server is permitted and gains a trapdoor tr. However, the cloud server will solely take a look at whether or not the 2 cipher texts contain an equivalent info and can't decode them exploitation the trapdoor. The legitimate users access information per their attributes and may decode their cipher texts or take a look at the cipher texts. If the legitimate users satisfy the access structure for the take a look at, they will get the take a look at results of the cipher texts from the cloud server. If the legitimate users satisfy the access structure for the decoding, they will decode the cipher texts.

An integrated KP-ABEwET theme consists of six algorithms: Setup, Encrypt, KeyGen, and Trapdoor, decode and check. Here, we have a tendency to let M be plaintext house and C be cipher text area.

(1) Setup (k): It takes a security parameter k as input, so it outputs the general public parameters pp and pk and also the master mk.
(2) Encrypt (M; pk; S; S0): It takes a message M a pair of M, public key pk and 2 sets of attributes S; S0 as inputs, so it outputs the cipher text CT a pair of C.
(3) KeyGen(T; T0; S; S0; pp; mk): This rule takes as inputs the master mk, 2 access trees T; T0, and 2 sets of attributes S; S0 that satisfy T(S) D one and T0(S0) D one, and it later on outputs the personal key sk.
(4) Trapdoor (S0; T0; mk): It takes mk, T0 and S0 as inputs, and it outputs the trapdoor td.
(5) Decrypt(CT ; sk; S; S0): It takes as inputs a cipher text CT a pair of C; S; S0 and also the non-public key sk, and it outputs the message M if T(S) D

one and T0(S0) D one. Here, CT is encrypted victimization the sets S and S0.

(6) Test(CTA; CTB; tdA; tdB; S0): Suppose that CTA may be a cipher text of the sets of attributes reserves and SA0 which CTB may be a cipher text of the sets of attributes SB and SB0. This algorithm takes as inputs 2 cipher texts CTA; CTB, the trapdoors tdA; tdB and also the set S0 of attributes that satisfy T0A(S0) D one and T0B(S0) D one, so it outputs one if CTA and CTB contain an equivalent message; otherwise, it returns zero.

B. Security Model

Here, the protection model of the projected theme and also the security model of authorization area unit conferred.

First, we tend to American state ne unidirectional against chosen-cipher text attack (OW-CCA) for KP-ABEwET below a selected set of attributes, as follows.

Game 1: Suppose that A is that the soul. A announces a group of attributes that he needs to be challenged, shown as S.

(1) Setup. The competition C takes a security parameter k as input and outputs public parameters pp to A with the Setup formula of KP-ABEwET.

(2) Phase one. A performs the subsequent varieties of queries polynomials repeatedly.

Key retrieve queries: A performs any queries for personal keys for several access structures Ti, wherever S 2= Ti for all i. C sends sk to A.

Decryption queries: A performs several queries for cipher texts. C runs the rewrite formula and out-puts the plaintext reminiscent of the cipher text or? to A.

Trapdoor queries: C runs the Trapdoor formula and outputs td to A.

(3) Challenge: C indiscriminately chooses a message M a pair of M, sets CT D Encrypt (pk; M) and sends CT to A as his challenge cipher text.

(4) part 2: Phase one is perennial. The constraints area unit that CT doesn't seem within the coding queries.

(5) Guess: A outputs a guess M two M and wins the sport if M D M.

The advantage of A is First State ned as Pr[M D M].

De nation 5: The KP-ABEwET theme is OW-CCA secure if the advantage of all polynomial time adversaries is negligible within the on top of game.

Finally, we tend to First State ne a testable against chosen-cipher text attack (T-CCA) of authorization for KP-ABEwET below the chosen sets of attributes, as follows:

Game 2: Suppose that A2 is associate degree individual. A2 announces 2 sets of attributes S and S0 that he desires to be challenged. Here, (S \ S0) D ?, S is employed for coding, and S0 is employed for the trapdoor.

(1) Setup. The competition, C, takes a security parameter k as input and outputs public parameters pp to A2 by mistreatment the Setup formula of KP-ABEwET.

(2) Phase one. A2 performs the subsequent kinds of queries polynomials over and over. Key retrieve queries: A2 performs several queries for personal keys for any access structures Ti and T0j, where

S 2= Ti for all i and S0 2= T0j for all j. C sends sk to A2. Decoding queries: A2 performs several queries for cipher texts. C runs the decode algorithmic rule and out-puts the plaintext akin to the cipher text or

? to A2.Trapdoor queries: C runs the Trapdoor algorithmic ruleand outputs td to A2.

Test queries: C runs the check algorithmic rule and outputs 1 for equality cipher texts and 0 for unequal cipher texts or?.

(3) Challenge: C chooses a random variety # two f0; 1g. If # D 1, then C chooses one message M, sets CT1 D Encrypt (pk; M); CT2 D Encrypt (pk; M)

and sends CT1 ; CT2 to A2 as his challenge cipher texts. If # D 0, C chooses 2 unequal messages, money supply and M2; sets

CT1 D Encrypt (pk; M1); CT2 D Encrypt(pk; M2) and sends CT1 ; CT2 to A2 as his challenge cipher texts.

(4) Part 2: Phase one is recurrent with the conditions that CT1 and CT2 don't seem in decoding queries and CT1 and CT2 don't seem in check queries.

(5) Guess: A2 outputs a guess # and wins the sport if

# D #, which means one for money supply D M2 or zero for money supply 6DM2.

The advantage of A2 is First State need as jPr[# D #] 1=2j. First State nation 6: The KP-ABEwET theme is T-CCA secure in terms of authorization if the advantage of all polynomial time adversaries is negligible within the previously mentioned game.

IV. OUR CONSTRUCTIONS

The following section presents the projected KP-ABEwET theme. Setup (k): It takes a security parameter k as input and outputs public parameters pp as follows:

(1) Generate linear teams, G1; G2 and jG1j D alphabetic character; jG2j D q, and select a random generator g 2 G1. Then, let e V G1 G1 ! G2 be a linear map.

(2) Let A be a universe of properties of attributes. For simplicity, we have a tendency to take the rst A parts of Zq because the universe,formally as 1; 2; jAj(mod q).

(3) Let H1 V f0; 1gjAj G2! f0; 1gkCl, H2 V f0; 1gjAj G2 ! G1, and H3 V 5G1 f0; 1gkCl! f0; 1gk be hash functions, wherever l is that the length of the weather of Zq.

(4) Choose x1; x2; ; xjAj; y1; y2 two Zq arbitrarily, then output public keys pk,

X1 D gx1 ; ; XjAj D gxjAj ; Y1 D e(g; g)y1 ; Y2 D e(g; g)y2 , and also the passkey mk, (x1; x2; ; xjAj; y1; y2). Encrypt (M; pk; S; S0): It takes a message M, public key pk and 2 sets of attributes S; S0 as inputs, wherever (S \ S0) D ;,S is used for coding, and S0 is employed for testing. Then, it outputs the cipher text as follows:

Choose r1; r2; r3 a pair of Zq at random, and so formulate the following:

CT D (S; S0; C1 D gr1 ; C2 D M k r1 H1(S; Y1r2); C3 D Mr1 H2(S0; Y2r3); C4 D fEi D Xir2 gi2S ;
C5 D fEj D Xjr3 gj2S0 ; C6 D H3(Mr1 ; C1; C2; C3; C4; C5))

KeyGen (T; T0; S; S0; pp; mk): This algorithmic program takes the passkey mk, 2 sets of attributes S; S0 satisfying T(S) D one and T0(S0) D one and (S0 TS) D ? as inputs, and it outputs the non-public key as follows:

(1) The algorithmic program chooses a polynomial qx for every node x within the tree T. The polynomials area unit chosen from prime to bottom, ranging from the basis node r. the small print area unit conferred as follows:

For each node x in T, it sets the degree dx of the polynomial qx to be one but the edge price kx of that node, which suggests that dx D kx.

V. SECURITY ANALYSIS

The following section provides the protection proof of the conferred KP-ABEwET theme.

Theorem 2: Our projected theme is OW-CCA secure against the resister World Health Organization is permitted with a trapdoor supported the BDH assumption within the random oracle model.

Proof: Suppose that A is that the resister that may break the bestowed KP-ABEwET theme. Then, there's AN algorithmic ruleC to solve the BDH drawback with a non-negligible advantage. Given a 4-tuple (g; A; B; C) D (g; ga; gb; gc), the target of algorithmic rule C is to calculate e(g; g)abc. Init Suppose that there's a universe. A chooses a group of Paste your text here and click on "Next" to look at this text editor do it's issue.

Don't have any text to check? don't have any text to check? Click "Select Samples". Phase 1 A performs the subsequent sorts of queries poly-nominally times.

H1-query: A could issue queries to the random oracle H1. to retort to those queries, C maintains a listing of tulles H1. every component within the list may be a tulle of the shape (S ; ;). The list is at first empty. Responding to question (S ;), C runs as follows:

If the question (S ;) already seems within the H1 list within the type (S ; ;), then C responds to A with H1(S ;) D.

Otherwise, C simply takes 2 G2, so it responds to A with H1(S ;) D . C adds the tulle (S ; ;) to the H1 list.

Key retrieve queries: A performs several queries for private keys for several access structures T, wherever S doesn't satisfy T. C sends sk to A as follows:

(1) C builds 2 algorithms: SatT and DNSatT, as follows:

SatT(Tx ; S; vx): This algorithmic program constructs the polynomials for the nodes of associate degree access sub-tree with a sates dysfunction root node once Lone-Star State (S) D one. It takes as inputs a group of attributes S, associate degree access tree Lone-Star State and a random range vx 2 Zp, and it outputs a polynomial qx of degree dx for the foundation node x as follows:

Let qx (0) D vx and indiscriminately select dx different points of the polynomial qx to construct qx . The algorithmic program constructs polynomials for every kid node x0 of x by death penalty the algorithmic program SatT(Tx0 ; S; qx (index(x0))).

DNSatT(Tx ; S; gvx): This algorithmic program constructs the polynomials for the nodes once Lone-Star State (S) D zero. It takes a group of attributes S, associate degree access tree Lone-Star State and a random part gvx a pair of G1, wherever vx a pair of Zp, and it outputs a polynomial qx of degree dx for the basis node x as follows:

Because Lone-Star State (S) D 0, the foundation node has but dx satis disjunction kids. Suppose that sx is that the range of sates disjunction kids of x, which means that sx < dx . The algorithmic program chooses a random range vx0 a pair of Zp for every satis disjunction kid x0 of x. Let qx (index(x0)) D vx0 and indiscriminately select different dx sx points of the polynomial qx to construct qx We acquire qx ( ) for every node in T as follows.

VI. PERFORMANCE EVALUATION

We in theory analyze the straight line quality of the projected theme and alternative PKEwET schemes in Table one. we have a tendency to describe the process quality in terms of the involution operation E and also the pairing operation P. we tend to denote the quantity of attributes needed within the cipher-text by jSC j and jSC0 j. In Table 1, CEnc, CDec and C Test represent the cryptography algorithms, decoding algorithms and check algorithms, severally. Lollop said genus represents the proof of authorization. From the second to the fourth columns, we tend to gift the process complexities of CEnc, CDec and C Test. The 5 column indicates whether or not the underlying schemes area unit attribute primarily based. The sixth column shows whether or not the schemes have the proof of authorization. The seventh column highlights the safety levels of the schemes. The last column presents the underlying assumptions for guaranteeing the safety.

From Table one, we have a tendency to observe that the process com-laxity of our theme depends on the amount of attributes needed by the cipher text. as a result of our theme incorporates the ABE state of affairs, it's going to not be as client because the current works. The trade off is adjusted whereas providing the protection of user identities. Moreover, in distinction to previous works, our theme additionally permits the users to get ne-grained authorization of cipher texts. To the simplest of our information, Ma et al. rest given four varieties of authorizations in [29]. we tend to nd that our projected theme will perform the authorization and take a look at in an exceedingly additional edible manner as a result of in our theme, we are able to perform the authorization mistreatment the attributes of users. moreover, for the time, the proof of authorization is evidenced supported the tDBDH assumption.

VII. CONCLUSION

In this paper, a replacement cryptosystem known as key-policy attribute-based encoding with equality check (KP-ABEwET) is pre-scented. To the most effective of our information, KP-ABEwET is that the rst commit to mix the general public key encoding supporting equality check with key-policy attribute-based secret writing. The planned theme are often viewed as AN extension of attribute-based encoding with keyword search (ABEwKS) with the distinction that it will check whether or not the cipher texts contain a similar info that were encrypted by completely different public keys. In distinction to previous schemes with equality check, the new theme supports testing the cipher texts with ne-grained authorization and additionally hides the identity of the user. Moreover, the projected theme is unidirectional secure against chosen-cipher text attack (OW-CCA) supported the linear Dif e-Hellman (BDH) downside. Moreover, a replacement computational downside known as twin-decision additive Dif e-Hellman downside (tDBDH) is projected and is proved to be as laborious because the DBDH downside. Finally, the protection model of authorization is conferred, and therefore the security of authorization supported the tDBDH assumption is proved within the random oracle model. To the simplest of our information, this work is that the RST to prove the protection of authorization in such a state of affairs.

VIII. REFERENCES

[1] KP- A. Boldyreva, S. Fehr, and A. O'Neill, ``On notions of security for deter-ministic encryption, and ef cient constructions without random oracles,'' in Proc. Annu. Int. Cryptol. Conf., 2008, pp. 335 359.

[2] A. Sahai and B. Waters, ``Fuzzy identity-based encryption,'' in Proc. Annu. Int. Conf. Theory Appl. Cryptograph. Techn., 2005, pp. 457 473.

[3] C. Wang, W. Li, Y. Li, and X. L. Xu, ``A ciphertext-policy attribute-based encryption scheme supporting keyword search function,'' in Proc. CSS, 2013, pp. 377 386.

[4] M. Bellare, M. Fischlin, A. O'Neill, and T. Ristenpart, ``Deterministic encryption: De nitional equivalences and constructions without random oracles,'' in Advances in Cryptology CRYPTO (Lecture Notes Com-put. Science), vol. 5157. Berlin, Germany: Springer-Verlag, Aug. 2008, pp. 360 378.

[5] M. Bellare, A. Boldyreva, and A. O'Neill, ``Deterministic and ef - ciently searchable encryption,'' in Proc. Annu. Int. Cryptol. Conf., 2007,pp. 535 552.

[6] M. Nishioka, ``Perfect keyword privacy in PEKS systems,'' in Provable Security. Berlin, Germany: Springer, 2012, pp. 175 192.

[7] J. Lai, X. Zhou, R. H. Deng, Y. Li, and K. Chen, ``Expressive search on encrypted data,'' in Proc. 8th ACM SIGSAC Symp. Inf., 2013, pp. 243 252.

[8] J. Li and L. Zhang, ``Attribute-based keyword search and data access control in cloud,'' in Proc. 10th Int. Conf. Comput. Intell. Secur. (CIS), Nov. 2014, pp. 382 386.

[9] J. Han, W. Susilo, Y. Mu, and J. Yan, ``Privacy-preserving decentralized key-policy attribute-based encryption,'' IEEE Trans. Parallel Distrib. Syst., vol. 23, no. 11, pp. 2150 2162, Nov. 2012.

[10] S. Li and M. Z. Xu, ``Attribute-based public encryption with keyword search,'' Chin. J. Comput., vol. 37, no. 5, pp. 1017 1024, 2014.

[11] P. Liu, J. Wang, H. Ma, and H. Nie, ``Efficient verifiable public key encryption with keyword search based on KP-ABE,'' in Proc. 9th Int. Conf. Broadband Wireless Comput., Commun. Appl. (BWCCA), Nov. 2014, pp. 584 589.

[12] A. Lewko and B. Waters, ``Decentralizing attribute-based encryption,'' in Proc. Annu. Int. Conf. Theory Appl. Cryptogr. Techn., 2011, pp. 568 588.

[13] S. Hohenberger and B. Waters, ``Online/offline attribute-based encryp-tion,'' in Proc. Int. Workshop Public Key Cryptogr., 2014, pp. 293 310.

[14] P. Datta, R. Dutta, and S. Mukhopadhyay, ``Fully secure online/offline predicate and attribute-based encryption,'' in Proc. ISPEC, 2015, pp. 331 345.

[15] G. Yang, C. H. Tan, Q. Huang, and D. S. Wong , ``Probabilistic public key encryption with equality test,'' in Proc. Cryptogr.-Track RSA Conf., 2010, pp. 119 131.

[16] S. Ma, Q. Huang, M. Zhang, and B. Yang, ``Efficient public key encryption with equality test supporting flexible authorization,'' IEEE Trans. Inf. Forensics Security, vol. 10, no. 3, pp. 458 470, Mar. 2015.

[17] S. Ma, ``Identity-based encryption with outsourced equality test in cloud computing,'' Inf. Sci., vol. 328, pp. 389 402, Jan. 2016.

[18] L. Wu, Y. Zhang, K.-K. R. Choo, and D. He, ``Efficient and secure identity-based encryption scheme with equality test in cloud computing,'' Future Generat. Comput. Syst., vol. 73, pp. 22 31, Aug. 2017.

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3| ISSN: 2456-3307

NS in Digitalization: Attacks and Defence

M. Soundharya[1],Mrs. E. Bhuvaneswari[2]

[1]Research Scholar, Department of Computer Science,Kamban College of Arts and Science for Women, Thiruvannamalai, Tamil Nadu, India

[2]Assistant Professor,Department of Computer Science,Kamban College of Arts and Science for Women, Thiruvannamalai, Tamil Nadu, India

bhuvaname2008@gmail.com[1],msriya333@gmail.com[2]

ABSTRACT

NS is become a gambol in our whole world. In today's world the volume of data increases every second and threats continue to transform and grow, weakening your ability to attacks. The business world is going digital as a result to bypass these things and we are adopting different methods. Network administrator has to keep track, has to update with all current advances in both the software and hardware fields to avert the user's data. Digitalization is playing an important role and integrated of digital technologies into our daily life. This paper precise similar method which are used to attack as well as various mechanisms against to defense them.

Keywords : DOS attacks, Encryption, Firewalls, Port Scanning, SHTTP, SSL, VPN

I. INTRODUCTION

Network security broach towards protect the website severs in various forms of attack. Network security has become foremost in every field of world such as military education, government, business and our day to day lives. we can better defend ourselves by keeping track of all the knowledge know how the attacks are attained. Modifying the network architecture we can avert these type of attacks, many companies employ firewall and diverse polices to safeguard them. Security network has immense field which was expanded stage and as per today's criteria. To understand the contemporary analysis being done, one should have knowledge of its background should have working in our present world internet is accessible everywhere in our house, in our work area, cars and mobiles everything is connected to the internet, if any unknown person is able to acquire access to this network they can not only spy on us but they can easily mishmash up our lives. Network comprises of routers from which information can easily be stolen by the use of malwares such as "Trojan Horses". Network security mainly focused on the data in the networks and devices which are used to the internet. A synchronous network consists of switches, since they not do buffer any of the data and they do not required to be protected. Digitalization is playing a leading role in everyone's daily life, so secure for network is the main issue to be organized. As prediction goes for the network security field, as some new trends are emanating and based on old trends such as biometric scanning while others are completely new and revolutionary. Social network sites are widely used services of day to day and it is also contain many serious shortfall, some of them do not have system of authenticating the sender as well as the receiver, during transmission as it is stored in multiple places which can be easily snatched and modified. A network contains many impuissant but most of them can be fixed by the following simple techniques. Such as updating the software, configuring network accurately, rules for firewall, by using a good anti-virus software etc. The basic information concerned with network security which would be outlined such searching and ending impuissant, preventing network from attacks and also security measures which are currently being used. Digital India is a crusade sprint by the Indian government to make our country a digitally authorized country. This enterprise was initiated to connect people from the rural areas with high-speed internet networks to blaze any information as per their requirement. Three important segments of digital India are like erection of digital infrastructure. Digital literacy and convey digital services to an all over the country.

II. DIFFERENT TYPE OF SECURITY ATTACKS

A. Passive Attacks

A passive attack is a network attack in which a system is monitored and sometimes scanned for open ports and vulnerabilities. This type of attacks incorporate the attempts to break the system using perceive data. One of the examples is plain text attack, where both the plain text are already well known to the attacker.

Properties of passive attacks are:

- **Interception:** This can be either an active or passive process.The data passing through a network can easily be snuffled and thus attacking the Confidentiality of the user, such as eavesdropping, "Man in the middle" attacks
- **Traffic analysis:** Traffic analysis is the process of interception and examine message in order to deduce information from patterns in communications. This is also a confidentiality attack. It can embrace trace back on a specific network like a CRT radiation.
-

B. Active Attacks

An active attack is a network exploit in which a hacker attempts to make changes to data on the target. In this type of attack the attacker sends data stream to one or both the groups involved or they can also be completely cut off the streams of data. It imputes are as follows:

- **Interruption:** It averts authenticated user form accessing the site. It attacks availability, such as DOS attacks.
- **Modification:** In this the data is altered mostly during the transmission. It's an integrity attacks.
- **Fabrication:** Creating spurious items on a network without genuine authorization.

C. DOS Attacks

Today a DOS attack has become a major threat for network security all over the world. They can easily be launched by any people with the basic knowledge of the network security. In a distributed denial-of-service exploit, large numbers of compromised system attack a single target. They don't require much time and planning as compared to other attacks, in short they are most cheaper and efficient method for network attacking. . They can shutdown the company network by cram-full as of with requests and thus affects network availability. With the help of network tools such as Torino, we can easily download from the internet by this any normal user can initiate an attack. DOS attacks usually works by enervate the targeted network of bandwidth, buffering of TCP connections. Application buffer, service buffer, CPU cycles, etc. DOS attacks uses many users connection to a network known as zombies, most of the time users are heedless of that their computer is infected.

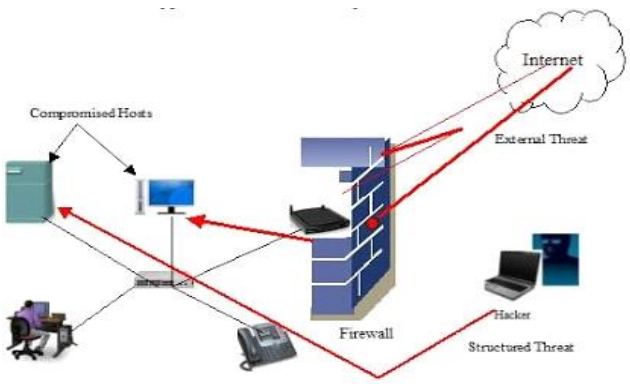
D. Different Types of DOS Attacks

Many attacks are used to accomplish a DOS attack so as to impair service. Some of them are as follows: TCP SYN Flooding which act as whenever a client wants to connect to the server, the client first has to sends to an SYN message to the server. Then the server responds to the client by sending a SYN-ACK message. Later the client consummates the connection by sending an ACK message. These grasp the system resources and the server has to wait till the end of the date. The person utilizing the server will never send the ACK message and will keep on sending a new connection request, until the server is overloaded and thus they cannot dispense access.

ICMP (Ping) Flood: ICMP flood overwhelms the target resources with ICMP echo request (ping) packets, generally sending packets as fast as possible without waiting for replies. This type of attack can consume both outgoing and incoming bandwidth, since the victim's servers will often attempt to respond with ICMP echo reply packets, resulting significant overall system slowdown.

UDP Flood: A UDP flood, by definition, is any DOS attack that floods a target with user datagram protocol (UDP) packets. Now many networks employ TCP and ICMP protocols to avert DOS attacks but a hacker can send large number of packages, so as UDP overloading the victim and averting any new connection.

E. Types of Network Security



III. DEFENCE AGAINST NETWORK ATTACKS

An inherent fragility in the system may be with by design, configuration or may be with implementation which contribute it to a threat. But extent of the vulnerabilities are not because of inoperative design but some may be caused due to sudden disasters both naturally and by human made or some maybe cause by the same persons trying to defend the system. Most of the Vulnerabilities are caused due to poor design, poor configuration, poor implementation, poor management, destitute physical vulnerabilities with hardware and software, information interception and human vulnerabilities. Most of the closely and applying the entire latest reinforcement available from the vendor to their software. However this cannot avert most of the attacks, to avert them each network requires configurations such as:

A. Configuration Management

The main weapon in network attack defence is tight configuration management. It is important for having a dive or slump firewall to avert the system. Anyone can use the remittance login to permit access to the network and as it can put the entire network at risk. All your configuration files in your operating systems or applications should have enough security. The machines inside the core of network must be running the run-up to update the copies of O and all the patches especially the security patches must be installed as soon as they are accessible, configuration files shall not have any known security holes, all the data is backed away in a secure manner, it allows us to allot with nine out of the ten topmost attacks. Several tools are also available which allows patches to range simultaneously and keep things tight.

B. Firewalls

Firewall is a device and/or software that stands between a local network and the internet, and filters traffic that might be harmful. It is the most extensively sold and accessible network security tool convenient in the market. This is the wall which upend between the local network and the internet, which filters the traffic ad averts most of the attacks in the network. There are three divergent types of firewalls be contingent on filtering at the IP level, Packet level, TCP level or application level. Firewalls help in averting unauthorized network traffic through an unsecured network through a private network. They can alert the user when an untrusted application is requisite access to the internet. They also devise a log for all the connections made to the system. These logs can be very damageable in case of any attempt in hacking. If the firewall lay down, it is not able to connect through the network as in a case of DOS attack. Firewall also diminishes the speed of network performance as it investigates both incoming and outgoing traffic. Firewall does not control any sort of internal traffic where most of the attacks arrive. Many companies are under flaw assumptions that by just employing a firewall its safe, but the truth is they are not under safe condition, firewall can be easily be bypassed. The best thing while configuring firewall is to contradict anything which is not allowed.

C. Encryption

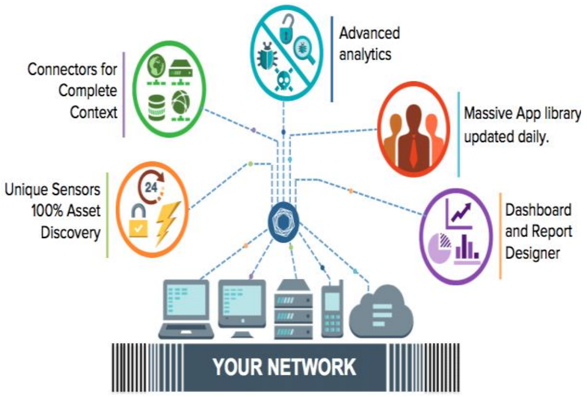
Encryption is another great weapon used in defence against network attacks. Using encryption mechanism one can avert hacker listening to the data because without the equitable key it will be debris to him. Different encryption mechanism such as HTTPS or SHTTP during the data transmission between the client and server, will avert man in the middle attack (MIM), this will also avert any disinter of data and thus any wiretap. Using VPN, which will encrypt all the data going through the network; it will also enhance the privacy of the user. Encryption also has pitfalls as all the encrypted mail and web pages are allowed through firewall they can also embrace malware in them. Encrypting data grasp processing power from the CPU. This in turn diminishes the speed at which data can be sent, as stronger the encryption it takes more time to decrypt.

D. Defence against DOS Attacks

To avert DDoS attack many technologies have been evolved such as intrusion detection systems (IDSs), enhanced routers, firewalls etc.Thesethings which are used between the servers and the internet. They overseer incoming connections plus outgoing connections and which automatically take steps to fortify the network. They have traffic inspection access control and redundancies are built into them. IDSs have been logged into both the incoming and outgoing connections. Later these logs can be compared with the baseline traffic to recognize potential DOS attacks. If there is any unusual lofty traffic on the server it also circumspect possible ongoing DOS attack such as TCP SYN flooding. With the required configuration, the Firewalls can also use as defence against DOS attacks. Firewalls are used to allow or deny certain ports, packets, IP addresses etc. Firewalls can also accomplish real time assessment of the traffic and take the necessary steps to avert the attack. Security measures can also be deployed in routers which can generate another defence line away from the target, so even if a DOS attack arises it won’t affect the internal net Service providers can also escalate the service quality of infrastructure.

E. Vulnerability Testing

A vulnerability testing is any mistakes or weakness in the system security procedures, design and implementation that may result in the violation of system's security policy. To avert any attacks on the network, one must notice any sort of open vulnerability in the network and close them; these might embrace open ports, defectiveness and outdated software with known vulnerabilities, outdated firewall regulations etc. One such method is used for port scanner which can be worn to probe a server and identify any open ports. This is used by many administrators to verify rules, policies of their servers and also can be used by attackers on a network to detect exploits. Some such tools which are obtained for free on the internet are Nmap, Super Scan. These tools are permitted to download by everyone and each comes with a detailed respective tutorial to use them. Different types of port scans are as follows below:


Connectors for Complete Context
Advanced analytics
Massive App library updated daily.
Unique Sensors 100% Asset Discovery
24
Dashboard and Report Designer
YOUR NETWORK

IV. ENCRYPTING THE WORLD WIDE WEB (WWW)

The objectives of privacy, confidentiality and availability our communications on the web should be consistently encrypted this will reduces the number of attacks and averts anyone to view the ongoing transmissions. These can be attained by putting all together for a system of encryption and deploying a system of digital certificates which is used in our digitalization techniques. The most vital way of encryption is the SSL protocol. Network security can also be contrast to human system. The human system can be clasped as analogy, providing a preservation at each point just like a body we can greatly refine the security. Using this mechanism we can extend our resources and avert dependent on one system.

A. Secure Sockets Layer It employs both asymmetric and symmetric keys encryption which transfers data in a secure mode over a consistent network. When SSL is deployed in a browser it initiates a secure connection between the browser application and the server. It's like an encrypted subway in which the data can proceed securely. Anyone listening on the network can't decode the data passing in the subway. It yields integrity using hashing algorithms and confidentiality using encryption. The session is tackled with an asymmetric encryption. The server sends public key to the client. After the asymmetric connection both sides are switched to a symmetric connection. Asymmetric algorithms are slow and accomplish more CPU power than symmetric. While symmetric encryption, CPU load is elevated, servers can only handle a fragment of connections as compared to servers with no encryption.

B. Secure HTTP (SHTTP) It's an substitution to HTTPS, it has the same working principles as HTTPS and is plotted to secure web pages and their messages. There is a differentiation between SHTTP and SSL protocol such as SSL is a connection oriented protocol and it works on the transport level by dispensing a secure subway for transmission whereas SHTTP works on the application level and here we are encrypting each message separately, but secure subway is created. SSL can be employed for secure TCP/IP protocols like FTP but SHTTP works only on HTTP.

C. VPN Virtual Private Network (VPN) is a mechanism to carry traffic on an unsecured network. It employs a combination of encrypting, authentication and subway. VPN empowers a user to secure its privacy, as it's very difficult to detect the location of the user as the network data may be dispelled through multiple locations expand across the world before reaching its final destination.

D. E-Mail Security Both sender and the receiver of the email must be distressed about the diplomatic of the information in the mail; it has been perspective by unauthorized users, being altered in the storage or in the middle. Email can be easily be simulated therefore one must always be authenticate its source.

V. CONCLUSION

As internet has become a herculean part of our daily life, so necessitate of network security has also extended exponentially from the previous decades. As much as the users are connecting to the internet it fascinates a lot of criminals attracts. Now a day's according to the Digital India, each and everything is connected to internet from simple grocery shopping to the defence confidentially, so as an outcome there is herculean need of security to the network. Transaction over Billions of dollars is happening every hour over the internet, at any cost this has to be protected. Even a minute unobserved vulnerability in a network can have devastating effect, if companies records are emanated, it can lay the users data such as their banking details, credit card, debit card information at threat, there are innumerable software's such as intervention in detection which have been averting these attacks, but on most of the occasion it's all because of a human oversight that these attacks transpire. Most of the attacks can be easily be averted, by re tendering many simply methods as outlined in this paper. As new and more complicated attacks prevail, researchers across the world are finding new methods to avert them. Numerous elevations are being mould in the field of network security both in the field of hardware and software, it's like a continual cat and mouse game between network security analyst and hackers/cracker, so per the requirement of internet shows no signs of diminishing it's only going to acquire much harder.

VI. REFERENCES

[1] R.E.Mahan, "Introduction to Computer & Network Security," Washington State University, 2000.

[2] M.Kassim, "An Analysis on bandwidth Utilization and Traffic Pattern." IA CSIT Press, 2011.

[3] M.M.B.W Picoulas J, “Software Agents and Computer Network Security." Napier University, Scotland, UK.

[4] A.R.F. Hamedani, "Network Security Issues, Tools for Testing," School of Information Science, Halmstad University, 2010.

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3| ISSN: 2456-3307

Digital Signal Processing Of Coherent and Generation Detection of QPSK Signal

Mrs. G.Sangeetha Lakshmi[1],Mrs. C.Vinodhini[2]

[1]Assistant Professor, Department of Computer Science and Applications, D.K.M College for Women (Autonomous), Sainathapuram, Vellore, Tamilnadu, India

[2]Research Scholar, Department of Computer Science and Applications, D.K.M College for Women (Autonomous), Sainathapuram, Vellore, Tamilnadu, India

ABSTRACT

We demonstrate an optical quadrature phase-shift keying (QPSK) signal transmitter and an optical receiver for demodulating optical QPSK signal with homodyne detection and digital signal processing (DSP). DSP on the homodyne detection scheme is employed without locking the phase of the local oscillator (LO). In this paper, we present an extracting one-dimensional array of down-sampling method for reducing unwanted samples of constellation diagram measurement. Such a novel scheme embodies the following major advantages over the other conventional optical QPSK signal detection methods. First, this homodyne detection scheme does not need strict requirement on LO in comparison with linear optical sampling, such as having at spectral density and phase over the spectral support of the source under test. Second, the LabVIEW software is directly used for recovering the QPSK signal constellation without employing complex DSP circuit. Third, this scheme is applicable to multilevel modulation formats such as M-ary PSK and quadrature amplitude modulation (QAM) or higher speed signals by making minor changes.

Keywords: Optical Coherence; Quadrature Phase-Shift Keying; Digital Signal Processing

I. INTRODUCTION

There has been a remarkable change in the area of high-speed optical fiber communications in recent years. Traditional binary ON{OFF signalling systems are being replaced with other more sophisticated modulation formats carrying more than one bit per symbol.1 Although mostly on{o_ keying (OOK) formats have been used in commercial applications so far, M-ary phase-shift keying (M-PSK) and M-ary quadrature amplitude modulation (M-QAM) have gained renewed attention to improve the spectral efficiency2 and meet the never-ending increasing demand for bandwidth in optical transmission systems.3 Advanced modulation formats, such as differential quadrature phase-shift keying (DQPSK)4,5 and differential 8-ary phase-shift keying (D8PSK),6 have already been demonstrated. Among various modulation formats that carry 2 bits of information per symbol, quadrature phase-shift keying (QPSK) is the most promising one because of its superior transmission characteristics.7,8 Several experiments have investigated the performance of QPSK systems with optical differential detection, where the receiver contains two sets of Mach{Zehnder interferometers and balanced photodetectors.9 However, although synchronous detection of QPSK signals requires an SNR per bit of about 2 dB lower than differential detection, the synchronous coherent receiver needs either to use a local oscillator (LO) locked to the carrier phase or to recover the carrier phase after homodyne detection.

In addition, the optical phase-locked loop(PLL) is still difficult to achieve because the practical voltage-controlled oscillator (VOL) operating at the optical stage is not available. Coherent optical communication systems widely attract people's attention mainly because of their improved sensitivity over direct detection, their ability to receive complex modulation formats such as n-PSK and QAM as well as the ability to access the full information of the optical field in the electrical domain.9 In Ref. 10, the authors have suggested to utilize a digital implementation of a phased-lockedloop for phase or even frequency actuation of LO laser as early as 1991. In fact, all of the recently published works11{13 focused on coherent receivers with employing digital signal processing (DSP) techniques for transmitter-local oscillator laser synchronization instead of using traditional optical phase-locked techniques. However, the sampling rate of the analog-to-digital convertors (ADCs) on the performance of QPSK signals was rarely studied. Combining advanced modulation and coherent detection has been regarded as an emerging key technology for fulfilling the excepted bandwidth demands of future optical networks. In this paper, to retrieve the constellation diagram of the QPSK signals, we demonstrate an optical receiver that uses homodyne detection and DSP based on the LabVIEW software. In order to overcome the restriction on oversampling in the ADCs embedded within a LeCroy serial data analyzer (SDA; SDA 825Zi-A), we propose a novel extracting one-dimensional array of down-sampling method for reducing unwanted samples of constellation diagram measurement. This enables us to monitor, simultaneously, the amplitude and phase modulations of ultrafast optical QPSK signals.

II. PRINCIPLE

Coherent detection performs the measurement of the electrical field of an optical QPSK signal by the interference with the local oscillation. The electric fields of the QPSK signal source under test and the LO are written as

"D(t) = √ P_D aIQM(t)_ exp(j(!0t + '0 + 'IQM(t))) ; (1)
"LO(t) = √ P_ LO exp(j(!0t + '0)) ; (2)

where PD, PLO aIQM(t) and 'IQM(t) describe the signal source power, LO power, amplitude and phase modulations of the IQM, respectively. In this experiment, we implement a data-encoded optical source by modulation of a continuous wave (CW) laser and choose w0 as the optical frequency of this CW laser. For LO source, the choice of its frequency is the same as the signal source because they come from the same CW laser. As depicted in Fig. 1, the source under test and the LO are coupled into the optical 90° -hybrid. The optical 90°-hybrid contains a pair of interferometers with a relative phase difference of _=2. It outputs orthogonal quadratures (cosine and sine elements) to the I- and Q-arms, respectively. The four outputs of the 90°-hybrid are received employing homodyne detection with two balanced photodetectors. Assuming an identical response R(t) for the four detectors, the output signals of the two balanced photodetectors are

$$\begin{aligned} S_A &= \mathrm{real}\left[\exp[i(\varphi_A)] \cdot \int_{-\infty}^{+\infty} R(t) \cdot \varepsilon_D(t) \cdot \varepsilon_{\mathrm{LO}}^{*}(t) dt\right] \\ &= R(t) a_{\mathrm{IQM}}(t) \sqrt{P_D P_{\mathrm{LO}}} \cos(\varphi_A + \varphi_{\mathrm{IQM}}(t)) \,, \qquad (3) \\ S_B &= \mathrm{imag}\left[\exp[i(\varphi_A)] \cdot \int_{-\infty}^{+\infty} R(t) \cdot \varepsilon_D(t) \cdot \varepsilon_{\mathrm{LO}}^{*}(t) dt\right] \\ &= R(t) a_{\mathrm{IQM}}(t) \sqrt{P_D P_{\mathrm{LO}}} \sin(\varphi_A + \varphi_{\mathrm{IQM}}(t)) \,, \qquad (4) \\ S_N &= \exp[i(\varphi_A)] \cdot \int_{-\infty}^{+\infty} R(t) \cdot \varepsilon_D(t) \cdot \varepsilon_{\mathrm{LO}}^{*}(t) dt \\ &= R(t) a_{\mathrm{IQM}}(t) \sqrt{P_D P_{\mathrm{LO}}} \exp(\varphi_A + \varphi_{\mathrm{IQM}}(t)) \,. \qquad (5) \end{aligned}$$

The electrical signals SA and SB contain information on the amplitude and the phase of the optical QPSK signal. The signals SA and SB are simultaneously sampled once after every symbol period T with ADCs of the oscilloscope. However, for asynchronous sampling, the signal must be sampled at twice the symbol rate and then resampled to keep one sample per symbol. Therefore, the electrical ADCs require a sampling rate of

fs,el = MelRs = 2Rs ; (6)

whereRs is the symbol rate of the received optical signal and Mel = 2 is the desired rational oversampling factor.

III. EXPERIMENTAL SETUP

Our proposed photonic QPSK signal transmitter is indicated in Fig. 1. A CW laser is split into two paths using a coupler. The upper branch is connected to a DPMZM via a polarization controller for optical carrier modulation, and the lower branch is utilized as a local oscillation for homodyne coherent detection. A DPMZM contains two children Mach{Zehnder modulators (MZ-C1 and MZ-C2) nested within a third Mach{Zehnder modulator (MZ-P). There are three independent DC bias voltages and two RF inputs. VRF1 and VRF2 represent the RF modulating electrical voltages of MZM1 and MZM2. Vbias1, Vbias2 and Vbias3 which are controlled by the bias voltage controller represent the DC bias voltages applied to MZ-C1, MZ-C2 and MZ-P, respectively. As illustrated in Fig. 1, the incoming light in the DPMZM is equally split into two arms, i.e. the in-phase (I) and the quadrature (Q) arms. The in-phase and quadrature components of the electric field are modulated independently by two pseudorandom binary sequences of length 231 □ 1, which are generated by a signal quality analyzer (MP 1800A). In both paths, a field amplitude modulator is achieved by operating the MZMs in the push{pull mode at the minimum transmission point. Furthermore, a relative phase shift of _=2 in both paths can be adjusted by the bias voltage controller. The resulting NRZ-QPSK data are imprinted onto light from the CW laser at 1562 nm with a linewidth of 10 kHz according to the manufacturer's specification. The choice of the NRZ formats stems from its experimental simplicity, since no pulse-carving stage is needed at the transmitter. Both the optical signal and the local oscillation are coupled using an optical 90°-hybrid. Then the in-phase and quadrature components of the optical QPSK signal are retrieved with two balanced photodiodes BPDA and BPDB, and converted to the electrical signals I(t)BPDA and I(t)BPDB, respectively. The signals I(t)BPDA (SA) and I(t)BPDB (SB) are simultaneously sampled at a rate of 80 Gsamples/s with

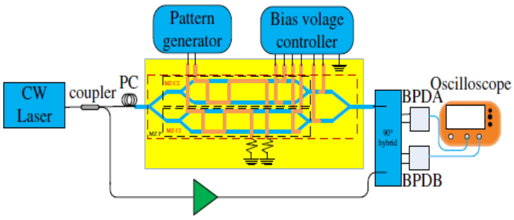


Fig. 1. Schematic of proposed photonic QPSK signal generator and homodyne demodulation (CW Laser: continuous wave laser, MZM: Mach–Zehnder modulator and PC: polarization controller).

Generation and coherent detection of QPSK signal using a novel method of DSP ADCs embedded within a LeCroy SDA oscilloscope. The sampled signals are then processed with the LabVIEW software. For all measurements, the output power of the EDFAs is adjusted to yield a signal power of 0 dBm and an LO power of 6 dBm at the input ports of the optical 90° -hybrids. The states of polarization of signal and LO are matched with manual polarization controllers.

IV. SIGNAL PROCESSING

A block diagram of the used signal-processing algorithm for the configuration of homodyne detection is shown in Fig. 2. The output signals of the BPDA and BPDB are asynchronously acquired and digitized by a LeCroy SDA with two channels (LeCroy SDA 825Zi-A), a sampling rate of 80 Gb/s per channel and 20 GHz band-width. The acquired signals of the two channels of SDA are real and imaginary parts of the complex samples IBPDA(n) + jIBPDB(n), where n denotes the number of samples, and then we calculated and estimated the phase of the acquired samples using these complex samples. To improve the performance of this detection scheme, we establish a phase reference by taking more than one symbol into account, e.g. by averaging the phase over a number of consecutive symbols. The process of the phase estimation is as follows: the reconstructed signal samples E(n) = [IBPDA(n) + jQBPDB(n)] to the four power cancel the phase modulation _(= _=4, 3_=4, 5_=4 and 7_=4), since E41 exp[j(4_)]. The complex amplitudes Calculate phase of the samples: ΦN The phase estimate Φe

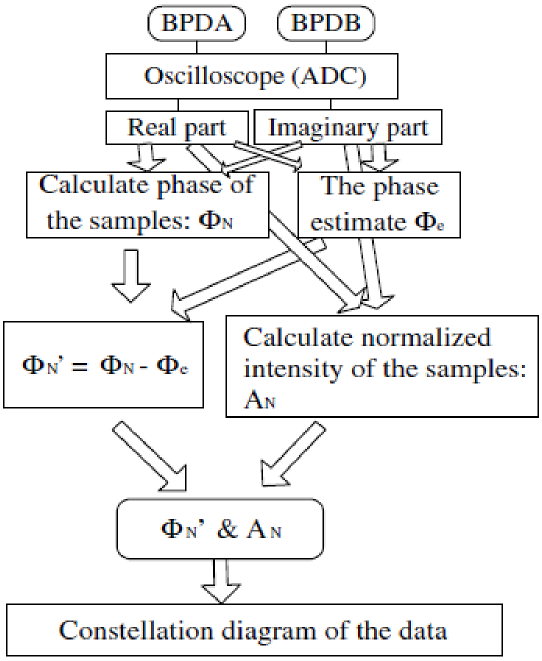

BPDA
BPDB
Oscilloscope (ADC)
Real part
Imaginary part
Calculate phase of the samples: ΦN
The phase estimate Φe
ΦN' = ΦN - Φe
Calculate normalized intensity of the samples: AN
ΦN' & AN
Constellation diagram of the data

Fig. 2. A block diagram of the signal-processing algorithm.

E4 are summed, so that the phase is averaged over the entire block. The phase of the resulting complex amplitude is divided by 4, leading to a phase estimation given as the resulting phase _0 n = [arg(E(i)) □ _] and the normalized intensity An of the nth sample are regarded as the phase and amplitude of reconstructed signal samples, respectively. Unfortunately, the employed SDA does not provide an option to externally clock its ADCs in order to set a desired sampling rate. Therefore, the received samples then need to be post-processed offline in a computer to recover the constellation diagram of the QPSK data signal.

V. RESULTS AND DISCUSSION

We test the described algorithm in Fig. 2 using the measurement setup depicted in Fig. 1. A LiNbO3-based QPSK transmitter is used to generate an optical QPSK signal by using differentially precoded PRBS sequences with a word length of 231□1 at a data rate of either 1 Gb/s or 10 Gb/s. The representation of the samples SN can be made using a constellation diagram in the complex plane, which shows the amplitude and phase of the samples of the QPSK signals at a given position in the bit slot. For asynchronous sampling, the signal must be sampled at twice the symbol rate and then resampled to keep one sample per symbol. Therefore, the electrical ADCs require a sampling rate of

fs,el = MelRs = 2Rs : (7)

When the symbol rate of QPSK is 1 Gsymbols/s, the required electrical ADCs' frequency is fs,el = 2Rs = 2 GHz. Unfortunately, the ADCs embedded within a LeCroy SDA oscilloscope do not provide an option to externally clock them in order to set a desired sampling rate and its default sampling rate is 80 Gsamples/s. The leading edge and trailing edge of NRZ-QPSK signals are sampled since the sampling rate of SDA is 40 times higher than the required electrical ADCs' frequency. The samples for waveform diagrams are displayed in Figs. 3(a) and 3(b). There exist a lot of redundant samples since the sampling rate of ADCs is too high. These samples which can become a source of noise will display in the constellation diagram of QPSK signals and reduce the performance of the QPSK signals. Two different symbol rates of QPSK signals are used in order to validate this problem. The results are displayed in Fig. 3. Figures 3(a) and 3(b) display such waveform diagrams for the samples measured on I- and Q-paths of 1-Gsymbol/s NRZ-QPSK signal generated with an IQMZM. We discover that there exist several over-sampled samples in the waveform diagram of the reconstructed QPSK signal samples. Therefore, we must further down-sample the samples of the reconstructed QPSK signal. The array including the normalized intensity An and the phase _0 n is extracted for every 20 elements to ensure the required electrical ADCs' frequency by down-sampling the samples. Figures 3(c) and 3(d) display the constellation diagram for the samples of 1-Gsymbol/s QPSK signal before and after extraction. Figures 3(e) and 3(f) display the constellation diagrams for the samples of 10-Gsymbol/s QPSK signal before and after extraction. We _nd that the performance of the QPSK signal is enhanced by down-sampling of the samples. The amplitude and phase deviation are 0.131 and 1.718, respectively.

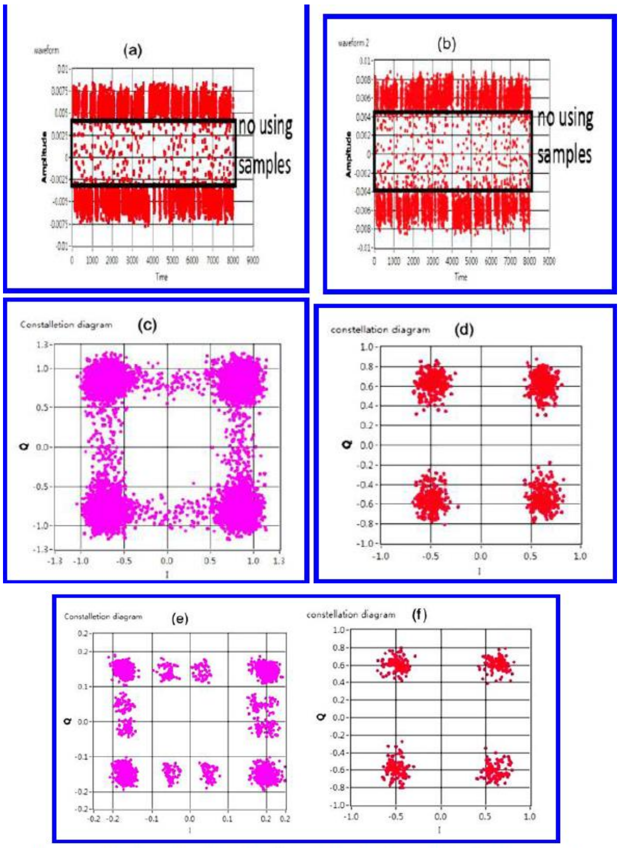


Fig. 3. (a) Waveform diagram for the samples measured on *I*-path of 1-Gsymbol/s NRZ-QPSK signal generated with an IQMZM. (b) Waveform diagram for the samples measured on *Q*-path of 1-Gsymbol/s NRZ-QPSK signal generated with an IQMZM. (c) The constellation diagram for the samples of 1-Gsymbol/s QPSK signal before extraction. (d) The constellation diagram for the samples of 1-Gsymbol/s QPSK signal after extraction. (e) The constellation diagram for the samples of 10-Gsymbol/s QPSK signal before extraction. (f) The constellation diagram for the samples of 10-Gsymbol/s QPSK signal after extraction.

VI. CONCLUSION

We have demonstrated a simple and effective method to directly characterize the amplitude and phase of the QPSK signals in the constellation diagram. The constellation diagrams of the 1-Gsymbol/s and 10-Gsymbol/s QPSK signals are clearly displayed, showing that the performance of the QPSK signal is enhanced by using our DSP method.

VII. REFERENCES

[1] M. Skold, M. Westlund, H. Sunnerud and P. A. Andreson, J. Lightwave Technol. 27 (2009) 3662.

[2] C. Dorrer, C. R. Doerr, I. Kang, R. Ryf, J. Leuthold and P. Winzer, J. Lightwave Technol. 23 (2005) 178.

[3] J. M. Gao, X. X. Xu, Q. J. Chang and Y. K. Su, Chin. Opt. Lett. 7 (2009) 109. 4. P. S. Cho, V. S. Grigoryan, Y. A. Godin, A. Salamon and Y. Achiam, IEEE Photonics Technol. Lett. 15 (2003) 473.

[4] J. Gao, Q. Chang, T. Wang and Y. Su, Chin. Opt. Lett. 6 (2008) 550.

[5] J. Kahn and K. P. Ho, IEEE J. Sel. Top. Quantum Electron. 10 (2004) 259. 8.

[6] D. S. Ly-Gagnon, S. Tsukamoto, K. Katoh and K. Kikuchi, J. Lightwave Technol. 24 (2006) 12.

AUTHOR PROFILE

	Mrs. G.Sangeetha Lakshmi Assistant Professor, Department of Computer Science and Applications,D.K.M College for Women (Autonomous), Vellore, Tamilnadu, India.
	Mrs. C.Vinodhini Research Scholar, Department of Computer Science and Applications,D.K.M College for Women (Autonomous),Vellore, Tamilnadu, India.

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3| ISSN : 2456-3307

Image Watermarking is the Process of Embedding an Imperceptible Data (Watermark) Into Cover Image

Mrs. B. Arulmozhi[1], Mrs. S. Kavitha[2], Mrs. S. Shanthi[3]

[1]Head of the Department (BCA), Dept of Computer Science and Applications, D.K.M College for Women (Autonomous), Vellore, Tamilnadu, India

[2]Research Scholar, Dept of Computer Science and Applications, D.K.M College for Women (Autonomous), Vellore, Tamilnadu, India

[3]Assistant Professor, Dept of Computer Science and Applications, D.K.M College for Women (Autonomous), Vellore, Tamilnadu, India

ABSTRACT

Image watermarking schemes are used to protect the digital images. Image watermarking is the process of embedding an imperceptible data (watermark) into cover image. The image watermarking schemes have been widely used to solve the copyright protection problems of digital image related to illegal usage or distribution. Several image watermarking schemes are proposed, considering different view points. The image Watermarking schemes are classified into different types based on domain of processing, visibility of watermark and rigidity of scheme. Not very many watermarking plans have been proposed for characterizing the copyrights of shading picture. To resolve the copyright protection problem of color image, we propose an effective, robust and imperceptible colorimage watermarking scheme. This plan implants the watermark into cover picture in (Red, Green, Blue) RGB space. The combination of Discrete Wavelet Transformation (DWT) and Singular Value Decomposition (SVD) of Blue channel is used to embed the watermark. The singular values of different subband coefficients of Blue channel are modified using different scaling factors to embed the singular values of the watermark. The copy of the watermark is embedded into four subbandcoefficients which are very difficult to remove or destroy. The combination of DWT and SVD increases the security, robustness and imperceptibility of the scheme.

Keywords: DWT, Digital Water Mark, Perceptivity, Image, DCT-Domain, SVD, Data Embedding

I. INTRODUCTION

Illegal copying, modifying, tampering and copyright protection have become very important issues with the rapid use of internet. Digital Watermarking is the process of hiding or embedding an imperceptible signal(data) into the given signal(data).This imperceptible signal(data) is called watermark or metadata and the given signal(data) is called cover work. This cover work can be an image, audio or a video file. A watermarking algorithm consists of two algorithms, an embedding and an extraction (or detection) algorithm. Watermarking techniques can be broadly classified into two categories: Spatial and Transform domain methods. Spatial domain methods are less complex and not robust against various attacks as no transform is used in them.A singular value decomposition (SVD) is usedas a new transform for watermarking. The SVD based water marking algorithm is introduced by Liu et al. SVD transform was again applied on the resultant matrix for finding the modified singular values.

II. CHOINCE OF WATERMARK-OBJECT

The main inquiry we have to ask with any watermarking or stenographic framework is what shape will the implanted message take? The most straight-forward approach is insert Content strings into a picture, enabling a picture to straightforwardly convey data, for example, creator, title, date… et cetera. The drawback however to this approach is that ASCII text in a way can be considered to be a form of LZW compression, which each letter being represented with a certain pattern of bits. By compressing the watermark-object before insertion, robustness suffers.

Due to the nature of ASCII codes, a single bit error due to an attack can entirely change the meaning of that character, and thus the message. It would be very simple for even a basic undertaking, for example, JPEG pressure to lessen a copyright string to an arbitrary accumulation of characters. Or maybe then characters, for what reason not install the data in an as of now profoundly excess frame, for example, a raster picture? Not exclusively do pictures loan themselves to picture watermarking applications, yet the properties of the HVS can without much of a stretch be misused in acknowledgment of a corrupted watermark.

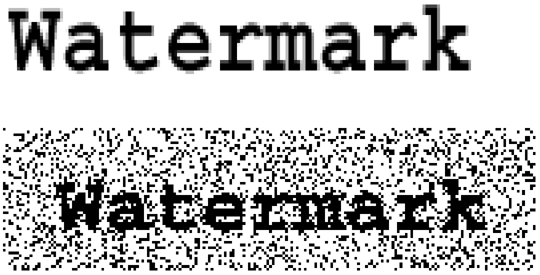


Figure 1. Ideal Watermark-Object vs. Object with 25%Additive Gaussian Noise

Note that despite the high number of errors made in watermark detection, the retrieved watermark is still highly recognizable.

III. OVERVIEW OF MULTIMEDIA DATA HIDING

The ideas of information hiding can be traced back to a few thousand years ago. In many rivalry environments, concealing the existence of communication is desirable to avoid suspicion from adversaries.

The word "steganography", which originated from Greek and is still in use today, literally means "covered writing". Stories of covert communications have been passed for generations, but they were mainly used by military and intelligence agencies. It is until the recent decade that information hiding began receiving wide attention from research community and information technology industry, with hundreds of publications and dozens of patents coming out in the past few years.

- Authentication or Tampering Detection: a set of secondary data is embedded in the multimedia source beforehand, and later is used to determine whether the host media is tampered or not. The robustness against removing the watermark or making it undetectable is not a concern as there is no such incentive from attacker point of view. However, forging a valid authentication watermark in an unauthorized or tampered media source must be prevented.
- Fingerprinting or labeling: the watermark in this application is used to trace the originator or recipients of a particular copy of multimedia source. For example, different watermarks are embedded in different copies of multimedia sources before distributing to a number of recipients. The robustness against obliterating and the ability to convey a non-trivial number of bits are required.

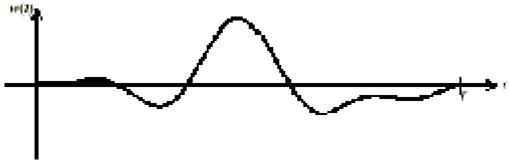
Using Least Significant Bit manipulation, a huge amount of information can be hidden with very little impact to image quality. This technique is performed in the spatial domain. The embedding of the watermark is performed choosing a subset of image pixels and substituting the least significant bit of each of the chosen pixels with watermark bits. The extracted bits do not have to exactly match with the inserted bits.

A correlation measure of both bit vectors can be calculated. If the correlation of extracted bits and inserted bits is above a certain threshold, then the extraction algorithm can decide that the watermark is detected. The implementation of this algorithm is quite simple. However, some policy decisions should be made. For example, how should the set of pixels to be modified be selected? One way to select these elements is by using a pseudorandom number generator also; the watermark extractor should have access to these selected elements.

IV. WAVELET TRANSFORM

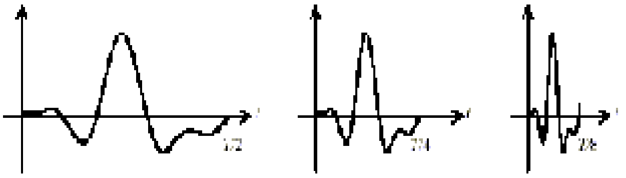
Wavelets are scientific capacities characterized over a limited interim and having a normal estimation of zero that change information into various recurrence parts, speaking to every segment with a determination coordinated to its scale.

The basic idea of the wavelet transform is to represent any arbitrary function as a superposition of a set of such wavelets or basis functions. These preface limits or newborn child wavelets are gotten from a singular model wavelet called the mother wavelet, by growths or withdrawals (scaling) and translations (shifts). They have advantages over traditional Fourier methods in analyzing physical situations where the signal contains discontinuities and sharp spikes. Numerous new wavelet applications, for example, picture pressure, turbulence, human vision, radar, and seismic tremor expectation are produced lately. In wavelet transform the basis functions are wavelets. Wavelets tend to be irregular and symmetric. All wavelet functions, $w(2kt - m)$, are derived from a single mother wavelet, $w(t)$.



Mother wavelet w(t)

Normally it starts at time $t = 0$ and ends at $t = T$. The shifted wavelet $w(t - m)$ starts at $t = m$ and ends at $t = m + T$. The scaled wavelets $w(2kt)$ start at $t = 0$ and end at $t = T/2k$. Their graphs are $w(t)$ compressed by the factor of $2k$ as , when $k = 1$, the wavelet is If $k = 2$ and 3, they are shown in (b) and (c), respectively.



Scaled wavelets

(a)w(2t) (b)w(4t) (c)w(8t)

The wavelets are called orthogonal when their inner products are zero. Wide wavelets are comparable to low-frequency sinusoids and narrow wavelets are comparable to high-frequency sinusoids.

V. SINGULAR VALUE DECOMPOSITON

A. The Watermark Embedding Procedure

To utilize the characteristics of the SVD domain for embedding a watermark, the coefficients of the D and U components were explored . In our observation, two important features of the D and U components were found. In the first feature, the number of non-zero coefficients in the D component could be used to determine the complexity of a block (matrix).Generally, the greater number of non-zero coefficients would indicate greater complexity. For a block-based watermarking scheme, a more complex block was favored for embedding a watermark with perceptibility. IN the second feature, the relationship between the coefficients in the first column of the U component could be preserved when general image processing was performed. Both features were supporting the idea to develop a robust SVD-based watermarking scheme.

In the proposed watermarking scheme , the host image was a gray-level image. The watermark W was a binary image consisting of wxh bits, where W=(w $_1$,w $_2$ ……w $_{nxh}$) and w $_1$ ε (0,1).

The host image was first partitioned into blocks with nxn pixels. And then the blocks were transformed by SVD. The number of non-zero coefficients in the D component of each block was calculated to determine the complexity of this block .A set of greater complexity blocks was selected according to pseudo random number generator (PRNG).And the feature of the D component. Using the PRNG increases the watermarking security. Applying the feature of the D component prevents the smooth blocks from being selected and benefits the perceptibility of the watermarked image.

On each selected block, the relationship between the first column coefficients in the **U** component was examined. This relationship could be taken as the magnitude difference between the neighboring coefficients. The magnitude difference could be either a positive or non-positive value. When the positive difference was computed, a positive relationship would be assigned. Otherwise, a negative relationship would be assigned. The relationship could be preserved when general image processing was performed. In other words, when one coefficient had a larger magnitude than the other, the positive relationship was not easily affected by image processing.

An example shown in Table 1 illustrates the relationship between the **U** component coefficients. Table 1(a) and (b) show the original block and the **JPEG** compressed block, respectively .Both the **SVD** transformed u components of Table 1(a) and (b) are shown in table 1(c) and (d) ,it can be observed that the positive relationships (i.e.), between the coordinates (1,1) and (1,2) were still preserved. As in Table 1(c) and (d) even though the compression processing was performed.

According to the features of the **U** component, it would seem that if a positive relationship is found, bit value of 1 would be hidden. Otherwise, a bit value of zero would be embedded. From that , the coefficients of the **U** component might be modified for embedding a watermark (e.g. positive relationship matching a bit value of 1 or negative relationship matching a bit value of 0), the coefficients are retained . Second, if the magnitude difference does not match the embedding watermark, the coefficients must be modified. However, the modification of **U** component coefficients may alter the original pixel values and degrade the quality of watermarked image. The larger the modification of the **U** component, the more the distortion of image quality and the stronger the robustness. On the other hand, the smaller modification implies that a better image quality and a weaker resistance have been achieved. There is, in other words, a tradeoff between robustness and quality.

To hold the picture quality and give a more grounded strength of a watermarking plan, the coefficient alteration is additionally considered. For each selected greater complexity block, the magnitude difference matched the watermark but was smaller than the predefined magnitude difference threshold, both coefficients had to be further modified. Both coefficients modification not only reduce the image perceptibility but also enhance the robustness to resist attacks. In addition, if the second scenario described in the above accounted , the magnitude difference between two modified coefficients must greater than or equal to the predefined magnitude difference threshold. It means that the gap between two modified coefficients must larger enough to against attacks.

VI. CONCLUSION

The Project presented a DWT- SVD based non-blind watermarking scheme. The SVD is an efficient tool for watermarking in the DWT domain. To embed the watermark into cover image the scaling factor is chosen from a wide range of values for all subbands. The transformed image has the approximation and detailed information. The same watermark is embedded into four subbands which are LH, HL, HH subbands because it is very difficult to remove or destroy. The extraction of logo image from the watermarked image is the opposite of embedding technique. The unbending nature of the proposed conspire is dissected by thinking about different sorts of picture handling assaults. The scheme was found robust to various types of image processing attacks. Finally, the performance validation shows the value of mean square error between the input and embedded image and peak signal to noise ratio factor for justifying the image quality.

VII. REFERENCES

[1] S. Kay and E. Izquierdo, "Robust content based image watermarking," in Proc. Workshop on Image Analysis for Multimedia Interactive Services, WIAMIS' 2001, Tampere, Finland, May 2001.

[2] S. Kak and A. Chatterjee, "On decimal sequences," IEEE Trans. Information Theory, vol. IT—27, pp. 647–652, 1981.

[3] S. Kak, "Encryption and error-correction coding using D sequences," IEEE Trans. Computers, vol. C-34, pp. 803–809, 1985.

[4] N. Mandhani and S. Kak, "Watermarking using decimal sequences," Cryptologia, vol. 29, pp. 50–58, 2005, arXiv: cs.CR/0602003.

[5] R. B. Wolfgang and E. J. Delp, "A watermark for digital images," in Proc. 1996 Int. Conference on Image Processing, Lausanne, Switzerland, Sept. 1996, vol. 3, pp. 219–222.

[6] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," IEEE Trans. Image Processing, vol. 6, pp. 1673–1687, Dec. 1997.

[7] J. J. K. Ó Ruanaidh, W. J. Dowling, and F. M. Boland, "Watermarking digital images for copyright protection," Proceedings of the IEEE, Vision, Image and Signal Processing, vol. 143, pp. 250–256, Aug. 1996.

[8] G. Schuller, "Time-Varying Filter Banks with Low Delay for Audio Coding," in Proc. 105th Conv. Aud. Eng. Soc.,preprint #4809, Sep. 1998.

[9] F. Baumgarte, "Evaluation of a Physiological Ear Model Considering Masking Effects Relevant to Audio Coding," in Proc. 105th Conv. Aud. Eng. Soc., preprint #4789, Sep. 1998.

[10] Y. Huang and T Chiueh, "A New Forward Masking Model and Its Application to Perceptual Audio Coding," in Proc.ICASSP-99, Mar. 1999.

[11] C. Lanciani and R. Schafer, "Subband-Domain Filtering of MPEG Audio Signals," in Proc. ICASSP-99, Mar. 1999.

[12] C. Neubauer and J. Herre, "Digital Watermarking and Its Influence on Audio Quality," in Proc. 105th Conv. Aud. Eng. Soc., preprint #4823, Sep. 1998.

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3| ISSN : 2456-3307

To Augment Protection of data in Cloud Computing using Digital Signature

Ms. B. Kauser Banu[1], Mr. V. Sakthivel[2]

[1]Research Scholar, Dept. of Computer Science, Shanmuga Industries Arts and Science College, Thiruvannamalai, Tamil Nadu, India

[2]Research Supervisor, HOD, Dept. of Computer Application, Shanmuga Industries Arts and Science College, Thiruvannamalai, Tamil Nadu, India

ABSTRACT

Cloud computing is an information technology for thedecade. It allows user to store large amount of data in cloud storage and use as and when required, from any part of the world, via any terminal equipment. Since cloud computing is rest on internet, security issues like privacy, protection of data, confidentiality, and authentication is encountered. In order to get rid of the same, a variety of encryption algorithms and mechanisms are used. Many researchers choose the best they found and use it in different combination to provide security to the data in cloud. On the similar terms, we have chosen to make use of a combination of authentication technique and key exchange algorithm blended with an encryption algorithm. This combination is referred to as "Three way mechanism" because it ensures all the three-protection scheme of authentication, protection of data and verification, at the same time. In this paper, we have proposed to make use of digital signature and Diffie Hellman key exchange blended with (AES) Advanced Encryption Standard encryption algorithm to protect confidentiality of data stored in cloud. Even if the key in transmission is hacked, the facility of Diffie Hellman key exchange render it useless, since key in transit is of no use without user's private key, which is confined only to the legitimate user. This proposed architecture of three-way mechanism makes it tough for hackers to crack the security system, thereby protecting data stored in cloud.

Keywords: Cloud Computing, AES Algorithm, Data Confidentiality

I. INTRODUCTION

Cloud computing simply means Internet computing generally the internet is seen as collection of clouds; thus the word cloud computing can be defined as utilizing the internet to provide technology enabled services to the people and organizations[5]. Cloud computing is new utility of the century, which many enterprises wants to incorporates in order to improve their way of working. It implies sharing of computing resources to handle applications. Cloud computing offers reduced capital expenditure, operational risks, complexity and maintenance, and increased scalability while providing services at different abstraction levels, namely Software-as-a-Service (SaaS), Platform-as-a-Service (PaaS), and Infrastructure-as-a-Service (IaaS)[6]. It is used in consumer-oriented applications such as financial portfoliosdelivering personalized information, or power immersive computer games. It is a pay as peruse kind of service, hence has become very popular in very less time.

Since cloud computing is a utility available on net, so various issues like user privacy, data theft and leakage, eaves dropping, unauthenticated access and various hackers' attacks are raised. These unsolved security issues of authentication, privacy, data protection and data verification are main hindrance for widespread adoption of cloud computing. Hence to get a overwhelmed acceptance to cloud computing in finance, market and industry as well, we have proposed a secure architecture for it. Under the above mentioned title, I am incorporating three security control mechanisms viz authentication, Encryption and data verification technique in to a single stand-alone system. Hence it is a three ways protection scheme wherein digital signature provides authentication, encryption algorithm provides session encryption key and is used to encrypt user data file as well, which is to be saved in cloud and lastly trusted computing to verify integrity of user data.

II. PROBLEM STATEMENT

With cloud computing, organizations can use services and data is stored at any physical location outside their own control. This facility raised the various security questions like privacy, confidentiality, integrity etc. and demanded a trusted computing environment wherein data confidentiality can be maintained. To induce trust in the computing, there is need of a system, which performs authentication, verification and encrypted data transfer, hence maintaining data confidentiality.

Name of Attack	**Description**
Tampering	An attacker may alter information either stored in local files, database or is sent over public network.
Eavesdropping Information Disclosure	This type of attack occurs when attacker gains access in the data path and gains access to monitor and read the messages.
Repudiation	Sender tries to repudiate, or refute the validity of a statement or contract which is sent by him/her.
Elevation of Privileges	An attacker may access unauthorized to information and resources
Man-in-the-Middle Attack	This type of attack occurs when an attacks infiltrates the communication channel in order to monitor the communication and modify the messages for malicious purposes
Replay Attack	A replay attack is defined as when an attacker or originator sends a valid data with intention to use it maliciously or fraudulently.
Identity Spoofing	Identity spoofing occurs when an attacker impersonates the users as the originator of the message in order to gain access on a network.
Differential Analysis Threat	When new versions are released, a differential analysis of the new and old version would indicate where differences in the code exist.
Viruses and Worms	Viruses and worms are very common and well known attacks. These are piece of code that decrease the performance of hardware and application even these malicious codes corrupts files on local file system.

Table 1. Types of Attacks [8]

III. RELATED WORKS

As per Uma Somani, Kanika Lakhani and Manish Mundra [1]: In Cloud computing, we have problem like security of data, files system, backups, network traffic, host security .They have proposed a concept of digital signature with RSA algorithm, to encrypt the data while transferring it over the network. This technique solves the dual problem of authentication and security. The strength of their work is the framework proposed to address security and privacy issue.

Volker Fusenig and Ayush Sharma [2]: states a new approach called cloud networking which adds networking functionalities to cloud computing and enables dynamic and flexible placement of virtual resources crossing provider borders. This allows various kinds of optimization, e.g., reducing latency or network load. This paper presents a security architecture that enables a user of cloud networking to define security requirements and enforce them in the cloud networking infrastructure.

As per Deyan Chen and Hong Zhao [3] from the consumers' perspective, cloud computing security concerns are specially protection of data and privacy protection issues which remain the primary inhibitor for adoption of cloud computing services. They provided a concise but all-round analysis on protection of data and privacy protection issues associated with cloud computing across all stages of data life cycle. Then they proposed to protect data using various scheme and policies like airavat etc. This system can prevent privacy leakage without authorization in Map-Reduce computing process. The weakness is that it just a theory which depends on other scheme and policies for its implementation.

As per EmanM.Mohamed and Hatem S. Abdelkader [8] Cloud computing moves the application software and databases to the large data centers, where the management of the data and services may not be fully trustworthy. This unique feature, however, raises many new securitychallenges. Every cloud provider solves this problem by encrypting the data by using encryption algorithms. Thus their paper investigates the basic problem of cloud computing protection of data. They presented the protection of data model of cloud computing based on the study of the cloud architecture. They implemented software to augment work in a protection of data model for cloud computing. Finally they applied this software in the Amazon EC2 Micro instance for evaluation process.

G. Jai Arul Jose, C. Sajeev, and Dr. C. Suyambulingom [9]: proposed to generate RSA Public keys and Private Keys for public and private access to overcome the problem of protection of data. Certificate Binary file is used inside control node configuration file to make sure cloud data flow securely. The control node sends data through Secure Socket Layer after certificate activation. Finally AES algorithm is used for encryption .This unique combination makes this solution best to prevent different types of attacks. The strength of their work is strong protection of data against various attacks. If a user is attempt to login falsely for many times, the system automatically slowing the service and temporarily stop the account service for the particular user.

IV. PROPOSED SYSTEM

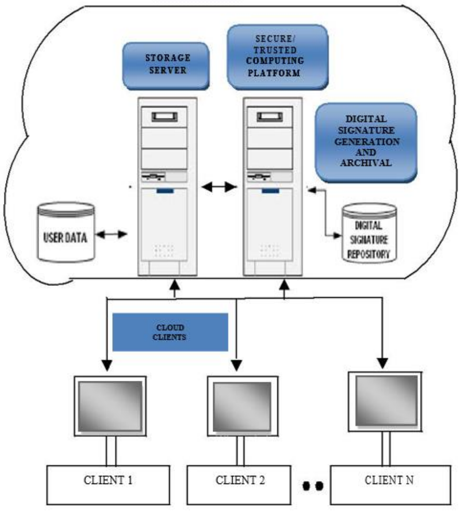

STORAGE SERVER
SECURE/ TRUSTED COMPUTING PLATFORM
DIGITAL SIGNATURE GENERATION AND ARCHIVAL
USER DATA
DIGITAL SIGNATURE REPOSITORY
CLOUD CLIENTS
CLIENT 1
CLIENT 2
CLIENT N

Figure 1. Proposed Architecture

In our proposed architecture, we are using three ways protection scheme. Firstly Diffie Hellman algorithm is used to generate keys for key exchange step. Then digital signature is used for authentication, thereafter AES encryption algorithm is used to encrypt or decrypt user's data file. All this is implemented to provide trusted computing environment in order to avoid data modification at the server end. For the same reason two separate servers are maintained, one for encryption process known as (trusted) computing platform and another known as storage server for storing user data file.

When a user wants to upload a file to the cloud server, first key are exchanged using Diffie Hellman key exchange at the time of login, then the client is authenticated using digital signature. Finally user's data file is encrypted using AES and only then it is uploaded to another (cloud) Storage server. Now when client is in need of same file, it is to be downloaded from cloud server. For that purpose, when user login, first encryption keys are exchanged, file to be downloaded is selected, authentication takes place using digital signature then, AES is used to decrypt the saved file and client can access the file.

A. Execution Steps

1. Sign up
2. Login from TCP
2.1 Key Exchange – Diffie Hellman
2.2 Digital Signature –SHA-I
3. Uploading / Downloading Data Encryption- AES
4. Data is stored / retrieved from Storage server
5. Logout.

B. Hardware specification

The system running the application should have following minimum requirements:

1. Pentium Core.
2. RAM Size 128mb.
3. Processor 1.2GHz.

C. Software Specification

The system running the application must have the following:

1. Supporting OS: Windows XP, VISTA, LINUX: Red Hat, Ubuntu, Fedora.
2. Java Development Kit - jdk1.6.0_02.
3. Java Runtime Environment - jre1.6.0_06.
4. Web Browser like Google chrome with Java Plug-in installed.
5. Wireless connectivity driver.

D. Technology Specific Tools used

In this work we use following technology tools:

1. Java Development Kit - jdk1.6.0_02/ Java Runtime Environment - jre1.6.0_06.
2. Java.awt package for layout of the applet.
3. Java.net package for connection settings and message passing.
4. Netbeans
5. Java Web Start.
6. SOAP.
7. Glassfish Server.

Socket Options interface of methods to get/set socket options.

V. REFERENCES

[1] Yoichiro Ueno, NoriharuMiyaho, Shuichi Suzuki,MuzaiGakuendai, Inzai-shi, Chiba,Kazuo Ichihara, 2010, "Performance Evaluation of a Disaster Recovery System and Practical Network System Applications," Fifth International Conference on Systems and Networks Communications,pp 256-259.

[2] Chi-won Song, Sungmin Park, Dong-wook Kim, Sooyong Kang, 2011, "Parity Cloud Service: A Privacy-Protected Personal Data Recovery Service," International Joint Conference of IEEE TrustCom-11/IEEE ICESS-11/FCST-11.

[3] Y.Ueno, N.Miyaho, and S.Suzuki, 2009, "Disaster Recovery Mechanism using Widely

Distributed Networking and Secure Metadata Handling Technology", Proceedings of the 4th edition of the UPGRADE-CN workshop, pp. 45-48.

[4] Giuseppe Pirr´o, Paolo Trunfio, Domenico Talia, Paolo Missier and Carole Goble, 2010, "ERGOT: A Semantic-based System for Service Discovery in Distributed Infrastructures," 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing.

[5] VijaykumarJavaraiah Brocade Advanced Networks and Telecommunication Systems (ANTS), 2011, "Backup for Cloud and Disaster Recovery for Consumers and SMBs," IEEE 5th International Conference, 2011.

[6] Lili Sun, Jianwei An, Yang Yang, Ming Zeng, 2011, "Recovery Strategies for Service Composition in Dynamic Network," International Conference on Cloud and Service Computing

[7] Xi Zhou, Junshuai Shi, YingxiaoXu, Yinsheng Li and Weiwei Sun, 2008, "A backup restoration algorithm of service composition in MANETs," Communication Technology ICCT 11th IEEE International Conference, pp. 588-591.

[8] M. Armbrust et al, "Above the clouds: A berkeley view of cloud computing,"http://www.eecs.berkeley.edu/Pubs/TechRpts/2009//EEC S-2009-28.pdf.

[9] F.BKashani, C.Chen,C.Shahabi.WSPDS, 2004, "Web Services Peerto Peer Discovery Service ," ICOMP.

[10] EleniPalkopoulouy, Dominic A. Schupke, Thomas Bauscherty, 2011, "Recovery Time Analysis for the Shared Backup Router Resources (SBRR) Architecture", IEEE ICC.

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 3| ISSN : 2456-3307

Robotics and Computer Integrated Manufacturing of Key Areas Using Cloud Computing

Ms. A. Sivasankari[1], Ms. Radhika S[2], Mrs. P. Mohanalakshmi[3]

[1]Head of the Department, Dept of Computer Science and Applications, D.K.M. College for Women (Autonomous), Vellore, Tamilnadu, India

[2]Research Scholar, Dept of Computer Science and Applications, D.K.M. College for Women (Autonomous), Vellore, Tamilnadu, India

[3]Assistant Professor, Dept of Computer Science and Applications, D.K.M. College for Women (Autonomous), Vellore, Tamilnadu, India

ABSTRACT

Cloud computing is changing the way industries and enterprises do their businesses in that dynamically scalable and virtualized resources are provided as a service over the Internet. This model creates a brand new opportunity for enterprises. In this paper, some of the essential features of cloud computing are briefly discussed with regard to the end-users, enterprises that use the cloud as a platform, and cloud providers themselves. Cloud computing is emerging as one of the major enablers for the manufacturing industry; it can transform the traditional manufacturing business model, help it to align product innovation with business strategy, and create intelligent factory networks that encourage effective collaboration. Two types of cloud computing adoptions in the manufacturing sector have been suggested, manufacturing with direct adoption of cloud computing technologies and cloud manufacturing—the manufacturing version of cloud computing. Cloud computing has been in some of key areas of manufacturing such as IT, pay-as-you-go business models, production scaling up and down per demand, and flexibility in deploying and customizing solutions. In cloud manufacturing, distributed resources are encapsulated into cloud services and managed in a centralized way. Clients can use cloud services according to their requirements. Cloud users can request services ranging from product design, manufacturing, testing, management, and all other stages of a product life cycle.

Keywords: Cloud Computing,Cloud Manufacturing,Service-Oriented Business Model

I. INTRODUCTION

Collaboration, Internet of things and cloud has been identified as key business technology trends that will reshape enter-rises worldwide. The manufacturing industry is undergoing a major transformation enabled by IT and related smart technologies. Cloud computing is one of such smart technologies. The main thrust of Cloud computing is to provide on-demand com-putting services with high reliability, scalability and availability in a distributed environment. The National Institute of Standards and Technology (NIST) defined cloud computing as "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."

In Cloud computing, everything is treated as a service (i.e. Xara's), e.g. SaaS (Software as a Service), PaaS (Platform as a Service) and IaaS (Infrastructure as a Service). These services define a layered system structure for cloud computing. At the Infrastructure layer, processing, storage, networks, and other fundamental computing resources are defined as standardized Services over the network. Cloud providers' clients can deploy and run operating systems and software for their underlying infrastructures. More and more businesses are taking advantage of cloud computing, one of which is NEC. Its Cloud-oriented Service Plat-form Solutions play an important role in transforming enterprise systems, contributing to cost reduction, agile deployment of services, expanded flexibility and improved productivity. Cloud computing is also being used in other business and science sectors, e.g. inline commerce, conference origination, and biomedical information sharing. There are valid reasons and perhaps requirement for manufacturing businesses to embrace cloud computing and to "borrow" the concept of cloud computing to give rise to "cloud manufacturing", i.e. the manufacturing version of cloud computing. Such a lateral thinking is considered logical and natural as manufacturing businesses in the new millennium become increasingly IT-reliant, globalized, distributed-ted and Agile demanding. In the first-half of this paper, the essential requirements of a cloud computing system are briefly discussed. These considerations are useful for software architects and developers to design cloud-based applications. They also preface the main focus of this paper, i.e. cloud manufacturing, which forms the second-half of the paper. The rest of this paper is organized as follows. The key requirements of cloud computing systems. Cloud computing in the context of manufacturing businesses. In particular, Section 3.1 discusses utilization of cloud computing in manufacturing businesses and Section 3.2 presents the "manufacturing version" of cloud computing—cloud manufacturing. Section 4 concludes the paper.

II. CLOUD COMPUTING SYSTEMS

This section provides an abridged version of general architectural requirements for cloud computing as presented. Rimalet al. classified architectural requirements into cloud providers, the enterprises that use the cloud, and cloud users.

A. Provider requirements

From the service provider's perspective, highly efficient service architecture to support infrastructure and services is needed in order to provide virtualized and dynamic services. This section explains the requirements of a provider service delivery model and other key requirements.

The middle layer, i.e. PaaS provides abstractions and services for developing, testing, deploying, hosting, and maintaining applications in the integrated development environment. The application layer provides a complete application set of SaaS. The user interface layer at the top enables seamless interaction with all the underlying XaaS layers.

a) Service Delivery Models

Software as a Service, Platform as a Service, and Infrastructure as a Service are three common types of service delivery models. These services are usually delivered through industry standard interfaces, such as Web services, service-oriented architecture (SOA) or Representational State Transfer (REST) services. Software-as-a-Service is sometimes referred to as Application as a Service (AaaS). It offers a multi-tenant platform whereby common resources and a single instance of both the object code of an application and the underlying database are used to support multiple customers simultaneously. To this end, SaaS is also referred to as the Application Service Provider (ASP) model. Examples of the key providers are the Sales force Customer.

As the name implies, Platform-as-a-Service provides developers with a platform including all the systems and environments comprising the life cycle of development, testing, deployment and hosting of sophisticated web applications as a service delivered by a cloud-based platform. Commonly found PaaS includes Facebook F8, Sales forge App Exchange, Google App Engine, Bunzee connect and Amazon EC2. PaaS may offer a number of readily available services, which means that PaaS can support multiple applications on the same platform.

Infrastructure-as-a-Service is sometimes called Hardware as a Service (HaaS). IaaS promotes a usage-based payment scheme, meaning that customers pay as they use. This service is extremely useful for enterprise users as it eliminates the need for investing in building and managing their own IT systems. Another important advantage is the ability of having access to, or using, the latest technology as it emerges. On-demand, self-sustaining or self-healing, multi-tenant, customer segregation are the key requirements of IaaS. Go Grid, Mosso/Rackspace, MSP On-Demand, and masterIT are some of the pioneer IaaS providers.

b) Other Essential Requirements

Other essential requirements are to do with service-centric issues, quality of service, interoperability, fault-tolerance, load balancing and virtualization management.

- **Service-centric issues**

Cloud architecture needs to have a unified service-centric approach. The cloud services should have the ability to dynamically adapt to changes with minimum human assistance. Services need to be self-describing so that they can notify the client exactly how they should be called and what type of data they will return.

- **Quality of Service (QoS)**

Like many services on offer, QoS provides a guarantee of performance, availability, security, reliability and dependability. QoS requirements are associated with service providers and end-users. Service Level Agreements (SLAs) are an effective means for assuring QoS between service providers and end-users. QoS may entail systematic monitoring of resources, storage, network, virtual machine, service migration and fault-tolerance. In the context of a Cloud service provider, QoS should emphasize the performance of virtualization and monitoring tools.

- **Interoperability**

Interoperability is about creation of an agreed-upon frame-work/ontology, open data format or open protocols/APIs that enable easy migration and integration of applications and data between different cloud service providers. It is an essential requirement for both service providers and enterprises. Services with interoperability allow applications to be ported between clouds, or to use multiple cloud infrastructures before business applications are delivered from the cloud.

- **Fault-tolerance**

Fault-tolerance presents the ability of a system to continue to operate in the event of the failure of some of its components. Application-specific, self-healing, and self-diagnosis mechanisms are for example enabling tools for cloud providers to detect failure. Once detected, fault is isolated and revision mode is activated.

- **Load balancing**

Load balancing represents the mechanism of self-regulating the workloads within the cloud's entities (e.g. servers, hard drives, network and IT resources). Load balancing is often used to implement failover in that the service components are monitored continually and when one becomes non-responsive, the load balancer stops sending traffic, de-provisions it and provisions a new service component. A load balancer is another key requirement to build dynamic and stable cloud architecture.

- **Virtualization management**

Virtualization refers to abstraction of logical resources from their underlying physical characteristics in order to improve agility, enhance flexibility and reduce cost . Virtualization in the cloud may concern servers, client/desktop/applications, storage (e.g. Storage Area Network), network, and service/ application infrastructure. Quality of virtualization determines the robustness of a cloud infrastructure. Good virtualization can effectively assist sharing of cloud facilities, managing of complex systems, and isolation of data/application.

B. Enterprise Requirements

Enterprises are being constantly reminded about the services they are paying in terms of the service quality, service levels, privacy matters, compliances, data ownership, and data mobility. This section describes some of the cloud deployment requirements for enterprises.

a) Cloud Deployment For Enterprises

There are four types of cloud deployment models, public, private, community and hybrid clouds. These cloud services are ubiquitous as a single point of access. Different types of deployment models suit different situations. Public cloud realizes the key concept of sharing the services and infrastructure provided by an off-site, third-party service provider in a multi-tenant environment . Private cloud entails sharing services and infra-structure provided by an organization or its specified service provider in a single-tenant environment. Enterprises' mission-critical and core-business applications are often kept in a private cloud. Community cloud is shared by several organizations and is supported by a specific community that has shared interests and concerns . Hybrid cloud consists of multiple internal (private) or external (public) clouds. Added complexity of determining how to distribute applications across both private and public clouds can be challenging. Clearly, enterprises need to strategically leverage all four cloud deployment models.

b) Security

The notion of entrusting data to information systems that are managed by external entities on remote servers "in the cloud" causes varying levels of anxiety. This is because corporate information often contains data of customers, consumers and employees, business know-how and intellectual properties. Popo-Vic΄ and Hocenski discussed security issues and challenges in detail. The above discussed service models (i.e. SaaS, PaaS and IasS) place different levels of security requirements in the Cloud environment. IaaS is the foundation of all cloud services, with PaaS built upon it and SaaS in turn built upon PaaS. Just as capabilities are inherited, so are the information security issues and risks .

c) Business Process Management (BPM)

Typically, a business process management system provides a business structure, security and consistent rules across businessProcesses, users, organization and territory. Some of the examples of BPM applications include customer relationship management (CRM), workforce performance management (WPM), enterprise resource planning (ERP) and e-commerce portals. Cloud-based BPM (e.g. combining SaaS with a BPM application) enhances flexibility, deploy-ability and affordability for complex enterprise applications.

C. User Requirements

Users' requirements are the third key factor for a willing and successful adoption of any cloud system in an enterprise. For users, trust is often a major concern. Trust-based cloud is therefore an essential and must-have feature. This section describes user consumption-based billing and metering requirements, user-centric privacy requirements, service level agreements and user experience requirements.

a) User Consumption-Based Billing and Metering

When it comes to individual end-users and consumption-based billing and metering in a cloud system, an analogy can be drawn with the consumption measurement and allocation of water, gas or electricity on a consumption unit basis. Cost management is important for making planning and controlling decisions. Cost breakdown analysis, tracing the utilized activity, adaptive cost management, transparency of consumption and billings are also important considerations.

b) User-Centric Privacy

In cloud computing, some of the users' data (regarded as his/ her personal intellectual property) are stored at mega-data centers located in the cyber space. In such an environment, privacy becomes a major issue. There is strong resistance and reluctance of an enterprise storing any sensitive data on the cloud. Thankfully, there are various technologies that can enhance data integrity,confidentiality, and security in the clouds, e.g. data compressing and encrypting at the storage level, virtual LANs and network middle-boxes (e.g., firewalls and packet filters).

c) Service Level Agreements (SLAs)

Service level agreements are mutual contracts between providers and users for the assurance of a cloud provider to deliver the services that are agreed-upon. Currently, many cloud providers offer SLAs, but they are rather weak on user compensations on outages. Some of the important architectural issues are measurement of service delivery, method of monitoring performance, and amendment of SLA over time.

d) User Experience (UX)

The notion of UX is to provide an insight into the needs and behaviors of an end-user so as to maximize the usability, desirability and productivity of the applications. UX-driven design and deployment is the next logical step in the evolution of Cloud Computing. Cloud-based application/systems should be easy to use, capable of providing faster and reliable services, easily scalable, and customizable to meet the goal of localization and standardization. Human-Computer Interaction, ergonomics and usability engineering are some of the key technologies that can be used for designing UX-based Cloud applications.

III. ACLOUD COMPUTING IN THE CONTEXT OF MANUFACTURING

In recent years, the philosophy of "Design Anywhere, Manu-facture Anywhere (DAMA)" has emerged [17–19]. The DAMAapproach demands the ability to exchange design and manufacturing data across multiple sites. DAMA also helps establish links between manufacturing resource planning, enterprise resource planning, engineering resource planning and customer relation-ship management. It is believed that cloud computing may play a critical role in the realization of DAMA. In general, there are two types of cloud computing adoptions in the manufacturing sector, manufacturing with direct adoption of some cloud computing technologies and cloud manufacturing—the manufacturing ver-sion of cloud computing.

a) Smart manufacturing with cloud computing

Cloud computing is rapidly moving from early adopters to mainstream organizations. It has become one of the top priorities of many CIOs in terms of strategic business considerations. Some manufacturing industry starts reaping the benefits of cloud adoption today, moving into an era of smart manufacturing with the new agile, scalable and efficient business practices, replacing traditional manufacturing business models.

In terms of cloud computing adoption in the manufacturing sector, the key areas are around IT and new business models that the cloud computing can readily support, such as pay-as-you-go, the convenience of scaling up and down per demand, and flexibility in deploying and customizing solutions.

The adoption is typically centered on the BPM applications such as HR, CRM, and ERP functions with Sales force and Model Metrics being two of the popular PaaS providers (refer to Section (b)).

The cost benefit of adopting clouds in a typical manufacturing enterprise can be multiple. The savings obtained from the elimination of some of the functions that were essential in traditional IT can be significant. With cloud-based solutions, some application customizations and tweaks that the company needs at the process level may be dealt with by the company's IT sector along with some of the smart cloud computing technologies. When adifferent way of executing a process is initiated, the IT staff can make the change happen seamlessly and in less time . Elkay Manufacturing Company, a world leader in stainlesssteel sinks, water coolers and kitchen cabinets, is one of the manufacturing companies that have successfully adopted and benefited from cloud computing technologies .

When it comes to supporting smart business processes, cloud computing can be effective in offering Business-to-business (B2B) solutions for commerce transactions between businesses, such as between a manufacturer and a wholesaler, or between a wholesaler and a retailer. Cloud-based solutions enable better integrated and more efficient processes.

Cloud computing can also be used to enhance many other aspects of manufacturing businesses by moving a traditional process to the cloud for improved operational efficiency. For example, cloud computing can assist the development of an application for customer on-boarding process that is more efficient than the traditional process of company on-boarding customers. The procedure for a company to on-board customers may involve a salesperson visiting a prospective customer, the customer filling in a form, company credit checking etc. A Cloud-based customer on-boarding process may do all of these automatically via cloud resources on the Internet.

Collaboration at scale using cloud technology is an emerging business trend according to McKinsey. Adopting cloud technologies, enterprise collaboration can happen at a much broader scale. Within the organization, demand planning and supply chain organization can be tied into a cloud-based system, allowing different parts of the organization to take a peek into the opportunities that their sales teams are working on. In a more traditional environment, that would involve a few sit-down meetings, several face-to-face discussions, or phone conversations.

The cloud in this case provides a collaborative environment that can give people agility, more transparency, and empowerment through more effective collaborations.

Typically, there are some parts of the manufacturing firm that can quickly and easily adopt cloud-based solutions, whereas other areas are better to remain traditional. Hence, what a cloud-adopting manufacturing enterprise also requires is a smart mechanism to deal with integration. Solutions such as Cast Iron are addressing some aspects of such integration; vendors such as Model Metrics are pitching in as well.

B. Cloud manufacturing

With cloud manufacturing, what comes into one's mind first is the existing networked manufacturing concept, or sometimes called Internet-based manufacturing or distributed manufacturing. However, today's networked manufacturing mainly refers to integration of distributed resources for undertaking a single manufacturing task. What is lacking in this type of manufacturing regime are the centralized operation management of the services, choice of different operation modes and embedded access of manufacturing equipment and resources, without which a seam-less, stable and high quality transaction of manufacturing resource services cannot be guaranteed. In a typical distributed manufacturing environment, the resource service provider and resource service demander have little coordination. Thus, adoption of the networked manufacturing concept has been slow and less effective.

In Cloud manufacturing, distributed resources are encapsulated into cloud services and managed in a centralized way. Clients can use the cloud services according to their requirements. Cloud users can request services ranging from product design, manufacturing, testing, management and all other stages of a product life cycle. A cloud manufacturing service platform performs search, intelligent mapping, recommendation and execution of a service. Fig. 2 illustrates a cloud manufacturing system framework, whichconsists of four layers, manufacturing resource layer, virtual service layer, global service layer and application layer.

a) Manufacturing resource layer

The manufacturing resource layer encompasses the resources that are required during the product development life cycle. These manufacturing resources may take two forms, manufacturing physical resources and manufacturing capabilities. Manufacturing physical resources can exist in thehardware or software form. The former includes equipment, computers, servers, rawmaterials, etc. The latter includes for example simulation soft-ware, analysis tools, "know-hows", data, standards, employees, etc. Manufacturing capabilities are intangible and dynamic recourses representing the capability of an organization under-taking a particular task with competence. These may include product design capability, simulation capability, experimentation, production capability, management capability, and maintenance capability. The types of service delivery models that may exist at this layer are IaaSs and SaaSs.

b) Manufacturing virtual service layer

The key functions of this layer are to (a) identify manufacturing resources, (b) virtualized them, and (c) package them as cloud manufacturing services. Comparing with a typical cloud computing environment, it is much more challenging to realize these functions for a cloud manufacturing application.

A number of technologies can be used for identifying (or tagging) manufacturing resources, e.g. RFID, computational RFID, wireless sensor networks (WSN), Internet of things, Cyber Physical Systems, GPS, sensor data classification, clustering and analysis, and adapter technologies.

Manufacturing resource virtualization refers to abstraction of logical resources from their underlying physical resources. Quality ofvirtualization determines the robustness of a cloud infrastructure. Different manufacturing resources are virtualized in different ways. Computational resources and manufacturing knowledge can be virtualized in similar ways as are the general Cloud computing resources. Manufacturing hardware is usually mapped to become virtual machines that are system-independent. Virtualization man-agers (e.g. Virtual Machine Monitor and Virtual Machine Manager (VMM)) are responsible for communicating with the lower level devices, and coordinating and allocating virtual machines.

Agent can be an effective tool for virtualization. Take MTConnect as an example. MTConnect is a standard based on an open protocol for data integration. Although it is for enhancing data acquisition capabilities of machine tools, the use of agent technology provides a plug-and-play environment for manufacturing facilities, which has the potential to support cloud manufacturing. Fig. 3 shows a schematic of a factory system with three machine tools that are virtualized and integrated via MTConnect agents. It needs to be pointed out though that MTConnect mainly supports monitoring processes.

The next step is to package the virtualized manufacturing resources to become cloud manufacturing services. To do this, resource description protocols and service description languages can be used. The latter may include different kinds of ontology languages,In a STEP-enabled, networked manufacturing process planning environment, a STEP resource locator (STRL for short) represents a simplest form of cloud manufacturing service. STRL consists of a URL, an Action and a Query (Fig. 4). STRL is similar to the concept of giving a URL address through the Web. The URL gives the name (location) of the system and data identification. It can therefore be used as a link to a particular manufacturing resource, e.g. a data file, working step (machining step), program, etc. For example, if an STRL as shown in Fig. 4 is activated, we will know that the client has requested the job (as described by file manifold.238) to be machined (run) from the first line until the end for all the three working steps.

To describe manufacturing capability, Zhang et al. used a four-dimension array: (Task, Resource, Participator, Knowledge). Task denotes a manufacturing job; Resource denotes the manufacturing resources that are needed to do the task; Participator represents human resources needed for the job; and Knowledge represents all the knowledge required to do the job.

c) Encapsulating Manufacturing Resources With Mapping

The one-to-many mapping concerns with a single resource that appears to a client as a multiple resource. The client interfaces with the virtualized resources as though he/she is the only consumer. In fact, the client is sharing the resource with other users. For example, ANSYS software can provide structure analysis, thermal analysis, magnetic analysis, and computational fluid dynamics analysis. Therefore, ANSYS software can be encapsulated by many different services.

d) Enterprise requirements—Global Service Layer

The Global Service Layer relies on a suite of cloud deployment technologies (i.e. PaaS). Internet of things has advanced to a new level with RFID, intelligent sensors, and Nano-technology as the supporting technologies. Interconnections between physical devices or products are made easier because of Internet of things. Having said this, a centralized and effective management regime needs to be in place to provide manufacturing enterprises with agile and dynamic cloud services. Based on the nature of the provided cloud resources and the user's specific requirements, two types of cloud manufacturing operation modes can take place at the Global Service Layer, complete service mode and partial service mode.

In a complete service mode, the Global Service Layer takes full responsibility of the entire cloud operational activities. The type of cloud service that suits this mode is virtualized computing resources, e.g. CPU, RAM, and network. These cloud services can be dynamically monitored, managed and load-balanced with ease. Application software is also suitable for the complete service mode in that running and execution of software can take place in a distributed computing environment taking advantage of grid computing and parallel computing. Knowledge, human resources, and manufacturing capabilities may also be managed at the Global Service Layer in a complete service mode.

It is possible and sometimes necessary to partially hand over an activity to the cloud manufacturing service a partial service mode. In such a mode, the service provider provides additional input and operational activities. Typically, manufacturing hardware (e.g. machine tools and experiment devices) is this type of cloud services. The Global Service Layer is mainly responsible for locating, allocating, fee-calculating and remote monitoring the manufacturing resources. The hardware providers are still responsible for executing the manufacturing tasks and ensuring the quality of the manufacturing job.

In order to meet the above enterprise requirements, some critical technologies are needed. For example, optimal resource selection and allocation methods are needed to guarantee an effective cloud manufacturing service. Theories such as Intuitionistic Fuzzy Set, Partial Swarm Optimization, and Quantum Multi-agent Evolutionary Algorithm can be handy when developing an enabling technology. Evaluation and management of QoS is another important exercise at this layer.

IV. RESEARCH CONTRIBUTIONS TO THE CONCEPT OF CLOUD MANUFACTURING

Although the concept of cloud manufacturing is new, virtual enterprise and distributed manufacturing concepts have been around for a while and some of the proposed systems and frame-works bear visible traces of cloud manufacturing or make contributions to a cloud manufacturing system. This section discusses some of these research outcomes.

A. Service-Oriented Manufacturing Environment

Brecher, et al. recognized that applications in an information-intensive manufacturing environment can be organized in a service-oriented manner. They proposed a module-based, configurable platform for interoperable CAD-CAM-CNC planning. The goal is to combat the problems of software inhomogeneity along the CAD-CAM-NC chain. The approach is called open Computer-Based Manufacturing (openCBM) in support of co-operative process planning. To implement the architecture and integrate inspection tasks into a sequence of machining operations, STEP standard is utilized to preserve the results of manufacturing processes that are fed back to the process planning stage . The openCBM platform is organized through a service-orientarchitecture providing the abstractions and tools to model the information and connect the models . It is much like the Platform as a Service concept and resembles an Application Layer, where applications are not realized as monolithic programs, but as a set of services that are loosely connected to each other, guaranteeing the modularity and reusability of a system. The module providers as shown in the figure form the Manufacturing Virtual Service Layer and the module database forms a Global Service Layer.

B. SaaS for Engineering Simulations

To achieve a run-time configuration integration environment for engineering simulations, van der Velde reported a plug-and-play framework for the construction of modular simulation software. In this framework the user (at the Application Layer as in a cloud manufacturing system) is allowed to select a target of simulation and assign the performer of the simulation called "component" before running the selected components. These components are effectively software entities (or otherwise known as SaaS as in cloud computing/manufacturing). They are modulated, self-contained, mobile and pluggable. After the simulation, the output is post-processed through the components. In such architecture, software modules are detected, loaded and used at run-time with the framework (i.e. the Global Service Layer) needing no prior knowledge of the type and availability of components, thus providing true plug-and-play capabilities.

V. CONCLUSION

Cloud computing is changing the way industries and enterprises do their businesses. With wider cloud adoption, access to business-critical data and analytics will not just help enterprises Stay ahead, it will also be crucial to their existence. There are three architectural features of cloud computing in terms of the requirements of end-users, enterprises that use the cloud as a platform, and cloud providers themselves. These architectural features play a major role in the adoption of the cloud computing paradigm as a mainstream commodity in the enterprise world.

Cloud computing is emerging as one of the major enablers for the manufacturing industry, transforming its business models, helping it align product innovation with business strategy, and creating intelligent factory networks that encourage effective collaboration. This pay-by-use scenario will revolutionize manufacturing in the same way that the Internet has already revolutionized our everyday and business lives. Manufacturing shops are starting to take advantage of cloud computing because it simply makes good economic sense. Two types of cloud computing adoptions in the manufacturing sector have been suggested, manufacturing with direct adoption of cloud computing technologies and cloud manufacturing—the manufacturing version of cloud computing.

In terms of direct adoption of cloud computing in the manufacturing sector, the key areas are around IT and new business models, e.g. pay-as-you-go, production scaling up and down per demand, and flexibility in deploying and customizing solutions.
The HR, CRM, and ERP functions may benefit from using some emerging PaaS. Cloud computing can be effective in offering Business-to-Business solutions for commerce transactions between businesses, such as between a manufacturer and a wholesaler, or between a wholesaler and a retailer.

Moving from production-oriented manufacturing to service-oriented manufacturing, cloud manufacturing seems to offer an attractive and natural solution. In cloud manufacturing, distributed-ted resources are encapsulated into cloud services and managed in a centralized way. Clients can use cloud services according to their requirements. Cloud users can request services ranging from product design, manufacturing, testing, management and all other stages of a product life cycle. The cloud manufacturing service platform performs search, mapping, recommendation, and execution of a service. Two main types of manufacturing resources can be considered at the manufacturing resource layer, manufacturing physical resources, and manufacturing capabilities.
Although the concept of cloud manufacturing is relatively new, virtual enterprise, and distributed manufacturing concepts have been around for a while and some of the proposed systems and frameworks bear the trace of cloud manufacturing, e.g. development of a service-oriented manufacturing environment anddifferent SaaS for engineering applications. There are also some embryonic cloud manufacturing systems developed in the past 2–3 years. In response to concerns about cloud computing adoption for manufacturing businesses, enterprises need to address them in constructive and positive ways. The IT professionals as well as personnel from other manufacturing departments need to work hand in hand to look for solutions to the problems. It can be anticipated that cloud manufacturing will provide effective solutions to the manufacturing industry that is becoming increasingly globalized and distributed. Cloud manufacturing means a new way of conducting manufacturing businesses, that is everything is perceived as a service, be it a service you request or a service you provide.

VI. REFERENCES

[1] Bughin J, Chui M, Clouds Manyika J. Big data, and smart assets: ten tech-enabled business trends to watch. McKinsey Quarterly. McKinsey Global Institute; August 2010.

[2] Mell P, Grance T. Perspectives on cloud computing and standards. National Institute of Standards and Technology (NIST). Information Technology Laboratory; 2009.

[3] Pallis G. Cloud computing: the new frontier of internet computing. IEEE Internet Computing 2010. [14:5: 5562494:70-73].

[4] Foster I, Zhao Y, Raicu I, Lu S Cloud computing cloud computing and grid computing 360 degree compared. In: grid computing environments work-shop; 2008.

[5] Kunio T. NEC cloud computing system. NEC Technical Journal 2010;5(2): 10–5.

[6] Dikaiakos MD, Katsaros D, Mehra P, Pallis G, Vakali A. Cloud computing: distributed internet computing for IT and scientific research. IEEE Internet Computing 2009;13(5):10–1.

[7] Ryan MD. Viewpoint cloud computing privacy concerns on our doorstep. Communications of the ACM 2011;54(1):36–8.

[8] Rosenthal A, Mork P, Li MH, Stanford J, Koester D, Reynolds P. Cloud computing: a new business paradigm for biomedical information sharing. Journal of Biomedical Informatics 2010;43:342–53.

[9] Rimal BP, Jukan A, Katsaros D, Goeleven Y. Architectural requirements for cloud computing systems: an enterprise cloud approach. Journal of Grid Computing 2011;9(1):3–26.

[10] The Open Group, Service Oriented Architecture (SOA). Available online at: /http://www.opengroup. org/projects/soa/S; 2010 [accessed on May 2011].

[11] Fielding RT The REpresentational State Transfer (REST). PhD dissertation. Irvine: Department of Information and Computer Science, University of

[12] California. Available online at: /http://www.ics.uci.edu/ fielding/pubs/dissertation/top.htmS; 2000.

[13] Golden B. Virtualization for dummies. Inc.Wiley Publishing; 2008.

[14] Wu L, Yang C. A solution of manufacturing resources sharing in cloud computing environment. Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics). In: Luo Y, editor. 6240 LNCS. Berlin Heidelberg: Springer-Verlag; 2010. p. 247–52.

[15] Popovic´ K, Hocenski Z. Cloud computing security issues and challenges. In: MIPRO 2010—33rd International convention on information and communication technology, electronics and microelectronics; 2010. p. 344–9 [art. no. 5533317].

[16] Subashini S, Kavitha V. A survey on security issues in service delivery models of cloud computing. Journal of Network and Computer Applications 2011;34(1): 1–11.

[17] Cavoukian A. Privacy in the clouds—a white paper on privacy and digital identity: implications for the Internet. Information and Privacy Commission of Ontario 2008.

[18] Heinrichs W. Do it anywhere. IEE Electronics Systems and Software 2005;3(4):30–3.

[19] Venkatesh S, Odendahl D, Xu X, Michaloski J, Proctor F, Kramer T. Validating portability of STEP-NC tool center programming. In: Proceedings of the 2005 ASME International design engineering technical conferences and computers and information in engineering conference; September 24–28 2005. p. 285–90.

[20] Manenti P. Building the global cars of the future. Managing Automation 2011;26(1):8–14.

[21] Shalini S. Smart manufacturing with cloud computing. Sramana Mitra: /http://www.sramanamitra.comS [accessed May 2011].

[22] Li B-H, Zhang L, Wang S-L, Tao F, Cao J-W, Jiang X-D, Song X, Chai X-D. Cloud manufacturing: a new service-oriented networked manufacturing model. Computer Integrated Manufacturing Systems CIMS 2010;16(1):1–7.

[23] Tao F, Hu YF, Zhang L. Theory and practice: optimal resource service allocation in manufacturing grid. Beijing: China Machine Press; 2010.

[24] Huang GQ, Zhang YF, Jiang PY. RFID-based wireless manufacturing for walking-worker assembly islands with fixed-position layouts. International Journal of Robotics and Computer Integrated Manufacture 2007;23(4):469–77.

[25] Huang GQ, Zhang YF, Jiang PY. RFID-based wireless manufacturing for real-time management of job shop WIP inventories. International Journal of Advanced Manufacturing Technology 2007;7–8(36):752–64.

[26] Huang GQ, Wright PK, Newman ST. Wireless manufacturing: a literature review, recent developments and case studies. International Journal of Computer Integrated Manufacturing 2009;22(7):1–16.

[27] Michaloski J, Lee B, Proctor F, Venkatesh S, Ly S. Quantifying the performance of MTConnect in a distributed manufacturing environment, 2010. In: Proceedings of the ASME International design engineering technical confer-ences and computers and information in engineering conference, vol. 2 (PART A); 2009. p. 533–9.

[28] Vijayaraghavan A, Sobel W, Fox A, Warndorf P, Dornfeld DA. Improving machine tool interoperability using standardized interface protocols: MTConnect. In: Proceedings of the International symposium on flexible automation; June 23–26 2008.

[29] Vijayaraghavan A, Huet L, Dornfeld DA, Sobel W, Blomquist B, Conley M. Addressing process planning and verification issues with MTConnect. Trans-actions of the North American Manufacturing Research Institution of SME 2009;37:557–64.

[30] Campos JG, Mı´guez LR. Manufacturing traceability data management in the supply chain. International Journal of Information Technology and Manage-ment 2009;8(3):321–39.

[31] Hardwick M, Loffredo D. Lessons learned implementing STEP-NC AP-238. International Journal of Computer Integrated Manufacturing 2006;19(6):523–32.

[32] Xu X, Wang H, Mao J, Newman ST, Kramer TR, Proctor FM, Michaloski JL, STEP—Compliant NC. Research: the search for intelligent CAD/CAPP/CAM/ CNC integration. International Journal of Production Research. 2005;43(17): 3703–43.

[33] Xu X, Nee AYC, editors. Advanced design and manufacturing based on STEP. Springer Verlag; 2010. ISBN: 978-1-84882-738-7.

[34] Tao F, Hu Y, Zhou Z. Application and modeling of resource service trust-QoS evaluation in manufacturing grid system. International Journal of Production Research 2009;47(6):1521–50.

[35] Zhang L, Luo Y-L, Tao F, Ren L, Guo H. Key technologies for the construction of manufacturing cloud. Computer Integrated Manufacturing Systems, CIMS 2010;16(11):2510–20.

[36] Guo H, Zhang L, Tao F, Ren L, Luo YL. Research on the measurement method of flexibility of resource service composition in cloud manufacturing. In: Proceedings of the International conference on manufacturing engineering and automation ICMEA; December 10–12 2010.

[37] Louth W. Metering the cloud: applying activity based costing (ABC) from code profiling up to performance and cost management of cloud computing. In: Proceedings of the International conference on JAVA technology; 2009.

[38] ISO 10303-1:1994. Industrial automation systems and integration—product data representation and exchange—part 1: overview and fundamental princi-ples. International Organization for Standardization, Geneva 20, Switzerland.

[39] ISO 10303–203:1994. Industrial automation systems and integration— configuration controlled 3D designs of mechanical parts and assemblies. International Organization for Standardization, Geneva 20, Switzerland.

[40] ISO 10303–238:2007. Industrial automation systems and integration— product data representation and exchange—part 238: application protocol: application interpreted model for computerized numerical controllers. Inter-national Organization for Standardization, Geneva 20, Switzerland.

[41] ISO 14649-11:2004. Industrial automation systems and integration— physical device control—data model for computerized numerical controllers— part 11: process data for milling. International Organization for Standardization, Geneva 20, Switzerland.

[42] Xu X, He Q. Striving for a total integration of CAD, CAPP, CAM, and CNC. Robotics and Computer-Integrated Manufacturing 2004;20(2):101–9.

[43] Brecher C, Lohse W, Vitr M. Module-based platform for seamless interoper-able CAD-CAM-CNC planning. In: XU XW, NEE AYC, editors. Advanced design and manufacturing based on STEP. London: Springer; 2009.

[44] Brecher C, Vitr M, Wolf J. Closed-loop CAPP/CAM/CNC process chain based on STEP and STEP-NC inspection tasks. International Journal of Computer Integrated Manufacturing 2006;19:570–80.

[45] Van de Velde PJMC. Runtime configurable systems for computational fluid dynamics simulations. PhD thesis. Auckland: Department of Mechanical Engineering, University of Auckland; 2009.

[46] Nassehi A, Newman ST, Xu XW, Rosso JR. RSU. Toward interoperable CNC manufacturing. Computer Integrated Manufacturing 2008;21:222–30.

[47] Newman ST, Nassehi A. Universal manufacturing platform for CNC machin-ing. Annals of the CIRP 2007;56:459.

[48] Mokhtar A, Houshmand M. Introducing a roadmap to implement the universal manufacturing platform using axiomatic design theory. Interna-tional Journal of Manufacturing Research 2010;5:252–69.

[49] Suh SH, Shin SJ, Yoon JS, Um JM. UbiDM: a new paradigm for product design and manufacturing via ubiquitous computing technology. International Journal of Computer Integrated Manufacturing 2008;21(5):540–9.

[50] Lee BE, Suh S-H. An architecture for ubiquitous product life cycle support system and its extension to machine tools with product data model. Interna-tional Journal of Advanced Manufacturing Technology 2009;42:606–20. Wang X, Xu X. DIMP: an interoperable solution for software integration and product data exchange. Enterprise information systems. TEIS-2010-0110, in press. doi:10.1080/17517575.2011.587544.

[51] Wang XV, Xu X, Haemmerle E. Distributed interoperable manufacturing platform based on STEP-NC. In: Proceedings of the 20th International flexible automation and intelligent manufacturing conference, FAIM; 2010.

IJSR CSEIT

www.ijsrcseit.com

A State Level Symposium & IT Meet - InnovIT'18

Organised By:

PG and Reaseach Department of Computer science,
Shanmuga Industries Arts and Science College,
Tiruvannamalai, Tamil Nadu, India

Published by :
TechnoScience Academy
www.technoscienceacademy.com

SHANMUGA INDUSTRIES ARTS AND SCIENCE COLLEGE (Co-Ed.,)
Certified under Section 2(f) & 12B of the UGC Act 1956
An ISO 9001: 2000 Certificated Institution
Permanently Affiliated to Thiruvalluvar University,
Vellore and Approved by the Government of Tamil Nadu & AICTE

Email: editor@ijsrcseit.com, ijsrcseit@gmail.com