International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2017 IJSRCSEIT | Volume 2 | Issue 2 | ISSN : 2456-3307

Analysis of Data Anonymization techniques for securing sensitive data in IoT Environment

Keerthiga*, Rajmohan, Narmadha

Computer Science and Engineering, Anna University/IFET College of Engineering, Villupuram, Tamil Nadu, India

ABSTRACT

The internet of thing is a network of internet-connected objects able to collect and exchange data using embedded sensors. Security and privacy in IoT is one of the most challenging areas. Sharing data in form of text is a wide range of activities but it raises a concern about privacy when sharing data that could be sensitive. In this work, we present an analysis of various data anonymization techniques for providing individual node centric privacy in IoT environments. Data anonymization is about providing the privacy of sensitive information in the data against a variety of attacks.

Keywords: Data Anonymization, Privacy, K-Anonymity, Generalization

I. INTRODUCTION

The misuse of personal data has increased in a slew of data privacy protection regulations by various governments across countries. The increasing trend of outsourcing software application development and testing has also increased the risk of misuse of sensitive data. Data anonymization is the process of de-identifying sensitive data while preserving its format and data type [1]. This process is important for sharing data without exposing to third parties any sensitive information contained in databases. Information disclosure is one by which we can reach to one's identity. By using Anonymization techniques, dataset privacy is preserved from various disclosures such as loss of information, membership disclosures etc. In this work, we analyse various data anonymization techniques to protect sensitive information created in an IOT environment.

In the forthcoming sections, section II is about Challenges for IOT data privacy, section III is IOT Data anonymization, section IV will show the Anonymization techniques, section V is Anonymization challenges and finally concluding with the conclusion in section VI.

II. CHALLENGES FOR IOT DATA PRIVACY

A. Traceability

The enlarged collection of data may raise issues of verification and trust in the objects [7]. In addition, it should also be noted that by using data collected about and from multiple objects related to a single person, that person may become more easily identifiable and better known.

B. Privacy and Security

Providing trust and quality-of-information in mutual information models to facilitate re-use across many applications [7]. Providing safe exchange of data between IoT devices and consumers of their information is more demanding.

C. Interoperability

With copious sources of data and heterogeneous devices, the use of standard interfaces between these miscellaneous entities becomes important [7]. This is especially so for applications that supports cross-organizational and various system boundaries.

Thus, the IoT systems need to handle high scale of interoperability. These challenges are to overcome by providing privacy for data. Data anonymization techniques can ensure the data privacy in IoT environment.

III. IOT DATA ANONYMIZATION

IoT should be considered as future internet as everything is going to be connected in a network so that one object can easily interact with each other objects, but still there are many issues, which are to be solved to make this a reality [3]. Modern technology generates such a huge amount of data that its security becomes an inevitable task. Nowadays it is necessary to provide security for personal data in order to achieve security, it is obligatory to preserve the privacy of personal data for that we use anonymization techniques [2]. More applicability of cryptographic models and security schemes but it requires detailed analysis, in order to be ensured that, they can be implemented in the specified resources of IoT [5]. The heart of IoT is on the data and information, rather than point-to-point communication [6]. Privacy protection in data can be achieved by using anonymization algorithms[4].We can use privacy models and transformation models in ARX tool which can transform the sensitive data of IoT applications into anonymized data.

IV. ANONYMIZATION TECHNIQUES

K-anonymity Approach [2]: - *k*-anonymity is a property obsessed by certain anonymized data. A relational data is said to have the *k*-anonymity property if the information for each person contained in the relational data cannot be distinguished from at least k1 individuals whose information also appear in the relational data. The following table is a nonanonymized relational table consisting of records of patients.

Table 1. Data table

Name	Age	Gender	State	Religion	Disease
Siva	24	Male	Tamil Nadu	Hindu	TB
Ram	29	Male	Kerala	Christian	Viral infection
Anita	24	Female	Tamil Nadu	Muslim	Cancer
Jothi	28	Female	Kerala	Hindu	Heart-related
Kajal	18	Female	Kerala	Christian	TB

There are 6 attributes and 5 records in this data. There are two common methods for achieving k-anonymity for some value of k.

1. Suppression: The certain values of the attributes are changed by '*'. All or some values of a column may be changed by '*'. In the anonymized table below, we have changed all the values in the 'Name' attribute and all the values in the ' Disease ' attribute with a '*'.
2. Generalization: The individual values of attributes are changed by with a broader category. For example, the value '18' of the attribute 'Age' may be changed by ' ≤ 20', the value '22' by '20 < Age ≤ 30' ,etc.

The next table.2 shows the anonymized table.

Table 2. Anonymized table

Name	Age	Gender	State	Religion	Disease
*	20 < Age ≤ 30	Male	Tamil Nadu	Hindu	*
*	20 < Age ≤ 30	Male	Kerala	Christian	*
*	20 < Age ≤ 30	Female	Tamil Nadu	Muslim	*
*	20 < Age ≤ 30	Female	Kerala	Hindu	*
*	Age ≤ 20	Female	Kerala	Christian	*

L-Diversity Approach [2]: - L-diversity is a form of cluster-based anonymization that is used to protect privacy in data sets .The l-diversity model is an extension of the k-anonymity model, which reduces the granularity of data representation using techniques including suppression and generalization such that any given record maps onto at least k1 individuals records in the data. The l-diversity model handles some of the weaknesses in the k-anonymity model where protected identities to the level of k-individuals is not equivalent to protecting the corresponding sensitive values that

were generalized or suppressed, especially when the sensitive values within a group exhibit homogeneity.

T-Closeness Approach [8]:-The existence of attacks where sensitive attributes may be incidental based upon the sharing of values for l-diverse data, the t-closeness method was produced to further l-diversity by additionally maintaining the sharing of sensitive fields. An equivalence class is said to have t-closeness if the distance between the sharing of a sensitive attribute in this class and the sharing of the attribute in the whole table is no more than a threshold t.

V. ANONYMIZATION CHALLENGES

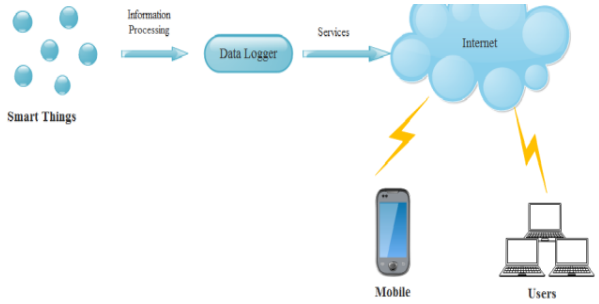

Figure 1. Anonymization Structure

The evolving nature of the IoT technologies leads to specific privacy threats and challenges. This section presents our classification of those threats. Figure 1 arranges them into different phases according to where they are most prone to appear. First, a definition and characterization of the threat category is given. Second denotes the analysis of IoT evolution impacts and threat. Third signifies the identification of approaches and counter-measures. Fourth, we present the main technical challenges and potential approaches to overcome those threats in the IoT. We majorly focus on anonymization challenges.

A. Identification-ID

The risk of identification is presently most prevailing in the data processing at the backend, where enormous amounts of information are concentrated in a vital place outside of the control.

B. Localization and Tracking

Localization and tracking is the risk of determining and recording a person's location through space and time .Tracking requires identification of some kind to attach continuous localizations to one individual.

C. Profiling

Profiling denotes the risk of compiling data about individuals in order to deduce interests by correlation with other profiles and data. Facebook and Twitter data can be used for such profiling.

D. Inventory attack

Inventory attacks refer to the illicit collection of data about the existence and characteristics of personal things. Here also facebook and twitter datasets can be used for specific individual oriented attack.

VI. CONCLUSION

In recent years, IoT is the rapid growing technology. In the field of security, which increases the capacity of storing and retrieving personal data by revealing sensitive, attribute information about individuals. In this paper, we analysed various anonymization techniques, which provide more accurate privacy protection. Anonymization techniques such as K-anonymity, L-diversity, and T-Closeness are well designed for improving accuracy in privacy preservation. Still the intruder can hack the private information of the user at some level. Data need analysis and membership disclosures are important aspects for providing privacy through anonymization. In future a detailed study on various anonymization techniques will be carried over which provide more accurate privacy preservation.

VII. REFERENCES

[1]. Francisco Dias, NunoMamede, Jorge Baptista "Automated Anonymization of Text Documents" 2016 IEEE Congress on Evolutionary Computation (CEC).

[2]. PreetChandanKaur*, TusharGhorpade+, Vanita Mane "Analysis of Data Security by using Anonymization Techniques" 978-1-4673-8203-8/16/$31.00_c 2016 IEEE 2016 6th International Conference-Cloud System and Big Data Engineering (Confluence).

[3]. Ruchi Parashar1, Abid Khan2, Neha3 "A SURVEY: THE INTERNET OF THINGS"

International Journal of Technical Research and Applications e-ISSN: 2320-8163, www.ijtra.com Volume 4, Issue 3 (May-June, 2016), PP. 251-257.

[4]. ShankarNayakBhukya, Dr.SureshPabboju, Dr. K Venkatesh Sharma "Implementing Privacy Mechanisms for Data using Anonymization Algorithms" International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT) – 2016.

[5]. Nicolas Sklavos, I. D. Zaharakis"Cryptography and Security in Internet of Things (IoTs): Models, Schemes, and Implementations" 978-1-5090-2914-3/16/$31.00 ©2016 IEEE".

[6]. AshviniBalte, AsmitaKashid, BalajiPatil "Security Issues in Internet of Things (IoT): A Survey" International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 4, 2015.

[7]. Keyur K Patel1, Sunil M Patel2 "Internet of Things-IOT: Definition, Characteristics, Architecture, Enabling Technologies, Application & Future Challenges" International Journal of Engineering Science and Computing, Research Article Volume 6 Issue No. 5, May 2016.

[8]. M.Saranya, R. Senthamil Selvi "Survey on Privacy Preservation for Anonymzing Data" International Journal of Emerging Engineering Research and Technology Volume 3, Issue 1, January 2015, PP 17-21 ISSN 2349-4395 (Print) & ISSN 2349-4409 (Online).