PRIVACY CHALLENGES IN SMART SURVEILLANCE

S.Sujith Roy[1], A.Kannammal[2]
[1]University of Canterbury, New Zealand
[2]Department of Computing
Coimbatore Institute of Technology, Coimbatore, India

Abstract : The increasing amount of integrated services in smart cities has led to collection of abundance data. Several services are being automated and optimized based on the data collected from smart cities. Analytics from the data collected from smart cities have resulted in data-oriented decision making. Smart surveillance is one such application in a smart city. This data being collected is subjected to privacy challenges. Most of the privacy issues could be addressed by pertinent technical measures and by following appropriate regulations. This paper highlights the adverse effects of smart surveillance with respect to user and data privacy. It also presents techniques currently being implemented to ensure privacy, drawbacks and additional techniques for privacy preservation in addition to providing recommendation to address privacy issues.

IndexTerms **- Cloud: smart city, data, decision making, analytics, surveillance, privacy, preservation techniques, user tracking.**

I. INTRODUCTION TO BUSINESS INTELLIGENCE

Business Intelligence (BI) can be termed as the process where actionable insights are drawn from the data extracted through various mediums. These insights help organizations make informed decisions. Everyday several billion of data is being generated through sensors, actuators, mobile phones and smart home devices. Business Intelligence through technologies such as Artificial Intelligence (AI), Internet of Things (IOT) and Machine Learning (ML) brings several value additions to end users. All these technologies make use of the data being generated.

1. SMART CITY AS AN APPLICATION OF BUSINESS INTELLIGENCE

Smart city is one of the applications of BI. Most metropolitan cities are accelerating towards making their city "smart". A city is a smart city, when an infrastructure including wireless sensor networks, cameras and a wide range of technologies is set that connects its people and things, tracks every single element in real time, and collects data to draw insights to enhance governance efficiency, smart infrastructure utilization, revenue growth and improved quality of life. Some of the use cases of a smart city include smart grid, smart waste management, smart traffic control, air quality management and smart surveillance (Righetti, Vallati, &Anastasi). Citizens of smart city need to install different applications in their smart devices. Through these applications and other devices, lot of data is been collected. Abundant availability of resources in terms of data, computing, technology, etc. facilitates the collection and processing of big data. This drives innovative intelligent business analytics at the same time raises the concern of privacy breaches. A use case is presented below.

2. Use Case: Smart Surveillance in a Smart City Leading to Privacy Issues

Smart surveillance systems facilitate real time monitoring to prevent crimes, terrorism, intrusions, and social unrest and to ensure national security. A smart city has well-built surveillance systems that are capable of sending automatic alarms, warnings, based on video and image analytics for faster rescue actions. These events are predicted by analyzing data collected from heterogeneous resources including cameras that are deployed at every nook and corner for real time monitoring, thus making the city "smart". This is to ensure safety by preventing possible crimes. Though this seems advantageous, it could really compromise the privacy of citizens. Essentially there exists a tradeoff between security and privacy. With the various methods of location data being collected, any user could be tracked down at any point of time. The behavioral patterns of a user can also be predicted using this system and the citizens are being monitored through spy-apps installed in their smart phones(Macaulay, 2017). Hence it nullifies the question of privacy. The questions like what data is being collected, how safe is the data, is my data being stored, am I being tracked, is still being a mystery to the end user. The question of user privacy still persists despite implementation of several regulations such as the General Data Protection Regulations (GDPR) that define the visibility and scope of the data being collected.

3. Business Intelligence/Analytics Vs Privacy

Big data collected from heterogeneous resources facilitates intelligent business analytics but leading to potential threats to privacy. The personal data that is being collected in smart city networks find their application in numerous ways. For example, facial recognition from CCTV's helps authorities to identify a person. This information is highly sensitive as it relates to one self. Such personal data for surveillance purposes should be regulated by governing bodies such as the GDPR for legitimate use of data. Most details of a citizen exist in one repository (database) or the other. These data are also used for surveillance through methods of aggregating and combination. This is highly useful in predictive policing as this data could be used to map the geographic location of previous occurrences of social unrests. Here postal codes help up in summarizing the possibility of higher crime rates (van Zoonen, 2016). All of these data can be analyzed to narrow down to specific individuals. Behavioral patterns of ex-convicts can also be studied to avenge outcomes. Cities are increasingly becoming monitored spaces and hence the user's identity is no longer anonymous. Also, the user privacy has become a paradox. The consent to collect, process and to share the data is no longer a question. Each and every move is being tracked by the Government through either surveillance systems or smart phones. The smart surveillance systems have been made the city safer in one perspective but have failed on the other. It is successful in realizing the goals of smart city, compromising the privacy of citizens. One such application of smart surveillance is real time traffic monitoring which was enabled using traffic intensity sensors that are capable of counting the number of vehicles entering and leaving the city. These were mounted on to building entrances or were buried underground (Righetti et al.). With the

deployment of smarter systems, surveillance cameras are now being used to monitor the same. These systems help us make more appropriate decisions and also help users with alternate routes. However, the data generated through cameras contain personal information and hence the anonymity is lost. The same case applies to the smart parking system which is again an application smart surveillance. Parking spots are constantly monitored using cameras and hence personally identifiable information (PII) is being recorded (van Zoonen, 2016). Again, anonymity is lost as location centric data is being tracked using PII. The effectiveness of a smart surveillance lies in the data collection method, interconnection and pervasiveness. The data extracted from different sources are stored in different databases and exhibit a perfect correlation. This correlation can be used to compile complete user profiles which is a threat again. This interconnectivity increases service quality, but the user privacy is still a concern (Eckhoff& Wagner, 2018). The main difference between the data being disclosed in social media to that being extracted through the smart surveillance devices is that there exists a willingness to disclose personal information in social media. But there doesn't exist user consent with regard to data being extracted through smart surveillance. Interactions with other people could be classified as one's social life. It contains sensitive data such as who a person interacts, how long and when. When such interactions are compromised, this data could be used to assess one's personal life. Also, the privacy of data from images and videos captured through smart surveillance is also specific to time. This leaves a digital foot print and is a serious violation of privacy. (Finn, Wright, &Friedewald, 2013) categorization of privacy and other researches in neuro science suggest that an individual's action at a certain location is the reflection of the state of mind body. These data are also subjected to exploitation. The violation of location privacy reveals personal information such as work space in addition to other data such as habits, purchases and social life.

4. Recommendations to Address Privacy Issues in Smart Surveillance

Smart city privacy can be categorized into three dimensions: data collection, user control, user awareness space (Hassan &Awad, 2018). Further, the privacy threats in smart city would fall under 6 categories: individual identity, user profiling, lifecycle transition, interactions, inventory attacks and linkage (Alabdulatif, Kumarage, Khalil, & Yi, 2017).
Some of the major factors to be considered when a smart city infrastructure is built:

- Privacy assurance has to be provided to the users at organizational and technical level.
- Enough control and transparency to what can be tracked and revealed should be provided to the user.
- The hierarchy of data ownership and data processing should be well defined as data is processed through various levels for various purposes.

Data privacy can be met by meeting the following requirements.
User Consent: In an IoT enabled smart city, data is extracted through various sources. Collection of sensitive data without the consent is a violation in privacy. A simpler solution would be requesting the consent from user (Dhungana, Engelbrecht, Parreira, Schuster, & Valerio).

Anonymization: Any Personal Identifiable information (PII) should be anonymized from the data repositories. Pseudonymization has to be enabled in a way that a PII cannot be linked between data sets (Sookhak, Tang, He, & Yu, 2018). Access control and customization: Simpler access control mechanism and personal anonymity should be provided as a choice instead of chance. The user should have the authority to grant and revoke access rights to service providers.

5. Techniques for Privacy preservation

Smart cities are accelerating towards open data. This means that the data being extracted is openly available to public. User related data cannot be redistributed without consent. However, this might result in breach of privacy and hence advanced frame work for privatization of data has to be set up. Some of the existing techniques are

Data Randomization is the process of subjecting a privacy sensitive data attribute to a random noise. Each data record is masked by this process. They have high levels of efficiency than most cryptographic algorithms but are less accurate due to the noise (Oliveira &Zaïane, 2004).

Homomorphic encryption (HE) is a privacy preservation technique that enables us to compute function on encrypted data. Sensitive data from smart city applications can be processed using homomorphic encryption as it allows third parties to process data without seeing the computational input or output.

Secure Multi Party Computation (SMC) is a frame work that allows evaluation of union of data sets while retaining the privacy of the data set. SMC is a based on a homomorphic encryption which means that the computation performed on the private input data set provides an encrypted result (Alabdulatif, Khalil, Kumarage, Zomaya, & Yi, 2018). SMC is accomplished through secret sharing which is an extended clustering analysis mechanism. In secret sharing each party gets a share of the data, but the confidentiality is retained. It is based on the fact that a single share cannot be used to recover the data. SMC don't scale up to large data sets and also require high computational costs (Alabdulatif et al., 2017).

6. Suggested methods for Data Oriented Privacy Protection

Data privacy could be achieved through the following methods. Data Minimization: Minimal data collection has to be enabled to solve the question of interest. Moderns sensors and devices emit tonnes of data. This is often referred as collateral data as more data than envisioned is collected for a particular task. Traffic monitoring system could be stated here as an example as it collects background information like people on the road in addition to monitoring vehicles. Data Anonymization: Most underlying databases contain identifiable features such as names. By using techniques like k-Anonymity, this data could be anonymized. In

this method each row of a database is grouped into equivalence classes. These rows remain indistinguishable by using separate quasi (non-unique) identifiers for each feature of the row. For example, assigning separate identifiers for age, gender and name of a person. But these identifiers reveal personal data only when all of the identifiers are co-related. Differential Privacy: It is similar to data randomization. Smart metering and location privacy are usually privatized using this methodology as a small amount of multiplicative factor is applied to each recorded observation. Encryption Techniques: This involves usage of a description key to decrypt the recorded data. They are classified as identity based and attribute-based encryption. Identity based encryption usually involves using an arbitrary string as public key. Attribute based encryption uses private keys and cipher texts. For a smart city-based application, attribute-based encryption is more suited as it gives better access control.

II. DRAWBACKS OF EXISTING SOLUTIONS

Techniques like Data Anonymization reveal underlying data when the identifiers are corelated. Several improvisations are being developed like the l-diversity (multiple classes for same identifier), but it still fails to anonymize completely when multidimensional data are used. Data Randomization and Differential Privacy doesn't formally guarantee the extent of privacy. Fully homomorphic cryptosystems are computationally expensive and allow only aggregation and union-based operations. Several optimizations were provided for Homomorphic Encryption (HE) based on public key compression technique, but it couldn't perform well on cloud-based applications such as a smart city.

III. CONCLUSION

The creation of privacy friendly smart city doesn't mean creating new technologies. It's the implementation of existing ones in larger scale by a holistic approach. The implementation of smart surveillance systems will greatly integrate lives. It will have a potential impact on the way of living. With the advancements in smart city applications, all models should specifically focus on retaining user privacy. Multiple technologies are integrated to build a smart city. Hence, privacy should be addressed for each component in each level. Increasing awareness for user privacy is a challenging task. Given the advantages, the data being collected should be deemed mandatory and the user should be made aware what is being tracked by the smart surveillance systems. Privacy policy should be clearly articulated.

REFERENCES

[1] Alabdulatif, A., Khalil, I., Kumarage, H., Zomaya, A. Y., & Yi, X. (2018). Privacy-preserving anomaly detection in the cloud for quality assured decision-making in smart cities. *Journal of Parallel and Distributed Computing*. doi:10.1016/j.jpdc.2017.12.011

[2] Alabdulatif, A., Kumarage, H., Khalil, I., & Yi, X. (2017). Privacy-preserving anomaly detection in cloud with lightweight homomorphic encryption. *Journal of Computer and System Sciences, 90*, 28-45. doi:10.1016/j.jcss.2017.03.001

[3] Dhungana, D., Engelbrecht, G., Parreira, J. X., Schuster, A., & Valerio, D. (2015). *Aspern smart ICT: Data analytics and privacy challenges in a smart city*.

[4] Eckhoff, D., & Wagner, I. (2018). Privacy in the Smart City—Applications, Technologies, Challenges, and Solutions. *IEEE Communications Surveys & Tutorials, 20*(1), 489-516. doi:10.1109/COMST.2017.2748998

[5] Finn, R. L., Wright, D., & Friedewald, M. (2013). Seven types of privacy. In.

[6] Hassan, A. M., & Awad, A. I. (2018). Urban Transition in the Era of the Internet of Things: Social Implications and Privacy Challenges. *IEEE Access, 6*, 36428-36440. doi:10.1109/ACCESS.2018.2838339

[7] Macaulay, T. (2017). Retrieved from https://www.techworld.com/apps-wearables/how-smartphone-apps-spy-on-users-by-listening-inaudible-signals-3662504/

[8] Oliveira, S. R. M., & Zaïane, O. R. (2004). Achieving privacy preservation when sharing data for clustering. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 3178*, 67-82.

[9] Righetti, F., Vallati, C., & Anastasi, G. (2018). *IoT Applications in Smart Cities: A Perspective Into Social and Ethical Issues*.

[10] Sookhak, M., Tang, H., He, Y., & Yu, F. R. (2018). Security and Privacy of Smart Cities: A Survey, Research Issues and Challenges. *IEEE Communications Surveys & Tutorials*, 1-1. doi:10.1109/COMST.2018.2867288

[11] van Zoonen, L. (2016). Privacy concerns in smart cities. *Government Information Quarterly, 33*(3), 472-480. doi:10.1016/j.giq.2016.06.004.