ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 9, Issue 3, March 2022

Privacy Preserved employee workspace Monitoring using Deep Learning

Thomas Abraham[1], Dr.Sajimon Abraham[3], Dr.Rajesh N[3]
School of Computer Sciences, Mahatma Gandhi University, Kottayam, Kerala, India[1,2,3]

Abstract: In this recent times a large volumes of video data's and information's are recorded by the closed-circuit television (CCTV) Surveillance and by increasing mobile devices and embedded sensors can be leveraged to answer queries of our lives, physical world and our evolving society. There is an exponential growth in video data volume and variety of data as due to diverse applications of computers in all domain areas. The growth has been achieved due to affordable availability of computer technology, storage, and network connectivity as well as Surveillance needs. There are many ways and techniques are suggested to calculate employees performance evaluation and all the techniques has their own limitations are there. In this paper we are going to discuss about a workspace monitoring method using deep learning approach. So that many of the employees and organizations are concern about privacy. We know that privacy is an important thing in this area. The organization development is based on the employee hard works as well as efficiency. The privacy-preserved workspace monitoring or performance evaluation has become much more important in these recent years because of the performance evaluation is very much important for the step by step development of any organization.

Keywords**—** Privacy preservation, CCTV Surveillance, Differential Privacy, Video splitting, Video encryption and monitoring.

I. Introduction

The pervasive cameras deployed in CCTV Surveillance devices, embedded sensors and so forth record large volumes of videos everyday. The recorded CCTV videos can be leveraged to answer queries about day by day activities of employees in a particular company or an organization, physical world and happenings at company surroundings. The **Closed Circuit Television** (CCTV) is also known as **video surveillance**. It is a system where all the elements like video camera, display monitors, recording devices are directly connected. It is used *to monitor a sensitive area* (A particular area which needs continuous observation and where there is no one to watch all the time). It is very helpful to prevent crime because it monitors all the activities and record them. It is also used for traffic monitoring by detecting congestion and notice accidents. It can be used in all areas for protection as well as monitoring of the employees.

Privacy preservation in data mining is an important concept, because when the data is transferred or communicated between different parties then it's compulsory to provide security to that data so that other parties do not know what data is communicated between original parties. Preserving in data mining means hiding output knowledge of data mining by using several methods when this output data is valuable and private. Mainly two techniques are used for this one is Input privacy in which data is manipulated by using different techniques and other one is the output privacy in which data is altered in order to hide the rules.

The process Data mining is the of extracting useful, interesting, and previously unknown information from large data sets. The success of data mining relies on the availability of high quality data and effective information sharing. We know that the collection of digital information by governments, organizations, corporations, and individuals has created a good environment that facilitates ultra large-scale data mining and data analysis in these years. More than this, driven by mutual understanding and benefits, or by rules and regulations that require certain data are to be published, there is a good demand for sharing data among various parties. For example, financed banks are required to submit specific demographic data and documents on every customer using their bank facility. Data publishing is equally ubiquitous in other domains too. For example, Netflix, it is a popular online movie rental service, few years back, published a data set containing movie ratings of 500,0000 subscribers, it in a drive used to improve the accuracy and effectiveness of favorable movie recommendations based on personal preferences and interests. But in case of video Surveillance is very much important for the efficient and proper working of an organization. The superiors can check employees works according with the videos that are already available and from other resources like CCTV etc. But these kind of informations like videos will disclosure the privacy of the employee to the superiors, while analyzing these kinds of video Surveillances. But, in an organization, the leaders or the supervisors can analyze and monitor employees workspace behavior and their working ability by using these kinds of video Surveillance, it can be used for the government organizations too and at the same time it cannot be published to the any other areas. If it is done it will leads to serious privacy isssues. But it can be published to the inside of the organization while the privacy of the employee must be protected.

Information sharing has a long history in information technology and development. Off course, the traditional type of information and data sharing refers to exchanges of data or information between a data holder and a data recipient. For example, the Electronic Data Interchange (EDI) is a successful implementation of electronic data transmission between organizations with the emphasis on the commercial and social sector. During these days, the terms "Information sharing" and "Data publishing" not only refer to the traditional one-to-one model, but also the more general models with multiple and many data holders and data recipients.

Now a days, privacy is very much important and have more concern than previous, The general public expresses and mentioned serious concerns on their privacy and the consequences of sharing their person-specific information. The current and recent privacy protection and preservation practice primarily relies on policies and guidelines to restrict the types of publishable data, and agreements on the use and storage of sensitive data. The limitation of this approach is that it either distorts data excessively or requires a trust level that is impractically high in many data-sharing and information sharing scenarios, Also, policies, rules and guidelines cannot prevent adversaries who do not follow rules in the first place. Contracts and agreements cannot guarantee and protect that sensitive data and information will not be carelessly misplaced and end up in the wrong hands.

A task of the utmost and very much importance is to develop methods, techniques and tools for analyzing and publishing data in a hostile environment so that the published data remain practically useful while individual employee privacy is preserved.

II. Review Of Literature

A group of surveys focuses on specific different families of privacy preserving algorithms and techniques. For instance, The author [1] in context-aware or perceptual applications which monitors or tracks the user activities and environment with cameras and other alternative sensors. The authors design a practically feasible privacy protection mechanism referred as DARKLY for an untrusted perceptual application executing on a trusted device by the third party. The proposed DARKLY is attached with the OpenCV library used to access visual inputs with the applications. The algorithm privacy transformation, access control and user audit were included as a multiple privacy protection mechanism in DARKLY. It implements diversified tasks such as object tracking, image recognition, and face recognition and security surveillance. The advantage of the proposed method is it implements all these tasks with minimal overhead compared to OpenCV. In all these cases the application accuracy and functionality are not decreased. Drawback of this is in some scenarios the privacy and utility should be automatically acceptable by considering the strong privacy protection. Supervised machine learning approaches are used to build filters and constructors for delicate and privacy-sensitive scenes which include gestures, movement patterns, physical proximity and text strings. It is failed to handle various types of audio data. To enable various security measures, the CCTV with intelligent security environment is required to improve the performance of video surveillance system. This leads to the privacy protection problem for surveillance system. The authors [2] addressed the leakage of personal information from the surveillance video which causes the violation in privacy to the severe social issues. This intelligent surveillance mechanism exhibits averse results like exposure of privacy, data manipulation, etc. The authors proposed block-chain based data processing approach for videos and synchronize huge data securely and in the delivery process the data is stored. The advantage of this approach is, it does not expose the privacy in the synchronized operation, because it protects from the forgery or falsification attacks efficiently. In [3] the authors proposed homomorphic encryption approach to study the moving object detection in cloud-based video surveillance system with encryption for privacy preserving. This homomorphic encryption approach validates the operations on between encrypted and encrypted data. The advantage of the proposed approach is more secure and efficient with real time analytics of encrypted video streams in motion detection or moving object detection. The drawback of the proposed method is it is not practical and requires high storage due to terrible computations. The proposed method uses Paillier cryptosystem which slowdowns the performance and consumes more time for processing, because it needs 1024 bits to represent each pixel rather than 8 bits in traditional approaches.

The author Simi et al. [4] provide an extensive study of works on k-anonymity and Dwork [5] focuses on differential privacy. Another group of surveys focuses on techniques that allow the execution of data mining or machine learning tasks with some privacy guarantees, e.g., Wang et al. [6] and Ji et al. [7]. In a more general scope, Wang et al. [8] analyze the challenges of privacy-preserving data publishing and offer a summary and evaluation of relevant techniques. Additional surveys look into issues around Big Data and user privacy. Indicatively, Jain et al. [9], and Soria-Comas and Domingo-Ferrer [10] examine how Big Data conflict with pre-existing concepts of privacy-preserving data management, and how efficiently k-anonymity and ε-differential privacy deal with the characteristics of Big Data. Others narrow down their research to location privacy issues. To name a few, Chow and Mokbel [11] investigate privacy protection in continuous LBSs and trajectory data publishing, Chatzikokolakis et al. [12] review privacy issues around the usage of LBSs and relevant protection mechanisms and metrics, Primault et al. [13] summarize location privacy threats and privacy-preserving mechanisms, and Fiore focus only on privacy-preserving publishing of trajectory microdata. Finally, there are some surveys on application-specific privacy challenges. But from all these papers we can find that no one has proposed any method, technique for analyzing video data especially CCTV video data with privacy preservation. So in this paper, suggesting a technique for privacy preservation in CCTV based workspace monitoring.

III. Data Collection

A typical representation scenario of data collection is described in the Fig (1.1). During the data collection phase, the video data is collected from CCTV footage of the company or from the organization. It is only for the monitoring and analyzing happening inside of the company and it is for their personal analyzing only. So that they can use their CCTV video footage for their own development purpose of the company. In this case we are considering some demo videos as the data set.

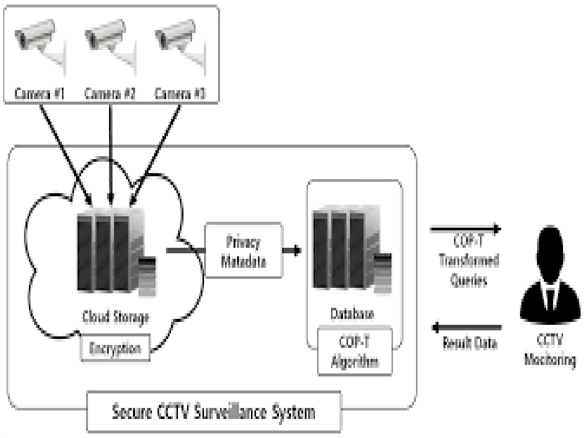

Fig (1.1) CCTV Video Data collection

The data collection is simple and not much conserved in sense but, the term or technique data publication is more concerned and important to the reputation of the company. There are two models of data holders. In the untrusted model, the data holder is not trusted and may attempt to identify sensitive information from record owners. Various cryptographic solutions, anonymous communications and statistical methods were proposed to collect records anonymously from their owners without revealing the owners' identity. In the trusted model, the data holder is trustworthy and record owners are willing to provide their personal information to the data holder; however, the trust is not transitive to the data recipient. In this paper we assume the trusted model of data holders and consider privacy issues in the data publishing phase. The video data monitoring and mapping is illustrated in the fig (1.2)

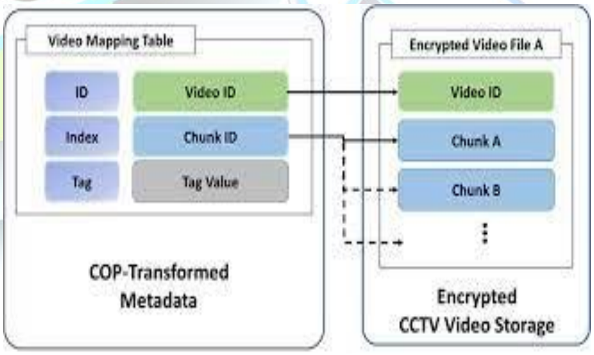

Fig (1.2) Video Mapping and Monitoring

III.A *Privacy-Preserved Video Data Monitoring*

Video surveillance, IP camera and CCTV become crucial for much government, private organizations or businesses and individual users to track the privacy of the individuals. Privacy preserving in these systems become a complex and challenging task. Nevertheless, researchers focus on the privacy aspects of such systems. The basic form of privacy preserved video data monitoring is shown below, the data holder has a table of the form,

D(Explicit Identifier, Quasi Identifier, Sensitive Attributes, Non-Sensitive Attributes)

where Explicit Identifier is a set of attributes, such as name and social security number (SSN), containing information that explicitly identifies record owners, Quasi Identifier is a set of attributes that could potentially identify record owners, Sensitive Attributes consist of sensitive person-specific information such as disease, salary, and disability status and Non-Sensitive Attributes contains all attributes that do not fall into the previous three categories. Most works assume that the four sets of attributes are disjoint. Most works assume that each record in the table represents a distinct record owner. There are many applications,

Web camera surveillance: one possible method for surveillance is with the use of publicly available webcams, which can be used to detect unusual activity. The approach can be made more privacy sensitive by extracting only facial count information from the images and using these in order to detect unusual activity. It has been hypothesized that unusual activity can be detected only in terms of facial count rather than using more specific information about particular individuals.

Video surveillance: in context of sharing video-surveillance data, a major threat is the use of facial recognition software, which can match the facial images in videos to the facial images in a driver license database. A balanced approach is to use selective downgrading of the facial information, so that it scientifically limits the ability of facial recognition software to reliably identify faces, while maintaining facial details in images. The algorithm is referred to as k-Same, and the key is to identify faces which are somewhat similar, and then construct new faces which construct combinations of features from these similar faces. Thus, the identity of the underlying individual is anonymized to a certain extent, but the video continues to remain useful.

To prevent linking attacks, the data holder publishes an anonymous table,

T (QID’ , Sensitive Attributes, Non-Sensitive Attributes)

QID’ is an anonymous version of the original QID obtained by applying anonymization operations to the attributes in QID in the original table D. Anonymization operations hide some detailed information so that multiple records become indistinguishable with respect to QID’ Consequently, if a person is linked to a record through QID’, the person is also linked to all other records that have the same value for QID’, making the linking ambiguous. Alternatively, anonymization operations could generate a synthetic data table T based on the statistical properties of the original table D, or add noise to the original table D. The anonymization problem is to produce an anonymous T that satisfies a given privacy requirement determined by the chosen privacy model and to retain as much data utility as possible. An information metric is used to measure the utility of an anonymous table. Note, the Non-Sensitive Attributes are published if they are important to the data mining.

IV. Proposed Methodology

There are many models, methods and techniques to achieve privacy preservation in video data monitoring and publishing especially in CCTV video footage. Many of them are efficient and convenient. In fact some of them has its on merits and demerits. So, in this chapter we are going to discuss about the techniques and its benefits over others, some of the recently used privacy preservation methods are mentioned below,

Generalization Techniques:

Generalization [14] is currently one of the main approaches and techniques for the anonymization of spatiotemporal trajectory datasets as well as in other type of data sets too. The generalization technique is predicated on two interrelated mechanisms. clustering and alignment: Clustering aims at finding the best grouping of trajectories that minimizes a predefined cost function and the alignment process aligns trajectories in each group.

K-anonymity:

Among from the anonymity models, k-anonymity [15] is the one of the most extensively studied and influenced one due to its intuitive and accurate anonymization process. Generally speaking, a dataset D satisfies k-anonymity if each QI value D(QI) appears in at least k records. K-anonymity counters record linkage by ensuring the indistinguishability of an individual within a k-anonymous group and meanwhile minimizing information loss.

l-diversity and t-closeness:

l-diversity [16][17][18] is proposed to ensure that an anonymity group contains at least l well-represented values for each sensitive attribute. Several definitions of well-represented values exist. For instance, a dataset D satisfies distinct l-diversity if the number of values for the sensitive attribute in D(QI) is at least l. Other definitions are based on entropy and frequency of values.

Differential Privacy

Differential privacy [19] ensures that the presence of a record in a dataset leaks a controlled amount of information ∈. An algorithm f satisfies ∈-differential privacy if for any two datasets D1 and D2 that differs on at most one record, and all sets S of values in the image of the algorithm (i.e., S ⊆ Range(f)), it has,

$$P\,r\,(f\,(D1) \in S) \geq e \in \cdot\, P\,r\,(f\,(D2) \in S)$$

In this algorithmic equation, where Pr is the probability to observe a specific output. Differential privacy is usually guaranteed by generating synthetic data from the original one with controlled amount of random noise. In the field of traditional relational database, several randomized mechanisms have been already utilized to achieve ∈-differential privacy.

Capturing the CCTV video footages and spilt into frames.

Step 1:

Initially, choose N number of relatively primes(GCD(I,J)=1 the I and J are relatively prime numbers) where N equal to number of shares.

Step 2:

Apply the Scaling (Scale the pixels with fixed integer: sc) and Randomization (adding random values to each pixel) for ensuring accuracy before sharing algorithm.

To obtain the shares from the secret S as share1= S mod $p1$, share2= S mod $p2$, share3=S mod $p3$ share N=S mod pN, where, $P1$,$P2$, … $.PNarerelativelyprime.$

Step 3: Each share is sent to individual computational severs and apply the affine transformation on each individual computational server.

Step 4: Affine Transformation on share1 as,

$p.\ di + q.scmodp1$, where p & q are fixed values which are used for privacy and di is pixel values of share1 and sc is scaling factor.

Step 5: Affine Transformation is applied to each independent server and keeps results for obtaining original result.

This is the algorithm for CCTV footage capture and splitting it into different frames. In this paper we are suggesting the Differential privacy is applied to this particular algorithm to achieve privacy in workspace monitoring of the employees using their CCTV video footage. From this Differential privacy method we can hide and protect the privacy of the employees during their work and also we can monitor their workspace abilities for the development of the organization. It is the efficient way to preserve privacy in CCTV based workspace monitoring of the employee.

V. Challenges

There are many challenges are facing this privacy preservation CCTV video data monitoring. In-fact this very difficult task to attain the privacy in every aspects. There are many privacy issues are facing due to privacy violations. Most of the general people do not follow these kinds of rules and regulations in all areas due to lack of knowledge but most of the times it happens due to excess knowledge and some typical kind of privacy violation tendencies. These kinds of things and acting’s are making serious challenges and serious concerns about privacy concepts. The main challenge of this study is to prevent these kinds of CCTV footage videos from violations and disclosure to the other areas. It will affect the employees works as well as company’s reputation also. It is very challenging opportunity for researchers to overcome all the concerns in privacy of the CCTV video footage.

VI. Future Directions

There are many opportunities in the field of privacy preservation. A very less number of researchers have been enrolled into the CCTV footage monitoring. So there is an area of research area with lots of interesting facts. Inference control methods for privacy-preserving data mining and publishing is an evolving area of research, Identifying a comprehensive listing of data uses, that would allow the definition of data unspecific information loss measures broadly accepted by the community. These new measures could complement and/or replace the general measures currently being used. This would help data attempt re-identification for each domain of application. This would help data protectors figuring out in more realistic terms which are the disclosure scenarios they should protect data against.

There are many applications which utilize some special functions of data. How to extend the utility-based privacy-preserving methods to various applications is highly interesting. For example, in the data set where ranking queries are usually issued, the utility of data should be measured as how much the dominance relationship among tuples is preserved. None of the existing models can handle this problem. Another interesting research direction would be to extend utility-based privacy preserving methods to other types of data, such as video facial data where the temporal characteristics are considered more important in analysis. Any way lots of research openings are there, but I would like to do this privacy preserved workspace monitoring of the employees using CCTV video data using machine learning algorithms with the combinations of other techniques.

VII. Conclusion

In this paper, we have discussed that a framework to preserve the privacy of the CCTV video footage while analyzing this video data's, Information's. We have also discussed about various model and algorithms for privacy preservation methods have been introduced and discussed in detail subsequently and also Some of the challenges that the real-world applications for privacy preserving video data mining and monitoring algorithms have been presented. We have discussed about some of the main applications and techniques that are used towards these privacy concerns. Most of them are commonly used techniques and algorithms. Then, we have discussed about most common challenges that are faced today. All of the areas concern about privacy in serious manner including government organizations too. All the personal things are now concern privacy. Many of the algorithms are used to prevent these type of serious issues and having many of the techniques and approaches including deep learning and machine learning.

References

1. Lee, Donghyeok& Park, Namje. (2020). "Blockchain based privacy preserving multimedia intelligent video surveillance using secure Merkle tree". Multimedia Tools and Applications. 10.1007/s11042-020- 08776-y.

2. Gallo P, Pongnumkul S, Nguyen UQ (2018) "BlockSee: Blockchain for IoT video surveillance in smart cities." In: 2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe). IEEE.

3. H. Sohn, et.al. "Privacy-preserving watch list screening in video surveillance system." In PCM (1), pages 622632, 2010.

4. SIMI, M. S., NAYAKI, K. S., AND ELAYIDOM, M. S. "An extensive study on data anonymization algorithms based on k-anonymity." In IOP Conference Series: Materials Science and Engineering (2017), vol. 225, IOP Publishing, p. 012279.

5. C. Differential privacy: A survey of results. In International Conference on Theory and Applications of Models of Computation (2008), Springer, pp. 1–19. doi:10.1007/978-3-540-79228-4,1.

6. WANG, J., LUO, Y., ZHAO, Y., AND LE, J. "A survey on privacy preserving data mining." In Database Technology and Applications, 2009 First International Workshop on (2009), IEEE, pp. 111–114.

7. JI, Z., LIPTON, Z. C., AND ELKAN, C. "Differential privacy and machine learning: a survey and review." arXiv preprint arXiv:1412.7584 (2014).

8. Fung, B., WANG, K., CHEN, R., AND YU, P."Privacy-preserving data publishing: A survey on recent developments." ACM Computing Surveys (2010). doi:10.1145/1749603.1749605.

9. JAIN, P., GYANCHANDANI, M., AND KHARE, N. "Big data privacy: a technological perspective and review." Journal of Big Data 3, 1 (2016), 25. doi:10.1186/s40537-016- 0059-y.

10. SORIA-COMAS, J., AND DOMINGO-FERRER, J." Big data privacy: challenges to privacy principles and models." Data Science and Engineering 1, 1 (2016), 21–28. doi:10.1007/s41019-015-0001-x.

11. CHOW, C.-Y., AND MOKBEL, M. F. "Trajectory privacy in location-based services and data publication." ACM Sigkdd Explorations Newsletter 13, 1 (2011), 19–29. doi:10.1145/2031331.2031335.

12. CHATZIKOKOLAKIS, K., ELSALAMOUNY, E., PALAMIDESSI, C., AND ANNA, P. "Methods for location privacy: A comparative overview. Foundations and Trends R in Privacy and Security," 1, 4 (2017), 199–257. doi:10.1561/3300000017.

13. PRIMAULT, V., BOUTET, A., MOKHTAR, S. B., AND BRUNIE, L. "The long road to computational location privacy: A survey" IEEE Communications Surveys & Tutorials (2018).

14. Sina Shaham, Ming Ding, Bo Liu, Shuping Dang, Zihuai Lin, and Jun Li, "Privacy Preserving Location Data Publishing: A Machine

15. Learning Approach," This work was submitted in part and accepted to appear in the proceedings of INFOCOM WORKSHOPS, 2019 [1].

16. A. Meyerson and R. Williams, "On the complexity of optimal k-anonymity," in Proc. PODS. ACM, 2004, pp. 223–228.

17. Fengmei Jin, Wen Hua, Matteo Francia, Pingfu Chao, Maria E Orlowska, Xiaofang Zhou, Fellow, "A Survey and Experimental Study on Privacy-Preserving Trajectory Data Publishing," IEEE.

18. A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam, "L-diversity: Privacy beyond k-anonymity," ACM Trans. Knowl. Discov. Data, vol. 1, no. 1, p. 3, 2007.

19. N. Li, T. Li, and S. Venkatasubramanian, "t-closeness: Privacy beyond k-anonymity and l-diversity," in Proc. ICDE. IEEE Computer Society, 2007, pp. 106–115.

20. C. Dwork, "Differential privacy," in Proc. ICALP, ser. Lecture Notes in Computer Science, vol. 4052. Springer, 2006, pp. 1– 12.