exeley INTERNATIONAL JOURNAL ON SMART SENSING AND INTELLIGENT SYSTEMS

Article | DOI: 10.21307/ijssis-2020-008 Issue 1 | Vol. 13 (2020)

A review of privacy-preserving human and human activity recognition

Im Y. Jung*

School of Electronics Engineering, College of IT Engineering, Kyungpook National University, Daegu, South Korea.

*E-mail: iyjung@ee.knu.ac.kr

The paper was edited by Subhas Chandra Mukhopadhyay and Nagender Kumar Suryadevara.

Received for publication March 30, 2020.

Abstract

Many automation technologies using software are making humans convenient. One of these technologies is to collect data through cameras and sensors that are common in personal life and automatically recognize human and human activities. The goal of automation is to analyze the various types of big data that are difficult to perform mechanical data mining. Raw data collected from cameras and sensors are nothing but big data before analysis. In this case, how to protect data by secure storage is the most important issue. However, when the context-aware semantic information such as a specific person and his behavior is extracted from the analysis, the security sensitivity is increased. In other words, the secondary information generated by interpreting and extracting personal location and behavioral information contained in images and videos is linked to other personal information, causing privacy infringement issues. Privacy issues become important because there is a lot of software that everyone can access. Therefore, it is necessary to study privacy protection methods in the automatic recognition of human and human activities. This paper analyzes the cutting-edge research trends, techniques, and issues of privacy-preserving human and human activity recognition.

Keywords

Human recognition, Human activity recognition, Privacy protection, Machine learning.

As more and more data is collected and the technology to process it develops, the importance of data is growing. In addition, technology is needed for sensitive data processing to protect privacy. All processes that process data, such as raw data, data being processed, and result data, require privacy. The general approaches to prevent privacy leakage adopted anonymity, access control, and transparency (Haris et al., 2014). With the introduction of machine learning (ML), big data processing is in full swing, but the task of privacy protection remains.

Machine learning technology has been actively introduced in big data processing, and applied in many applications where mechanical data mining is difficult. However, privacy concerns are raised in applications that extract information through deep learning (Tanuwidjaja et al., 2019). Privacy protection is essential as the application of deep learning is expanded from medical applications that process sensitive information such as patient diseases (Tanuwidjaja et al., 2019) to applications that analyze data collected by cameras and sensors to extract personal information (Wang et al., 2019).

There are several concerns that the machine learning approach can violate the user's privacy; Figure 1 shows a machine learning process (Osia et al., 2018), and Figure 2 shows privacy issues during the machine learning process in Figure 1. Privacy may be violated when: (i) data holder shares a public dataset: anonymity of individuals are threatened; (ii) data holders participate

© 2020 Author. This work is licensed under the Creative Commons Attribution-Non-Commercial-NoDerivs 4.0 License https://creativecommons.org/licenses/by-nc-nd/4.0/

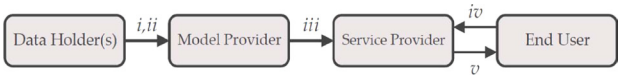

Figure 1: A machine learning process (Osia et al., 2018).

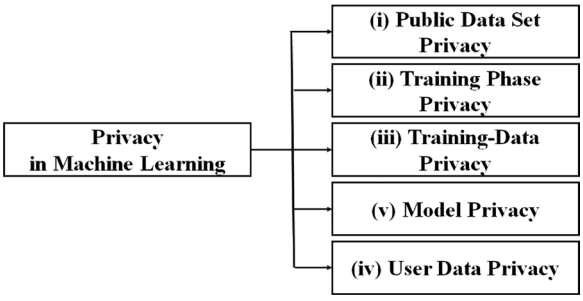

Figure 2: Privacy issues during the machine learning process in Figure 1.

in a model training procedure with their private data; (iii) a model provider shares a publicly learned model: the privacy of the individuals' data used for training is at risk; (iv) an end user shares his/her data with the service provider: private information can be revealed to the service provider; (v) a service provider shares query answers with the end user: an attacker can infer the model itself by launching repeated queries.

In the field of vision-based machine learning, many studies have shown serious privacy concerns as described in Table 1.

On the other hand, collaborative machine learning and federated learning allow multiple participants, each with his/her own training dataset, to build a joint model by training locally and periodically exchanging model updates (Melis et al., 2018). The updates can leak unintended information about participants' training data, and passive and active inference attacks can exploit this leakage as shown in Figure 3.

In addition, with big data processing, both population privacy and individual privacy become important (Cormode et al., 2012). Population privacy is violated by disclosing that some specific people are highly susceptible to a given genetic condition and individual privacy is violated by disclosing that a specific patient has that condition. In general applications, it is difficult to hide all the information in big data with cryptography because it is resource constraint and overly complex. Instead, many approaches have chosen to remove the sensitive parts of the information, while at the same time preserving the necessary information for further analysis (Osia et al., 2020).

Privacy protection is especially important in the field of dealing with human-related data. Machine learning has widely been applied to the recognition of human and human activity.

Human recognition is known as very useful in many application domains, for example, autonomous driving, post-disaster rescue, automated surveillance, military and robotics services (Gajjar et al., 2017). Human face recognition is another well-known application to search specific person in videos or in the list of images. Sensor data other than images or

Table 1. Privacy concerns in vision-based machine learning.

Research	Application domain	Privacy concerns
Chattopadhyay and Boult (2007)	Intelligent surveillance system	Conflict between the purpose of intelligent surveillance systems and the privacy of individuals
Wu et al. (2018)	Smart camera application	Private information leakage during device-captured visual data upload to centralized cloud for analysis
Gomathisankaran et al. (2013)	Medical image analysis on the Cloud	Private information leakage of medical data transmitted in the network and processed in the cloud
Shokri et al. (2017)	'Machine learning as a service' provided by Google and Amazon	Information leakage about training datasets
Speciale et al. (2019)	Augmented/Mixed reality (AR/MR) and autonomous robotic system	Confidential information disclosure about captured 3D scene

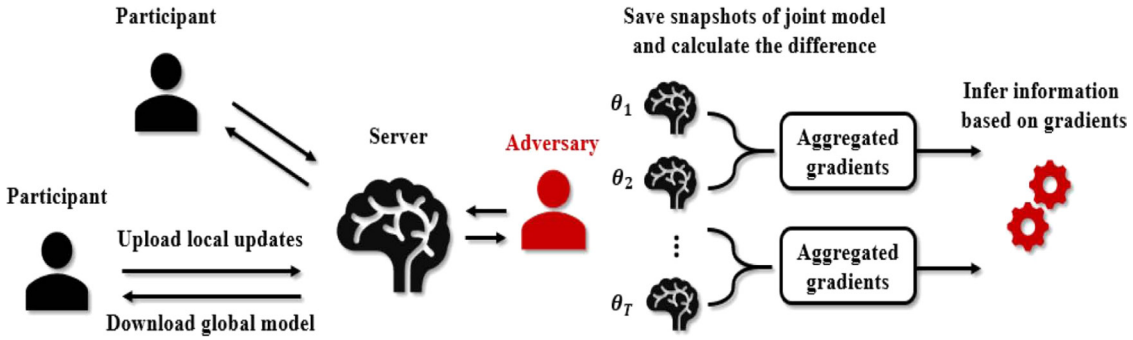

Figure 3: An inference attack model against collaborative learning (Melis et al., 2018).

videos can be used to determine if there is no human being. However, videos or images are commonly used to track and recognize specific persons and human diseases. Videos and images are sensitive data because they directly demonstrate human characteristics. Even if the original big data such as images or videos are stored safely, the secondary information extracted by the deep learning processing also has the risk of leakage of private information. In the medical field, automated analysis technology is utilized for disease diagnosis and tracking. Personal disease information is also concerned about the invasion of privacy because the secondary semantic information is extracted through deep learning for diagnosis.

Recently, human activity recognition has been utilized in many applications such as smart homes, healthcare, and manufacturing. Human activity is mainly recognized by cameras or sensors (Chen et al., 2020). Cameras and sensors are embedded in not only many large electronic systems, such as vehicles, home appliances, and surveillance systems, but also many portable Internet of Things (IoT) devices and wearable devices, such as smartphones, watches, and fitbits. These devices are spreading for anyone to access, so they can collect personal-area data easily and that data can be used to recognize the human activity. In human activity recognition, sensor-based approaches have been used more than video-based approaches due to privacy concerns when placing cameras in human personal spaces. However, data about user behavior that is continuously measured and generated by user-friendly IoT devices (Iwasawa et al., 2017) allow adversary to infer private information about the user such as age, gender (Lu et al., 2013; Jain and Kanhangad, 2016), or possibly levels of health (Iwasawa et al., 2017; Chen et al., 2018).

Privacy issues become important because there is much software for automatic processing of big data, and this software is easily accessible to anyone. And, anyone can collect big data about human and human activities. The human can be me or my family. Therefore, it is necessary to study the privacy protection methods in the automatic recognition of human and human activities. This paper analyzes the cutting-edge research trends, techniques, and issues of privacy-preserving human and human activity recognition.

The rest of this paper is organized as follows: in the second section, the human recognition is analyzed in terms of its applications, its approaches, and its privacy vulnerability. In the third section, the human activity recognition is analyzed in terms of its applications, its approaches, and its privacy vulnerability. In the fourth section, privacy-preserving approaches to human and human activity recognition are discussed. In the fifth section, we conclude this paper and briefly discuss the possible future work directions.

Human recognition and privacy issues

Table 2 shows recent research on human recognition and related privacy issues. In the deep learning-based approaches, privacy issues exist.

Most human detection tasks are still based on visual images (Hwang et al., 2015). Many intelligent and complex video surveillance systems show a double-edged sword, high performance in detection, and privacy protection. When photos or images are recorded and processed by the surveillance system, the individuals and groups taken in the photos or images may be exposed unintentionally and analyzed differently from the system's original purpose (Chattopadhyay and Boult, 2007; Ren et al., 2016; Wu et al., 2018) as shown in Figure 4.

Table 2. Human recognition.

Research	Object	Recognition methods	Data	Application domain	Privacy issue
Chattopadhyay and Boult (2007) Wu et al. (2018) Facial Network, Ren et al. (2016)	Human, object	Deep learning	Image, video	Video surveillance	Public dataset privacy User data privacy
Song and Shmatikov (2020) Nelus and Martin (2019)	Human face	Deep learning	Image	Binary gender classification	Model privacy
Haris et al. (2014) Gajjar et al. (2017) Nike, Malinowski (2010)	Human location	Deep learning	Sensor data	Location-based services, mobile heath applications	Public dataset privacy User data privacy
Gomathisankaran et al. (2013) Wang et al. (2014) Ertin et al. (2011)	Human disease, human health	Deep learning	Clinical records, image	Medical care	Public dataset privacy User data privacy

Song and Shmatikov (2020) and Nelus and Martin (2019) introduced overlearning. That is, a model trained for a seemingly simple objective implicitly learns to recognize attributes and concepts sensitive from privacy. For example, a binary gender classifier of facial images also learns to recognize races – even races that are not represented in the training data – and identities.

Gajjar et al. (2017) detected and tracked human in video surveillance using histogram of oriented gradients (HOG) features. Nike and Malinowski (2010) track user's activities by using GPS and other sensors on mobile devices.

Gomathisankaran et al. (2013) considered privacy leakage during medical image processing. Wang et al. (2014) detected the mental health, performance, and behavioral trends of the students by using sensing data from the smartphone. Ertin et al. (2011) tried to understand the psychological state of the user in real time by using the sensors to record physiological data.

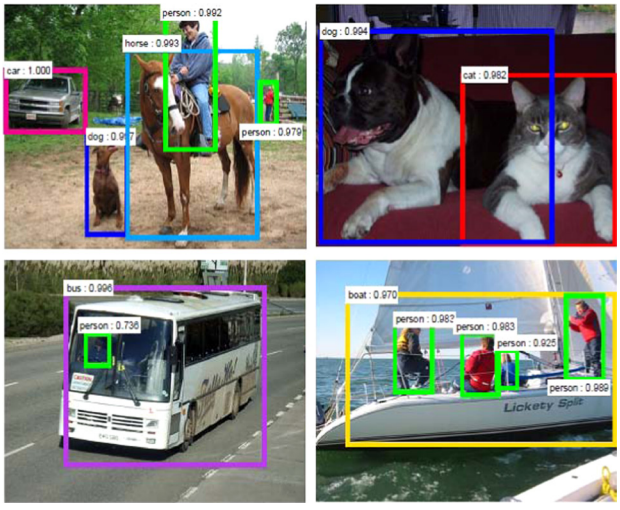

Figure 4: Object detection (Ren et al., 2016).

Human activity recognition and privacy issues

Table 3 shows recent research on human activity recognition and related privacy issues.

Iwasawa et al. (2017) showed that deep neural networks (DNN) can reveals user-discriminative features unintentionally. DNN has the black-box property, it is hard to predict what DNN learns from training data. In other words, DNN can learn about the user information, the application gets to disclose the information unintentionally without the user's consent. You et al. (2012) proposed Carsafe as an application that learns the driving behaviors of users by using the two cameras.

Chen et al. (2018), Hu et al. (2019), and Zhang et al. (2019) considered that the collected time series data are shared to infer the users' physical activities

Table 3. Human activity recognition.

Research	Object	Recognition methods	Data	Application domain	Privacy issue
Iwasawa et al. (2017) You et al. (2012)	Human activity	Deep learning	Sensor data from smart wearable devices	Daily activity investigation	Training data privacy User data privacy
Chen et al. (2018) Hu et al. (2019) Zhang et al. (2019)	Physical activities such as walking and running	Deep learning	Time series sensor data from smart wearable devices	Daily activity investigation	User data privacy
Phan et al. (2016)	Human activity	Deep learning	Physical activities, biomarkers, biometric measures	Health social network	Training data privacy

as shown in Figure 5, the personal information can also be inferred from the same data that is used for activity recognition. This is because people show characteristics of activity according to personal facts like age, gender, and so on (Lu et al., 2013; Jain and Kanhangad, 2016).

Phan et al. (2016) collected health social network data and considered privacy preservation.

Privacy-preserving approaches to human and human activity recognition

This section analyzes and summarizes privacy-preserving approaches in human and human activity recognition. Many studies utilize compound methods

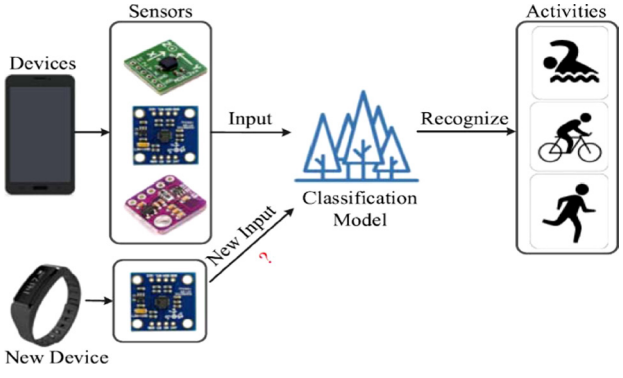

Figure 5: Sensor-based human activity recognition (Hu et al., 2019).

to protect human-related privacy. Recently, while utilizing deep learning in human and human activity recognition, the privacy-preserving approaches show the direction to address privacy leakage related to the characteristics of deep learning additionally. In Table 4, privacy issues and privacy-preserving approaches raised in recent studies are analyzed.

Garcia and Jacobs (2010) and Fontaine and Galand (2007) proposed complete data isolation using cryptography. Garcia and Jacobs (2010) and Gomathisankaran et al. (2013) adopted homomorphic encryption (HE). Sensitive data are processed encrypted and there is no information leakage in the process as shown in Figure 6 (El-Yahyaoui and Ech-Cherif El Kettani, 2019). There is an approach to encrypt sensitive areas in images or videos as shown in Figure 7 (Chattopadhyay and Boult, 2007).

On the other hand, anonymized videos are intentionally captured or processed to be in special low quality conditions that only allow for the recognition of some target events or activities (Butler et al., 2015; Dai et al., 2015; Ryoo et al., 2017; Ren et al., 2018) as shown in Figure 8. And, Winkler et al. (2014) introduced cartoon-like effects as shown in Figure 9. Speciale et al. (2019) protected confidential information about the captured 3D scene by lifting the map representation from a 3D point cloud to a 3D line cloud. In Figure 10, (a) shows that 3D point cloud reveals potentially confidential information in the scene. In contrast, (b) protects user privacy by concealing the scene geometry and preventing inversion attacks, while still enabling accurate and efficient localization.

Table 4. Privacy-preserving approaches.

Research	Privacy issue	Privacy-preserving approach	Protected object
Garcia and Jacobs (2010) Fontaine and Galand (2007) Gomathisankaran et al. (2013) Chattopadhyay and Boult (2007)	Private information leakage (Public dataset privacy) (User data privacy)	Cryptography	Private information (medical image, lifestyle, financial information, face, private location, biometric information, disease information), Human activity (daily life activity, movement)
Butler et al. (2015) Dai et al. (2015) Ryoo et al. (2017) Ren et al. (2018) Winkler et al. (2014) Speciale et al. (2019)	Private information leakage (Public dataset privacy)(User data privacy)	Anonymized videos	
Garcia Lopez et al. (2015)	Private information leakage from database (Public dataset privacy)(User data privacy)	Local processing	
Liu (2019) Bun and Steinke (2016)	Information leakage from large-scale database (Public dataset privacy)	Differential privacy	
Bian et al. (2020)	Information leakage in visual recognition(Public dataset privacy)(Training data privacy)	Secure inference by homomorphic encryptionSecret sharingHomomorphic convolution	
Iwasawa et al. (2017) Ajakan et al. (2015) Edwards and Storkey (2016) Malekzadeh et al. (2018, 2019) Osia et al. (2020)	Information disclosure by unintentional discriminating of user information during deep learning (Training data privacy)	Adversarial training	
Zhang et al. (2019)	Adversarial training which is effective on particular sensitive attributes (Training data privacy)	Image style transformation	
Phan et al. (2016) Abadi et al. (2016) Papernot et al. (2017)	Information leakage during deep learning (Training data privacy)	Differential privacy	
Tramèr et al. (2016) Wang and Gong (2018) Juuti et al. (2019) Kariyappa and Kariyappa (2019)	Information leakage during deep learning (Model privacy)	Analyze attacker's queries, Defense against attacks	

And, Garcia Lopez et al. (2015) proposed edge computing for processing isolation; processing is performed near the data collected as shown in Figure 11. Central computing is a technique for collecting data in a central data center and performing intensive processing. In contrast, edge computing is a technology that processes data from a user's device or near the point, where data are

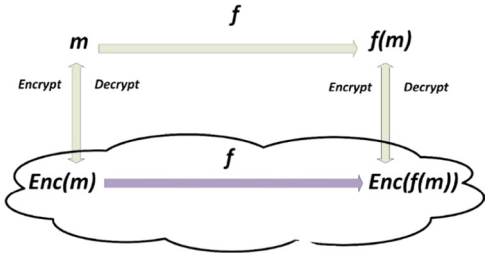

Figure 6: Homomorphic encryption (El-Yahyaoui and Ech-Cherif El Kettani, 2019).

Figure 9: A cartooning image (Winkler et al., 2014).

Figure 7: Sensitive areas are encrypted in the image (Chattopadhyay and Boult, 2007).

collected or where these are collected or generated. Because data are analyzed immediately at the edge, where data are collected and applied to the field, these are evaluated as a computing technology that can ensure immediate response and reliability rather than using a central data center such as the cloud (Xiao et al., 2019). That is, compared to central computing, edge computing supports a wide range of device mobility, has a low risk of data center hacking by distributed data processing, and has a short delay in data transmission and response for data processing. Central computing has an advantage in high-performance processing of big data, but edge computing is more efficient for applications that are sensitive to network failures or delays, such as autonomous vehicles, drones, or airplane engines.

Liu (2019) and Bun and Steinke (2016) provided a strong privacy guarantee by confusing a statistical query response drawn from a population-scale database by adding noise. They preserved that the presence or absence of a user in the database by differential privacy (Dwork, 2008) as shown in Figure 12 (Wood et al., 2018). By differential privacy, the general information for the entire population in a data set can be obtained without revealing individual information. In Figure 12, the difference between the analysis result for real-world data set and the analysis result for X's opt-out data set is at most ε. This means that private information can be shielded at statistical database analysis. It is based on ε-differential privacy (Dwork, 2008). By injecting random noise into the released statistical results computed from the underlying sensitive data, such that the distribution of the noisy results is relatively insensitive to any change of a single record in the original data set.

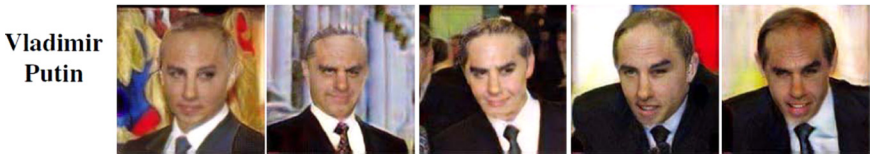

Figure 8: Anonymized images: different modified pictures of the same person (Ren et al., 2018).

Figure 10: A user privacy protection in image-based localization (Speciale et al., 2019).

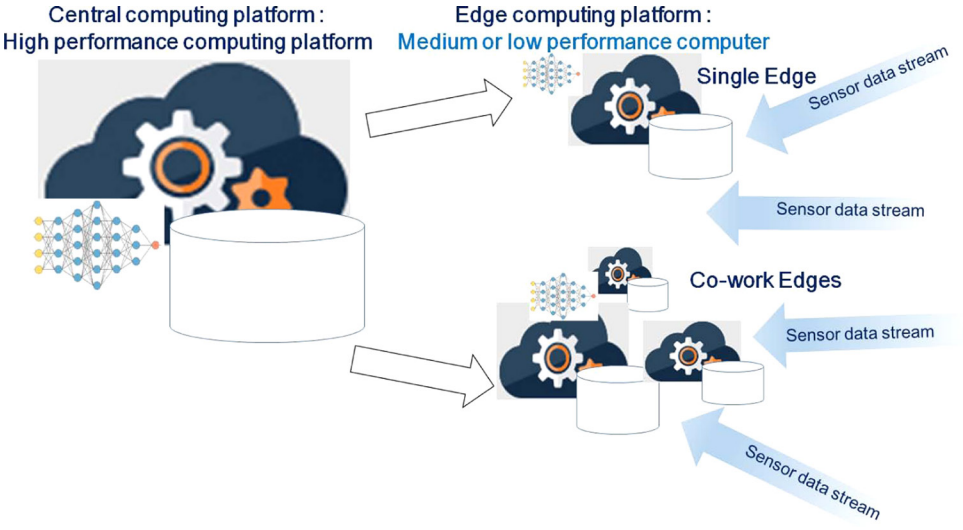

Figure 11: Central computing versus edge computing.

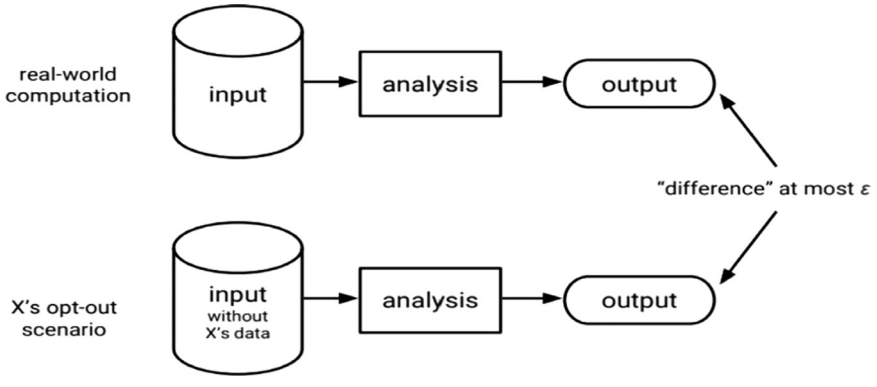

Figure 12: Differential privacy (Wood et al., 2018).

Bian et al. (2020) proposed a combined approach for privacy-preserving image recognition using homomorphic encryption, secret sharing protocol, and homomorphic convolution.

As privacy protection approaches for deep learning, Iwasawa et al. (2017), Malekzadeh et al. (2018, 2019), Ajakan et al. (2015), Edwards and Storkey (2016), and Osia et al. (2020) proposed adversarial trainings to suppress the information disclosure. Ajakan et al. (2015) and Edwards and Storkey (2016) introduced an adversarial training framework, domain-adversarial neural network (DANN), and adversarial learned fair representations (ALFR) to remove sensitive information from representations each. Adversarial training increases robustness by augmenting training data with adversarial examples because machine learning models are often vulnerable to adversarial examples, maliciously perturbed inputs designed to mislead a model at test time (Tramer et al., 2018). Iwasawa et al. (2017) proposed an adversarial training framework with information sources categorized into multiple features rather than binary features; DANN and ALFR used binary features. Malekzadeh et al. (2018, 2019) also proposed to integrate an adversarial loss with the standard activity classification loss. But, an adversarial loss function can only be used for protecting one kind of private information, such as user identity and gender. Iwasawa et al. (2017), Malekzadeh et al. (2018), and Osia et al. (2020) require the labels of private information for adversarial trainings. In Figure 13, z and y are sensitive variables. z-predictor just uses f_1, whereas y-remover uses both f_1 and f_2.

Zhang et al. (2019) adopted the image style transformation to protect all private information at once and maintain the desired information being inferred normally. The presented approach transforms raw sensor data into a new format that has a 'style' (sensitive information) of random noise and a 'content' (desired information) of the raw sensor data as shown in Figure 14. The pre-trained LossNet is used to define the loss functions that measure 'style' difference between transformed data and random noise and 'content' difference between transformed data and raw data.

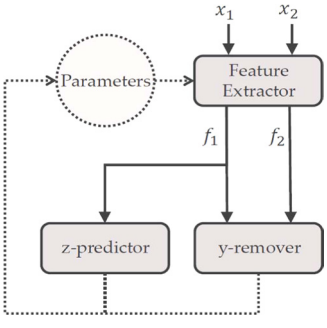

Figure 13: The private-feature extraction framework (Osia et al., 2020).

Phan et al. (2016), Abadi et al. (2016), and Papernot et al. (2017) used differential privacy for training data privacy. Instead of applying differential privacy to the query process for large statistical data sets, they used differential privacy during machine learning. Phan et al. (2016) proposed a privacy preservation encoder, deep private auto-encoder (dPA), by developing an ε-differential privacy-preserving deep learning model. That is, they enforced ε-differential privacy by perturbing the objective functions of the traditional deep auto-encoder. Abadi et al. (2016) proposed differentially private deep models and Papernot et al. (2017) utilized differential privacy, which is not specific to the learning model.

Tramèr et al. (2016) and Wang and Gong (2018) considered the learning model privacy. An adversary can infer the model parameters by making many queries to the learning model as shown in Figure 15. f is the train model data owner has. An attacker uses q queries to extract $\hat{f} \approx f$. Wang and Gong (2018) introduced hyper-parameter stealing attacks applicable to a variety of popular machine learning algorithms such as ridge regression, logistic regression, support vector machine, and neural network. Juuti et al. (2019) proposed PRADA to protect against the model stealing attack. It analyzes the distribution of consecutive API queries and raises an alarm when this distribution deviates from benign behavior. Kariyappa and Kariyappa (2019) substantially degrade the accuracy of the attacker's clone model by selectively sending incorrect predictions for attackers' queries.

The approaches in Table 4 protect privacy by securing the entire data set and data processing using encryption, anonymity, and isolation. For large statistical data sets, differential privacy is applied to prevent individual personal information from leaking. When neural network is adopted in data processing, adversarial training is proposed against adversarial examples misleading the learning models. Improving the effectiveness of privacy protection using data format conversion was also taken into account. In addition, differential privacy was applied to data and learning models by inserting noise to prevent the disclosure of private information during machine learning. There were approaches to prevent inference or theft of learning models in machine learning.

On the other hand, a membership inference attack was mentioned against machine learning models

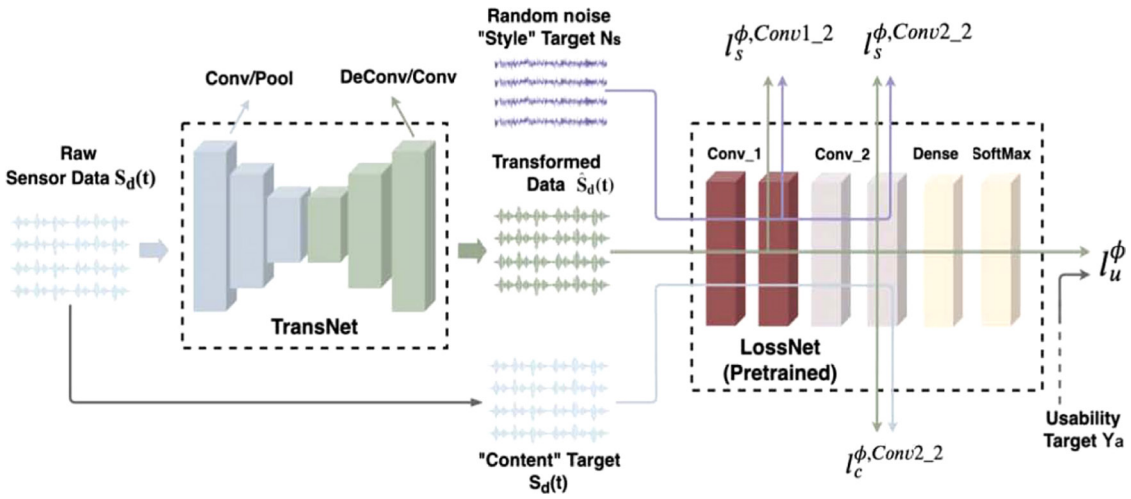

Figure 14: Collective protection of all sensitive information at once (Zhang et al., 2019).

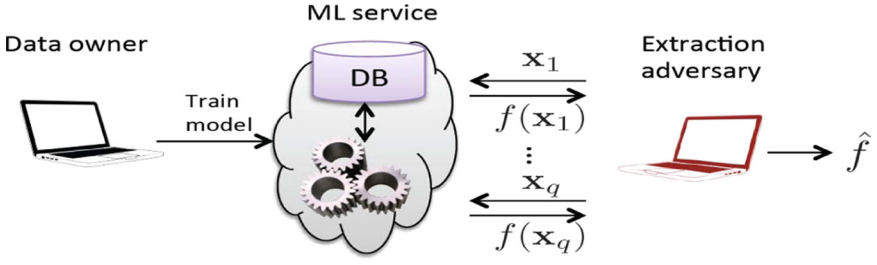

Figure 15: Model extraction attack (Wang and Gong, 2018).

(Shokri et al., 2017). At the membership inference, given a machine learning model and a record, an attacker can determine whether a record was used as part of the model's training dataset or not. Shokri et al. (2017) showed privacy can be breached by membership inference in supervised machine learning. They suggested overfitting reduction by regularization, and trivial structure in machine learning models to mitigate the privacy breach. Differentially private models are robust to this attack, but the models reduce prediction accuracy for small ε values. Adversary training makes learning models be robust by prohibiting the models from being biased to produce a certain result by adversarial attacks. Song et al. (2019) showed that adversarially trained models are vulnerable to membership inference attacks. Moreover, an increased robustness of the adversarially trained model is correlated with an increase in the success of the membership inference attack due to adversarial generalization. Nasr et al. (2018) and Hayes and Ohrimenko (2018) design privacy mechanisms to reduce adversarial generalization. However, member inference attacks require in-depth research related to adversarial training.

If the approach that controls privacy is static, it is difficult to ensure satisfactory privacy preservation for dynamic context-aware applications. Large amounts of sensor data and context-aware applications create new types of ambiguous privacy issues that make it difficult for users to determine sensitive data (Haris et al., 2014). Therefore, privacy control should be adjusted to the situation, and a method to protect privacy should be developed by adapting to data, application domain, and data processing technology.

Conclusion

In this study, we focused on privacy-preserving human and human activity recognition. With the development

of affordable, high-performance cameras and IoT devices equipped with various sensors, many applications are pouring out to provide convenience by analyzing various types of big data collected. Big data is difficult to find meaningful information until it is analyzed. However, the results analyzed by the development of computing technologies such as deep learning have new security problems. For human and human activity awareness, extracting information about a particular individual and his activities can cause problems that can infringe on privacy. Therefore, privacy-preserving approaches are important in human and human activity recognition. Because there is no single best solution for privacy protection, it should be studied in parallel with the expansion of deep learning applications. In this paper, privacy-preserving approaches and related issues were investigated in cutting-edge research.

Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education, Korea (2017R1D1A1B03034950).

Literature Cited

Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K. and Zhang, L. 2016. Deep learning with differential privacy. Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, ACM, pp. 308–318.

Ajakan, H., Germain, P., Larochelle, H., Laviolette, F. and Marchand, M. 2015. Domain-adversarial neural networks, available at: https://arxiv.org/abs/1412.4446

Bian, S., Wang, T., Hiromoto, M. and Shi, Y. 2020. ENSEI: efficient secure inference via frequency-domain homomorphic convolution for privacy-preserving visual recognition, available at: https://arxiv.org/abs/2003.05328

Bun, M. and Steinke, T. 2016. Concentrated differential privacy: simplifications, extensions, and lower bounds. Theory of Cryptography Conference, Springer, pp. 635–658.

Butler, D. J., Huang, J., Roesner, F. and Cakmak, M. 2015. The privacy utility tradeoff for remotely teleoperated robots. Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction, pp. 27–34.

Chattopadhyay, A. and Boult, T. E. 2007. Privacycam: a privacy preserving camera using uclinux on the blackfin DSP, 2007 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8.

Chen, K., Yao, L., Wang, X., Zhang, D., Gu, T., Yu, Z. and Yang, Z. 2018. Interpretable parallel recurrent neural networks with convolutional attentions for multi-modality activity modeling. IJCNN, IEEE, pp. 1–8.

Chen, K., Zhang, D., Yao, L., Guo, B., Yu, Z. and Liu, Y. 2020. Deep learning for sensor-based human activity recognition: overview, challenges and opportunities, available at: https://arxiv.org/abs/2001.07416

Cormode, G., Procopiuc, C. M., Srivastava, D. and Tran, T. T. L. 2012. Differentially private summaries for sparse data. International Conference on Database Theory, pp. 299–311.

Dai, J., Saghafi, B., Wu, J., Konrad, J. and Ishwar, P. 2015. Towards privacy-preserving recognition of human activities. 2015 IEEE International Conference on In Image Processing (ICIP), pp. 4238–4242.

Dwork, C. 2008. Differential privacy: a survey of results. TAMC, Springer, pp. 1–19.

Edwards, H. and Storkey, A. 2016. Censoring representations with an adversary. Proceedings of ICLR, San Juan, Puerto Rico, May 2–4.

El-Yahyaoui, A. and Ech-Cherif El Kettani, M. D. 2019. A verifiable fully homomorphic encryption scheme for cloud computing security. *Technologies* 7(21): 1–15.

Ertin, E., Stohs, N., Kumar, S., Raij, A., al'Absi, M. and Shah, S. 2011. Autosense: unobtrusively wearable sensor suite for inferring the onset, causality, and consequences of stress in the field. Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems, ACM, pp. 274–287.

FacialNetwork.com, Nametag application, available at: http: //www.nametag.ws/

Fontaine, C. and Galand, F. 2007. A survey of homomorphic encryption for nonspecialists. *EURASIP Journal on Information Security* 2007(1): 013801.

Gajjar, V., Khandhediya, Y. and Gurnani, A. 2017. Human detection and tracking for video surveillance a cognitive science approach. IEEE International Conference on Computer Vision Workshops (ICCVW), Venice, pp. 2805–2809, doi: 10.1109/ICCVW.2017.330.

Garcia, F. D. and Jacobs, B. 2010. Privacy-friendly energy-metering via homomorphic encryption. International Workshop on Security and Trust Management, Springer, pp. 226–238.

Garcia Lopez, P., Montresor, A., Epema, D., Datta, A., Higashino, T., Iamnitchi, A., Barcellos, M., Felber, P. and Riviere, E. 2015. Edge-centric computing: vision and challenges. *ACM SIGCOMM Computer Communication Review* 45(5): 37–42.

Gomathisankaran, M., Yuan, X. and Kamongi, P. 2013. Ensure privacy and security in the process of medical image analysis. 2013 IEEE International Conference on Granular Computing (GrC), pp. 120–125.

Haris, M., Haddadi, H. and Hui, P. 2014. Privacy leakage in mobile computing: tools, methods, and characteristics, available at: https://arxiv.org/abs/1410.4978.

Hayes, J. and Ohrimenko, O. 2018. Contamination attacks and mitigation in multi-party machine learning.

Conference on Neural Information Processing Systems (NeurIPS), pp. 6602–6614.

Hu, C., Chen, Y., Peng, X., Yu, H., Gao, C. and Hu, L. 2019. A novel feature incremental learning method for sensor-based activity recognition. *IEEE Transactions on Knowledge and Data Engineering* 31(6): 1038–1350.

Hwang, S., Park, J., Kim, N., Choi, Y. and Kweon, I. S. 2015. Multispectral pedestrian detection: benchmark dataset and baseline. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 7-12, pp. 1037–1045.

Iwasawa, Y., Nakayama, K., Yairi, I. E. and Matsuo, Y. 2017. Privacy issues regarding the application of DNNs to activity-recognition using wearables and its countermeasures by use of adversarial training. International Joint Conference on Artificial Intelligence (IJCAI-17), pp. 1930–1936.

Jain, A. and Kanhangad, V. 2016. Investigating gender recognition in smartphones using accelerometer and gyroscope sensor readings. ICCTICT, IEEE, pp. 597–602.

Juuti, M., Szyller, S., Marchal, S. and Asokan, N. 2019. PRADA: protecting against DNN model stealing attacks. 2019 IEEE European Symposium on Security and Privacy (EuroS&P), Stockholm, pp. 512–527.

Kariyappa, S. and Kariyappa, S. 2019. Defending against model stealing attacks with adaptive misinformation, available at: https://arxiv.org/abs/1911.07100

Liu, F. 2019. Generalized Gaussian mechanism for differential privacy. *IEEE TKDE* 31(4): 747–756.

Lu, J., Wang, G. and Moulin, P. 2013. Human identity and gender recognition from gait sequences with arbitrary walking directions. *IEEE TIFS* 9(1): 51–61.

Malekzadeh, M., Clegg, R. G., Cavallaro, A. and Haddadi, H. 2018. Protecting sensory data against sensitive inferences. Proceedings of the 1st Workshop on Privacy by Design in Distributed Systems, ACM, p. 2.

Malekzadeh, M., Clegg, R. G., Cavallaro, A. and Haddadi, H. 2019. Mobile sensor data anonymization. Proceedings of the International Conference on Internet of Things Design and Implementation, pp. 49–58.

Malinowski, E. (2010). Adidas miCoach app sets sights square on nike+. *Wired Magazine*, available at: https://www.wired.com/2010/08/adidas-micoach-app/

Melis, L., Song, C., De Cristofaro, E. and Shmatikov, V. 2018. Exploiting unintended feature leakage in collaborative learning, available at: https://arxiv.org/abs/1805.04049

Nasr, M., Shokri, R. and Houmansadr, A. 2018. Machine learning with membership privacy using adversarial regularization. ACM Conference on Computer and Communications Security (CCS), Toronto, Canada, October 15–19.

Nelus, A. and Martin, R. 2019. Privacy-aware feature extraction for gender discrimination versus speaker identification, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, May 12–17.

Nike, Nike+ running, available at: http://www.nike.com/us/enus/c/running/nikeplus/gps-app

Osia, S. A., Taheri, A., Shamsabadi, A. S., Katevas, K., Haddadi, H. and Rabiee, H. R. 2018. Deep private-feature extraction, available at: https://arxiv.org/abs/1802.03151

Osia, S. A., Taheri, A., Shamsabadi, A. S., Katevas, K., Haddadi, H. and Rabiee, H. R. 2020. Deep private-feature extraction. *IEEE Transactions on Knowledge and Data Engineering* 32(1): 54–66.

Papernot, N., Abadi, M., Erlingsson, U., Goodfellow, I. and Talwar, K. 2017. Semi-supervised knowledge transfer for deep learning from private training data. Proceedings of the International Conference on Learning Representations (ICLR), Toulon, France, April 24–26.

Phan, N., Wang, Y., Wu, X. and Dou, D. 2016. Differential privacy preservation for deep auto-encoders: an application of human behavior prediction. Thirtieth AAAI Conference on Artificial Intelligence, Phoenix, AZ, February 12–17.

Ren, S., He, K., Girshick, R. and Sun, J. 2016. Faster R-CNN: towards real-time object detection with region proposal networks, available at: https://arxiv.org/abs/1506.01497

Ren, Z., Lee, Y. J. and Ryoo, M. S. 2018. Learning to anonymize faces for privacy preserving action detection. European Conference on Computer Vision (ECCV), 620–636.

Ryoo, M. S., Rothrock, B., Fleming, C. and Yang, H. J. 2017. Privacy-preserving human activity recognition from extreme low resolution. AAAI Conference on Artificial Intelligence, San Francisco, CA, February 4–9.

Shokri, R., Stronati, M., Song, C. and Shmatikov, V. 2017. Membership inference attacks against machine learning models. 2017 IEEE Symposium on Security and Privacy (SP), pp. 3–18.

Song, C. and Shmatikov, V. 2020. Overlearning reveals sensitive attributes, Proceedings of International Conference on Learning Representations (ICLR). Virtual Conference, April 26–30.

Song, L., Shokri, R. and Mittal, P. 2019. Membership inference attacks against adversarially robust deep learning models. IEEE Security and Privacy Workshops (SPW).

Speciale, P., Schonberger, J. L., Kang, S. B., Sinha, S. N. and Pollefeys, M. 2019. Privacy preserving image-based localization. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5493–5503.

Tanuwidjaja, H. C., Choi, R. and Kim, K. 2019. A survey on deep learning techniques for privacy-preserving, machine learning for cyber security. ML4CS 2019. *Lecture Notes in Computer Science* 11806: 29–46.

Tramer, F., Kurakin, A., Papernot, N., Goodfellow, I., Boneh, D. and McDaniel, P. 2018. *Ensemble adversarial training: attacks and defenses* Proceedings of International Conference on Learning Representations (ICLR). Vancouver, Canada, April 30–May 3.

Tramèr, F., Zhang, F., Juels, A., Reiter, M. K. and Ristenpart, T. 2016. Stealing machine learning models

via prediction APIs. USENIX Security Symposium, pp. 601–618.

Wang, B. and Gong, N. Z. 2018. Stealing hyperparameters in machine learning. IEEE Symposium on Security and Privacy, Hyatt Regency, San Francisco, May 21–23.

Wang, J., Chen, Y., Hao, S., Peng, X. and Hu, L. 2019. Deep learning for sensor-based activity recognition: a survey. *Pattern Recognition Letters* 119: 3–11.

Wang, R., Chen, F., Chen, Z., Li, T., Harari, G., Tignor, S., Zhou, X., Ben-Zeev, D. and Campbell, A. T. 2014. Studentlife: assessing mental health, academic performance and behavioral trends of college students using smartphones. Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, ACM, pp. 3–14.

Winkler, T., Erd´elyi, A. and Rinner, B. 2014. TrustEYE. m4: protecting the sensor-not the camera. 2014 11th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), pp. 159–164.

Wood, A., Altman, M., Bembenek, A., Bun, M., Gaboardi, M., Honaker, J., Nissim, K., OBrien, D. R., Steinke, T. and Vadhan, S. 2018. Differential privacy: a primer for a non-technical audience. *Vanderbilt Journal of Entertainment & Technology Law* 21(1): 209–275.

Wu, Z., Wang, Z., Wang, Z. and Jin, H. 2018. Towards privacy-preserving visual recognition via adversarial training: a pilot study, Proceedings of the European Conference on Computer Vision (ECCV), pp. 606–624.

Xiao, Y., Jia, Y., Liu, C., Cheng, X., Yu, J. and Lv, W. 2019. Edge computing security: state of the art and challenges. *Proceedings of the IEEE* 107(8): 1608–1631.

You, C.-W., Montes-de Oca, M., Bao, T. J., Lane, N. D, Lu, H., Cardone, G., Torresani, L. and Campbell, A. T. 2012. Carsafe: a driver safety app that detects dangerous driving behavior using dual-cameras on smartphones. Proceedings of the 2012 ACM Conference on Ubiquitous Computing, ACM, pp. 671–672.

Zhang, D., Yao, L., Chen, K., Long, G. and Wang, S. 2019. Collective protection: preventing sensitive inferences via integrative transformation. 19th IEEE International Conference on Data Mining (ICDM), IEEE, pp. 1–6.