Workshops at 18th International Conference on Intelligent Environments (IE2022)
H.H. Alvarez Valera and M. Luštrek (Eds.)

doi:10.3233/AISE220038

Ensuring Privacy in Natural Language Processing: A PRISMA Oriented Literature Review

Sima ALAHYARI [a], Matthias HAFNER [a,1] and Silvia KNITTL [b]
[a] *Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany*
[b] *PwC GmbH WGP, Germany*

Abstract. Using Natural Language Processing (NLP) as a discipline of machine learning, organizations can better organize their data in order to better represent their internal knowledge. To make NLP models easier to re-use in other contexts, they should be protected accordingly. This raises the question of the appropriate privacy-ensuring technology (PET). To be able to address this question this paper conducts a literature review regarding the ensuring of privacy in NLP following the PRISMA framework. After following the identification process of significant sources, 22 valuable studies were selected. Some of these studies have shown promising results, however the field of privacy ensuring in NLP is still uncategorized with different approaches difficult to compare.

Keywords. Privacy Techniques, Natural Language Processing, NLP, PET

1. Introduction

Natural Language Processing (NLP) is a field in machine learning with a computer's ability to understand, analyze, manipulate, and potentially generate human language [1]. NLP is proliferating due to its wide range of usages in text analysis; however, it also enables highly sophisticated attacks as an NLP algorithm can identify and replicate the defining characteristics of an individual's communication patterns [2]. NLP's significant challenges involve the recognition of natural speech, the understanding of the spoken natural language, and in the last step, the generation of sensible natural language. Furthermore, privacy is one of the biggest concerns to train the data of sensitive information [3]. There is an effort and associated cost to train these models. In order to be able to further use these models in another context, it is important that private content can be identified from these models. For this, it is necessary that PET (privacy-ensuring technologies) can also be applied to these models. The goal of this paper is to identify the state of the art literature on how to ensure NLP's privacy to lay the foundations for future research. Considering that data privacy in NLP is an interdisciplinary field, reviewing all the implementations done on privacy-ensuring technologies is not in this study's scope. NLP implies all kinds of tools humans use to communicate with other humans,

[1]Corresponding Author: FAU, Erlangen-Nürnberg, Germany; E-mail: matthias.hafner@fau.de

such as speech, text, and images. Therefore we are focusing our research on textual data and have reviewed work done in this area. This paper's primary purpose is to be able to identify PET applied to textual datasets based on their applications on reviewed papers. Therefore, we exclude any PET on Non-NLP domains such as cloud computing. Another inclusion criterion is the studies in the area of Predictive Analysis, such as text classification and sentiment analysis, meaning we leave out the PET that are used for data verification. As an introduction to the topic, the next section 2 provides an overview of relevant technologies in the scope of our article. Then we present our research methodology in section 3, show the outcome of our analysis in section 4, discuss our findings in section 5, and close with a short summary of the paper in section 6.

2. Overview of Privacy Techniques

Pavlou or Smith et al. describe data or information privacy as a branch of data security where data handling from collection and storage of data to whether or how data is shared with third parties is involved [4,5]. The main techniques under investigation in our literature review are Federated Learning (FL), Anonymization, Differential Privacy (DP), Secure Multi-Party Computation (SMPC), Homomorphic Encryption (HE).

The reason of the selection of SMPC and HE for this research is that we aimed to widen our research results by means of reviewing not only anonymization techniques but also cryptography-based techniques. FL was selected since it preserves privacy during the gathering and analysis of the data and does not wait to be applied on the results only. DP is a perturbation-based technique which is one of the very most recent introduced PET. It was chosen since it can quantify the level of preserved privacy by using a magic number called Epsilon(ε). The selected techniques are briefly introduced in the following.

FL: FL is a decentralized machine learning approach for building a prediction model. FL enables the owner of the data collaboratively learn a shared prediction model while maintaining their ownership of data without sending them to a machine that they do not control [6]. FL works based on the general approach of "bringing the code to the data, instead of the data to the code" and aim the vital problems of privacy, ownership, and locality of data [7].

Anonymization: Anonymization deals with the automatic identification of sensitive information such as proper names, sensitive phrases, and numeric values e.g., credit card number in a given text document, and remove or replace them with dummy entities. The purpose of this technique is to hide the sensitive information and allow the text to be used for analysis purposes [8].

DP: DP was introduced by Dwork, Dinur, Nassim, Smith and others (c. f. [9,10,11]) and got several expansions like in [12]. Desfontaines and Pejó provide an elaborate taxonomy of DP in [13]. DP guarantees that the delivery of any analysis on a database is not affected significantly by the entity of any individual. It is therefore difficult for an adversary to extract any individual data row. The idea revolves around the popular indistinguishability concept in semantic security. The value for ε is considered to be an indicator of the level of required privacy protection while also affecting utility or accuracy of this privacy mechanism. Smaller values of ε indicate a stronger privacy protection; because the scenario where the output is generated without the input data of each individual 'X' is mimicking the real world scenario. Smaller ε values come with less accuracy [14].

SMPC: SMPC is another cryptography approach. SMPC allows multiple parties to implement a joint computation on a series of inputs without revealing their own input [15]. Based on each private data multiple parties cooperate in order to perform computations [16]. They compute a function $f(\alpha 1, .., \alpha n)$. The $\alpha 1, .., \alpha n$ are private values and the result $\alpha = f(\alpha 1, .., \alpha n)$ is revealed to everyone. To foster the privacy of every parties' input secret sharing is used [17]. Each party distributes shares of her secret among all other parties in the sharing phase. In the computation phase, addition and multiplication are used as the two secure operations. The case with two secrets would be: Addition, i.e. each party locally adds shares of the two secrets and multiplication, i.e. that the party locally multiplies shares of the two secrets by each other. Then further computations allow the collaboration. In the reconstruction step, the parties reveal the shares of the function value so that they all learn it (see also [18]).

HE: HE is a cryptography-based technique which works on distributed databases [19]. An encryption scheme consists of a method of transforming plaintexts into ciphertexts (encrypted texts) and vice versa, by means of proper keys. These keys are essential to make these transformations effective. Formally, these transformations are performed by corresponding algorithms: an encryption algorithm that transforms a given plaintext and an adequate (encryption) key into a corresponding ciphertext, and a decryption algorithm that given the ciphertext and an adequate (decryption) key recovers the original plaintext. As a third algorithm the probabilistic algorithm used to generate keys (i.e., a key-generation algorithm) needs consideration. This algorithm must be probabilistic (or else, by invoking it, the adversary obtains the very same key used by the receiver). We stress that the encryption scheme itself (i.e., the aforementioned three algorithms) may be known to the adversary, and the scheme's security relies on the hypothesis that the adversary does not know the actual keys in use [20].

Fully homomorphic encryption enables computation on encrypted data without leaking any information about the underlying data and thus, enables preserving privacy also for NLP algorithms. A party can encrypt some input data, while another party that does not have access to the decryption key can blindly perform some computation on this encrypted input. The final result is also encrypted, and it can be recovered only by the party that possesses the secret key [21].

3. Method

We have integrated the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) framework [22] for the literature study. The PRISMA protocol primary was introduced to provide a framework for a systematic review of medical publications to support researchers to identify, select, synthesize and summarize articles in a way that avoids any potential bias. We adapted this framework with some minor modifications, i. e. we followed the systematic reviews and not the meta-analysis since we did not need to perform quantitative synthesis for our research as it is conducted for medical records.

Our study has two main components: the PRISMA checklist and PRISMA statement. The PRISMA Statement consists of a 27-item checklist and a four-phase flow diagram. The checklist includes items deemed essential for transparent reporting of a systematic review. When reviewers selectively choose which information to include in a

report based on the direction and significance of findings, they risk biasing the readers with their personal opinions [23]. Therefore, we followed PRISMA in order to avoid any personal bias regarding the selection of the most relevant previous studies in this area of research. Our process of literature selection for inclusion consisted of four steps: (1) Selection of the literature sources, (2) Selection of the keywords, (3) Definition of inclusion criteria, and (4) Selected articles.

(1) Selection of literature sources: As a starting point for our selection of the literature sources we looked in the top conferences and journals in the field of data and cyber security. The Top Journals and Conferences we considered are TDSC: IEEE Transactions on Dependable and Secure Computing and TOPS: ACM Transactions on Privacy and Security. Additionally, we looked as well in the viable web search engines: Institute of Electrical and Electronics Engineers (IEEE), Elsevier's abstract and citation database (Scopus) and Google Scholar.

(2) Selection of keywords: To be able to compare the use of privacy-preserving techniques in the domain of textual data, as a next step, we restrict our study with the help of keyword search. Our scope is to review literature for the use of PET on textual databases, namely unstructured or semi-structured databases (only text and not images, speech, etc.), and we do not consider their applications on numerical datasets namely structured databases. We performed a database search by looking for the keywords of the above introduced techniques: Differential Privacy, Homomorphic Encryption, Secure Multiparty Computation, Anonymization and Federated Learning in combination with the second keyword "Text".

The outcome of this phase was the number of articles found in the mentioned literature resources containing the defined keywords, which resulted in 265 articles in total. The Figure 1 illustrates the overview of the article selection process and outcome.

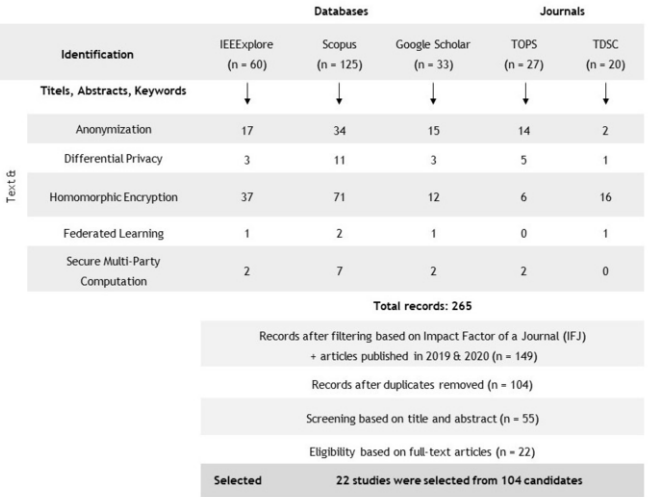

Figure 1. Overview of selection process

(3) Definition of inclusion criteria: In the next step the most relevant articles were identified by defining inclusion criteria. Therefore, we defined the time frame (2010 - 02/2020), the language (English) and the impact factor ("$IFJ \geq 3$"). In the next step

duplicates were removed. As a result, 104 unique data records are left. Subsequently, we carefully read the abstract of each article precisely and performed the secondary evaluation of the articles. The most relevant articles are included, and the other records which did not contain these criteria were excluded. The total number of studies after this step was 55 records.

In this section we have outlined the approach of our literature study. According to [24], an effective literature review provides a new basis for advancing knowledge and identifying areas where further research is needed. Our findings for both areas are presented in the next section.

4. Result

The Table 1 outlines the concept matrix of our results. The conducted literature review indicated that applications of some PET, such as Anonymization and DP on textual data, are addressed in the literature; however, the implementations of Cryptography-based PET on textual data still need to be covered in the scientific research. Regarding FL, initially, the combination of this method with the term "textual data" or "text" returned few results; however, after full-text reading of the found articles, it turned out these five initially found articles have no empirical research done on textual data, therefore, none was relevant to the scope of our research. FL is not categorized in PET, although, it is considered a method to enhance privacy while performing an ML task on distributed databases. In the last phase of our research, we detected 22 of the most relevant contributions out of 265 records that initially met our criteria. After reading through these final selected articles, we summarized the highlights of these papers based on three metrics[attributes]: privacy, utility, and execution time. Data utility is a measure of information loss or loss in the functionality of data in providing the results and privacy is protecting the safety of individual data or sensitive knowledge without sacrificing the utility of the data [25]; however, PET should provide privacy as long as there is a meaningful trade-off between privacy and utility. These metrics were identified and extracted from the studied papers as the measurements to make us able to compare PET.

Anonymization on Textual Data: One of the benefits of Anonymization is that it is scalable toward privacy preservation, it means we can anonymize a sensitive attribute in a way that it reveals a part of the information or reveals nothing. Regarding the preserved utility, there is always a matter of fact that more privacy with this technique leads to less accuracy. Therefore, one has to find a meaningful trade-off between these two attributes. The execution time was not understudied, which means this technique is efficient regarding resource utilization.

DP on Textual Data: The significant advantage of DP is the parameter ε that shows the provided level of privacy. There is no experimental guide on setting ε as it strongly depends on the dataset; however, it was observed that if DP is applied on raw data, ε is set more close to zero, but if DP is applied on the word embedding, normally, it is set between 10 and 30. Selected studies show that the utility of the data after DP applications completely depends on the value of ε. Feyisetan et al. indicate that $\varepsilon = 0.125$ leads to decreased utility, and $\varepsilon = 8$ leads to higher accuracy scores [39]. The execution time was not understudied, which means this technique is efficient regarding resource utilization.

HE on Textual Data: HE provides an entirely secure process of text-data analysis with an almost intact utility of the data.But what can encryption do to preserve privacy

in applying ML or NLP algorithms? The answer is that fully homomorphic encryption enables computation on encrypted data (e.g., ciphertexts) without leaking any information about the underlying data. briefly, a party can encrypt some input data, while another party that does not have access to the decryption key can blindly perform some computation on this encrypted input. The final result is also encrypted, and it can be recovered only by the party that possesses the secret key [21]. However, the biggest problem is its

Author(s)	Year	Anonymization	Differential Privacy	Homomorphic Encryption	Federated Learning	Secure Multiparty Computation
Anandan et al. [26]	2012	X				
Kim et al. [27]	2014	X				
Rahmani et al. [28]	2014			X		
Mamede et al. [29]	2016	X				
Maeda et al. [30]	2016	X				
Li and Qin [31]	2017	X				
Costantino et al. [32]	2017			X		
Coavoux et al. [33]	2018		X			
Weggenmann and Kerschbaum [34]	2018		X			
Naqvi et al. [35]	2018			X		
Darivandpour and Atallah [36]	2018					X
Fernandes et al. [37]	2019		X			
Zhao et al. [38]	2019		X			
Feyisetan et al. [39]	2019		X			
Beigi et al. [40]	2019		X			
Mosallanezhad et al. [41]	2019	X				
Jiang et al. [42]	2019	X				
Hassan et al. [43]	2019	X				
Al Badawi et al.. [44]	2020			X		
De Cock et al. [45]	2019					X
Romanov et al. [46]	2020	X				
Feyisetan et al. [47]	2020		X			
Sum		**9**	**7**	**4**	**0**	**2**

Table 1. Concept Matrix for Privacy-Preserving Methods Literature Review

computational time. Costantino et al. perform a classification task on tweets and emails; their study shows that 19 minutes is needed to analyze a single tweet and 78 minutes to analyze an email, which is considered as extremely high [32].

SMPC on Textual Data: As SMPC is another type of cryptography-based PET, it provides an entirely secure process of data analysis and an intact data utility. As these techniques have the same characters, implementing SMPC needs a high amount of time.

5. Discussion

Since data privacy in NLP is an interdisciplinary field, it consists of a wide range of research and literature. Applying a systematic review helped us study the state-of-the-art of PET on document-based databases and reviewed the history of these methods' applications on real-world data sets. Although some studies have shown promising results, applying PET for NLP tasks is still infancy. There is much to be done, especially for cryptography-based methods, since they provide the entire secure process with remaining the privacy of the data intact for predictive analysis as it showed in the identified works such as [28], the high computational time is still their significant disadvantage which can be examined in more research. During our research, we faced some challenges that can be a potential for future research:

Modified field of application: We searched for the keywords of PET and text; however, after reading the abstracts of 55 matching articles, it appeared that 19 records are concerned with the implementations of PET on other domains such as cloud computing and not on NLP; therefore, they are out of this research scope.

New field of application: Training models with FL are increasing, especially among mobile devices. It has opened a new area of research, for instance, on personalized word suggestions, however, we could not find articles examining PET on FL, and we suggest this to be a possible future research domain.

Hybrid mechanism: A hybrid mechanism consisting of a suitable combination of two or more PET is used or could be used in the real-world. However, investigating and studying the combinations of these technologies was beyond the scope of this research but should be considered in the future.

Comparative studies: Another challenge was that most of the studies introduced their approaches or systems with the respective utility and privacy measures, making it difficult for us to compare the results based on unified metrics. Conducting more comparative studies that allow us to have and compare the metrics can be potential future research.

Comparison metrics: Furthermore, we defined three metrics, namely, guaranteed privacy, level of preserved utility, and the execution time. These metrics act as indicators and allow us to evaluate a given privacy-preserving technique. However, it can be said that there is no privacy-preserving technique that can excel in all the metrics and overcomes the other methods and protrude. However, the underlying literature analysis is a good start for further research in this field to come.

6. Summary

Summing up, the goal of this paper was to identify literature of implementations done on PET applied to textual datasets and to introduce a taxonomy that is fitting the content. We achieved these via integration of the PRISMA protocol. After matching and reading through our findings, a total number of 22 studies which fit our context were evaluated in detail and tested for our metrics (privacy, utility, execution time). At the middle of our research we noticed that two PETs are not suitable for applying on textual data, therefore, we crossed them out from the further work. We identified four different implementations for PET, namely anonymization on textual data, differential privacy on textual data, homomorphic encryption on textual data, and secure multiparty computation on textual data. It is to say that none of these given privacy-preserving techniques can excel in all the metrics. Hence, some future research and work has to be done in this area specifically regarding cryptographic methods such as Homomorphic encryption and SMPC since they still have high potential and provide high security level. However, our developed concept matrix for PET in NLP can function as a good foundation for the work to come.

Critical appreciation of the work: Our primary objective was to investigate the latest technologies and their applications. Therefore, we limited our time frame to the last ten years, which can and should be extended in future research, due to fast evolution of technologies. A significant barrier for our research was how to compare techniques based on the defined metrics. They were not comparable as each technique has its own measurement of provided privacy or the level of preserved utility. For instance, DP provides Epsilon which represents the level of trade-off between privacy and utility, whereas the other techniques do not provide a specific number for it. Since the available information in the future will constantly grow in size and complexity, it is therefore suggested that future work looks at defining appropriate metrics and applying them to the respective PETs. As a starting point for this, the attributes identified in this paper ”privacy”, ”utility” and ”execution time” can be used and should be further enhanced.

References

[1] Shetty B. Natural Language Processing(NLP) for Machine Learning. Towards data science. 2018 11. Available from: https://towardsdatascience.com/natural-language-processing-nlp-for-machine-learning-d44498845d5b.

[2] Shropshire J. Natural language processing as a weapon. In: Proceedings of the 13th Pre-ICIS Workshop on Information Security and Privacy. vol. 1; 2018. .

[3] Xu H, Gupta S. The effects of privacy concerns and personal innovativeness on potential and experienced customers' adoption of location-based services. Electronic Markets. 2009;19(2-3):137-49.

[4] Smith H, Dinev T, Xu H. Information Privacy Research: An Interdisciplinary Review. MIS Quarterly. 2011 12;35:989-1015.

[5] Pavlou P. State of the Information Privacy Literature: Where are We Now And Where Should We Go? MIS Quarterly. 2011 12;35:977-88.

[6] McMahan B, Ramage D. Federated Learning: Collaborative Machine Learning Without Centralized Training Data; 2017. Available from: https://ai.googleblog.com/2017/04/federated-learning-collaborative.html.

[7] Bonawitz K, Eichner H, Grieskamp W, Huba D, Ingerman A, Ivanov V, et al. Towards Federated Learning at Scale: System Design. CoRR. 2019. Available from: http://arxiv.org/abs/1902.01046.

[8] Saygin Y, Hakkini-Tur D, Tur G. Sanitization and anonymization of document repositories. In: Web and information security. IGI Global; 2006. p. 133-48.
[9] Dinur I, Nissim K. Revealing Information While Preserving Privacy. In: Proceedings of the Twenty-Second ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems. PODS '03. New York, NY, USA: Association for Computing Machinery; 2003. p. 202–210.
[10] Dwork C. Differential privacy: A survey of results. In: International conference on theory and applications of models of computation. Springer; 2008. p. 1-19.
[11] Dwork C, McSherry F, Nissim K, Smith A. Calibrating noise to sensitivity in private data analysis. In: Theory of cryptography conference. Springer; 2006. p. 265-84.
[12] Jain P, Gyanchandani M, Khare N. Differential privacy: its technological prescriptive using big data. Journal of Big Data. 2018;5(1):15.
[13] Desfontaines D, Pejó B. SoK: Differential privacies. Proceedings on Privacy Enhancing Technologies. 2020;2:288-313.
[14] Wood A, Altman M, Bembenek A, Bun M, Gaboardi M, Honaker J, et al. Differential privacy: A primer for a non-technical audience. Vand J Ent & Tech L. 2018;21:209.
[15] Malkhi D, Nisan N, Pinkas B, Sella Y, et al. Fairplay-Secure Two-Party Computation System. In: USENIX Security Symposium. vol. 4. San Diego, CA, USA; 2004. p. 9.
[16] Du W, Atallah MJ. Secure Multi-Party Computation Problems and Their Applications: A Review and Open Problems. In: Proceedings of the 2001 Workshop on New Security Paradigms. NSPW '01. New York, NY, USA: Association for Computing Machinery; 2001. p. 13–22.
[17] Gennaro R, Ishai Y, Kushilevitz E, Rabin T. The round complexity of verifiable secret sharing and secure multicast. In: Proceedings of the thirty-third annual ACM symposium on Theory of computing; 2001. p. 580-9.
[18] Nojoumian, Mehrdad. Novel Secret Sharing and Commitment Schemes for Cryptographic Applications. University of Waterloo; 2012. Available from: http://hdl.handle.net/10012/6858.
[19] Bairagi SI, Jawandiya PM. Cloud Computing: Ensuring Data Storage Security in Cloud. International Journal of Engineering Development and Research. 2016;4:1141-9.
[20] Goldreich O. Foundations of cryptography: volume 2, basic applications. Cambridge university press; 2009.
[21] Minelli M. Fully Homomorphic Encryption for Machine Learning. Paris, France: PSL Research University; 2018. Available from: https://tel.archives-ouvertes.fr/tel-01918263v2/document.
[22] Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg. 2010;8(5):336-41.
[23] Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. Bmj. 2015;349.
[24] Webster J, Watson RT. Analyzing the past to prepare for the future: Writing a literature review. MIS quarterly. 2002:xiii-xxiii.
[25] Majid Bashir Malik MAG, Ali R. Privacy Preserving Data Mining Techniques: Current Scenario and Future Prospects. 2012 Third International Conference on Computer and Communication Technology. 2012:26-32.
[26] Anandan B, Clifton C, Wei Jiang MM, Pastrana-Camacho P, Si L. t-Plausibility: Generalizing Words to Desensitize Text. Transactions on Data Privacy. 2012 12;5:505–534.
[27] Kim SH, Kwon DG, Cho HG. Privacy-enhanced string matching with wordwise positional sampling. In: Proceedings of the 8th International Conference on Ubiquitous Information Management and Communication; 2014. p. 1-8.
[28] Rahmani A, Amine A, Mohamed RH. A Multilayer Evolutionary Homomorphic Encryption Approach for Privacy Preserving over Big Data. In: 2014 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery; 2014. p. 19-26.
[29] Mamede N, Baptista J, Dias F. Automated anonymization of text documents. In: 2016 IEEE Congress on Evolutionary Computation (CEC); 2016. p. 1287-94.
[30] Maeda W, Suzuki Y, Nakamura S. Fast text anonymization using k-anonyminity. In: Proceedings of the 18th International Conference on Information Integration and Web-based Applications and Services; 2016. p. 340-4.
[31] Li XB, Qin J. Anonymizing and sharing medical text records. Information Systems Research. 2017;28(2):332-52.

[32] Costantino G, La Marra A, Martinelli F, Saracino A, Sheikhalishahi M. Privacy-Preserving Text Mining as a Service. IEEE Symposium on Computers and Communications (ISCC). 2017:890-7.
[33] Coavoux M, Narayan S, Cohen SB. Privacy-preserving Neural Representations of Text. In: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Brussels, Belgium: Association for Computational Linguistics; 2018. p. 1-10.
[34] Weggenmann B, Kerschbaum F. Syntf: Synthetic and differentially private term frequency vectors for privacy-preserving text mining. In: The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval; 2018. p. 305-14.
[35] Naqvi N, Abbasi AT, Hussain R, Khan MA, Ahmad B. Multilayer partially homomorphic encryption text steganography (MLPHE-TS): a zero steganography approach. Wireless Personal Communications. 2018;103(2):1563-85.
[36] Darivandpour J, Atallah MJ. Efficient and secure pattern matching with wildcards using lightweight cryptography. Computers and Security. 2018 8;77:666-74.
[37] Fernandes N, Dras M, McIver A. Generalised Differential Privacy for Text Document Processing. In: Nielson F, Sands D, editors. Principles of Security and Trust. Cham: Springer International Publishing; 2019. p. 123-48.
[38] Zhao F, Ren X, Yang S, Yang X. On Privacy Protection of Latent Dirichlet Allocation Model Training. CoRR. 2019;abs/1906.01178. Available from: http://arxiv.org/abs/1906.01178.
[39] Feyisetan O, Diethe T, Drake T. Leveraging hierarchical representations for preserving privacy and utility in text. In: 2019 IEEE International Conference on Data Mining (ICDM). IEEE; 2019. p. 210-9.
[40] Beigi G, Shu K, Guo R, Wang S, Liu H. I Am Not What I Write: Privacy Preserving Text Representation Learning. CoRR. 2019;abs/1907.03189. Available from: http://arxiv.org/abs/1907.03189.
[41] Mosallanezhad A, Beigi G, Liu H. Deep Reinforcement Learning-based Text Anonymization against Private-Attribute Inference. In: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). Hong Kong, China: Association for Computational Linguistics; 2019. p. 2360-9.
[42] Jiang D, Shen Y, Chen S, Tang B, Wang X, Chen Q, et al. A Deep Learning-Based System for the MEDDOCAN Task. In: IberLEF@ SEPLN; 2019. p. 761-7.
[43] Hassan F, Sánchez D, Soria-Comas J, Domingo-Ferrer J. Automatic Anonymization of Textual Documents: Detecting Sensitive Information via Word Embeddings. In: 2019 18th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/13th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE). IEEE; 2019. p. 358-65.
[44] Al Badawi A, Hoang L, Mun CF, Laine K, Aung KMM. Privft: Private and fast text classification with homomorphic encryption. IEEE Access. 2020;8:226544-56.
[45] De Cock M, Dowsley R, Nascimento AC, Reich D, Todoki A. Privacy-Preserving Classification of Personal Text Messages with Secure Multi-Party Computation: An Application to Hate-Speech Detection. Advances in Neural Information Processing Systems. 2019 06;32:3752-64.
[46] Romanov A, Kurtukova A, Fedotova A, Meshcheryakov R. Natural text anonymization using universal transformer with a self-attention. CEUR-WS: 3rd International Conference on R Piotrowski's Readings in Language Engineering and Applied Linguistics, PRLEAL 2019. 2020;2552:22-37.
[47] Feyisetan O, Balle B, Drake T, Diethe T. Privacy-and utility-preserving textual analysis via calibrated multivariate perturbations. In: Proceedings of the 13th International Conference on Web Search and Data Mining; 2020. p. 178-86.