EXTENDING STATISTICAL DISCLOSURE CONTROL MODELS WITH ADAPTIVE METRICS AND EVALUATION FRAMEWORK

Abstract

1. Motivation

Frequent occurrences of sensitive data leakage from systems with high-risk profiles have shown that even, or perhaps primarily, large organizations whose investments in IT assets protection are certainly substantial, must deal with critical security issues which not only affect their finances but also their reputation and may lead to legal issues. As many predicted in the past, currently we live in a landscape in which the question: *will there be an attack on the organization's resources* has currently changed into the question: *when will an attack on the organization's resources take place?* Currently, the problem of data leaks challenges not only the classical understanding of the leak of the whole or a part of a database, sensitive organizational information, or internal documents, but also errors in the process of anonymization and obfuscation of data and inadequate methods of securing statistical data collected in official surveys as well as those collected in user profiling processes.

The analysis of the previous and current works in the area of data privacy shows that there are many singular solutions for individual data privacy breaches, but what lacks is a root cause analysis with a holistic understanding of handling statistical data security during the whole lifespan of a system. A modern approach ruled by *security-by-design* and *security-in-depth* premises implicates applying protection measures right at the beginning of the system's planning phase is conceptually bulletproof, however in information-based systems in which data volume increases and characteristics fluctuate significantly over time, it is an almost impossible task to design a future-proof database architecture. The complexity of such a task significantly increases in scenarios in which the datasets are meant to be interchanged between multiple entities, e.g., the data is sent for statistical analyses from the original data owner to an external entity.

The *security-by-design* theoretically could be applied for newly developed information-driven systems with dynamic data sets by designing and applying a dedicated sensitive disclosure control method. However, the bigger issue would be designing such a disclosure control method that could be effectively applied to satisfy the data privacy in the currently existing systems storing present and historical data. This work intents to elaborate a security mechanism that could be treated as a statistical disclosure control method that would remain effective in terms of protecting data privacy even for already existing databases in the maintenance phase, not only for those which are still in the planning phase and can easily employ and adapt *security-by-design* measures.

2. Problem

Existing solutions for protecting information-driven production environments processing statistical data revolve around either anonymization of the data which, as has been proven in my research does not always provide a satisfying level of privacy if used incorrectly, or strongly perturbative methods which affect the statistical data in such a way that they lack the accuracy needed for further analyses.

One of the most widely used and most dynamically evolving perturbative statistical disclosure control models which intend to tackle statistical disclosure is ***differential privacy***. According to the author of the concept (Dwork & Roth, The Algorithmic Foundations of Differential Privacy, 2014), the model she developed is a concept, not a complete solution. Because of that, the base model needs to be extended to ensure:

- minimizing the risk associated with inference attacks that exploit several data sets as well as historical and future changes in data sets,
- quantification method of the privacy loss, and thus determining the level of security of the method used,
- adaptability for cases in which the original model is too strong.

The research covered in this thesis is primarily aimed at addressing those aspects of the original differential privacy model by extending it with additional components – both – algorithmic and procedural. The research intents to extend the possibilities of application of the differential privacy in changing data environments and propose adaptive methods which can improve the tradeoff between the accuracy of the results at the same time maintaining the privacy of the records with the use of this perturbative statistical disclosure control method.

3. Decreasing the inference attacks risk

The elaborated method is based on the ***differential-privacy*** model, but because it must satisfy the privacy requirements of a particular database and the accuracy of the retrieved data additional factors are introduced and should be used together to provide the best results in terms of privacy and accuracy:

- ***Risk-accuracy metric*** – an adjustable metric that allowed for the analysts to manipulate the statistical noise in such a way that the result would provide a selected level of a tradeoff between the accuracy and the security, which could be easily adapted for the real-life usages of statistical databases, e.g., the different tradeoffs for data processed only internally by an entity and sent to external entities or providing role-based tradeoffs at the database level
- ***Information score metric*** – an incremental metric that by design automatically changes the noise of the same or similar query retrieved multiple times. This metric intends to mitigate the risk of statistical disclosure in case of dynamic inference attacks while the results remain statistically integral. This factor potentially can be expanded to provide a multi-source inference control mechanism, i.e., in attack scenarios where more than one database is used.

The elaborated method assumes that neither the data collection method must apply any statistical control, nor the data model of the database must satisfy the statistical *security-by-design* measures. The method must work as a part of the statistics derivation method which employs the raw data sets from the database. This way it can be easily applied as an additional end-user interface for existing databases during a retrofitting process or incorporated into a database system in case of the systems yet in the planning phase.

4. Evaluation

The efficiency of the elaborated method was determined by the threat modeling an the statistical security criteria defined in the VIOLAS Framework, elaborated in the process of this research:

- ***Statistical confidentiality*** - statistical confidentiality is determined by evaluating the risk of the primary and the secondary data identification.
- ***Statistical integrity*** - statistical integrity is achieved when the retrieved results from different queries asking for the same data are statistically identical.
- ***Statistical accuracy*** - statistical accuracy is a measure of the quality of the retrieved statistics.
- ***Statistical transparency*** - an additional measure is privacy transparency which can be defined as protecting the selected critical data sets by elimination from the results.

As a result of the analysis, it was possible to establish the criteria by which the extended model provides better results than the classically approached differential privacy.

Selected references:

1. Denning, D. (1982). *Cryptography and Data Security.* Boston: Addison-Wesley Publishing Company, Inc.
2. Dwork, C. (2006). Differential Privacy. *In Proceedings of the 33rd international conference on Automata, Languages and Programming - Volume Part II (ICALP'06)*, 1-12.
3. Dwork, C., & Roth, A. (2014). The Algorithmic Foundations of Differential Privacy. *Foundations and Trends in Theoretical Computer Science, 9*(3-4 (2014)), 211-407.
4. Dwork, C., Kenthapadi, K., McSherry, F., Mironov, I., & Naor, M. (2006). Our data, ourselves: Privacy via distributed noise generation. *In Advances in Cryptology-EUROCRYPT 2006*, 486–503.
5. Dwork, C., Kenthapadi, K., McSherry, F., Mironov, I., & Naor, M. (2006). Our data, ourselves: Privacy via distributed noise generation. *Advances in Cryptology-EUROCRYPT*, 486–503.
6. Dwork, C., McShelly, F., Nissim, K., & Smith, A. (2006). Calibrating Noise to Sensitivity in Private Data Analysis. *Halevi S., Rabin T. (eds) Theory of Cryptography. TCC 2006. Lecture Notes in Computer Science, vol 3876. Springer, Berlin, Heidelberg*, pp. 265-284.
7. Domingo-Ferrer, J. (2002). Advances in Inference Control in Statistical Databases: An Overview, Inference Control in Statistical Databases: From Theory to Practice. *Lecture Notes in Computer Science*(2316), pp. 1-7.
8. Domingo-Ferrer, J. (2008). A Survey of Inference Control Methods for Privacy-Preserving Data Mining. *Advances in Database Systems*(34), pp. 53-80.
9. Sweeney, L. (2000). *Simple Demographics Often Identify People Uniquely.* Pittsburgh: Carnegie Mellon University.
10. Narayanan, A., & Shmatikov, V. (2006). *How to Break Anonymity of the Netflix Prize Dataset.* arXiv:cs/0610105v2 [cs.CR].
11. Ohm, P. (2010). Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. *UCLA Law Review*(57), pp. 1701–1777.
12. Adam, N., & Worthman, J. (1989). Security-control methods for statistical databases: a comparative study. *ACM Comput. Surv. 21*(4).
13. Dzięgielewska, O. (2017). Anonymization, tokenization, encryption. How to recover unrecoverable data. *Computer Science and Mathematical Modelling*, pp. 9-13.
14. Dzięgielewska, O. (2020). Defeating Inference Threat with Scoring Metrics. 36th IBIMA Conference on 4-5 November 2020 Granada, Spain.: Conference proceedings (ISBN: 978-0-9998551-5-7, Published in the USA).
15. Dzięgielewska, O. (2020). Evaluating Quality of Statistical Disclosure Control Methods – VIOLAS Framework. *Lecture Notes in Computer Science, Springer, Cham. , 12276*(Privacy in Statistical Databases. PSD 2020.).
16. Dzięgielewska, O., & Szafrański, B. (2016). A brief overview of basic inference attacks and protection controls for statistical databases. *Computer Science and Mathematical Modelling*(4), pp. 19–24.