ROZSZERZENIE METOD OCHRONY DANYCH STATYSTYCZNYCH O METRYKI ADAPTACYJNE I MECHANIZMY OCENY

Streszczenie

1. Kontekst

Liczne wycieki danych wrażliwych pokazują, że nawet, a może przede wszystkim duże organizacje, muszą liczyć się ze znaczącymi konsekwencjami występowania krytycznych kwestii bezpieczeństwa w systemach IT. Pytanie: *czy nastąpi atak na zasoby organizacji,* już dawno zmieniło się na: *kiedy nastąpi atak na zasoby organizacji*. Problem wycieków danych dotyczy nie tylko klasycznego rozumienia wycieku całej bazy danych czy wrażliwych dokumentów organizacji, ale również błędów w procesie animizacji i zaciemniania danych, a także nieodpowiednich metod zabezpieczających dane statystyczne zbierane zarówno w oficjalnych badaniach jak i takich zbieranych w procesach profilowania użytkowników.

Analiza prac w obszarze prywatności danych pokazuje, że istnieje wiele pojedynczych rozwiązań dla poszczególnych problemów naruszenia prywatności danych, ale brakuje w nich podstawowej analizy przyczyn z holistycznym zrozumieniem bezpieczeństwa danych podczas całego cyklu życia systemu. Aktualnie wykorzystywane podejście dla systemów rzeczywistych, które dyktuje zasadę *security-by-design* i implikuje projektowanie metod ochrony już na samym początku fazy planowania systemu są koncepcyjnie niezawodne, jednak w systemach opartych na informacji, w których objętość danych wzrasta, a charakterystyki zmieniają się znacznie w czasie, prawie niemożliwe jest zaprojektowanie bezpiecznej architektury bazy danych uwzględniającej przyszłe zmiany. Zadanie staje się jeszcze bardziej problematyczne w scenariuszach, w których zbiory danych mają być wymieniane między wieloma podmiotami, np. dane są przesyłane do analiz statystycznych od pierwotnego właściciela danych do podmiotu zewnętrznego.

Architektura bezpieczeństwa zaprojektowana dla nowo powstałych dynamicznych systemów informacyjnych mogłaby zostać potencjalnie osiągnięta, ale w dalszym ciągu nie adresowałaby problemu prywatności danych w obecnie obsługiwanych systemach przechowujących aktualne i historyczne dane. Badania mają na celu opracowanie rozwiązania, które mogłoby być skuteczne pod względem ochrony prywatności danych, nawet w przypadku już zarządzanych baz danych.

2. Problem

Istniejące rozwiązania w zakresie ochrony danych statystycznych wykorzystują metody anonimizacji danych, które nie zapewniają satysfakcjonującego poziomu prywatności w przypadku niewłaściwego użycia lub silnie perturbacyjnych metod, które silnie zaburzają dokładność otrzymywanych statystyk. Obecnie jednym z najczęściej używanych podejść do zabezpieczania danych statystycznych jest model prywatności różnicowej (*differential privacy*). Autorka koncepcji (Dwork & Roth, The Algorithmic Foundations of Differential Privacy, 2014), wskazuje wiele otwartych zagadnień, z których w moich badaniach zajmuję się głównie:

- minimalizacją ryzyka związanego z atakami wnioskowaniem, które wykorzystują kilka zbiorów danych oraz historyczne i przyszłe zmiany w zbiorach danych,
- metodą kwantyfikującą utratę prywatności, a tym samym określenia poziomu bezpieczeństwa stosowanej metody.

Prowadzone przeze mnie badania mają na celu wzbogacenie modelu prywatności różnicowej o dodatkowe elementy, które stanowią propozycję rozwiązania powyższych zagadnień.

3. Minimalizacja ryzyka związanego z atakami wnioskowaniem

Opracowana przeze mnie metoda zabezpieczania danych statystycznych opiera się na modelu prywatności różnicowej (*differential privacy*) rozszerzając go o dodatkowe, następujące metryki, które zwiększają poziom bezpieczeństwa oryginalnego modelu:

- metrykę dokładności - metryka, która pozwala manipulować szumem statystycznym w taki sposób, że wynik zapewniłby wybrany poziom kompromisu pomiędzy dokładnością a bezpieczeństwem, który można łatwo dostosować do rzeczywistych potrzeb dla przeprowadzanych analiz,
- metrykę wyniku informacyjnego - przyrostowa miara, która automatycznie modyfikuje dodany szum statystyczny w zapytaniach o takie same lub podobne dane pobierane wielokrotnie. Ta miara ma na celu minimalizację ryzyka ujawnienia wrażliwych danych w przypadku dynamicznych ataków wnioskowaniem (*dynamic inference attacks*). Miara docelowo powinna być rozszerzona, aby minimalizować ryzyko ataków, w których wykorzystywana jest więcej niż jedna baza danych jako źródło wnioskowania.

Opracowana metoda zakłada, że ani metoda gromadzenia danych ani model danych bazy danych nie muszą stosować żadnej metody kontroli statystycznej, ponieważ mechanizmy kontroli będą stanowić część interfejsu pozyskującego dane z bazy danych. Mechanizm opracowany zgodnie z moją metodą musi działać jako część metody wyznaczania statystyk, która wykorzystuje nieprzetworzone zbiory danych z bazy danych. W ten sposób można ją łatwo zastosować jako dodatkowy interfejs użytkownika końcowego dla istniejących baz danych lub włączyć do systemu baz danych w przypadku systemów znajdujących się jeszcze w fazie planowania.

4. Ocena skuteczności metody

Skuteczność opracowanej metody została określona na podstawie modelowania zagrożeń oraz kryteriów bezpieczeństwa statystycznego określonych w opracowanym dla tego celu „frameworku" uwzględniającym:

- poufność statystyczną - poufność statystyczną określa się, oceniając ryzyko identyfikacji i reidentyfikacji danych pierwotnych i wtórnych,
- integralność statystyczną - integralność statystyczną uzyskuje się, gdy zapytania o te same dane, ale za pomocą różnych zapytań dają wyniki statystycznie identyczne,
- dokładność statystyczną - dokładność statystyczna jest miarą jakości pobranych statystyk, czyli jak bardzo zaburzony wynik odbiega od wyniku dokładnego,
- przejrzystość statystyczną – metryka zakłada zapewnienie poufności statystycznej nawet w przypadku, gdy ze zbiorów danych zostanie usunięty wpis dotyczący konkretnego podmiotu.

W wyniku przeprowadzonej analizy określono warunki w których opracowana metoda wykazuje lepsze charakterystyki niż klasyczna implementacja metody prywatności różnicowej.

Wybrana bibliografia:

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.