Dr. A. Peter (UT), SHARE

English scientific summary

The modern economy is largely data-driven and relies on the processing and sharing of data across organizations as a key contributor to its success. At the same time, the value, amount, and sensitivity of processed data is steadily increasing, making it a major target of cyber-attacks. A large fraction of the many reported data breaches happened in the healthcare sector, mostly affecting privacy-sensitive data such as medical records and other patient data. This puts data security technologies as a priority item on the agenda of many healthcare organizations, such as of the Dutch health insurance company Centraal Ziekenfonds (CZ). In particular when it comes to sharing data securely, practical data protection technologies are lacking as they mostly focus on securing the link between two organizations while being completely oblivious of what is happening with the data after sharing. For CZ, searchable encryption (SE) technologies that allow to share data in encrypted form, while enabling the private search on this encrypted data without the need to decrypt, are of particular interest. Unfortunately, existing efficient SE schemes completely leak the access pattern (= pattern of encrypted search results, e.g. identifiers of retrieved items) and the search pattern (= pattern of search queries, e.g. frequency of same queries), making them susceptible to leakage-abuse attacks that exploit this leakage to recover what has been queried for and/or (parts of) the shared data itself.

The SHARE project will investigate ways to reduce the leakage in searchable encryption in order to mitigate the impact of leakage-abuse attacks while keeping the performance-level high enough for practical use. Concretely, we propose the construction of SE schemes that allow the leakage to be modeled as a statistic released on the queries and shared dataset in terms of ε-differential privacy, a well-established notion that informally says that, after observing the statistic, you learn approximately (determined by the ε-parameter) the same amount of information about an individual data item or query as if the item was not present in the dataset or the query has not been performed. Naturally, such an approach will produce false positives and negatives in the querying process, affecting the scheme's performance. By calibrating the ε-parameter, we can achieve various leakage-performance trade-offs tailored to the needs of specific applications.

SHARE will explore the idea of differentially private leakage on different parts of SE with different search capabilities, starting with exact-keyword-match SE schemes with differentially private leakage on the access pattern only, up to schemes with differentially private leakage on the access and search pattern as well as on the shared dataset itself, allowing for more expressive query types like fuzzy match, range, or substring queries. SHARE comes with an attack lab in which we investigate existing and new types of leakage-abuse attacks to assess the mitigation-potential of our proposed combination of differential privacy with cryptographic guarantees in searchable encryption.
To stimulate commercial exploitation of SHARE-results, our consortium partners CZ and TNO will take the lead on applying and evaluating our envisioned technologies in various healthcare use-cases.

English public summary

The SHARE project develops advanced encryption techniques that allow for the sharing of sensitive data in encrypted form while enabling the private search on this encrypted data without the need to decrypt nor to reveal what is being searched for. SHARE applies these techniques in healthcare to protect medical data.

Dutch public summary

Het SHARE project ontwikkelt geavanceerde vercijferingstechnieken die het mogelijk maken om gevoelige (bijvoorbeeld medische) gegevens in vercijferde vorm te delen. Deze technieken zorgen er hierbij voor dat de vercijferde data voor analyse- en zoekopdrachten niet ontcijferd hoeven te worden en dat niet wordt onthuld waar op gezocht is.