Publication: The Business Times, p 13
Date: 20 July 2016
Headline: Retrieving data in an encrypted world

A Special Feature Brought to You By Singapore Management University

Retrieving data in an encrypted world

Professor Pang Hwee Hwa, the new dean of SMU School of Information Systems, is tackling one of the thorniest problems in data security

As businesses run more of their operations through "cloud" applications provided by third parties on the Internet, those encrypting their data for security reasons face a practical concern: how do you retrieve just one item among an entire encrypted data collection?

With today's encryption techniques, a data set is typically mapped by a random mathematical function that makes the original data unreadable without a secret "key" to decrypt the information. It also means that to retrieve selected items, the entire collection needs to be decrypted to perform a search. But do you trust the third party application provider to be given the secret key to decrypt your entire data to perform a search on your behalf? Surely giving the secret key to a third party compromises security and defeats the whole purpose of encrypting your data.

What's the alternative? For the third party to return the entire encrypted collection for the user to decrypt and conduct his own search? Is that the desired solution?

"Say I have a list of 10,000 calendar entries or human resource data. If I want to retrieve one entry that concerns a meeting, or a specific record on a person, I wouldn't want to retrieve all 10,000 entries to my mobile device every time before I do my search," said Singapore Management University (SMU) Professor Pang Hwee Hwa, who works on data security issues.

A similar problem crops up when a bank wants to compare its own encrypted data with another bank to check for suspicious activities conducted by the same person. "The banks won't release to each other their entire list, so how do you find common patterns among them?"

The scenario is prevalent even within the same organisation, Prof Pang said. Different divisions can come under data laws of different jurisdictions. There will be occasions when different government agencies need to compare data with each other and with private companies. Cloud storage providers commonly operate data centres across different countries.

"If you are a business, you may be required by the laws of the country you operate in not to release your data outside of certain uses."

A possible solution

The issues Prof Pang described are very real in today's world and the tremendous application potential have inspired much research in information security circles.

The topic is known as searchable encryption – how one outsources a collection of data in a private manner to a third party and, when the need arises, enables the third party to retrieve specific data matching certain query values from the collection.

"One distinctive feature that we need to achieve is the third party must not obtain any information on the data or queries, even if the same value underlies multiple data or queries." In contrast, general searchable encryption techniques provide this safeguard for only the data or the queries, not both.

After a dozen failed attempts, Prof Pang, working with his colleague Associate Professor Ding Xuhua, came up with a solution to the encrypted data retrieval problem in 2014 that did not lead to security loopholes.

Their solution: a randomised mathematical function that provides an easy way for the third party to spot a match between the data value and query value. The data and search queries remain encrypted throughout, with the third party none the wiser on what information is being asked for and retrieved.

Yet, improvements needed to be made. "It was too inefficient because it involved a lot of computation. With our method, matching two values took milliseconds. While that may sound fast, we are talking about nanoseconds for normal search queries. Our solution was still a million times slower."

Recently, the two professors and two other researchers have made further advances on the problem. Besides speeding up the computation time for matching encrypted values, they have added new functionalities to their solution. Specifically, they formulated a mechanism for multiple users to independently encrypt data with their respective secret keys. Subsequently, the users may jointly enable a third party to perform matching across their encrypted data.

"Now, multiple organisations can retrieve data among themselves or allow a third party to do the matching," Prof Pang said.

"Going forward, my research will centre on building a secure cloud retrieval solution, or creating a system that can compare the private datasets of multiple organisations. These solutions can be applied to the financial, healthcare or defence sectors."

Professor Pang Hwee Hwa, new dean of SMU School of Information Systems

A dedicated computer scientist

Prof Pang received his PhD from the University of Wisconsin at Madison in 1994. Since then, he has worked in various predecessor organisations of the current A*STAR Institute for Infocomm Research.

He joined SMU in 2005. In May 2016, he was appointed as the next dean of SMU's School of Information Systems starting in July, selected from a global pool of candidates after an extensive and rigorous search that began last October.

"Prof Pang's strong commitment to research in information systems and a passion for excellence in education make him the ideal candidate to lead the school," SMU said.

Prof Pang's research interests include Data Security & Privacy, Database Systems, Information Retrieval & Multimedia Analytics, Spatial and Context Aware Data Management, Recommender Systems & Preference Analytics.

His strength in research is evident from the more than 20 high-quality papers he has published in top-tier refereed journals. He has also presented 50 papers at numerous conferences in the US, Europe, China, Korea and Japan. His research output also includes a number of commercial systems and 14 patents. He is highly regarded in his professional area and has served in the programme committees of prestigious academic conferences.

On top of his research, Prof Pang teaches two courses: an undergraduate course on search engine technologies and a PhD course that combines database systems and search engines.

As dean, he will be working to create the right culture and support system for the faculty to conduct interesting and impactful research, while training undergraduates and postgraduates to provide value to employers.

His interest in computing got him to where he is today. His original training is in database systems, which is about how one can effectively store and process various types of data.

Around 10 years ago, Prof Pang started picking up information security techniques, and began to work with colleagues trained in applied cryptography. This, coupled with his initial background in databases, enables him to bring together knowledge across the two academic disciplines in his research.

"The problem seemingly involved contradicting requirements. Can we encrypt data securely, yet permit controlled comparison operations? It wasn't clear at the beginning whether this was even possible. In research, you sometimes encounter such problems that are both intellectually very intriguing and have tremendous application potential."

The next step for his research is to implement a secure cloud retrieval solution, or create a system that can compare the private datasets of multiple organisations, he said. Promising applications can be found in the financial, healthcare or defence sectors.

This is a monthly series brought to you by the Singapore Management University. Next month's feature will discuss managers' short-termism and how companies can encourage long-term investments.