Identifying problems with national identifiers: Supposedly encrypted numbers can be easily decrypted

September 29 2015

In a pair of experiments that raise questions about the use of national identifying numbers, Harvard researchers have shown that Resident Registration Numbers (RRN) used in South Korea can be decrypted to reveal a host of personal information.

Led by Professor of Government and Technology in Residence Latanya Sweeney, a team of researchers in two experiments was able to decrypt more than 23,000 RRNs using both computation and logical reasoning. The findings suggest that, while such identifiers are encrypted to protect privacy, they remain vulnerable to attack and must be designed to avoid such weaknesses. The studies are described in a September 29 paper published in *Technology Science*.

"Like most data driven highly networked societies, South Korea uses personally identifying numbers as a linchpin to personal identity in employment, banking, taxation, and for social and medical services. In the United States, we use Social Security numbers similarly. When these numbers become easily accessible to others, whether through breaches or poor encryption in data sold or given away, the major institutions that rely on them become vulnerable. "

Sweeney and Ji Su Yoo, a Research Assistant at the Data Privacy Lab at Harvard and an author of the study, were able to show that each number in the RRN could be replaced with a letter in a recognizable pattern.

Using such a pattern, they were able to decrypt thousands of RRNs that could uncover personal information about their users.

They also found that, like credit cards, the final digit is a weighted sum of prior digits, meaning researchers were able to decrypt the numbers, then used arithmetic to confirm the accuracy of the information they uncovered.

"South Koreans depend on personally identifying numbers for numerous economic transactions and it is inconvenient for businesses and individuals alike to verify identities and track clients without these numbers. But in the end, it is the South Korean population that is receiving the short end of the stick. That is, when data is so easily de-anonymized, individual privacy, not company profits, are compromised. Our study shows that weak encoding systems, which refer to the very design of the number, render encryptions as poor methods of protecting privacy. South Koreans are aware of the vulnerabilities of the RRN encoding system - our study therefore urges a more robust redesign of these personally identifying numbers not only for the sake of the institutions and system that depend on them but also for the individuals who use them."

Sweeney and Yoo conducted the study using prescription data that was presumed to be anonymous because it did not include patient's name or address, and had encrypted their RRN. Similar data is often shared with corporations around the world who track health data - believed to be anonymous - on millions of South Koreans.

"Administrators often use simple schemes to encrypt personal information because it passes a face test -if it looks okay, it must be okay. Sometimes they use strong encryption but in a wrong way, leading to the same vulnerable outcome. If researchers like us don't provide scientific facts and insights into these practices, who will? Companies

that receive the data may exploit these same vulnerabilities to advantage. If so, they would hardly then turn around and tell administrators. It is up to researchers to give administrators and society the scientific knowledge needed to make better choices."

The findings are particularly timely, Sweeney said, because South Korea is currently debating a redesign of RRNs and other nations, including the United States, have discussed the use of a single identifier for medical records.

The study, she said, reveals that such identifiers - if not carefully designed and monitored - can be vulnerable to leaks, and must be carefully considered going forward.

"The problem is not unique to South Korea, it's a worldwide concern because we all rely on credit card numbers and other identifiers to function."

Provided by Harvard University

Citation: Identifying problems with national identifiers: Supposedly encrypted numbers can be easily decrypted (2015, September 29) retrieved 23 April 2024 from https://phys.org/news/2015-09-problems-national-supposedly-encrypted-easily.html