Calculating encryption schemes' theoretical security guarantees eases comparison, improvement

October 30 2014, by Larry Hardesty

Credit: Christine Daniloff/MIT

Most modern cryptographic schemes rely on computational complexity for their security. In principle, they can be cracked, but that would take a prohibitively long time, even with enormous computational resources.

There is, however, another notion of security—information-theoretic security—which means that even an adversary with unbounded computational power could extract no useful information from an encrypted message. Cryptographic schemes that promise information-theoretical security have been devised, but they're far too complicated to be practical.

In a series of papers presented at the Allerton Conference on Communication, Control, and Computing, researchers at MIT and Maynooth University in Ireland have shown that existing, practical cryptographic schemes come with their own information-theoretic guarantees: Some of the data they encode can't be extracted, even by a computationally unbounded adversary.

The researchers show how to calculate the minimum-security guarantees for any given encryption scheme, which could enable information managers to make more informed decisions about how to protect data.

"By investigating these limits and characterizing them, you can gain quite a bit of insight about the performance of these schemes and how you can leverage tools from other fields, like coding theory and so forth, for designing and understanding security systems," says Flavio du Pin Calmon, a graduate student in electrical engineering and computer science and first author on all three Allerton papers. His advisor, Muriel Médard, the Cecil E. Green Professor of Electrical Engineering and Computer Science, is also on all three papers; they're joined by colleagues including Ken Duffy of Maynooth and Mayank Varia of MIT's Lincoln Laboratory.

The researchers' mathematical framework also applies to the problem of data privacy, or how much information can be gleaned from aggregated—and supposedly "anonymized"—data about Internet users' online histories. If, for instance, Netflix releases data about users' movie preferences, is it also inadvertently releasing data about their political preferences? Calmon and his colleagues' technique could help data managers either modify aggregated data or structure its presentation in a way that minimizes the risk of privacy compromises.

Staying close

To get a sense of how the technique works, imagine an encryption scheme that takes only three possible inputs, or plaintexts—"A," "B," and "C"—and produces only three possible outputs, or ciphertexts. For each ciphertext, there is some probability that it encodes each of the three plaintexts.

The ciphertexts can be represented as points inside a triangle whose vertices represent the three possible plaintexts. The higher the probability that a given ciphertext encodes a particular plaintext, the closer it is to the corresponding vertex: Ciphertexts more likely to encode A than B or C are closer to vertex A than to vertices B and C. A secure encryption scheme is one in which the points describing the ciphertexts are clustered together, rather than spread out around the triangle. That means that no ciphertext gives an adversary any more information about the scheme than any other.

Of course, for most encrypted messages, there are way more than three possible corresponding plaintexts. Even a plaintext as simple as a nine-digit number has a billion possible values, so the probabilities corresponding to an encoded Social Security number would describe a point in a billion-dimensional space. But the general principle is the same: Schemes that yield closely clustered points are good, while schemes that don't are not.

An adversary wouldn't actually know the probabilities associated with any given ciphertext. Even someone with access to an encryption scheme's private key would have difficulty calculating them. For their analyses, Calmon, Médard, and their colleagues developed security metrics that hold for a wide range of distributions, and they augmented them with precise calculation of the worst cases—the points farthest from the center of the main cluster. But the mathematical description of the degree to which the probabilities cluster together is a direct indication of how much information an adversary could, in principle, extract from a ciphertext.

Targeted protection

In their first Allerton paper, in 2012, the researchers used this probabilistic framework to demonstrate that, while a ciphertext as a whole may not be information-theoretically secure, some of its bits could be. It should thus be possible to devise encryption schemes that can't guarantee perfect security across the board but could provide it for particular data—say, a Social Security number.

"Talking with cryptographers, they would always ask us, 'Oh, cool! You can guarantee that regardless of what you do, you can hide individual symbols. What about functions of the plaintext?'" Calmon says. "Standard cryptographic definitions of security care about that."

An encryption scheme might, that is, guarantee that an adversary can't extract an encoded Social Security number; but it might still allow the adversary to extract the last four digits of the number. Similarly, it might prevent an adversary from determining a subject's age; but it might allow the adversary to deduce that, say, the subject is between 30 and 40 years of age.

This is the problem that the researchers tackle in their last two Allerton papers. There, Calmon, Médard, and Varia show that if you can determine that a particular function is difficult or easy to extract from a ciphertext, then so are a host of correlated functions. In addition to addressing cryptographers' concerns about functions of the plaintext, this approach has the advantage of not requiring analysis of massively multidimensional probability spaces. Information about the security of a single function—which can often be determined through a fairly simple analysis—can provide strong guarantees about the security of an encryption scheme as a whole.

"Perfect secrecy is a very stringent requirement—essentially, the only way of guaranteeing that is to use a one-time pad, like they would in spy novels," says Maxim Reginsky, an assistant professor of electrical and computer engineering at the University of Illinois at Urbana-Champaign. "Instead, let's just accept the empirical fact that practical security systems we rely on every day do not deliver perfect secrecy. Some information about the data they try to protect will leak out. The work by Calmon, Varia, and Médard shows that there are limits to what an adversary can infer from this leaked information. Naturally, this is relevant in the age of big data."

The mathematical techniques that the MIT researchers employed "have been used in statistical analysis," Reginsky adds. "But the information-theoretic implications are all new. This will definitely lead to a great deal of interesting research activity."

*This story is republished courtesy of MIT News (*web.mit.edu/newsoffice/*), a popular site that covers news about MIT research, innovation and teaching.*

Provided by Massachusetts Institute of Technology

Citation: Calculating encryption schemes' theoretical security guarantees eases comparison, improvement (2014, October 30) retrieved 20 April 2024 from https://phys.org/news/2014-10-encryption-schemes-theoretical-eases-comparison.html