Chair of Communications Engineering
TUM Department of Electrical and Computer Engineering
Technical University of Munich

Master's Thesis

[security] Practical implementation of physical-layer semantic security

The goal of this project is to implement in Python/Sagemath the security functions (at least one of four) described in https://arxiv.org/abs/2102.00983
Sagemath contains libraries for mosaics, BIBDs, etc, that can be used for the project.

Motivation:
There are various types of security definitions.
The mutual information based types, in increasing order of security requirement are

1. Weak secresy asks that the average mutual information of the eavesdropper I(M:E)/n goes to 0 for a uniform message M (average here means averaged over the blocklength n, an additional average over M is implicit in the mutual information)
2. Strong secrecy asks that the total mutual information I(M:E) goes to 0,
3. Semantic security asks that the total mutual informaiton I(M:E) goes to 0 for any distribution of the message M (and thus in particular for all distributions that pick any of two chosen messages with 1/2 probabilty)

Then there are the almost-equivalent respective indistiguishablity types of security requirements (below |P-Q|_1 is the statistical distance and Exp_M is expectation value over M)

1. average indistinguishability 1/n Exp_M | P_{E|M} - P_E |_1 for a uniform message M goes to 0 (again average refers over the blocklegth n, clearly there is also the average over M)
2. total indistiguishability Exp_M | P_{E|M} - P_E |_1 for a uniform message M goes to 0
3. indistinguishability |P_{E|m} - P_{E|m'}|_1 for any two messages m and m' goes to 0.

Each of the indistiguishabilities can also be written using KL digvergence instead of statistical distance, in which case the conditions are exactly equivalent to their mutual information versions.

Strong secrecy is the standard security requirement considered in information-theoretic security, while semantic security is the minimum requirement considered in computational security. Information-theoretic (physical-layer) security differs from computational security in that the secrecy is guaranteed irrespective of the power of the adversary, while in computational security E is computationally bounded. Computational security also assumes that the message is at least of a certain length for the schemes to work, and thus if the message to be secured is too small it needs to be padded to a larger message.

In practice, information theoretic security is expensive, because the messages that can be secured can be only as long as the keys that can be generated. However, in identification only a very small part of the message needs to be secured, which in computational security triggers padding and thus waste, but on the other side makes information-theoretic security accessible and not so expensive.

At the same time, the security of identification implicitly requires semantic security. It has been known for a while that hash functions provide information-theoretic strong secrecy. However, because the standard for information-theoretic security has been strong secrecy, before https://arxiv.org/abs/2102.00983 no efficient functions where known to provide information-theoretic semantic security.
We need an implementation of these type of functions so that we can integrate information-theoretic security into our identification project.

Advisors

Roberto Ferrara