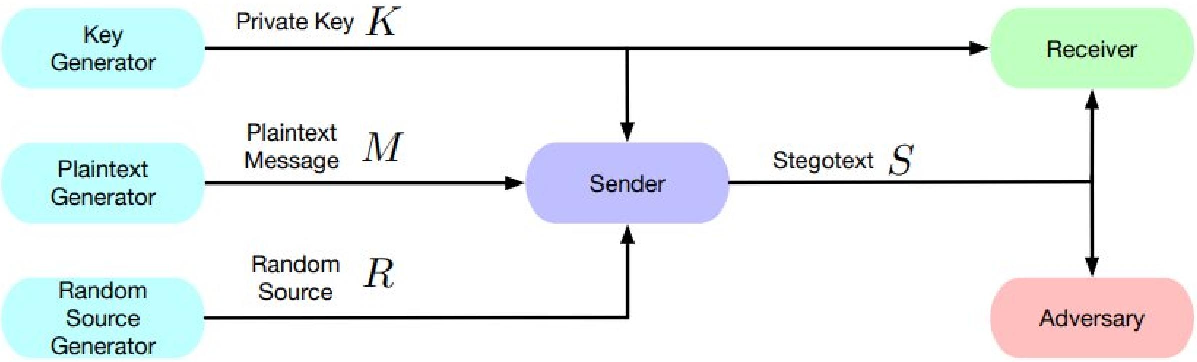
TechXplore

Breakthrough enables perfectly secure digital communications

March 7 2023



A graphical depiction of steganography. The sender receives a plaintext message, a source of randomness, and a private key, and outputs a stegotext. The receiver receives the same private key as the sender, along with the stegotext. The adversary also receives the stegotext. Credit: *arXiv* (2022). DOI: 10.48550/arxiv.2210.14889

A group of researchers has achieved a breakthrough in secure communications by developing an algorithm that conceals sensitive information so effectively that it is impossible to detect that anything has been hidden.

The team, led by the University of Oxford in close collaboration with Carnegie Mellon University, envisage that this method may soon be used widely in digital human communications, including social media and private messaging. In particular, the ability to send perfectly secure information may empower vulnerable groups, such as dissidents, investigative journalists, and humanitarian aid workers.

The algorithm applies to a setting called steganography: the practice of hiding sensitive information inside of innocuous content. Steganography differs from cryptography because the sensitive information is concealed in such a way that this obscures the fact that something has been hidden. An example could be hiding a Shakespeare poem inside an AI-generated image of a cat.

Despite having been studied for more than 25 years, existing steganography approaches generally have imperfect security, meaning that individuals who use these methods risk being detected. This is because previous steganography algorithms would subtly change the distribution of the innocuous content.

To overcome this, the research team used recent breakthroughs in information theory, specifically minimum entropy coupling, which allows one to join two distributions of data together such that their mutual information is maximized, but the individual distributions are preserved.

As a result, with the new algorithm, there is no statistical difference between the distribution of the innocuous content and the distribution of content that encodes sensitive information.

The algorithm was tested using several types of models that produce auto-generated content, such as GPT-2, an open-source language model, and WAVE-RNN, a text-to-speech converter. Besides being perfectly secure, the new algorithm showed up to 40% higher encoding efficiency than previous steganography methods across a variety of applications, enabling more information to be concealed within a given amount of data. This may make steganography an attractive method even if perfect security is not required, due to the benefits for data compression and storage.

The research team has filed a patent for the algorithm, but intend to issue it under a free license to third parties for non-commercial responsible use. This includes academic and humanitarian use, and trusted third-party security audits. The researchers have published this work as a preprint paper on *arXiv*, as well as open-sourced an inefficient implementation of their method on Github. They will also present the new algorithm at the premier AI conference, the 2023 International Conference on Learning Representations in May.

AI-generated content is increasingly used in ordinary human communications, fueled by products such as ChatGPT, Snapchat AI-stickers, and TikTok video filters. As a result, steganography may become more widespread as the mere presence of AI-generated content will cease to arouse suspicion.

Co-lead author Dr. Christian Schroeder de Witt (Department of Engineering Science, University of Oxford) said, "Our method can be applied to any software that automatically generates content, for instance probabilistic video filters, or meme generators. This could be very valuable, for instance, for journalists and aid workers in countries where the act of encryption is illegal. However, users still need to exercise precaution as any encryption technique may be vulnerable to side-channel attacks such as detecting a steganography app on the user's phone."

Co-lead author Samuel Sokota (Machine Learning Department, Carnegie Mellon University) said, "The main contribution of the work is showing a deep connection between a problem called minimum entropy coupling and perfectly secure steganography. By leveraging this connection, we introduce a new family of steganography algorithms that have perfect security guarantees."

Contributing author Professor Jakob Foerster (Department of Engineering Science, University of Oxford) said, "This paper is a great example of research into the foundations of machine learning that leads to breakthrough discoveries for crucial application areas. It's wonderful to see that Oxford, and our young lab in particular, is at the forefront of it all."

More information: Christian Schroeder de Witt et al, Perfectly Secure Steganography Using Minimum Entropy Coupling, *arXiv* (2022). DOI: 10.48550/arxiv.2210.14889

Provided by University of Oxford

Citation: Breakthrough enables perfectly secure digital communications (2023, March 7) retrieved 27 April 2024 from https://techxplore.com/news/2023-03-breakthrough-enables-digital-communications.html