Information Hiding

The aims of steganography are:

1. for a sender to communicate using a public channel such that
2. encoded messages can be decoded by only the intended recipient, but
3. eavesdroppers can't detect the presence of hidden information.

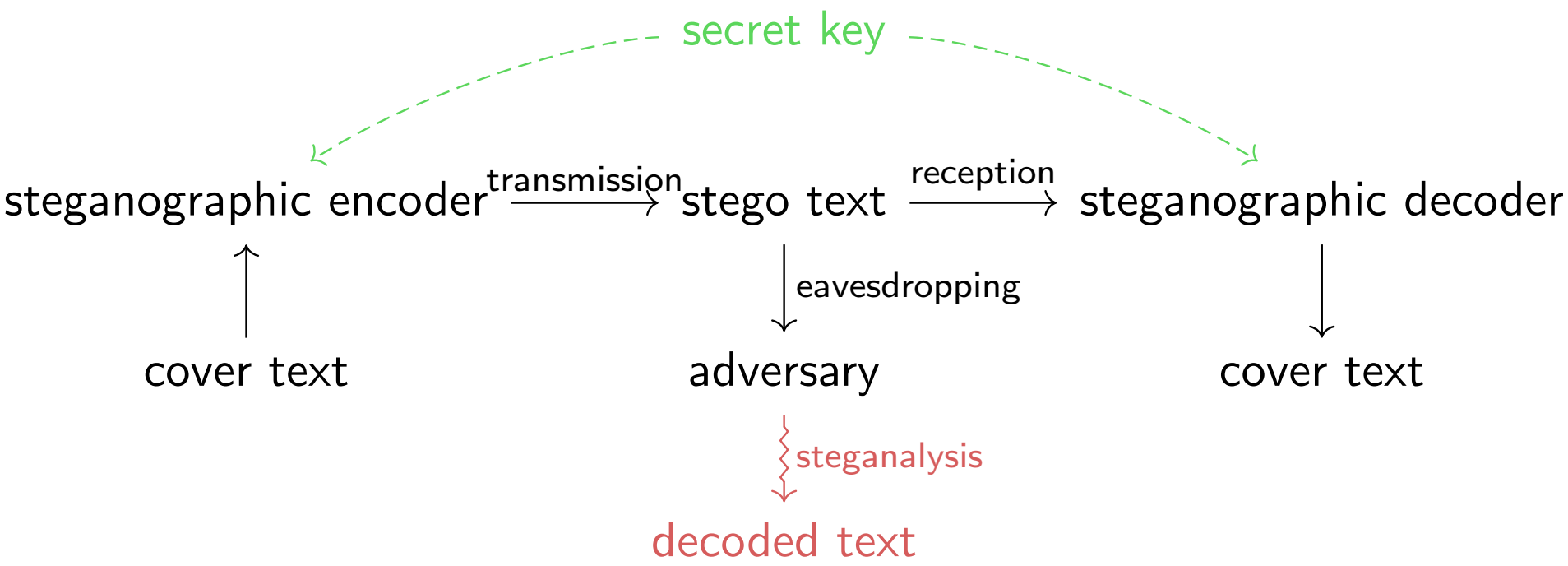

Figure: Text steganalysis aims to detect the presence of hidden information.

Results: Receiver Operating Characteristic (ROC)

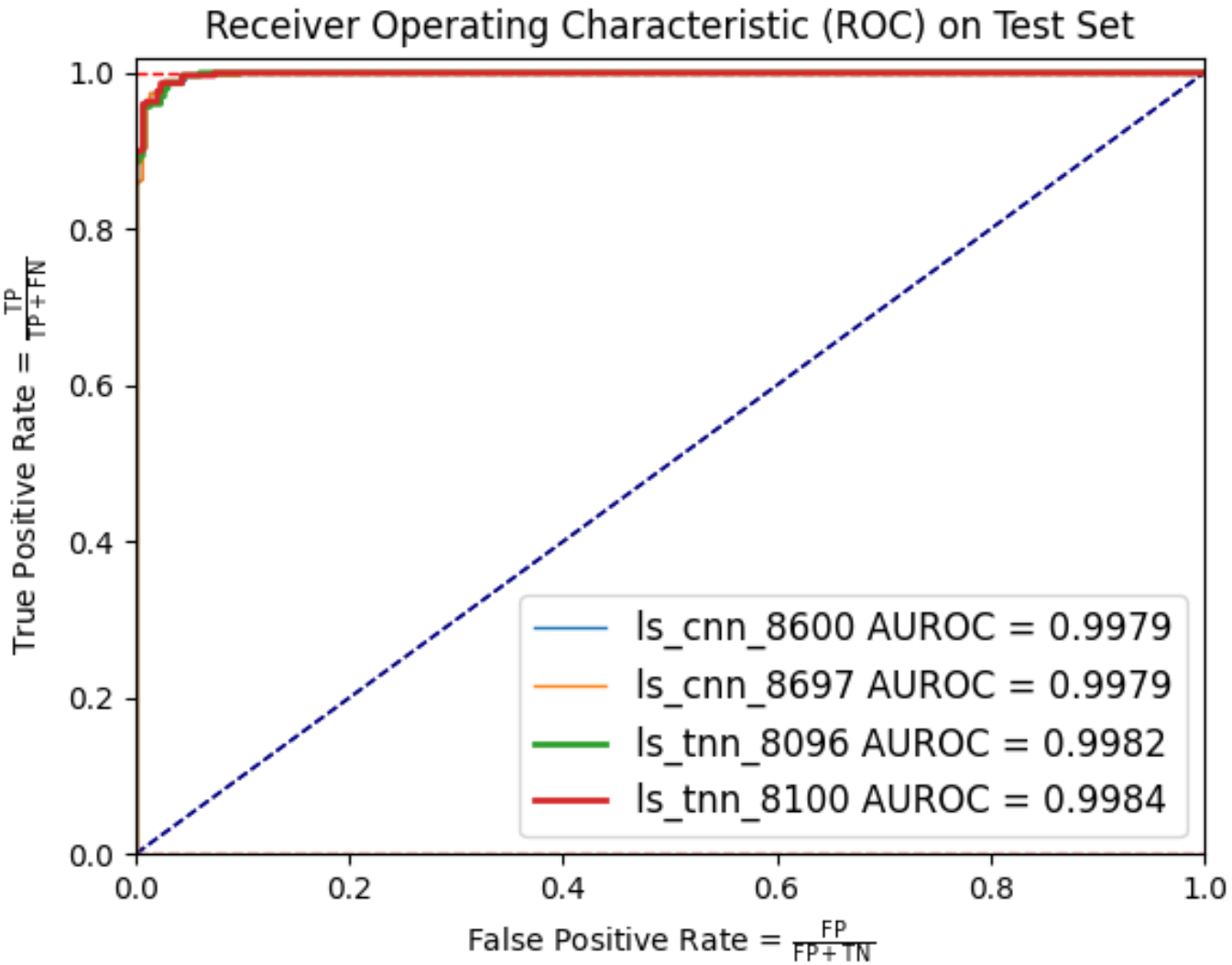

Figure: Areas Under ROC curves range from 0.5 (random) to 1.0 (perfect)

Implications for National Security

Unlike cryptography which conceals only the content of information, steganography conceals the very presence of information.

Potential uses of text steganalysis include:

1. open-source intelligence collection from publicly available information
2. refinement of communications intelligence and other collected data
3. detection of data leakage via covert or subliminal channels
4. transfer learning related tasks:
    1. attribution of pseudonymous authors on X, formerly known as Twitter
    2. artificial intelligence safety problem of watermarking[1] generative AI
5. informing the development of methods for steganalysis[2]
6. informing the development of methods for steganography due to the "arms race" between capabilities for concealment and detection.[3]

[1] My AI Safety Lecture for UT Effective Altrusim by Professor Scott Aaronson

[2] E. g., steganalysis of text, images, video, graphs, network traffic, and other carriers

[3] I. e., defend against known methods of steganalysis.