TrueSeeing: Defensive software against text-encoding based invisible attacks

CAN304 Group 16 Presentation

Chunyu Jiang Guanyuming He Hanyu Zhang Kemu Xu Yifei Du Yilu Shi

Department of Computing, School of Advanced Technology
Xi'an Jiaotong-Liverpool University

21 May, 2024 / SC169

Xi'an Jiaotong-Liverpool University 西交利物浦大學

Introduction:
Do these two words look the same to you?

United states

United states

This happens because...

- Texts, like all other digital data, are stored in binary.
- We humans see texts visually.
- But some programs, like ChatGPT, compilers, search engines, see them in binary.
- And there are many ways to make texts that look the same but different in binary.

Xi'an Jiaotong-Liverpool University 西交利物浦大學

State-of-the-art attack methods

Invisible Characters You cannot see some characters that a program can see.

Homoglyphs Different characters that look the same (the introduction).

Reordering Texts whose order looks different to you and to a computer program.

Deleteion Something you can see cannot be seen by a computer program.

Xi'an Jiaotong-Liverpool University 西交利物浦大學

How to stop this...

- Make the programs see as you see (hard).
- Make you see the evil binary characters (easy).
- And remove them.

Therefore we present TrueSeeing

- It accepts texts from clipboard and files.
- It renders them in two ways simultaneously:
  1. What you would normally see.
  2. Expose all the evil characters.
- Then you are allowed to make modifications to the text until all evil characters are gone.

Prevent the text from being exploited again

TrueSeeing also allows you to give a digital signature to the neutralised text: RSA-FDH

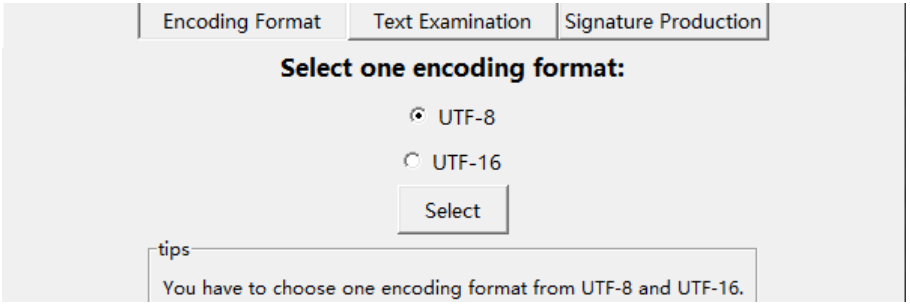
- Gen: RSAGen: public key (N, e), private key (N, d)
- Sign: $\sigma = H(m)^d \bmod N$
- Vrfy:

$$Vrfy_{pk}(m, \sigma) = \begin{cases} 1, & \text{if } \sigma^e = H(m) \bmod N \\ 0, & \text{otherwise} \end{cases}$$

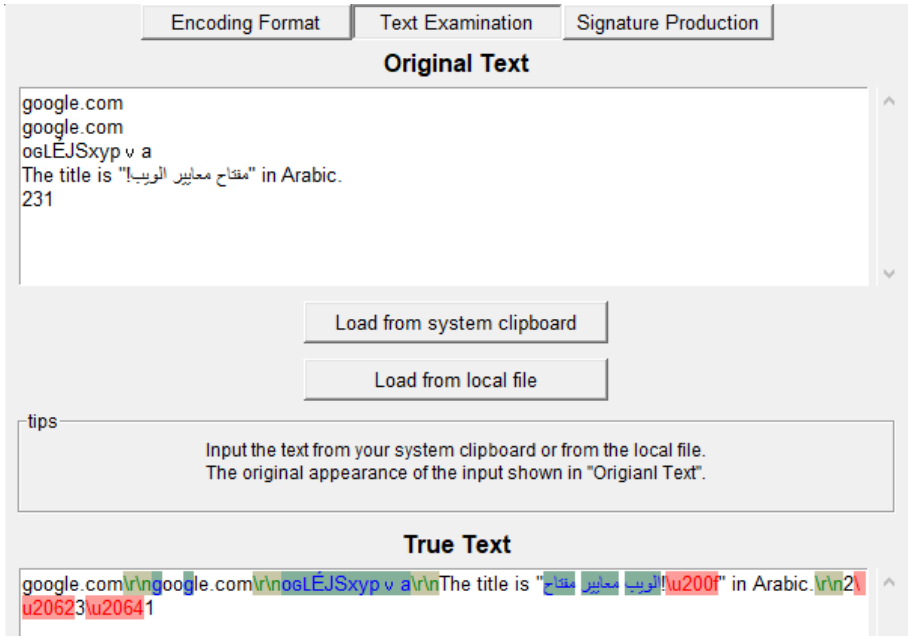
Xi'an Jiaotong-Liverpool University 西交利物浦大學

Demonstration: TrueSeeing - Encoding Format

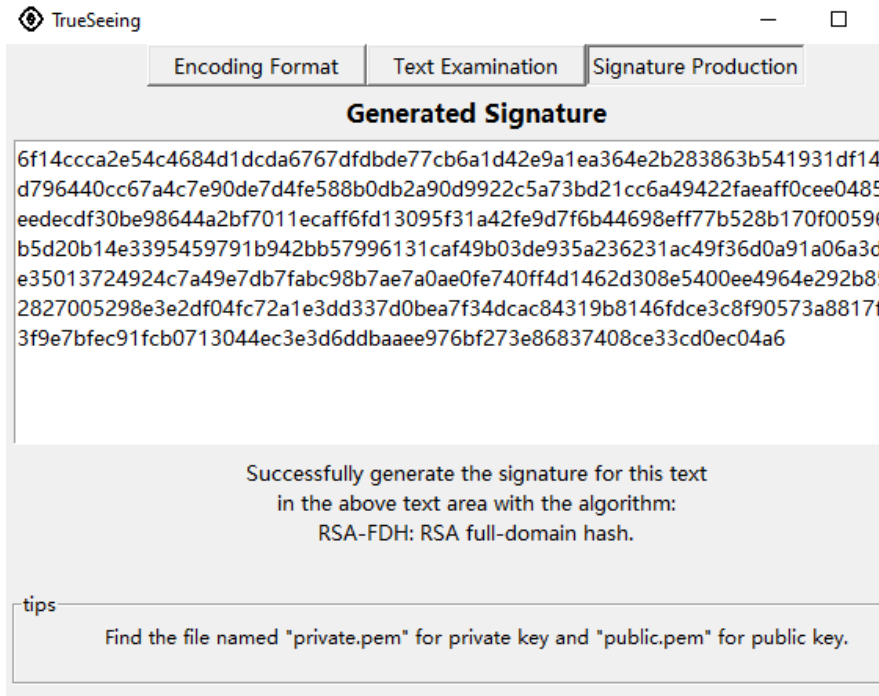
Choose Encoding format



Revealing Malicious Characters



Digital Signature



Xi'an Jiaotong-Liverpool University 西交利物浦大學

Thank You

Feel free to ask Questions.

Xi'an Jiaotong-Liverpool University
西交利物浦大學