Attacks against stream ciphers

CATALDO BASILE

< CATALDO.BASILE@ POLITO.IT >

POLITECNICO DI TORINO

Agenda

- attacks
  - keystream reuse attacks
- learning objectives
  - hints on how to mount statistical attacks
    - some manual/interactive way may work well in most cases
  - additional information on what ‘known plaintext’ means
    - also knowing the structure leaks important data
  - appropriateness of stream ciphers in many context
    - at least without taking advantage of nonces

Stream ciphers

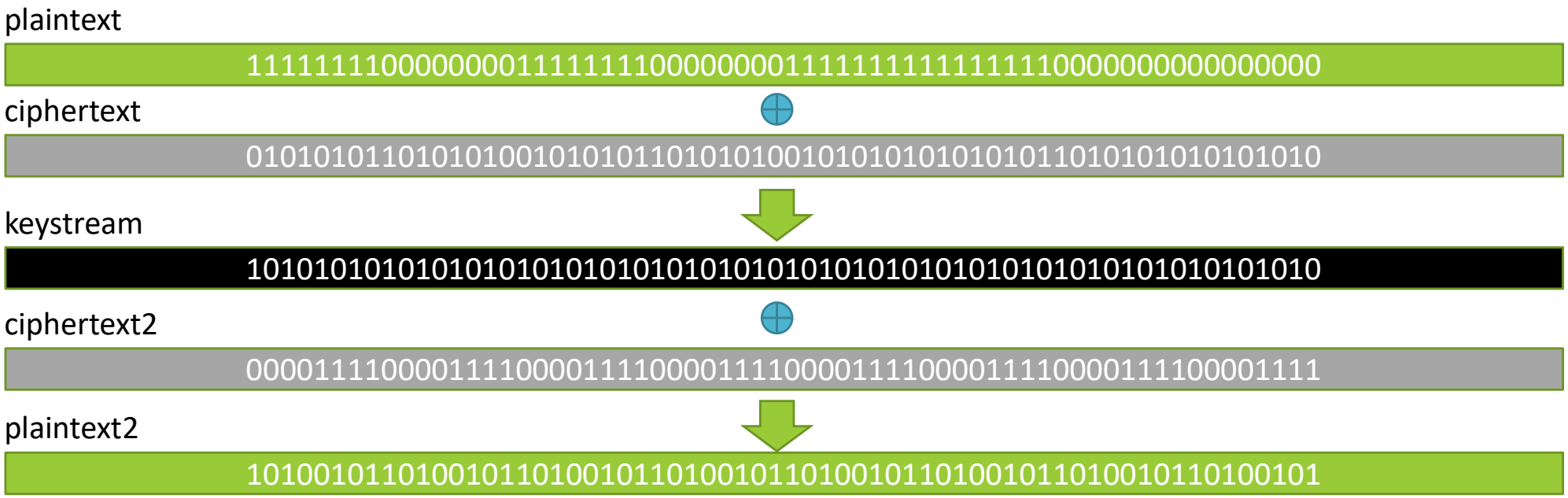
- keystream generated from a **key** (e.g., RC4)
  - and for more modern and CBC-based also a **nonce**
  - NOTE: it also work for all the block cipher modes of operation that simulate streaming
    - CTR, OFB, CFB, …
- ideally, we should be as close as possible to the Shannon's one-time pad algorithm
  - never reuse the nonce (= number used once)
  - change the key quite often
    - easy when paired with public crypto key exchange algorithms

nonetheless…

Keystream reuse

first trivial attack

- known plaintext attack
- as soon as both plaintext and ciphertext are available
  - derive the keystream!



Keystream reuse

properties of the plaintext propagate to the ciphertext

- same mathematical deductions apply to ciphertext blocks

ciphertext

0101010110101010010101011010101001010101010101011010101010101010

ciphertext

0010010100101010101010101001010010100101001010100101010101010010

ciphertext

1001010010010010001001010101011110101010101010101101010111101010 0

ciphertext

0101101111111111111111111111100000000000000000000000000000000000

ciphertext

0101010010101010100101001001010101010101010010101010101010010101

ciphertext

1100100100101110101111011010111101010101010110101010101010101011

keystream

0000000110101010101010101010101010101010101010101010101010101010

Encrypted text messages

- …means ASCII or UTF-8 characters
  - this information allows guessing **individual bytes of the keystream**
    - i.e., …discard candidate bytes of the keystream
      - if they do not produce ASCII characters

- when several ciphertexts are collected statistical attacks are possible
  - outside the scope of this course to investigate the ways to collect statistical data
  - 'Etaoin shrdlu': approximate order of frequency of the 12 most commonly used letters in the English language (case insensitive)
    - 'Etaoin srhldcu' after a recent work from Google…
  - 'tsamcin brped': approximate order of frequency of the 12 most commonly used letters in the English language as a first letter of sentences
  - http://norvig.com/mayzner.html
  - http://storage.googleapis.com/books/ngrams/books/datasetsv2.html
  - http://www.fitaly.com/board/domper3/posts/136.html

Alternative approaches

xortool

- pip install xortool
- https://github.com/hellman/xortool

MTP

- pip install mtp
- https://github.com/CameronLonsdale/MTP
- https://asciinema.org/a/204705

cryptonita

- https://github.com/cryptonitas/cryptonita