CyberSecurity: Principle and Practice

BSc Degree in Computer Science
2023-2024

Lesson 3: Encoding

Prof. Mauro Conti
Department of Mathematics
University of Padua
conti@math.unipd.it
http://www.math.unipd.it/~conti/

Teaching Assistants
Tommaso Bianchi
tommaso.bianchi@phd.unipd.it
Riccardo Preatoni
riccardo.preatoni@studenti.unipd.it

UNIVERSITÀ DEGLI STUDI DI PADOVA

SPRITZ SECURITY & PRIVACY RESEARCH GROUP

DIPARTIMENTO MATEMATICA

Motivations

- The world is full of different "environment"

  - e.g., nation, music notation

- Each "environment" follows its own standard

  - e.g., language, rules

- We need interaction strategies

Encoding

Why Encoding data?

- **Transform** data so that it can be properly (and safely) consumed by a different type of system
- The goal is <u>not</u> to keep information secret
  - e.g., **keys** are not required
- Encoding can be <u>easily</u> be reversed
  - easy to recognize an encoding strategy
- Decoding is the inverse operation

decimal   binary

→ 65 → 01000001
→ 66 → 01000010
→ 67 → 01000011

Encoding

Some examples of encoding:
- base64

  - web communication

- Hexadecimal

  - simplified version of binary in CS

- Uuencoding

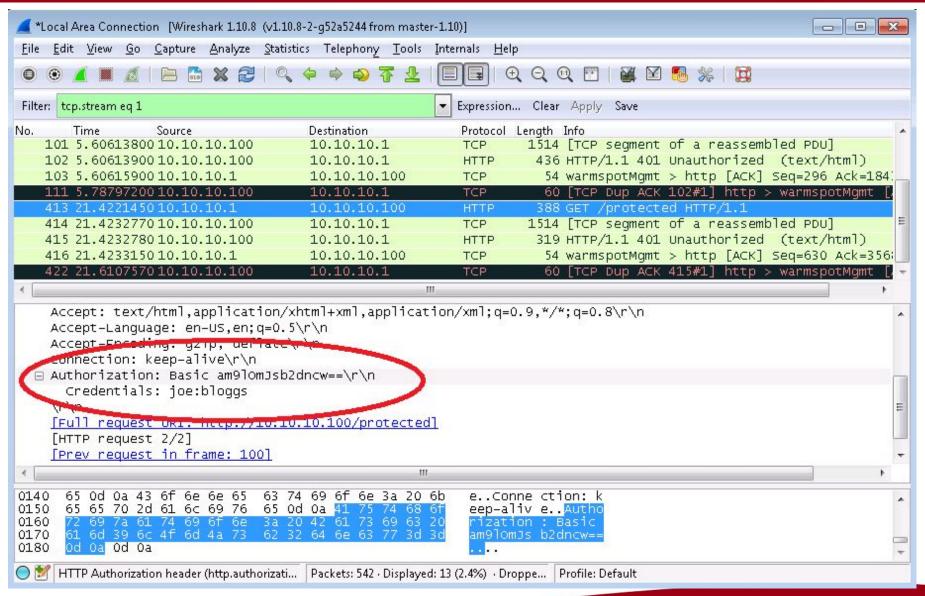 - binary to text encoding for UNIX

Base64

- Very common in web communications
- Lengths of final messages **always** a multiple of 4
- Unique alphabet
  - [A-Z, a-z, 0-9, +, /, =]
  - 0 = A, 1 = B, …, 26 = a, 27 = b, …
  - 52 = 0, 53 = 1, …, 62 = +, 63 = /
- Padding
  - Might end with "==" or "="

Base64

Base64

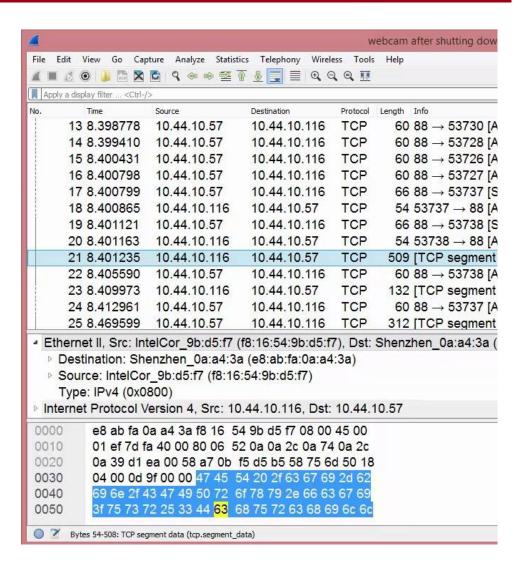Some examples are:

- pleasure ->   cGxlYXN1cmU=

- leasure ->   bGVhc3VyZQ==

- easure ->   ZWFzdXJl

- asure ->   YXN1cmU=

- sure ->   c3VyZQ==

Hexadecimal

- Similar to base64

- The alphabet is [A-F, 0-9]

- Widely used to represent

  ○ MAC Address

  ○ Memory dumps

UUEncoding

- **Always start** with _begin_ followed by the _mode_ and the _file name_

- **Always end** with both

  - " ' "

  - _end_

```
begin 600 test.txt
M5&&AI<R!I<R!A('1E<W0@9FEL92!F;W(@:6QL=7-T<F%T:6YG('1H92!V87)I
M;W5S"F5N8V]D:6YG(&UE=&A09',N($$@<75A:V4@=&%I<R!T97AT(&%Q0
M;;F=E<!!T:&%%N"!C4W(&!)Y=&5S("'1O('='R=87'@;&EN97,L0&=VET:"!"87-E
M-9&%T82!P0@=&]0+@I'<F5E=&EN9W,L(($9R86YK(%%!09F5R"!@''
'
end
```

Encoding in Practice

- When we analyze data, it might be represented in

  an unknown encoding

- How to identify the proper encoding?

  - experience

  - alphabet

  - patterns

  - origin of data

UNIVERSITÀ
DEGLI STUDI
DI PADOVA