

CyberSecurity: Principle and Practice

*BSc Degree in Computer Science
2020-2021*

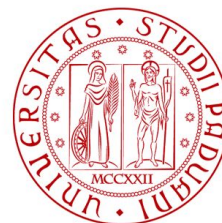
Lesson 3: Encoding

Prof. Mauro Conti

Department of Mathematics
University of Padua
conti@math.unipd.it
<http://www.math.unipd.it/~conti/>

Teaching Assistants

Luca Pajola
pajola@math.unipd.it.
Pier Paolo Tricomi
pierpaolo.tricomi@phd.unipd.it



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



SPRITZ
SECURITY & PRIVACY
RESEARCH GROUP



DIPARTIMENTO
MATEMATICA

- 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

Why Encoding data?

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

<i>decimal</i>	<i>binary</i>
→ 65	→ 01000001
→ 66	→ 01000010
→ 67	→ 01000011

Some examples of encoding:

- base64
 - web communication
- Hexadecimal
 - simplified version of binary in CS
- Uuencoding
 - binary to text encoding for UNIX

- 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 "="

*Local Area Connection [Wireshark 1.10.8 (v1.10.8-2-g52a5244 from master-1.10)]

File Edit View Go Capture Analyze Statistics Telephony Tools Internals Help

Filter: tcp.stream eq 1 Expression... Clear Apply Save

No.	Time	Source	Destination	Protocol	Length	Info
101	5.60613800	10.10.10.100	10.10.10.1	TCP	1514	[TCP segment of a reassembled PDU]
102	5.60613900	10.10.10.100	10.10.10.1	HTTP	436	HTTP/1.1 401 Unauthorized (text/html)
103	5.60615900	10.10.10.1	10.10.10.100	TCP	54	warmspotMgmt > http [ACK] Seq=296 Ack=184
111	5.78797200	10.10.10.100	10.10.10.1	TCP	60	[TCP Dup ACK 102#1] http > warmspotMgmt
413	21.4221450	10.10.10.1	10.10.10.100	HTTP	388	GET /protected HTTP/1.1
414	21.4232770	10.10.10.100	10.10.10.1	TCP	1514	[TCP segment of a reassembled PDU]
415	21.4232780	10.10.10.100	10.10.10.1	HTTP	319	HTTP/1.1 401 Unauthorized (text/html)
416	21.4233150	10.10.10.1	10.10.10.100	TCP	54	warmspotMgmt > http [ACK] Seq=630 Ack=356
422	21.6107570	10.10.10.100	10.10.10.1	TCP	60	[TCP Dup ACK 415#1] http > warmspotMgmt

Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8\r\n
Accept-Language: en-US,en;q=0.5\r\n
Accept-Encoding: gzip, deflate\r\n
Connection: keep-alive\r\n
Authorization: Basic am9lomjsb2dncw==\r\n
Credentials: joe:bloggs

[Full request URL: http://10.10.10.100/protected]
[HTTP request 2/2]
[Prev request in frame: 100]

Offset	Hex	ASCII
0140	65 0d 0a 43 6f 6e 6e 65	e..Connection: k
0150	65 65 70 2d 61 6c 69 76	keep-alive..Autho
0160	72 69 7a 61 74 69 6f 6e	rization: Basic
0170	61 6d 39 6c 4f 6d 4a 73	am9lomjs b2dncw==
0180	0d 0a 0d 0a	...

HTTP Authorization header (http.authorization...) Packets: 542 · Displayed: 13 (2.4%) · Droppe... Profile: Default

Some examples are:

- pleasure -> cGxIYXN1cmU=
- leasure -> bGVhc3VyZQ==
- easure -> ZWFzdXJl
- asure -> YXN1cmU=
- sure -> c3VyZQ==

- Similar to base64
- The alphabet is [A-F, 0-9]
- Widely used to represent
 - MAC Address
 - Memory dumps

webcam after shutting dow

File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help

Apply a display filter ... <Ctrl-/>

No.	Time	Source	Destination	Protocol	Length	Info
13	8.398778	10.44.10.57	10.44.10.116	TCP	60	88 → 53730 [A
14	8.399410	10.44.10.57	10.44.10.116	TCP	60	88 → 53728 [A
15	8.400431	10.44.10.57	10.44.10.116	TCP	60	88 → 53726 [A
16	8.400798	10.44.10.57	10.44.10.116	TCP	60	88 → 53727 [A
17	8.400799	10.44.10.57	10.44.10.116	TCP	66	88 → 53737 [S
18	8.400865	10.44.10.116	10.44.10.57	TCP	54	53737 → 88 [A
19	8.401121	10.44.10.57	10.44.10.116	TCP	66	88 → 53738 [S
20	8.401163	10.44.10.116	10.44.10.57	TCP	54	53738 → 88 [A
21	8.401235	10.44.10.116	10.44.10.57	TCP	509	[TCP segment
22	8.405590	10.44.10.57	10.44.10.116	TCP	60	88 → 53738 [A
23	8.409973	10.44.10.116	10.44.10.57	TCP	132	[TCP segment
24	8.412961	10.44.10.57	10.44.10.116	TCP	60	88 → 53737 [A
25	8.469599	10.44.10.57	10.44.10.116	TCP	312	[TCP segment

▸ Ethernet II, Src: IntelCor_9b:d5:f7 (f8:16:54:9b:d5:f7), Dst: Shenzhen_0a:a4:3a (e8:ab:fa:0a:a4:3a)
▸ Destination: Shenzhen_0a:a4:3a (e8:ab:fa:0a:a4:3a)
▸ Source: IntelCor_9b:d5:f7 (f8:16:54:9b:d5:f7)
Type: IPv4 (0x0800)
▸ Internet Protocol Version 4, Src: 10.44.10.116, Dst: 10.44.10.57

0000	e8 ab fa 0a a4 3a f8 16 54 9b d5 f7 08 00 45 00
0010	01 ef 7d fa 40 00 80 06 52 0a 0a 2c 0a 74 0a 2c
0020	0a 39 d1 ea 00 58 a7 0b f5 d5 b5 58 75 6d 50 18
0030	04 00 0d 9f 00 00 47 45 54 20 2f 63 67 69 2d 62
0040	69 6e 2f 43 47 49 50 72 6f 78 79 2e 66 63 67 69
0050	3f 75 73 72 25 33 44 63 68 75 72 63 68 69 6c 6c

Bytes 54-508: TCP segment data (tcp.segment_data)

- 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"F5N8VJD:6YG(&UE=&A09',N($QE="=S(&UA:V4@=&AI<R!T97AT(&Q0
M;F=E<B!T:&%N"C4W(&)Y=&5S('10('=R87'@;&EN97,@=VET:"!"87-E-C0@
E9&%T82P@=&]0+@I'<F5E=&EN9W,L($9R86YK(%!I;&A09F5R"@' '
'
end
```

- When we analyze data, it might be represented in an unknown encoding
- How to identify the proper encoding?
 - experience
 - alphabet
 - patterns
 - origin of data



Questions? Feedback? Suggestions?



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

