Network-based Ransomware Detection

D. Mülders & P. Meessen

April, 13th 2017

Technische Universiteit
Eindhoven
University of Technology

Introduction

Today we present the results of SpySpot research into Ransomware and Intrusion Detection Systems.

Technische Universiteit
Eindhoven
University of Technology

Eindhoven Institute for
the Protection of Systems
and Information

SpySpot
Seeing is believing

Master in
Cyber Security

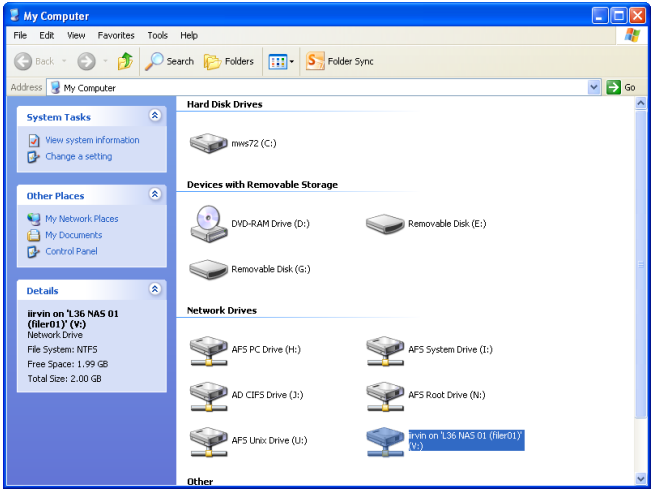
Ransomware

Ransomware is a class of malware, which interferes with the normal operation of a computer and aims to extort the owner of the computer into paying a ransom in order to undo or avoid further damage. - after: (Kharraz et al., 2015)

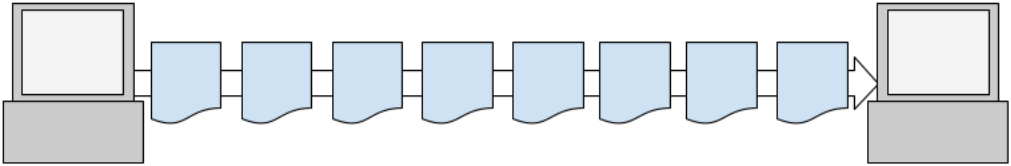
Ransomware using Encryption

This project focuses on ransomware that uses encryption (AES) to prevent victims accessing their files.

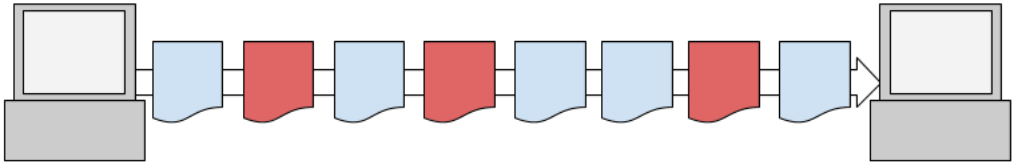
Windows Shares



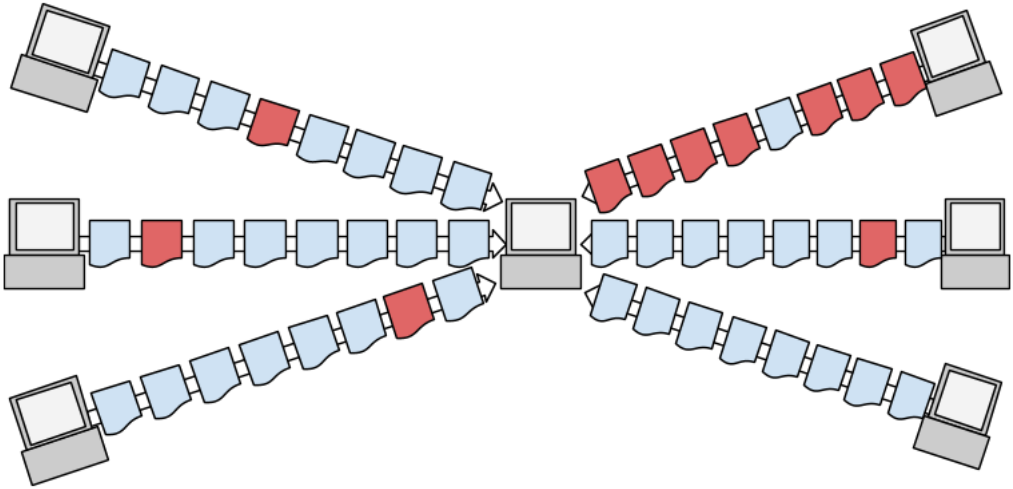
Example: SMB traffic



Example: SMB traffic

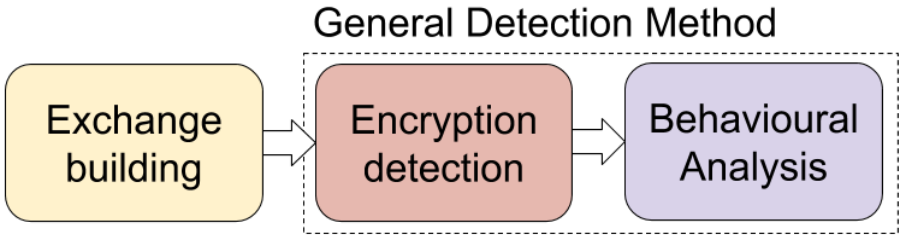


Example: SMB traffic

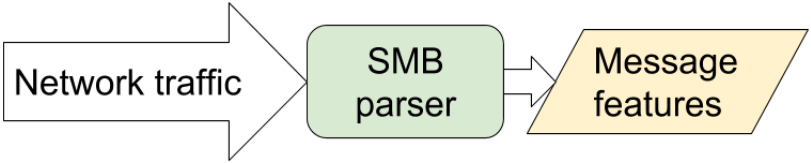


Ransomware detection

Some messages are recorded on the network traffic, that might contain ransomware. Extract crucial data, construct an exchange, detect encryption and analyze the general behaviour.



Message extraction

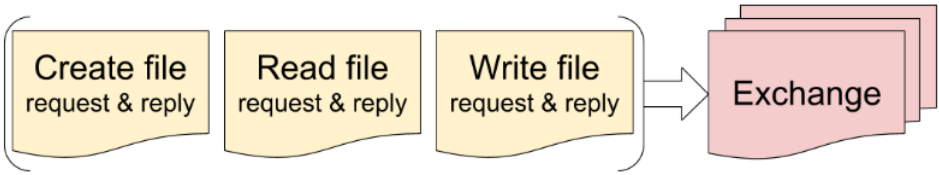


Message features

- file data
- data size
- name
- message & file identifiers
- etc.

Exchange building

By matching multiple related messages, based on their message identifiers & file manipulation patterns, we can build exchanges. The can contain encryption.



We can now calculate entropy, using the file data, which we can use to detect encryption.

Say we have two files:

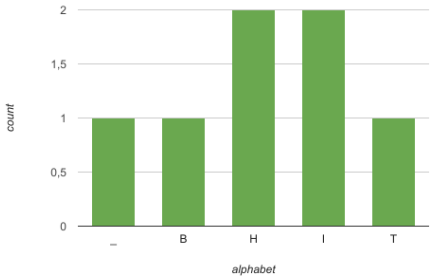
unencrypted file:
"HI HITB"

encrypted file:
"XMz5#a!"

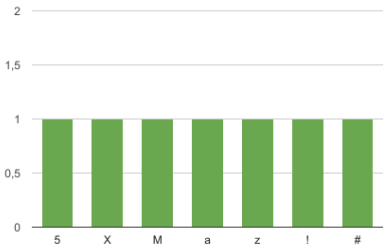
An $\mathbf{n}$-gram is the histogram of the substrings of length-$\mathbf{n}$ in a text.

1-grams

"HI HITB"

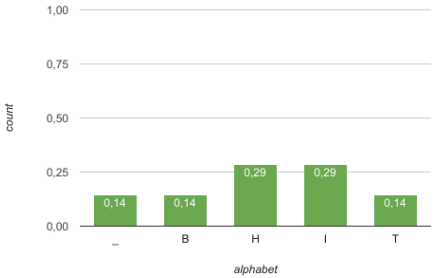


"XMz5#a!"

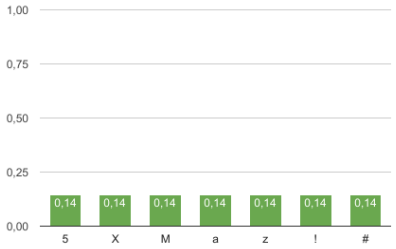


Distribution

"HI HITB"



"XMz5#a!"



From distributions to numbers

Lottery

	odds	payout
Lottery:	$\frac{1}{1.000.000}$	($100.000)
	$\frac{1}{1.000}$	($100)

$$\text{E(playing the lottery)} = \frac{1}{1.000.000} \times (\$100.000) + \frac{1}{1000} \times (\$100) = \$0,20$$

Expected Value

$$E(X) = \sum_{x \in X} p(x) \cdot f(x)$$

From Expected Value to Entropy

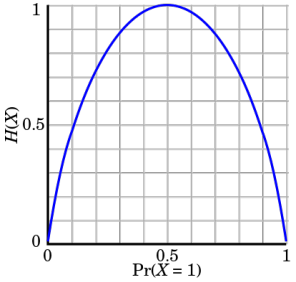
$$E(X) = \sum_{x \in X} p(x) \cdot f(p(x))$$

Payout function for Shannon Entropy:

$$f(p(x)) \Rightarrow \log_2(\frac{1}{p(x)})$$

Payout function for Shannon Entropy:

$$f(p(x)) \Rightarrow \log_2\left(\frac{1}{p(x)}\right)$$



https://commons.wikimedia.org/wiki/File:Binary_entropy_plot.svg

$$H(X) = \sum_{x \in X} p(x) \cdot \log_2 \frac{1}{p(x)}$$

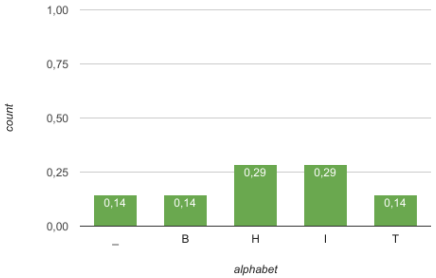
Shannon Entropy

"$\mathsf{H}(\mathsf{X})$ is the lower bound on the number of (yes/no) questions that you need to ask about $[\mathsf{X}]$ in order to learn the outcome $\mathbf{x}$."

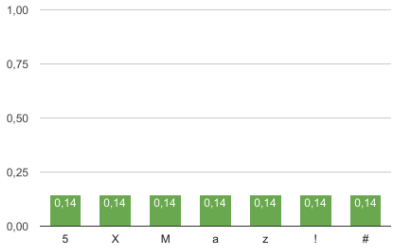
TU/e Course 2IMS10, Lecture Notes v1.7 2016

Distribution

"HI HITB"



"XMz5#a!"



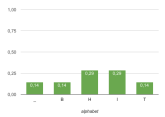
Calculating the Shannon Entropy for n-grams

"Text" $\Rightarrow$ 1-gram $\Rightarrow$ Distribution $\Rightarrow$ Entropy $\Rightarrow$ Number

Calculating the Shannon Entropy for n-grams

"HI HITB"

$$H\left(\ \right) =$$



$$3 \times (1/7) \cdot \log_2(7/1)$$
$$+$$
$$2 \times (2/7) \cdot \log_2(7/2)$$
$$= 2.2359\ldots$$

"XMz5#a!"

$$H\left(\ \right) =$$

$$7 \times (1/7) \cdot \log_2(7/1)$$

$$= 2.8073\ldots$$

Calculating the Shannon Entropy for n-grams

normal

"HI HITB"

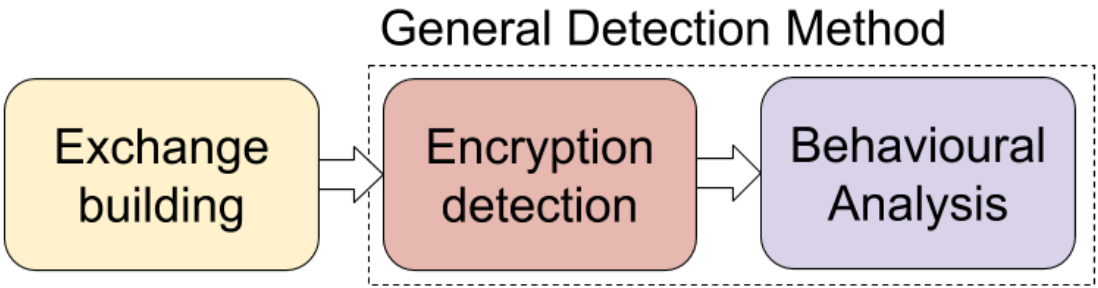
≈ 2.2

encrypted

"XMz5#a!"

≈ 2.8

Global system overview

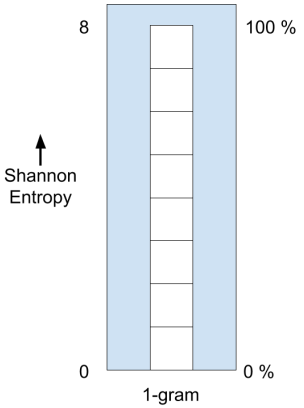


Encryption

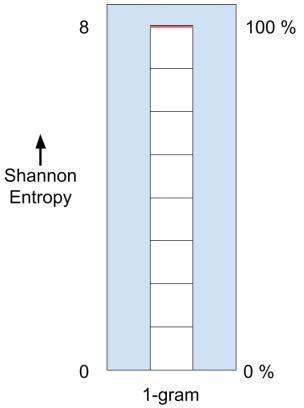
The 1-gram of an encrypted text should have:

- 8 bits of entropy, and
- use all the 256 characters in the ASCII alphabet.

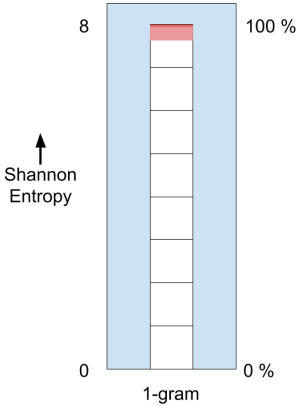
Simple Detector Rule



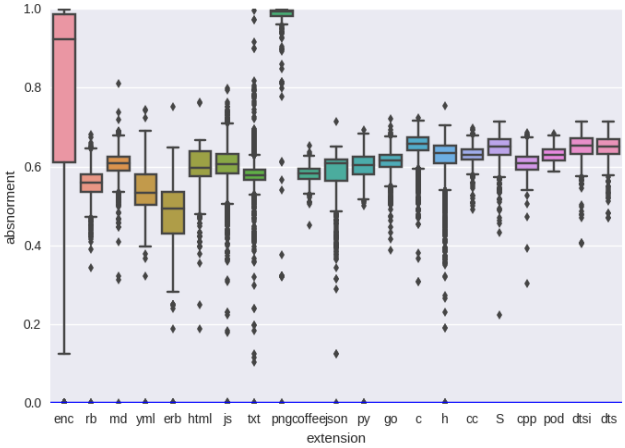
Simple Detector Rule



Simple Detector Rule



1-gram entropy for different file types



We need a new behavioral rule to remove the false-positives.

We can detect encryption on exchanges, using relative entropy and same size characteristics.

Encryption detection rate

Sample	FN	TP	TP rate	FP rate
CryptXXX	15847	2068	11.54%	0% !
CryptoWall	699	63	8.27%	0% !
JigSaw	19336	28887	59.90%	0% !
User	160	24	13.04%	0% !

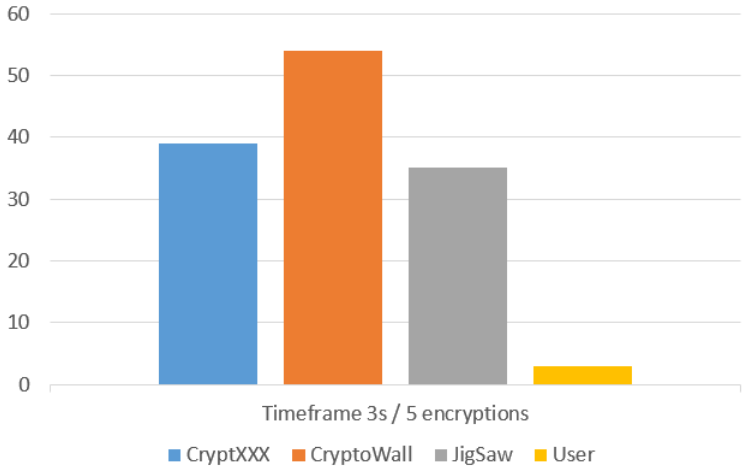
Using the detected exchanges, based on the rate of encryption, we can distinguish between ransomware & regular user traffic.

Analyse the rate of encryption, using varying time-frames and required number of encryptions.

Behavioural analysis

Sample	1s/5	1s/10	1s/15	3s/5	3s/10	3s/15	5s/15	5s/20	5s/25
CryptXXX	39	19	13	39	19	13	13	10	8
CryptoWall	54	27	18	54	27	18	18	14	9
JigSaw	35	17	12	35	17	12	12	9	7
User	0	0	0	3	0	0	0	0	0

Behavioural analysis results



Enterprise applications

- Ransomware & Intrusion detection system
- Blocking traffic from an infected client
- Backing up data that is being attacked

Questions

Please visit:
https://nomoreransom.org
http://security1.win.tue.nl/spyspot/

Special thanks to:

Tijmen van Dries, Sandro Etalle, Davide Fauri, Jerry den Hartog, Emil Nikolov, Erik Poll, Peter Wu, Rob Wu, Joe Joe Wong, Omer Yüksel.

References I

Kharraz, A., Robertson, W., Balzarotti, D., Bilge, L., and Kirda, E. (2015). Cutting the gordian knot: a look under the hood of ransomware attacks. In International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment, pages 3–24. Springer.

Nativ, Y. and Shalev, S. (2016-2017). thezoo. https://github.com/ytisf/theZoo.

Sikorski, M. and Honig, A. (2012). Practical Malware Analysis: The Hands-On Guide to Dissecting Malicious Software. No Starch Press Series. No Starch Press.

Stokkel, M. (2016). Ransomware detection with bro. Talk at BroCon '16, Austin, https://www.bro.org/brocon2016/slides/stokkel_ransomware.pdf.