WM0824 Economics of Cyber Security
Assignment 1

Maria Simidžioski Sam van Hooff Tom Slooff

September 2019

1 Introduction

1.1 What security issue does the data speak to?

The given data is collected by antivirus scanners on different web portals. When a scanner finds a virus in a web portal it collects data about the portal and the activity of the virus. The data captures how long the virus was seen in the portal, with a beginning and an end time. If known the type of virus is shown. And data about the portal / connection is given, such as the domain, the autonomous system it belongs to, the name server(s) used, the country the connection comes from, and more. This data can be used to create metrics which can be analyzed over time in order to evaluate the security level of these portals.

2 Methodology

2.1 What would be the ideal metrics for security decision makers?

The ideal metrics for security decision makers would include all of the four categories of metrics: controls, incidents, vulnerabilities and losses because each of these categories explores a different aspect of the threat environment.

2.2 What are the metrics that exist in practice?

The metrics that exist in practice are:

- the number of incidents over time.
- Control checks, e.g. HSTS policy, CSRF Tokens, etc.
- Vulnerability checks, e.g. remote javascript inclusion, X-XSS-Protection header

2.3 A definition of the metrics you can design from the dataset

- Length of time that a virus remained in a portal (domain)
- Number of viruses in a portal (domain) over time (e.g. per day)
- Ranking of the most attacked portals
- Virusname existence over time on different domains
- Percentage of portals that use SSL (checking http vs https)
- Virus preferences for specific countries

- Number of vulnerable URLs per domain (i.e. different urls where virus was found)
- Virus preferences for specific continents/areas (checked through source)
- inetnum used for normalization of domain, e.g. of number of vulnerable URLs per domain
- Number of domains in netname which are vulnerable
- Number of URLs in netname which are vulnerable
- Weighted rank of domain by ranking URL vulnerabilities using time. (e.g. vuln of 2 years weighs more heavily than vuln of 1 day)
- Weighted rank of netname by using ranked domain vulnerabilities

3 Results

3.1 An evaluation of the the metrics you have defined.

4 Discussion