Poster: Misuseablity Analysis for IT Infrastructure

Asaf Shabtai, Yuval Elovici
Ben-Gurion University of the Negev
{shabtaia, elovici}@bgu.ac.il

ABSTRACT

Today, organizations have limited resources available to allocate to the detection of complex cyber-attacks. In order to optimize their resource allocation, organizations must conduct a thorough risk analysis process so as to focus their efforts and resources on the protection of the organization's important assets. In this study we propose a framework that automatically and dynamically derives a misuseability score for every IT component (e.g., PC, laptop, server, router, smartphone, and user). The misuseability score encapsulates the potential damage that can be caused to the organization when its assets are compromised and misused.

Categories and Subject Descriptors

C.2.0 [**Computer-Communication Networks**]: General – *security and protection*

General Terms

Security.

Keywords

Security, misuseability score, risk analysis, insider threat.

1. INTRODUCTION

The vast number of cyber-attacks facing organizations today makes their detection a very difficult task. The challenges include: the need to analyze the massive amount of data that is collected by information technology (IT) infrastructure; the highly advanced cyber-attacks that are continuously becoming more and more sophisticated; the fact that attacks can be originated by insiders or external entities; the rapid introduction of new technologies integrated into the organization's infrastructure; and the variety and costs of security solutions.

Because of these challenges organizations must conduct a thorough risk analysis process in order to focus their efforts and resources on the protection of highly critical assets. In the risk analysis process the organization identifies the most important assets by a assigning a risk value to an asset; the risk value is a function of the asset's value and the likelihood of the threat to be realized. This process is time consuming and therefore ignores the dynamic nature of the IT infrastructure. Because of the dynamic nature of the IT infrastructure the value of assets may change over time and will not be reflected by the risk analysis process results.

In this study we propose a dynamic framework for supporting the risk analysis process. This is done by automatically and dynamically deriving a misuseability score for every IT component (e.g., PC, laptop, server, router, smartphone, and user). The misuseability score encapsulates the potential damage that can be caused to the organization when an asset is compromised and misused as part of a cyber-attack.

Harel *et al*. [1] initially addressed this challenge and presented a new concept, *Misuseability Weight*, which assigns a sensitivity score to data, thereby estimating the level of harm that might be inflicted upon the organization when the data is leaked. Assigning a misuseability weight to a given dataset is strongly related to the way the data is presented (e.g., tabular data, structured, or free text) and is domain-specific. Harel *et al*. [1] focus on mitigating leakage or misuse incidents of data stored in databases (i.e., tabular data) and presented the *M-Score*, a misuseability weight measure for tabular data.

Vartanian and Shabtai [2] proposed an extension of the misuseability weight concept and specifically focused on textual content. The main goal is to define a misuseability measure, termed *TM*-Score, for textual content. Using this measure it is possible to estimate the extent of damage that can be caused by an insider that is continuously and gradually exposed to documents. The extent of damage is determined by the amount, type, and quality of information to which the insider is exposed.

In this study we extend the misuseability concept to a full and comprehensive framework that is able to derive a misuseabilty score for each IT element including personal computers, servers, smartphones, databases, routers, switches, and users.

This concept is depicted in Figure 1. For each IT element a misuseability score is derived. By analyzing the connectivity level among the elements it is possible to identify clusters of elements that are not only highly connected but also have a high misuseability score as a group and therefore should be carefully analyzed and protected.

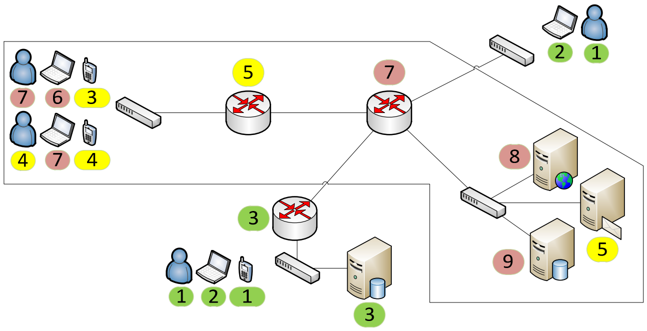

Figure 1. Example of the misuseability analysis framework. A set of IT elements that are highly collected are grouped together.

2. PROPOSED METHOD

2.1 Misuseablity Score

The misuseability score is a quantitative measure that is calculated for each IT element based on data collected from the element itself (using a dedicated agent) or externally from the network traffic.

It is derived by applying a predefined function on a set of relevant parameters that are calculated from the collected data. Thus, given

CCS'14, November 3-7, 2014, Scottsdale, AZ, USA
ACM 978-1-4503-2957-6/14/11.
http://dx.doi.org/10.1145/2660267.2662385

a set of parameters, p_1, p_2, …, p_n that are dynamically computed from the data that is collected from the IT element, the misuseability score of the IT element can be computed according to the following equation:

$$M - Score(e, c) = \sum_{i=1}^{n} \alpha_i \cdot p_i$$

where $\alpha_1, \alpha_2, \dots, \alpha_n$ are constants which define the importance of each parameter. These constants should be provided by the security officer or derived automatically. In the equation c refers to the specific context for which the misuseability score is computed. For example, the context can be determined by the day of week and part of day for which the misuseability score is derived.

In this section we present a set of possible misuseability measures.

Router *M*-Score (*RM*-Score)

A router misuseability score depends on the amount and quality of information that is handled by the router as well as on the activation of various capabilities of the router. For example, potentially, the more data that is handled by the router the higher the misuseability score. Data that is sent to/from a sensitive machine (e.g., a machine with a high misuseability score) via the router increase the misuseability score of the router. In addition, a specific capability that is activated on the router (e.g., a VPN, VLAN) increases the attractiveness of the router and the potential misuseability. In addition, the type and version of the router may influence the misuseability score because an old operating system or outdated firmware may result in a more vulnerable device.

Therefore the *RM*-Score is derived from the following parameters:

- Configuration
    - Vendor
    - OS type
    - OS version
    - Known vulnerabilities
    - Activated functionalities (e.g., segmentation, tunneling/encryption, quality of service)
- Purpose
    - Location (gateway or internal)
    - Importance of connecting networks
    - Sensitivity level of IP addresses (for example, according to the computed misuseability score of the IP address)
- Activity
    - Amount of data transmitted
    - Number of distinct IP addresses
    - Percentage of encrypted traffic

Server *M*-Score (*SM*-Score)

A server misuseability score depends on the volume of activity of the server, the number of connected users, the type and importance of the services that are provided by the server, and the properties of the server such as the open ports, running services and the type and version of the operating system.

Therefore, the *SM*-Score is derived from the following parameters:

- Configuration
    - OS type
    - OS version
    - Known vulnerabilities
    - Number and type of open ports
    - Running services
    - Number of users including administrators
- Purpose
    - Importance of the service provided by the server
    - Misuseability score of the connected hosts (*HM*-Score)
    - Location of server (e.g., internal, DMZ)
- Activity
    - Number of hosts served by the server
    - Volume of activity
    - Volume and type of network traffic

Host *M*-Score (*HM*-Score)

The Host *M*-Score is computed for each end-user device which may include personal computers, laptops, and smartphones.

Similar to a sever misuseability score, the host misuseability score depends on the volume of the activity, the number of users, the type and importance of the services that are used, and the properties of the host such as the type of the device, open ports, running services, services used such as secured remote connection, and the type and version of the operating system.

The host misuseability score may be derived from the following parameters:

- Configuration
    - Host type
    - OS type
    - OS version
    - Known vulnerabilities
    - Number and type of open ports
    - Running services
    - Number of users including administrators
- Purpose
    - Importance of the service provided by the server
    - Misuseability score of the users (*UM*-Score)
    - Location of host
- Activity
    - Running applications and services
    - Number of users using the host machine
    - Volume of activity
    - Volume and type of network traffic
    - Connected networks (both wired and wireless)

Tabular Data *M*-Score (*DM*-Score)

The *DM*-Score measure [1] estimates the extent of damage that can be caused by an insider that is continuously and gradually exposed to tabular data; i.e., datasets (e.g., result sets of relational database queries). The *DM*-Score is primarily influenced by the number of entities exposed to the insider (i.e., number of records), the number of properties available on each entity (i.e., number of attributes), the value of properties, and the anonymity level which is regarded as the effort that is required in order to fully identify a specific entity in the data.

Textual *M*-Score (*TM*-Score)

The *TM*-Score measure [2] estimates the extent of damage that can be caused by an insider that is continuously and gradually exposed to documents. The extent of damage is determined by the amount, type, and quality of information to which the insider is exposed. This is done by deriving an accumulated *TM*-Score each time that the user is exposed to a document (e.g., opening a file, printing a file, copying a file to a storage device). The *TM*-Score is cumulative in the sense that it considers the documents that the insider was exposed to in the past as well as the recently exposed document. The main challenge in deriving the cumulative *TM*-Score is the identification of residual information in the recently exposed document (i.e., identifying exactly what is the true new information in the document compared to that of the previously exposed documents) and its contribution to the cumulative *TM*-Score.

User *M*-Score (*UM*-Score)

In general, the user *M*-Score is derived from the type of services and machines that the user accesses as well the type of data and information that the user is exposed to. It can also be derived from general behavioral patterns of the user (e.g., a salesperson who travels a lot and connects to many WiFi networks). Therefore, the *UM*-Score is computed from the following parameters:

- Demographic features
  - Age
  - Role in the organization
  - Seniority
- Activity
  - Number of systems and services that the user accesses and the misuseability score of each systems
  - Volume of activity (i.e., how often the user accesses the systems and services)
  - Volume and type of generated network traffic (e.g., encrypted traffic)
  - Connected networks (both wired and wireless)
- Sensitivity of data
  - User's derived *DM*-Score
  - User's derived *TM*-Score

Figure 2 summarizes the connections among the misuseability scores of the various IT components and presents the derivation tree of misuseability scores.

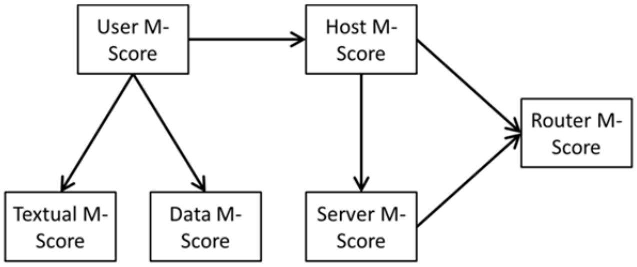

Figure 2. Derivation tree of misuseability scores of IT components.

2.2 Aggregation and analysis

As illustrated by Figure 1, the IT infrastructure of an organization can be represented as a graph. The nodes of the graph indicate an IT element and are assigned with attributes such as type of IT element and its computed misuseability score. The links may indicate the strength of the connection between elements. The strength of the connection between elements can be derived from the amount of traffic or activity between two elements (e.g., a server and router or a PC and server).

Given such a graph, various graph clustering and community detection algorithms may be applied in order to identify clusters of IT elements that as a group are highly connected and highly misuseable. Note that each individual element in the identified group may not necessarily have the highest misuseability score. The security officer of the organization should focus on those highly misuseable sets and increase protection of these sets, for example by raising the awareness of the users, adding monitoring security measures, and lowering the thresholds of existing security measures (although lowering the thresholds may come on the account of a higher false positive rate).

In addition, an anomaly detection process can be applied to the misuseability score of each IT element in order to learn, for each context, the acceptable misuseability score of the IT element and identify significant deviations from the misuseability level. For example, when the tunneling functionality of a router is disabled, the misuseability score of the router should be reduced significantly, a fact that should signal an alert to the security officer.

3. CONCLUSION AND FUTURE WORK

In this study we extend the misuseability weight concept to a full, comprehensive, and dynamic framework that derives a misuseabilty score for each IT element. The misuseability score encapsulates the potential damage that can be caused to the organization in case when an asset (i.e., IT element) is attacked or misused.

In future work we plan to implement the proposed framework and evaluate it on real data collected from actual IT infrastructure. In addition, we plan to propose methods for the automatic assignment of weights to each parameter of the misuseability score function.

4. REFERENCES

[1] Harel, A., Shabtai, A., Rokach, L., and Elovici, Y., 2012. *M*-score: A misuseability weight measure. IEEE Trans. on Dependable and Secure Computing, 9(3), 2012, 414-428.

[2] Vartanian, A., Shabtai, A., 2014. *TM*-Score: A Misuseability Weight Measure for Textual Content", IEEE Trans. on Information Forensics and Security.