AN ALTERNATIVE APPROACH FOR FORMULA MODELLING IN SECURITY METRICS

Rodrigo Sanches Miani, Felipe Marques Pires and Leonardo de Souza Mendes
Department of Communication, School of Electrical and Computer Engineering, State University of Campinas
Av. Albert Einstein, 400, Cidade Universitaria "Zeferino Vaz", Distrito Baro Geraldo, Campinas, SP, Brazil

Keywords: Security metrics, Network security, Security analysis.

Abstract: This paper proposes an alternative approach to modelling the formula attribute within the context of security metrics. This approach seeks to correct past errors by treating a security metric like a set, and inserting a component that addresses the set intersection between the security elements. The work consists in to define the model, explain the differences to the previous model and validate it, with examples from the metrics found in literature and also with the results of a case study applied in Metropolitan Broadband Access Network in Pedreira, a city located in the state of So Paulo, Brazil.

1 INTRODUCTION

A widely used concept in the information security scope, is the security metric. Metrics can be defined as a set of measures that can generate a quantitative approach about a problem. (Lowans, 2002).

The primary goal of a metric is to convert raw data into information capable of analysis. In the information security world, large organizations as CERT (Computer Emergency Response Team), SANS (SysAdmin, Audit, Network, Security) and NIST (National Institute of Standards and Technology) develop and recommend the implementation of security metrics.

The metrics are usually defined from a series of attributes. Among them we highlight: purpose, frequency, data source, measures and formula. The formula attribute, in particular, is important to describe the calculations that will be performed to quantify the metrics in a numerical expression (Swanson et al., 2003). From the result of the formula, the metrics value or indicator is obtained, and it is usually expressed in percentage terms.

This paper will present an alternative approach that was proposed in (Miani et al., 2008) for the modelling of security metrics formulas, aiming to increase the reliability degree from the results obtained by the formulas. We will show the differences between the two approaches and a comparative study of the research results of a case study performed in the Metropolitan Broadband Access Network of Pedreira.

This paper is organized as follows. Section 2 brings some related works of security metrics. The section 3 presents the mainly motivations and defines the concepts that will be studied in this work. In section 4 we will introduce the basis of the proposed model and the differences between the model that will be compared. In section 5 we present one model application example and the results of a metrics implementation case study developed in the Metropolitan Broadband Access Network of Pedreira. The section 6 brings the conclusion and some future works.

2 RELATED WORKS

The study of security metrics and their applications in the IT scenarios are targets of several discussions (Rosenblatt, 2008). An increase in failure rate of components, discovered vulnerability in software and communication network attacks may cause a big concern on questions related to information security. To deal with these problems it is necessary to invest in security controls implementation and security policies. The characteristics of these investments must be carefully accounted for and can be defined from measures and analysis of information security structure. This process is formalized using a security metrics application (Weiss et al., 2005). Through a combination
of predefined objectives, collection and data analysis, metrics can indicate the actual level of security we must aim at, directing the actions network administrators must take to secure the network (Payne, 2006).

Jaquith (Jaquith, 2007), Swanson et al. (Swanson et al., 2003), Payne (Payne, 2006) and the ISO/IEC 27002 (ISO, 2005) standard, contributed to the development and formalization of attributes that constitute a security metric.

Herrera (Herrera, 2005) examined the indicators development from the security metrics, affirming that there is no magical formula to establish the perfect indicator, each organization shall determine which indicators are useful according to their business and how to get to the results. However, the growth and the need for security metrics has created a gap in the concepts standardization. Every security framework provides its own security indicator method and this fact may affect the development of security indicators.

The CVSS (Common Vulnerability Scoring System) (Mell et al., 2007) is an initiative in this direction. Its goal is the creation of standard indicators for security vulnerabilities from the equations that are divided in three groups of measures: base, temporal and environmental. Popular vulnerability scanners such as Nessus, already use the CVSS in their database, as well as the NVD - National Vulnerability Database of the U.S. government, maintained by NIST.

Other efforts in the standardization of measures in the information security area can be found in (Jelen and Williams, 1998). In this work, Jelen and Williams argue that assurance is an integral part of the risk and security management process. They propose a formulation intended to be universal in the measuring assurance area. Weiss et al. (Weiss et al., 2005) proposes a model for the security level calculation of an organization through the percentage of lost assets.

Miani et al. (Miani et al., 2008) defines a model for formula calculation of security metrics. The attributes that constitute the model are: objective, metric, measure, data source, frequency, metrics classification and formula. The model standardizes the nomenclature of terms relative to security metrics, and proposes the formula calculation in a generic way contributing to the decreasing of subjective criteria in the metrics formulation. The model define the calculation only using the arithmetic mean. This work proposes a model that using set theory and the inclusion of a new component, the intersection component, that seeks to correct possible flaws in the model interpretation proposed by Miani et al.

### 3 METRICS FORMULATION

The security metrics model, proposed in (Miani et al., 2008) has a different characteristic from the traditional approaches, that is from the grouping of multiple metrics in a common group, calculate the security indicator of this group. For example, consider the following security metrics proposed by ISO/IEC 27002:

- \( P_1 = \) Percentage of communication channels controlled by the organization that have been secured in accordance with policy.
- \( P_2 = \) Percentage of mobile users who access enterprise facilities using secure communication methods.
- \( P_3 = \) Percentage of workstation firewalls, host firewalls, sub-network firewalls, and perimeter firewalls configured in accordance with policy.

Although the metrics are organized into a common group called “Network access control”, they are individually treated. Each metric has its own formula and there are no recommendations on how to analyze the whole group. The model proposed by Miani et al. consists in, to combine the three metrics in only one, aiming the overall group analysis by calculating a single formula representing the “Network access control” level.

The grouping is important because it unifies several results in only one number, easing the results interpretation of the non-technical organization staff. When necessary the calculation of individual metrics can also be performed.

In this case, even according to the proposal, the formula of the group “Network access control” it would be calculated as follows:

Consider the components, \( P_1, P_2 \) and \( P_3 \). The next step is to examine the component security. \( P_1 \) is secure, because when the number of secure communication channels increase, the risks of security problems decrease. Analogously, \( P_2 \) and \( P_3 \) are secure components too. Thus, the formula is given by the arithmetic mean between \( P_1, P_2 \) e \( P_3 \).

The simplicity and possible flaws in the interpretation of results, are the main motivating factors for the development of a new reliable model. Take the following example:

**Example 1.** Consider a metric \( M \) which aims to measure the security between the connections of the MBAN buildings. Let \( A_i \) be the set of buildings that constitute the network. Consider two secure components \( A_1 \) and \( A_2 \) of the metric such that: \( A_1 \) is a subset of \( A_i \), where \( A_1 \) represents the number of buildings that have firewall resources or another logical access control and \( A_2 \) is a subset of \( A_i \), where \( A_2 \) represents
the number of buildings that have ciphered connections.

Miani et al. states that the metrics formula, in this case would be the arithmetic mean between the components. However, we can consider the existence of another subset $A_3$ with $A_3 = A_1 \cap A_2$, in other words $A_3$ is the set of buildings that have firewall resources and encryption between connections. This new parameter must be part of the formula. For this consider the number of security resources in a system, in general, more quantity of security resources implies in more security. In our case, the set $A_3$ has more security resources then the sets $A_1$ and $A_2$ and hence, the set $A_3$ will have a greater weight in relation to other sets. Therefore, we can conclude that the intersection quantity of a component must affect its weight.

This is the major motivation of this work, to develop a standardized model capable of correcting the inaccuracies of the model proposed by Miani et al. and encourage its use in any kind of security metric.

4 MODEL DESCRIPTION

Take the set of the components $a_1, a_2, a_3, ..., a_n$. For each $a_i$, let $a_i$ be the maximum value that this measure assumes. Rewriting this sentence using the set theory notation, we have one set for each component $a_1, a_2, a_3, ..., a_n$. The correspondent $a_i$ will be the set that contains the respective $a_i$. Then, $a_1 \subset a_1, a_2 \subset a_2$ and so on.

The model objective is to increase the reliability of the security index calculation. For this, the components in the formula calculation will be balanced by using different weights and another factor will be presented: the intersection component between the sets.

The first step to update the calculation of the formula is to verify if the metric has a maximum component or maximum value set, with two or more subsets. Then, you should classify every metrics component.

The model proposed by Miani et al. states that, given a metric $M$: i) $M$ is composed only by secure components, ii) $M$ is composed only by insecure components and iii) $M$ is composed by insecure and secure components.

Let us first consider the case where the metric is composed only by secure components. The other cases will be deducted from this.

In cases where maximum value sets with two or more related subsets do not exists, the formula calculation is reduced to the mean between the components. Here we only consider the existence of maximum value sets with at least two related subsets.

We begin the formula construction for the case where the number of subsets of a maximum value set is equal to 2, then the case where such number is 3 and at last the formula will be generalized. Consider a metric $M$, consisting of a maximum value set $T$ and two sets $A_1$ and $A_2$ such that $A_1 \subset T$ and $A_2 \subset T$. Let $I_{1,2}$ be the set formed by the intersection between $A_1$ and $A_2$. The cardinality of these sets are: $a_1 = \#A_1, a_2 = \#A_2, i_{1,2} = \#I_{1,2}$ and $t = \#T$.

The formula for the case where the number of subsets is 2, will be built using a weighted mean between the subsets. The weights will be distributed as follows: 2 for the intersection component and 1 for the other components. Note that the weight 2 represents the number of subsets of $T$. Then the formula will be written as:

$$F_2 = \frac{(2 \times a_2) + (1 \times i_{1,2}) + (1 \times i_{1,2})}{2(2 + 1)}$$

It is important to make an analysis of the maximum and minimum of the formula. A metric with formula equal to 0 represents that no security requirements have been accomplished. Similarly, a metric with formula equal to 1 represents that the security requirements have been accomplished. However, formula equal to 1, does not mean that security is fully accomplished.

The maximum security is achieved when the number of elements of $I_{1,2}$ is equal to the number of elements of $T$, that is, $i_{1,2} = t$. On the other side this is only possible when $A_1 = A_2$. If $A_1 = A_2$ then $I_{1,2} = A_1 = A_2$ and $a_1 = a_2 = i_{1,2} = t$. Calculating the formula, we obtain 1.

The minimum security is achieved if no security requirements were met, that is, if $A_1 = A_2 = I_{1,2} = \emptyset$. Calculating the formula, we obtain 0.

In other words, the maximum and minimum analysis shows that the developed formula is consistent with the defined requirements.

Consider a metric $M$, composed by a maximum value set $T$ and three sets $A_1, A_2, A_3$ such that $A_1 \subset T, A_2 \subset T$ and $A_3 \subset T$ and the intersection sets $I_{1,2}, I_{1,3}, I_{2,3}$ and $I_{1,2,3}$. The cardinality of these sets are: $a_1, a_2, a_3, i_{1,2}, i_{1,3}, i_{2,3}, i_{1,2,3}$ and $t$.

The weights will be distributed as follows: 3 for the intersection set $I_{1,2,3}$, 2 for the other intersection sets, $I_{1,2}, I_{1,3}$ and $I_{2,3}$ and at last 1 for the other sets. Thus,

$$F_3 = \frac{(3 \times i_{1,2,3}) + (2 \times i_{1,2}) + (2 \times i_{1,3}) + (2 \times i_{2,3}) + (1 \times i_{1,2}) + (1 \times i_{1,3}) + (1 \times i_{2,3})}{3(3 + 2 + 2 + 1 + 1)}$$

The same analysis of maximum and minimum should be made here. Note that the results do not change because the requirements for the maximum security level is that $I_{1,2,3} = t$ that is, $A_1 = A_2 = A_3$.

Now we can generalize the formula calculation for the case where the number of subsets of $T$ is $n$. 

383
Consider a metric \( M \) which is composed by one maximum value set \( T \) with \( A_1, A_2, \ldots, A_n \) subsets of \( T \). We denote the cardinality of this sets as follows: \( a_j = \#A_j \).

The formula will be built using a general rule for obtaining each of the terms. The term of weight \( n \) is obtained by the ratio between the cardinality of the intersection of the \( n \) subgroups and the cardinality of the set \( T \). The term of weight \( n-1 \) is obtained adding all the ratios between the cardinality of the intersection of \( n-1 \) sets and the cardinality of the set \( T \). Continuing this process, all terms will be obtained. The denominator is formed by the sum of the weights of each of the terms. Each of the combinations \( C_n^k \) represents the number of subsets in each of 1 to \( n \) terms. A generalized version of the formula is:

\[
F_n = \frac{n}{n(2) + \ldots + (n-1)(2) + \ldots + 1(2)} + \ldots + \frac{n-n}{n(2) + \ldots + (n-1)(2) + \ldots + 1(2)} + \ldots + \frac{n-n}{n(2) + \ldots + (n-1)(2) + \ldots + 1(2)}
\]

However, we need to recall one last detail. The formula is valid only for one maximum value set. For \( m \) sets we should do the calculation for each one, and then calculate the arithmetic mean between the results.

### 4.1 Differences between the Models

This section aims to show that the inclusion of weights makes the model presented here more precise than the model proposed in (Miani et al., 2008). In other words, the formula results presented in this work are always smaller than the formula results proposed in (Miani et al., 2008). However, we need to recall one last detail. The formula proposed by Miani et al. is given by

\[
F_1 = \frac{a_{i.2}}{a_2}, \quad \text{and the formula proposed in this work is given by} \quad F_2 = \frac{2a_{i.2} + a_1}{a_2}.
\]

Therefore, we would like to demonstrate that

\[
a_1 + a_2 \geq \frac{3}{4}a_{i.2}.
\]

Using the proof by contradiction, we obtain \( a_1 + a_2 < 2a_{i.2} \). This contradicts our assumption because, if we sum the inequalities \( i_{1.2} \leq a_1 \) and \( i_{1.2} \leq a_2 \) we have \( a_1 + a_2 \geq 2i_{1.2} \). In this case, the formula proposed in this work is always less or equal than the formula proposed by Miani et al.

For the other cases the demonstration is analogue, using the proof by contradiction in the obtained inequalities. The case that the number of sets is equal to 3, for instance, the following inequality must be proved:

\[
a_1 + a_2 + a_3 \geq \frac{3}{4}(a_{1.2.3} + 2a_{i.2.3} + 2a_{i.3} + a_1 + a_2 + a_3)
\]

regarding the validity of the following inequalities: i) \( i_{1.2.3} \leq a_1 \) and \( i_{1.2.3} \leq a_2 \) and \( i_{1.2.3} \leq a_3 \), ii) \( i_{1.2} \leq a_1 \) and \( i_{1.2} \leq a_2 \), iii) \( i_{1.2} \leq a_1 \) and \( i_{1.2} \leq a_3 \) and iv) \( i_{2.3} \leq a_2 \) and \( i_{2.3} \leq a_3 \).

### 5 RESULTS AND APPLICATION EXAMPLE

In this section will be presented an application example of the proposed model in security metrics found in (ISO, 2005) and the case study, using the proposed model in the Metropolitan Broadband Access Network (MBAN) of Pedreira.

#### 5.1 Application Example

Consider the metrics of the Communication and Operations Management group, proposed in (ISO, 2005). Within this group we can identify three metrics that can address the same security control: backup. In our proposal, will be created a new group called “Backup Policy” containing such metrics. The metrics definition are,

1. Assets backed up: measures the percentage of systems with critical information assets that have been backed up in accordance with policy.
2. Assets backup validated: measures the percentage of systems with critical information assets where restoration from a stored backup has been successfully demonstrated.
3. Assets backup offsite: measures the percentage of backup media stored offsite in secure storage.

The metrics, separately, would be calculated as follows. Consider \( a_i \) = total number of assets, \( a_1 \) = number of assets backed up, \( a_2 \) = number of assets with backup procedures and validated, \( a_3 \) = number of assets whose backups are stored off-site. So, for the first metric, we have \( \frac{a_1}{a_i} \) for the second metric \( \frac{a_2}{a_i} \) and finally for the third metric \( \frac{a_3}{a_i} \). Note that the three components, \( a_1, a_2 \) and \( a_3 \), have the same total value component, number of assets, allowing that the proposed
model may be applied here. Applying the model presented in this work, the new formula of the group “Backup Policy” will be calculated like this:

\[
F = \frac{(3)\left(\frac{a_1}{t}\right) + (2)\left(\frac{a_2}{t}\right) + \frac{a_3}{t}}{\text{total number of servers}}
\]

Thus, we have a security indicator for the whole “Backup Policy” task. The model can be applied to any set of metrics within requirements, easing the elucidation of the results and producing an efficient and balanced overview of a security question.

5.2 Metrics Application in the MBAN of Pedreira

Metropolitan broadband access networks (MBAN) can be defined as the convergence of services, applications and infrastructure to create a community communications network of a city. This implements the public information highway, characterized for high bandwidth transmission capacity and data aggregation of several types (Mendes, 2006). Further informations about this kind of network can be found in (Alexiou et al., 2006).

The MBAN of Pedreira is a project that has been developed by the State University of Campinas (UNICAMP) and by the government of the city of Pedreira. The project started in 2005 and officially launched in 2007. The detection of security vulnerabilities in Pedreira’s network was the main motivation for the development of particular metrics that could quantify the high amount of data generated by technical reports and management software. In next will be presented the results of three security metrics applied in this period. Also will be showed a comparison between the two models discussed in this work. The metrics are: i) Security between the MBAN buildings, ii) Security requirements in the VoIP network and iii) Availability and reliability in the MBAN servers.

Security between the MBAN Buildings

The aim here is to analyze and to increase security level among the MBAN buildings. The formula components are the following: \(a_1\) = total number of buildings, \(a_2\) = number of buildings that use firewall resources or logical access control in their connections, \(a_3\) = number of buildings that use ciphered resources in their connections and \(i_1\) = number of buildings that have both firewall resources and encrypted connections. The formula also have a factor \(p\) that varies accordingly the size of the cryptographic protocol used, attached to the \(a_2\) component. The formula is given by:

\[
F_1 = \frac{2\left(\frac{a_1}{t}\right) + \frac{a_2}{t} + \frac{a_3}{t}}{\text{total number of servers}}
\]

Availability and Reliability in the MBAN Servers

The aim here is to evaluate the impact of the unplanned downtime in the services deployed by the MBAN servers. The formula components are the following: \(a_1\) = total number of servers, \(a_2\) = total number of hours, \(a_3\) = number of servers with redundancy resources, \(a_4\) = number of servers that are in the backup program, \(a_5\) = number of servers that stores the backups in security offsite, \(a_6\) = uptime mean of servers and \(i_1\) = number of servers with both redundancy and in the backup program. The formula is given by:

\[
F_2 = \frac{2\left(\frac{a_1}{t}\right) + \left(\frac{a_2}{t}\right) + \left(\frac{a_3}{t}\right)}{2} + \left(\frac{a_4}{t}\right)
\]

Table 1 shows the result of each one of the metrics and also compares with the model proposed by Miani et al. We denote Model 1 for the model proposed by Miani et al. and Model 2 for the model proposed in this work.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Formula Model 1</th>
<th>Formula Model 2</th>
<th>Decrease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security between the MBAN buildings</td>
<td>0.5411</td>
<td>0.3117</td>
<td>42.39%</td>
</tr>
<tr>
<td>VoIP security requirements</td>
<td>0.7296</td>
<td>0.6046</td>
<td>17.13%</td>
</tr>
<tr>
<td>Server’s availability and reliability</td>
<td>0.7496</td>
<td>0.7217</td>
<td>3.72%</td>
</tr>
</tbody>
</table>

According to what was showed in section 4.1, the model proposed here achieved lower results. The column “Decrease” illustrates the difference between the results of the two models. This difference is obtained
from the values of the component intersection. Higher values imply in the decrease of distance between the formulas. Similarly, lower values imply in the increase of distance between the formulas.

Besides the intersection component, the way that the metrics formula is obtained can also influence on the difference between the models. If the formula has components that require in its composition the insertion of additional arithmetic mean, such as have simultaneously secure and insecure components (metric 2) or have components outside the intersection (metric 3), these components will act as follows: values near to 1 decrease the difference and values near to 0 increase the difference.

6 CONCLUSIONS

Security metrics are modern tools and with high research potential. They are extremely important for the security level understanding of the organization when properly developed and applied.

A classic security metric has several components, including: objective, data source, frequency, classification and formula. The purpose of the formula, in particular, is to describe the calculations to be performed for quantify the metrics in a numerical expression. That is, the metrics results are investigated from the formula. It is important that this task be accomplished in a clear, robust and generic way.

The model proposed in this work sought to correct the inaccuracies of the model proposed by Miani et al. developing a new component, which deals with sets intersections of security measures. This component plays an important role in the model, distributing the weights in the proposed formula. Besides the formula, the whole nomenclature and the logic construction developed in this work can be reused to build other models in this area.

The model validation it was obtained in two ways: from the metrics application found in literature and with a case study. Classic security metrics as found in (Jaquith, 2007), (Swanson et al., 2003) and (ISO, 2005) are easily migrated to our model. One of the benefits is the aggregation of various measures in only one, easing the overview and the results interpretation of the non-technical organization staff. Besides that, the proposed model was used in three security metrics that were implemented in the MBAN of Pedreira. The results showed that the model proposed here achieved lower results when compared to Miani et al. model and could also explain how the numerical differences between the models are established.

Future works includes the model utilization in other security metrics, aiming to create its own catalog, such like what is developed in the Metrics Catalog Project (MetricsCenter, 2008) and the application of new case studies to refine the proposed model in private institutions, government and other MBANs enabling the development of a security metrics database.

REFERENCES


