

Smarter Computer Intrusion Detection Utilizing Decision Modeling

Christopher C. Valentino
The University of Maryland, Baltimore County
Department of Information Systems
1000 Hilltop Circle
Baltimore, Maryland 21250
cvalen1@umbc.edu

Intrusion Detection, Decision Modeling, Security

ABSTRACT

This paper addresses specific problems within the area of performing computer system intrusion detection, and presents the reader with an effective decision model to addressing these problems. Current intrusion detection analysis methods are reluctant to properly evaluate the results of decisions made based on their analysis outcomes. These analysis outcomes influence the decision making process involved in the response to an intrusion. Utilizing basic decision modeling methods we can develop a model that is both effective and easy to use. To form this model we must have the following within our environment; standard analysis procedure and the classification of information elements. These will feed into our structured decision model and aid in our final decision outcome.

INTRODUCTION

With the rapid growth of the Internet and the need for information to be publicly accessible and both private and public sector businesses and agencies rendering service via the web to both consumers and potential adversaries computer security has grown rapidly. Recently the area of Intrusion Detection has grown into a large and distinct discipline. Although alone it is nothing more than a measurement device, coupled with both human and computer analytical power it becomes a proactive tool in preventing current and future computer intrusions.

The challenge of intrusion detection is performing accurate and correct analysis of the presented data. Most often a decision to block a suspected attacker are made in hast and cause network outages. These outages result in the failure to deliver service to the end customer, and in some cases this is an attacker's objective. A proper process must be created and followed during the

analytical process to assure that decisions are (1) accurate and (2) unbiased.

Decision Modeling provides us with a structured approach that can be used to (1) formulate a standard analysis method and (2) formulate an overall decision model. Decision theory allows us to associate a set of probability distributions with each event to reflect the expectations or uncertainties of the decision maker. (Butler) Within this paper we will focus on this overall model, and the elements important to creating and utilizing it. To begin we will conduct a discussion of the overall problem, then define the environment specifically focusing on where we can obtain additional analysis data, and finally form and review the decision model. Our goal in applying the decision modeling method is to make "smarter" decisions.

UNDERSTANDING THE PROBLEM

At first glance computer intrusion detection appears to be a simple process of installing a commercial system, and then literally watching the screen for red blips. As the area of computer intrusion detection has evolved we have found that this analogy is not true, in fact it is dead wrong. Computer intrusion detection is a highly analytical heavy process, which lacks good overall structure and analysis processes. To often the literal is taken as the truth. There is a simple rule of thumb, don't believe what you see at first, and always follow the directions. Based on this statement we can deduce that three main problems exist within the analytical process (1) utilizing all available information sources, (2) verifying the validity of a suspected computer system intrusion, and (3) following a standard process.

UNDERSTANDING THE ENVIRONMENT

Understanding the environment in which you are operating within is probably the single most important thing you can do to be effective at intrusion detection operations. A rudimentary example can be the operation of a motor vehicle. If you are unaware of your current environment (e.g. don't know the highway system) you will be unable to operate the motor vehicle effectively. The same holds true for computer intrusion detection operations, if you are unaware of the environment you will not go very far.

The environment is all of the elements that are related to the events you are analyzing. It is important to note to achieve full validity of an intrusion event or incident we must accept that our environment goes beyond your local networked systems, and extends past the service provider and into the attackers network. This introduces the idea of both an external and internal environment. It is important to understand the relationship between the two and the information available for

our analysis and final decision. Both are great sources of information and compliment each other.

Information Elements

To be successful at computer intrusion detection operations we must have good valid analysis information. Both the external and internal environments can provide this data to our decision process. We refer to such environment components as information elements. We classify information elements into two major groups; live elements and captured elements. We can also refer to these as data in motion (real-time) and data at rest (captured). When defining information elements one must ask themselves. Where can I get more information?

Live Elements

Live elements are best defined as those sources on information that provide real time or real-time feedback. These systems include firewalls, routers, host audit logs, packet sniffers, and intrusion detection systems. Information collected from these elements includes connection logs, protocol analysis, and target system access logs.

These elements are typically feed into either an analysis framework or to a human analyst. The key though is collectively all of these devices can give you a clear and accurate real-time picture of your network.

Captured Elements

Captured elements consist of data that is at rest within the environment these elements are best broken into two categories; configuration data and knowledge base. It is important to make this distinction for configuration data consist of the current makeup and state of the environment and the knowledge base looks at data occurring in the past (i.e. previous intrusion attempts)

Configuration data includes network topology maps, vulnerability assessment data, firewall policy documents, router configurations, and current host configuration information. This information effectively describes the current state of the network both from a configuration view and a security posture view. Example given, with current vulnerability assessment data we can determine if a host is vulnerable to a given intrusion attempt. If the host is not vulnerable to a distinct attack for instance an attempted windows attack on a UNIX host, then we can move on to the next alarm to be analyzed.

An additional resource is the Knowledge Base of past intrusion data, configurations, and real threat data. Utilizing this database allows us to get a historical view on current intrusion attempts, to include previous decisions.

STRUCTURING THE PROBLEM

To structure the problem correctly we must achieve the following; understand our current standard analysis procedure, define our decision environment, and finally form our objectives.

Standard Analysis Procedure

Most organizations currently conducting intrusion detection operations will have some analysis procedure in place. If it is correct or even being utilized is another issue. For our model to be successful the analysis procedure must be consistent and strictly enforced. This will assure that all available information elements will be used to form the analysis outcome.

This model will consist of collecting, organizing, analyzing, and generating some form of analysis output. This output can include if it is a suspected intrusion or a begin trigger. A good process will utilize all available information elements for the final analysis outcomes. Again, the challenge is (1) implementing a standard analysis procedure and (2) consistently using the process.

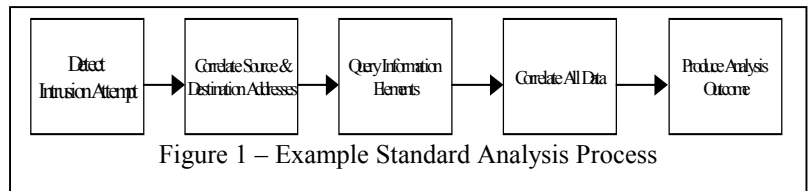


Figure 1 – Example Standard Analysis Process

The Decision Environment

The decision environment can be broken into three zones; uncertainty, risk, and certainty. (Forgionne 1999) Each analysis outcome must be grouped into one of these three situations. From a quantitative approach each information element would bear a certain weight, and the total number of elements would also affect the final decision outcome. We use the three-decision situations as an overall guide. Subsequent efforts will work towards quantifying the Information Elements and assigning fine confidence levels to each situation (i.e. $\alpha=.95$).

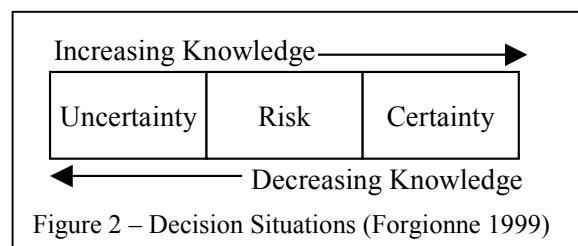


Figure 2 – Decision Situations (Forgionne 1999)

From figure 2 we can see that as our “knowledge” increases our level of certainty increases proportionally. For our model we can say that as our number of information elements increases our level of certainty increases also. It is important to remember that each information element will carry an individual weight. Thus one source of information may be more useful than another. For instance actual intrusion data from a target host system, that confirms the intrusion bears a higher weight than just information from an intrusion detection system.

Uncertainty

Of our three overall decisions situations uncertainty of course is most undesirable. At this extreme situation the decision maker can identify possible outcomes, but does not possess enough information to begin to determine if an intrusion has occurred. (Forgionne 1999) Decisions made under uncertainty bear the following characteristics; little correlation with other information elements, elements used are weighted extreme low, it is unknown if an intrusion has actually taken place, and the true source of the intrusion or attempted intrusion is unknown.

Even with the level of unknown associated with this situation the true fact is most decisions on handling an intrusion or attempted intrusion are made within this zone. This is a costly and most risky situation to find yourself in, Especially for large public or private sector agencies and businesses. Although in some situations decision-making under uncertainty is unavoidable, the decision maker will now be aware of the state in which they are making their decision.

Risk

Once the decision model is applied to the overall analytical process, most decisions will be made under the decision situation of risk. Within is decision situation a maximum number of information elements have been used to correlate intrusion data, excluding information elements from the attacker. Given this we can say that (1) we know that in intrusion or attempted intrusion has occurred and (2) maximum number of local information elements have been used. However we still do not know the true source of the attacking network.

It is a true statement to make that we may never know the true source of an intrusion. Depending on the objective of the decision maker this may be acceptable. If our goal is to be proactive and prevent further intrusions than knowing the true source becomes less important. On the other hand if we must know the true source of an

intrusion or attempted intrusion, decision-making under risk is unacceptable.

Certainty

Decisions made under complete certainty are and will be a rarity. Under certainty we (1) know an intrusion has occurred, (2) know the true source of the intrusion, and (3) exhausted our local and global information elements. Thus a decision maker can feel 100% confident that the decision they make is based on accurate information. Thus in statistical terms the confidence level α is equal to 1. Without a presence on the attacking network achieving this decision situation is impossible.

Objectives

To apply decision-modeling techniques we must establish a clear list of objectives. At the most rudimentary points we want to establish that the analysis outcomes we are providing are both accurate and unbiased. Accuracy relates to the intrusion or attempted intrusion data being correlated with all available and applicable data elements. The question of biases directly relates to identifying the true source of the intrusion or attempted intrusion.

The outcome of our decision model will aid the decision maker in issuing a proper response. These responses could include ignore the event, block traffic from the suspected source, or at the most extreme issue an attack against the source in retaliation. With this said our clear objective is to provide a confidence level that the decision maker can then use to select the appropriate action.

FORMULATING THE MODEL

We have now defined our environment and presented a structure to our problem. From this point on we can begin to formulate our decision model for computer intrusion detection. From a quantitative view our model will use scoring and weights that relate to our information elements to provide a confidence level to the decision maker. In addition to weights and scores statistical inferences can be made regarding certain data sets, particularly in the area of threat data.

Weighting & Scoring

The difficulty of applying proper weights to information elements is one the operational organization must make. Industry sources and internal organizational studies can be used to produce their level of confidence within their information elements. After weights are assigned to individual information elements statistical inferences can be

used to determine the proper sample size or proper number of elements needed to make a decision at a give decision situation.

A scoring system is applied to the overall collection of information elements. The score is based on the overall weights and number of elements used; this in turn will determine the overall confidence level or the decision situation.

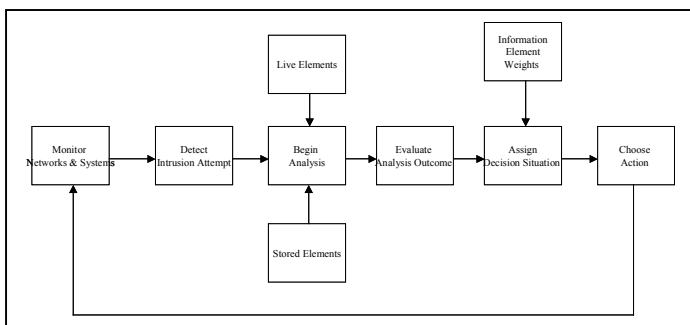
Statistical Inferences

Statistical inferences can be used particularly with threat data. Threat data is current or archived information regarding attempted intrusions, successful intrusions, knowledge of state-sponsored information warfare programs. It can be either local to the organization or global to it. Example give the SANS Institute provides hourly updates on the number of intrusion attempts occurring on the Internet. This of course is just a subset of the overall situation, but it allows us to view what events are occurring from a worldwide view. We can make statistical inferences and assign probabilities that either a certain source or group of sources are attempted to intrude our computer network.

The Model

Once the environment is defined, problem is structured, weighting and scoring system created, and probabilities assigned to our select information elements the model itself becomes simple.

At the base our model integrates into our analysis process. In fact it oversees that the analytical process is completed correctly. From start to finish we begin with



monitoring our networks, once an attempted intrusion is detected we begin the event analysis process. This process uses our predefined information elements as input in determining the analysis outcome. Once the analysis outcome is determined a decision situation is assigned based on a set of predetermined weights. This overall situation is then used to make the final reactive decision

to the intrusion or attempted intrusion. The process can be repeated and used as either a single-stage decision or in a series of sequential decisions.

CONCLUSION

Within this paper we have discussed applying decision modeling and decision theory to computer intrusion detection. We find that the structure of decision modeling allows us to make sure and sound decisions based accurate and true information collected utilizing our information elements. Decision Theory concepts allow us to assign appropriate weights and probabilities to our information elements. Thus, in the end we are able to provide “smarter” decisions based on our analytical outcomes.

REFERENCES

Forgionne, G. A. 1999. *Management Science*. Wiley Custom Services, United States of America.

Butler, S.A. 2001. *Improving Security Technology Selections with Decision Theory*.

BIOGRAPHY

Mr. Valentino is a senior at The University of Maryland, Baltimore County studying Information Systems. His studies are concentrated in area of structured systems analysis and design.

In his professional capacity he is the Lead Technologist of a Maryland based Information Assurance company named The Windermere Group. In this capacity at Windermere he designs and integrates complex Security Architectures, to include enterprise computer intrusion detection systems. Mr. Valentino’s customers include the United States National Security Agency, United States Patent and Trade Mark Office, and the United States Army.