



Defence Research and
Development Canada

Recherche et développement
pour la défense Canada



Event prioritisation using a fuzzy risk analysis approach

Maxwell Dondo

Defence R&D Canada – Ottawa

Canada

Technical Memorandum
DRDC Ottawa TM 2009-287
March 2010

Event prioritisation using a fuzzy risk analysis approach

Maxwell Dondo
DRDC Ottawa

Defence R&D Canada – Ottawa

Technical Memorandum

DRDC Ottawa TM 2009-287

March 2010

Principal Author

Original signed by Maxwell Dondo

Maxwell Dondo

Approved by

Original signed by J. Lefebvre

J. Lefebvre
Head/NIO Section

Approved for release by

Original signed by C. Boulet

C. Boulet
Head/Document Review Panel

© Her Majesty the Queen in Right of Canada as represented by the Minister of National Defence, 2010

© Sa Majesté la Reine (en droit du Canada), telle que représentée par le ministre de la Défense nationale, 2010

Abstract

Analysts handle multitudes of computer network security events on a daily basis. They must make an assessment on the potential impact these events have on their organization's assets. As the number of events increases, it becomes increasingly difficult for the analyst to make an assessment as to which events to handle first. This can be resolved by calculating a potential risk metric associated with each event, and then prioritizing the events based on the calculated risk values. Most risk analysis approaches available are based on models which require historical data. In many cases, numerical data related to uncertainty factors about the risk calculations is not available, but the experiential expertise of analysts is. This experiential expertise can be modeled as linguistic variables and functions about an event, and be used to model the risk value associated with each event. In this paper, we present an approach to determine the potential risk value associated with each computer security event by modeling the experiential expertise of analysts through fuzzy linguistic declarations about an event. We then rank these events based on the relative calculated risk values for each. We test our approach on a prototype network using real vulnerability data.

Résumé

Les analystes traitent quotidiennement une multitude d'incidents de sécurité qui se produisent sur leur réseau informatique. Ils doivent évaluer les répercussions potentielles de ces incidents sur les équipements de leur organisation. À mesure que le nombre d'incidents augmente, il devient graduellement plus difficile pour l'analyste de déterminer quel incident il doit traiter en premier. Pour y arriver, on peut calculer un paramètre de risque potentiel pour chaque incident et ensuite placer les incidents en ordre de priorité, en fonction des valeurs de risque calculées. La plupart des approches liées à l'analyse du risque sont basées sur des modèles qui nécessitent des données historiques. Dans de nombreux cas, les données numériques liées aux facteurs d'incertitude concernant le calcul du risque ne sont pas disponibles, mais l'expertise expérimentielle de l'analyste l'est. Cette expertise expérimentielle peut être modélisée sous forme de variables et de fonctions linguistiques liées à l'incident et être utilisée pour modéliser la valeur de risque associée à chaque incident. Dans le présent document, nous présentons une approche qui a pour but de déterminer la valeur de risque potentiel associée à chaque incident de sécurité en modélisant l'expertise expérimentielle des analystes par l'intermédiaire d'énoncés linguistiques flous concernant l'incident. Nous classons ensuite ces incidents en fonction des valeurs de risques relatives calculées pour chaque événement. Nous testons notre approche sur un prototype de réseau qui fait appel à des véritables données de vulnérabilité.

This page intentionally left blank.

Executive summary

Event prioritisation using a fuzzy risk analysis approach

Maxwell Dondo; DRDC Ottawa TM 2009-287; Defence R&D Canada – Ottawa; March 2010.

Background: This work is a result of an identified need by analysts at Canadian Forces Network Operations Centre (CFNOC). Analysts deal with many events on a daily basis. To enable them to handle these events effectively, they need a way to be able to prioritise the events so that they can attend to the ones that need immediate attention while putting aside those that are not as important, so that they could be looked at later. That way, they will be able to prioritise and schedule their work accordingly.

In this work, we propose an approach to rank events based on the possible risk to which they expose network assets. In computer network security events, it is difficult, if not impossible to get historical data on events in order to make statistical inferences about an event. We therefore propose to use fuzzy systems theory to model analysts' experiential linguistic declarations about security events. We then use information fusion techniques implemented through fuzzy inference systems (FISs) to combine the security event attributes to give a relative risk value for each event on a given asset. Finally, we rank the events in order of this risk value at asset and network levels.

Principal results: Our approach is applied to real vulnerability data from National Vulnerability Database (NVD), and Common Vulnerabilities and Exposures (CVE) on a typical 10-host network. We simulated events using vulnerabilities from these databases. We show that we can rank the events based on the potential risk they pose to the network. Our approach ranks events at an organisation, network, asset, and node levels.

Significance of results: The proof-of-concept model results show that we can translate analysts' experience into a model that can be used to prioritise network security events on a wide scale. The results are a promising replacement of the current "gut-feel" approach by network security analysts.

Future work: The long-term goal is to test this approach in an operational environment for possible implementation and deployment of the model at the CFNOC. In the short-term, we will develop a portable demonstrator to be integrated into the ongoing

work on a defensive posture demonstrator. That will enable us to make side-by-side comparisons of our method with other methods before any possible deployment can be implemented.

Sommaire

Event prioritisation using a fuzzy risk analysis approach

Maxwell Dondo ; DRDC Ottawa TM 2009-287 ; R & D pour la défense Canada – Ottawa ; mars 2010.

Contexte : Ces travaux sont le résultat d'un besoin identifié par les analystes du Centre d'opérations des réseaux des Forces canadiennes (CORFC). Les analystes sont confrontés quotidiennement à de nombreux incidents. Afin de leur permettre de traiter ces incidents de façon efficace, ils doivent pouvoir les placer en ordre de priorité afin de déterminer lesquels nécessitent leur attention immédiate et mettre de côté pour plus tard ceux qui ne sont pas aussi importants. Ainsi, ils sont en mesure de prioriser leur travail et d'établir leur horaire en conséquence.

Dans ce document, nous proposons une approche permettant de classer les incidents en fonction du risque potentiel qu'ils représentent pour les installations réseaux. Dans le domaine des incidents de sécurité sur les réseaux informatiques, il est difficile, voire impossible, d'obtenir des données historiques sur les incidents qui permettent de faire des inductions statistiques sur un incident. En conséquence, nous proposons l'utilisation de la théorie des systèmes flous afin de modéliser les énoncés linguistiques expérimentiels des analystes sur les incidents de sécurité. Nous utilisons ensuite les techniques de fusion des informations élaborées par les systèmes d'inférence floue (SIF) afin de combiner les attributs des incidents de sécurité et ainsi déterminer une valeur de risque relative à chaque incident sur un équipement donné. Enfin, nous classons les incidents en fonction de cette valeur de risque, aux niveaux de l'équipement et des réseaux.

Principaux résultats : Notre approche utilise de véritables données de vulnérabilité provenant de la base de données nationale sur la vulnérabilité (BDNV) et les expositions et vulnérabilités communes (EVC) sur un réseau typique à 10-hôtes. Nous avons simulé des incidents en utilisant des vulnérabilités provenant de ces bases de données. Nous montrons que nous pouvons classer les incidents en fonction du risque potentiel qu'ils représentent pour le réseau. Notre approche classe les incidents à différents niveaux : organisationnel, réseau, équipement et nœud.

Signification des résultats : Les résultats du modèle de validation de principe montrent que nous pouvons traduire l'expérience des analystes en un modèle pouvant être utilisé à grande échelle pour classer par ordre prioritaire les incidents de sécurité touchant les réseaux. Ces résultats sont un substitut prometteur par rapport

à l'approche basée sur l'instinct qui est actuellement utilisée par les analyses de la sécurité des réseaux.

Travail À venir : À long terme, le but est de tester cette approche dans un environnement opérationnel en vue de son éventuelle mise en œuvre et du déploiement du modèle au CORFC. À court terme, nous allons mettre au point un démonstrateur portable qui sera intégré dans les travaux en cours sur un démonstrateur de posture défensive. Cela va nous permettre de faire une comparaison en parallèle entre notre méthode et d'autres méthodes avant que tout déploiement potentiel ne soit mis en œuvre.

Table of contents

Abstract	i
Résumé	i
Executive summary	iii
Sommaire	v
Table of contents	vii
List of figures	ix
List of tables	x
Acknowledgements	xi
1 Introduction	1
1.1 Background	1
1.2 Challenges	1
1.2.1 Existing Approaches	2
1.2.2 Proposed Approach	3
2 Definitions	4
2.1 Model Definitions	4
2.2 Attribute Variability	5
2.2.1 Asset Attributes	5
2.2.2 Vulnerability Attributes	6
3 Fuzzy Risk Analysis	9
3.1 Basic Fuzzy Systems Theory	9
3.2 Fuzzy Ranking System	11

4	Experimentation and Results	12
4.1	Fuzzification of Attributes	12
4.2	Fuzzy Rules	15
4.3	Model Implementation	16
4.3.1	Dependency Factor	17
4.3.2	Simulation Data	19
4.4	Results	24
5	Conclusion	29
	References	30
	Acronyms and Abbreviations	32
	Annex A: Numerical Significance of Results	35
	A.1 Monotonic Increasing Risk Function	35
	A.2 Monotonic Decreasing Risk function	36
	A.3 Time Variability	37
	Annex B: Fuzzy Rules	39
	B.1 FIS BaseValue Rules	39
	B.2 FIS EV Rules	39
	B.3 FIS ImpactValue Rules	40
	B.4 FIS Likelihood Rules	40
	Annex C: Common Vulnerability Scoring System (CVSS)	45

List of figures

Figure 1:	Illustration of the collection of network objects.	4
Figure 2:	A fuzzy set showing $\mu(x)$ on the vertical axis and x on the horizontal axis.	9
Figure 3:	Typical membership functions.	10
Figure 4:	A fuzzy dependency input, reflecting the two asset types of “Data” and “Application”.	14
Figure 5:	A fuzzy dependency input. This reflects the asset’s importance as “Low”, “Medium”, “High”, and “Very High”.	15
Figure 6:	The fuzzy dependency output.	15
Figure 7:	Experimental implementation.	17
Figure 8:	The fuzzy dependency factor.	19
Figure 9:	A Typical 10 host network.	20
Figure 10:	All the vulnerabilities affecting the 10 node network	24
Figure 11:	The affected networks in order of priority.	24
Figure 12:	All the events on the 10 node network	25
Figure 13:	The output events from the three networks.	26
Figure 14:	The output events from three selected nodes.	27
Figure 15:	The affected assets on node 24.	27
Figure 16:	The output events from two selected assets.	28
Figure A.1:	The risk value as the security posture of an asset deteriorates.	36
Figure A.2:	The risk value as the security posture of an asset improves.	37
Figure A.3:	38

List of tables

Table 1:	Rules Overview.	16
Table 2:	Dependency rules.	18
Table 3:	Asset identification for 10-host network (see Figure 9).	21
Table 4:	Network vulnerabilities.	22
Table 5:	Simulated events for the 10-Host network.	23
Table 6:	Asset dependencies for the 10-Host network.	23

Acknowledgements

I would like to thank Dr. Peter Mason for the extra effort that he put into extensively reviewing and providing advice on the contents and presentation of this document.

This page intentionally left blank.

1 Introduction

1.1 Background

On a daily basis, analysts handle hundreds or even thousands of events on a network or a group of networks under their watch. They are supposed to analyse each one of these events and take appropriate action, which can be a tedious process. This can be made worse if the network they are dealing with is very big or consists of many groups of networks.

To enable analysts to handle these events more effectively, it would be beneficial to be able to prioritize them. That way, analysts would know where to focus their energy by allocating resources where they are most needed. In this work, we use a risk analysis approach to rank events. In information technology, risk is defined as the loss of confidentiality, integrity or availability (CIA) due to a specific threat [1]. We determine the risk attributes of each event or group of events. We then associate each event to a registered vulnerability (or set of vulnerabilities). Then, using a vulnerability prioritisation approach, we rank these events. The ranking provides the priority for which event, asset or network to pay attention to first.

In the event that there is no vulnerability associated with any event, we extract the event attributes as the analyst understands them at the time. We then apply them to the vulnerability prioritization algorithm. This has been modeled to handle events in a similar way to vulnerabilities. As will be explained later, our approach makes use of the experiential input from analysts to determine the level of risk associated with an event.

The approach determines the risk associated with each vulnerability (and event) on a given asset (and therefore network) by determining the potential loss in value of a given asset when a threat exploits a vulnerability on that asset. We then rank the calculated risk values in order of priority.

This section presents a general discussion on risk analysis and an analysis of other methods in use. It is followed by a brief description of our proposed approach.

1.2 Challenges

Events occur on computer network infrastructure; the effect of these events on the infrastructure are investigated in the algorithms presented here. In this section, we define the different parts of the network infrastructure as used in this work.

1.2.1 Existing Approaches

Inasmuch as event prioritisation is very important to network analysts, there is not much research work going on in this area. Part of the reason for this may be that the currently used network management tools incorporate some form of event prioritisation algorithms that are often based on proprietary algorithms not publicly known. The scarcity of research in the area may also be an indication of the success of these tools. There are, however, some research efforts that have been put in this area, and we will look at some of them, in particular, those that are closely related to our work here.

The most basic approach prioritizing events is the *Delphi* approach [2] which has been in use since organisations started setting up computer network analyst teams. In this technique several raters (analysts) estimate priority based on predetermined metrics like the likelihood of exploitation. Individual raters are given the opportunity to change their ratings after considering those given by the other raters. This process is repeated until the ratings are reasonably consistent, in which case the results are adopted, otherwise, the raters meet to discuss the different ratings, until an agreement is reached. However, the resultant ratings are based on a limited number of metrics which can be applicable to individual assets, and it would be difficult, if not impossible, to use this method on a network or a group of networks.

In another approach, Mosleh *et. al.* [3], develop and implement a Bayesian probabilistic model to assess risks associated with large computer systems. They model the potential loss due to the occurrence of a threat as a family of normal distributions. They go on to model the probability of loss which they solve by numerical methods. While this approach provides a way to compare different vulnerabilities by the value of the risk they expose assets to, it makes assumptions on the statistical distributions of asset losses (they used a normal distribution for asset loss and and gamma distribution for frequency of a given threat). In the absence of enough statistical data, which is usually the case in these types of problems, it is difficult to make an inference on the statistical distributions of asset losses, and therefore the likelihood of attack.

The Vulnerability Assessment and Mitigation (**VAM**) methodology [4, 5], developed for the military defence environment, takes a systems approach to risk analysis. The approach employs steps to identify vulnerabilities and their attributes and then matches them with safeguards in a way that reduces risk. In theory, this is a very good approach to identify safeguards for a full and complete protection of system vulnerabilities. However, **VAM** risk analysis calculations still need to be carried out to give an indication of priority.

There are also a number of other approaches, such as Fault Trees Analysis (**FTA**), Event Trees Analysis (**ETA**), and Markov Analysis [6–8], that are used in risk analysis and decision making. These methods determine the likelihood of attack through

sequences of steps. They use these values to determine the relative risk of vulnerabilities. Although these methods could give relatively accurate rankings for individual assets, it is not trivial to handle a network or a group of networks.

Fuzzy systems have also been widely used in risk analysis [9–12]. In these approaches, researchers used fuzzy logic to determine the probability of failure or likelihood of an attack. Chen *et. al.* [9] go further by improving on previous fuzzy systems' approaches while introducing dependencies to component failures. Their fuzzy models are based on the severity and likelihood fuzzy numbers (FNs). Shah [11] used several key risk indicators (KRIs) (operational variables that provide the basis for estimating losses corresponding to risk) to determine risk based on their linguistic descriptors. All these approaches need substantial modification for them to be applicable to prioritising vulnerabilities based on the risk they pose to a network or a set of networks.

The biggest shortcoming in traditional approaches is the incomplete representation of KRIs. They make estimates of the likelihood of an attack, but they do not model the relative importance of each to the final risk value. In our earlier work [17], we used fuzzy systems to model all known asset and vulnerability KRIs as well as analysts' experiential input to determine the fuzzy risk value associated with each vulnerability on a given asset. We ranked the the vulnerabilities based on the calculated risk values. As we will show in this work, we will build on this work and include attribute interdependencies to reflect real-life scenarios in prioritising events.

1.2.2 Proposed Approach

We propose to tackle this problem by using experiential knowledge. In a way similar to the Delphi approach [2], we capitalize on analyst experience and intuition. We go beyond the Delphi approach by preserving each metric that is used to determine the potential risk of an event on an asset. We go further by fusing this information and applying the the theory of fuzzy systems to the state, attributes, experience and intuition, to come up with a model that reflects the relative importance of the every event that an analyst sees and therefore helps the analyst to make the correct judgement call when called upon to.

In the rest of this document, we start in Section 2 by giving definitions of the objects used in our modeling. We define the fuzzy theory and fuzzy risk calculations used in this work in Section 3. This is followed by experimentation and results in Section 4, and finally conclusions in Section 5.

2 Definitions

In this section, we look at the definitions of the objects used in this work. We also define the attributes or KRIs that we use in our model.

2.1 Model Definitions

Our network model is made up of nodes and assets as illustrated in Figure 1. This shows an enterprise with N networks or subnets. In Network i , we have M nodes. Similarly Node ij has K assets.

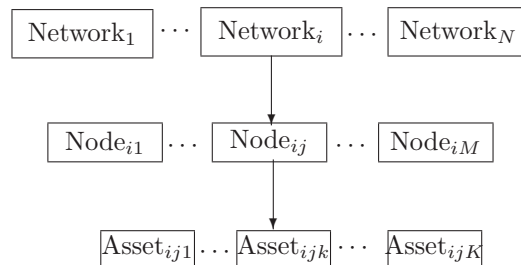


Figure 1: Illustration of the collection of network objects.

The definitions of these network objects are summarized as follows:

Node We define a node as a connection point, either a redistribution point or a communication endpoint. A node may be connected to many assets.

Network A collection of interconnected nodes make up a network.

Asset We define an asset as any object connected to the node. This may be physical objects such as switches and routers, or services. Assets may also be extended to include concepts like reputation (although we don't model these in this work), and data. In this work, our assets are exclusively software applications or data; physical objects like switches and routers are controlled by the real-time operating systems (RTOs), like CISCO's IOS, which are software applications. Physical damage to physical objects is not modeled in this work¹.

Asset Value In this work, we assign an asset value to each asset. This asset value is a reflection of its importance to the organisation relative to other assets of the same organisation. The value may also be reflective of the mission that the asset is involved in or the business reputation at stake. An asset's replacement

¹Modeling the RTO on the physical object and assigning a correct asset value will be reflective of any potential loss of a physical object.

costs may be relatively lower in value to other assets in the organisation, but its asset value, for the purposes of this prioritisation may be very high. An asset value, is therefore not constant, but dependent on the function on hand.

Asset Dependency In a network, assets are inter-dependent of each other. A portion of the asset value is added to all other assets that depend on it. We will show how we calculate the inter-dependency factors in Section 3.2.

Safeguards Safeguards protect assets (and networks) against threats. They reduce the security risks on an asset by a factor of μ .

Vulnerability A weakness in an asset (or network) that can be exploited by threatening agents and therefore increases the security risks on the asset (or network). In this work we associate a vulnerability v with an asset; thus, an asset, node, or network can have many vulnerabilities. Conversely, in a network with many similar, identical, or dependent assets, one vulnerability may be associated with many assets, nodes, or networks.

2.2 Attribute Variability

Our model is based on an earlier approach [17] which ranks vulnerabilities based on the potential risk they pose to the organization. To model and distinguish between vulnerabilities, we collect a general set of attributes that will unambiguously characterize the vulnerabilities' KRIs and assist us in determining a relative risk value for comparison purposes. Since our model is based on the linguistic-declaration [13, 14] model of fuzzy logic, we split the attributes into the different qualifiers that will enable us to implement the fuzzy model.

2.2.1 Asset Attributes

As mentioned above, for the purposes of this work, we categorize assets as applications or data. The “applications” category includes general purpose applications, Operating Systems (OSs), servers (mail, database, file) and their derivatives. The “Data” category covers any data-file that can be considered a an asset worth protecting. This includes databases, data files of any form or their data derivatives.

The third attribute that we assign to the asset is “Importance”. This is a value between 0 and 1 which reflects how important, to the best understanding of the analyst, an asset is. This is a subjective attribute in the sense that different analyst may enter different values. However, our model will be able to handle this variability.

In addition to the above, we also introduce the dependency factor. This factor combines all the attributes above to give a fractional factor that reflects how one asset depends on another.

2.2.2 Vulnerability Attributes

In this work we consider those attributes of vulnerabilities that can be used to give an indication of the potential risk if the vulnerability is exploited. These attributes, which we also call key risk indicators (KRIs), are used in our algorithm to determine the potential risk value which we use to prioritize the vulnerabilities and eventually the individual events on a given asset or network.

Many of these attributes are similar to the CVSS attributes and have similar meaning. The vulnerability attributes are as follows:

Access Complexity The access complexity (AC) attribute gives an indication of the complexity of attack required to exploit the vulnerability once an attacker has access to the target system. The possible levels for this metric are as follows:

1. *High*: Specialized access conditions exist
2. *Low*: System always exploitable; this indicates a higher potential risk value

Authentication This attribute indicates whether or not an attacker needs to be authenticated to the target system in order to exploit the vulnerability. The possible levels for this attribute are as follows:

1. *Required*: Authentication required to exploit the vulnerability
2. *Not Required*: Authentication not required to exploit the vulnerability; this indicates a higher potential risk value

Access Vector This attribute indicates how the vulnerability can be exploited. The possible levels for this are:

1. *Local*: Indicates an attacker with physical access to the system
2. *Remote*: Indicates that the vulnerability may allow remote access; this indicates a higher potential risk value, since there is an unlimited number of remote attackers, whereas in “local” access, the set of attackers may be limited and control (restrictive) measures can reduce the possibility of an attack.

Confidentiality Impact This attribute indicates the level of impact on confidentiality if the vulnerability is exploited. The possible levels for this are:

1. *None*: Indicates no impact on confidentiality
2. *Partial*: Indicates partial impact on confidentiality
3. *Complete*: Indicates a complete information disclosure on the asset; this indicates the highest potential risk value

Integrity Impact This attribute indicates the level of impact on information integrity if the vulnerability is exploited. The possible levels for this are:

1. *None*: Indicates no impact on integrity
2. *Partial*: Indicates partial impact on integrity
3. *Complete*: Indicates a complete compromise of asset integrity; this indicates the highest potential risk value

Availability Impact This attribute indicates the level of impact on availability if the vulnerability is exploited. The possible levels for this are:

1. *None*: Indicates no impact on availability
2. *Partial*: Indicates partial impact on availability
3. *Complete*: Indicates a complete denial of service (DOS) on the asset; this indicates the highest potential risk value

Exploitability This attribute indicates whether the exploit techniques or code are prevalent. The possible levels for this are:

1. *Unproven*: No known exploit code or technique
2. *Proof-of-concept*: Only a proof-of-concept method has been demonstrated and may not be useable on all types of assets affected
3. *Functional*: Functional code which works on most systems exists
4. *High*: Details are widely available on how to exploit the vulnerability; this indicates the highest potential risk value

Report Confidence This attribute indicates the degree of confidence on what is known about this vulnerability. The possible levels for this are:

1. *Unconfirmed*: Little confidence in report validity
2. *Uncorroborated*: Higher confidence about the existence of the vulnerability than the "unconfirmed"
3. *Confirmed*: Details are available from the vendor on how to exploit the vulnerability; this indicates the highest potential risk value

Remediation Level This attribute indicates whether there are ways to patch against this vulnerability. The possible levels for this are:

1. *Official Fix*: A fix is available from the vendor
2. *Temporary Fix*: A temporary fix is available from the vendor
3. *Workaround*: Temporary, nonofficial fix can be implemented
4. *Unavailable*: No known fix is available; this indicates the highest potential contribution to the risk value

In this attribute, only the "Official-fix" can be declared with certainty. The rest need to be tested with Mulval to make sure that whatever selection is made is correct and unambiguous.

Safeguards This attribute indicates the strength of the installed safeguards and how effective they are against this vulnerability on a given asset. The possible levels for this are:

1. *Complete*: The safeguard is effective in protecting the asset against the exploitation of this vulnerability
2. *Partial*: Partially effective safeguards
3. *None*: Completely ineffective safeguards; this indicates the highest potential risk value

Time This attribute indicates the time that has elapsed since the vulnerability was known to exist. The different levels are based on what past experience has shown it to be. The Symantec Internet Security report [15] states that the average number of days for exploit development was 6.0 for the period of Jan-June 2005. The risk value increase with time elapsed, but not in a linear fashion. The longer it takes to handle a vulnerability affecting the organization's assets, the higher the likelihood that the vulnerability could be exploited.

When data is entered into Vulnerability Prioritization Assistant (VPA), there will be inevitably be discrepancies from analyst to analyst. To mitigate this, we streamline our attributes and clearly define what the entry should be when certain conditions prevail.

3 Fuzzy Risk Analysis

The fuzzy systems approach used in this work capitalises on the application of fuzzy theory. In the absence of enough crisp and deterministic data to make a reasonable mathematical inference, it has been shown that linguistic declarations by experts about a process, can be modeled using fuzzy logic. In this section we give a brief description of relevant fuzzy logic theory. We go on to show how this theory is applied in this work.

3.1 Basic Fuzzy Systems Theory

A fuzzy set is defined as an extension of a classical (crisp) set with each member of the set assigned a membership value which represents the degree to which that member belongs to the set (the degree of truth) [14]. If X is the universe of discourse consisting of elements denoted by x , then the fuzzy set \tilde{A} in X is defined as a set of ordered pairs given by:

$$\tilde{A} = \{x, \mu_{\tilde{A}}(x) \mid x \in X\} \tag{1}$$

where $\mu_A(x)$ is called the membership function (MF) of x in \tilde{A} . It represents the *degree of truth* that x belongs to A . It is bounded in $[0, 1]$; elements x with $\mu(x) = 0$ are not listed.

A typical MF is shown in Figure 2. It characterises a fuzzy set. There are many

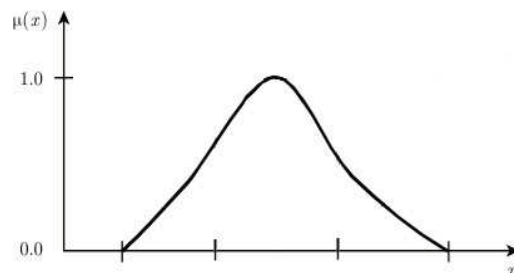


Figure 2: A fuzzy set showing $\mu(x)$ on the vertical axis and x on the horizontal axis.

types of MFs, with the simplest being those formed by straight lines, the *straight line* MFs. In Figure 3, we show two of the most commonly used MFs, the triangular and trapezoidal. A convex and normalized fuzzy set is called a fuzzy number (FN) if its MF is at least segmentally continuous and has the functional value $\mu_{\tilde{A}}(x) = 1$ at only one member element [14]. However, in everyday research work, straight line MFs have been used as close approximations to functions representing real-life situations; the trapezoidal MF is a special case of the triangular MF in that it has several values of x with $\mu_{\tilde{A}}(x) = 1$ as shown in Figure 3.

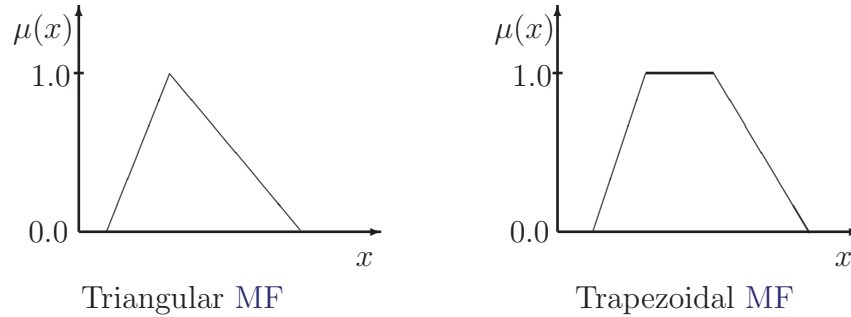


Figure 3: Typical membership functions.

These straight line MFs have the advantage of simplicity and are widely used in research work. Other MFs include the Gaussian, generalised bell, and various polynomial based curves. These latter MFs are popular because of their smoothness and concise notation, but they are not required for a good fuzzy inference system (FIS) [16].

The power of fuzzy systems lies in their ability to model vague and imprecise concepts, such as “this safeguard is weak” or “this vulnerability is very exploitable”. Everyday language that cannot be easily quantified can be converted into fuzzy variables through the fuzzification process. Using fuzzy functional relations, inferences (FISs) can be made about the relationships between the fuzzy variables.

Some of the fuzzy logic (FL) properties and relationships that we will use in this work are presented here. We assume two fuzzy sets \tilde{A} and \tilde{B} for ease of explanation.

- The union operator, which implements the fuzzy *OR* function, has a membership function that is pointwise defined by:

$$\mu_{\tilde{A} \cup \tilde{B}}(x) = \text{Max}[\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)] \mid x \in X \quad (2)$$

- The intersection operator, which implements the fuzzy *AND* function, has a membership function is pointwise defined by:

$$\mu_{\tilde{A} \cap \tilde{B}}(x) = \text{Min}[\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)] \mid x \in X \quad (3)$$

- The support of a fuzzy set is the set of elements in the set whose MF values are greater than zero.
- Fuzzy relations are defined through fuzzy *if-* and *then-* rules; for example *If A AND B then C*.

Most set theory relations are also applicable to fuzzy systems, although the implementation may be different. Coverage of these relationships is beyond the scope of our current work.

To revert back to the crisp domain (for the final decision), a fuzzy set needs to be defuzzified. Although this results in the loss of information originally represented by the fuzzy set, it is a necessary step. The most commonly used defuzzification method is the centroid method [14] which returns the center of area under the curve. This is represented as follows:

$$x = \frac{\sum_i x_i \mu(x_i)}{\sum_i \mu(x_i)} \quad (4)$$

where x is the crisp defuzzified value which can then be used for decision making.

3.2 Fuzzy Ranking System

As already stated earlier, this work ranks events based on the potential risk they pose to the network. We use fuzzy logic to determine this potential risk value. To determine this risk using fuzzy logic, we need to fuzzify the vulnerability and asset attributes that characterize the threats that exploit vulnerabilities giving rise to these events. We then look at the rules that puts all these attributes together to determine the risk level.

4 Experimentation and Results

In this section, we present the model implementation and experimental results of our work.

4.1 Fuzzification of Attributes

To fuzzify the attributes and model the fuzzy rules that put these attributes together, we assume the role of an experienced analyst whose linguistic declarations about the security events and vulnerabilities is translated into a fuzzy model from which the relative risk-based ranking system is derived. In this section, we present the fuzzy attributes and show examples of the fuzzification process for three attribute sets. Detailed fuzzification processes for the rest of the attributes are given in [17], and we will use them in this work unchanged.

The fuzzy attributes used in this work are as follows:

Access Complexity The fuzzy **AC** attribute has the following fuzzy FNs:

1. *High*
2. *Low*

Authentication This fuzzy attribute has the following FNs:

1. *Required*
2. *Not Required*

Access Vector This fuzzy attribute has the following FNs:

1. *Local*
2. *Remote*

Confidentiality Impact This fuzzy attribute has the following FNs:

1. *None*
2. *Partial*
3. *Complete*

Integrity Impact This fuzzy attribute has the following FNs:

1. *None*
2. *Partial*
3. *Complete*

Availability Impact This fuzzy attribute has the following FNs:

1. *None*
2. *Partial*
3. *Complete*

Exploitability This fuzzy attribute has the following FNs:

1. *Unproven*
2. *Proof-of-concept*
3. *Functional*
4. *High*

Report Confidence This fuzzy attribute has the following FNs:

1. *Unconfirmed*
2. *Uncorroborated*
3. *Confirmed*

Remediation Level This fuzzy attribute has the following FNs:

1. *Official Fix*
2. *Temporary Fix*
3. *Workaround*
4. *Unavailable*

Safeguards This fuzzy attribute has the following FNs:

1. *Complete*
2. *Partial*
3. *None*

Time This fuzzy attribute has the following FNs:

1. *Very Low*
2. *Low*
3. *Medium*
4. *High*
5. *Very High*

The second set of attributes relate to asset dependencies. We assigned two broad asset categories of “Data” and “Application”. To each of these categories, we assign a fuzzy variable, “Importance”, which represents the fuzzy relative importance of an asset over others. The linguistic declaration about the “Data” type would be as follows:

The importance of the dependency of an asset on another asset of type "Data" is about 50%².

At the same time the importance of the "Application" asset type would be:

The importance of the dependency of an asset on another asset of type "Application" is about 80%, and is higher than that of an asset of type "Data".

The membership functions are shown in Figure 4.

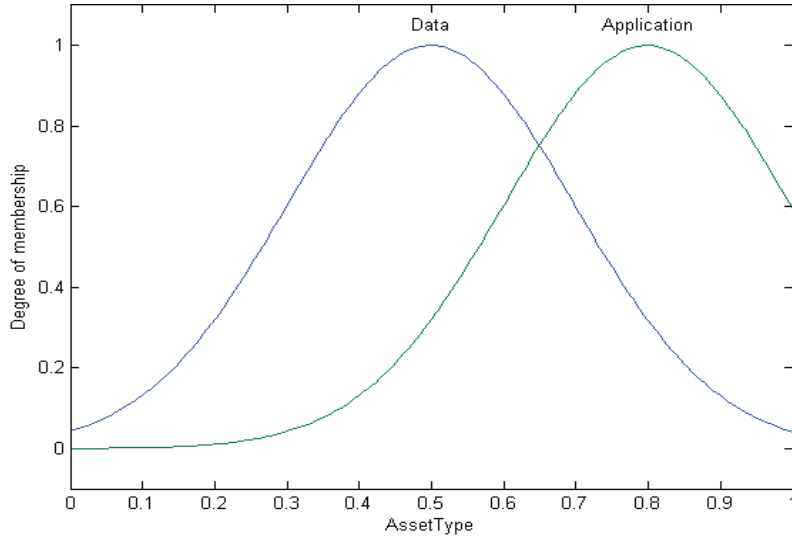


Figure 4: A fuzzy dependency input, reflecting the two asset types of "Data" and "Application".

The importance factor has four fuzzy numbers; namely "Low", "Medium", "High", and "Very High". Their linguistic declarations are as follows:

The importance of an asset is termed "Low" when it is around 25%. Similarly, the respective importance values for "Medium", "High", and "Very High" are 50%, 80%, and 100%.

The membership functions are shown in Figure 5.

In fuzzy theory, the implementation requires that we define the range of the output. The output from this inference is the fuzzy dependency variable. The dependency output is a fuzzy factor in $[0, 1]$. It represents the output dependency factors in the fuzzy domain. The fuzzy numbers of "Low", "Medium", and "High" have the following linguistic declarations:

²The allocation of numerical values is based on the relative linguistic comparison by an experience expert.

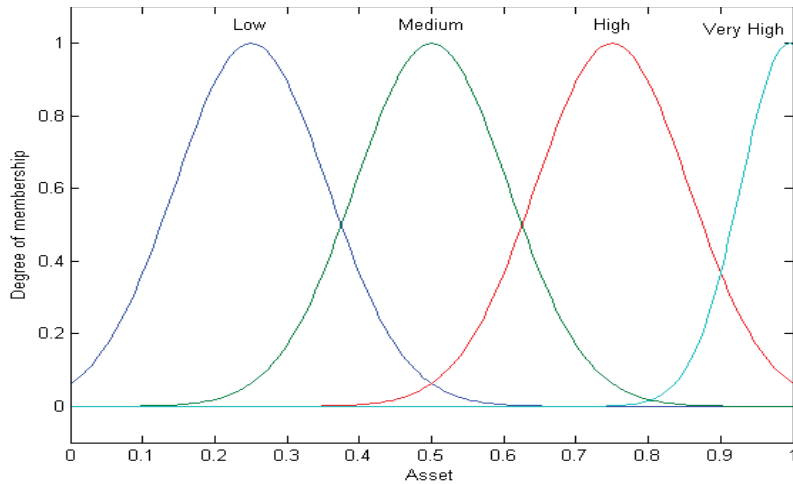


Figure 5: A fuzzy dependency input. This reflects the asset’s importance as “Low”, “Medium”, “High”, and “Very High”.

The dependency factor of an asset is termed “Low” when it is around 25%. Similarly, the respective dependency factors for “Medium” and “High” are 50% and 100%.

These membership functions are shown in Figure 6.

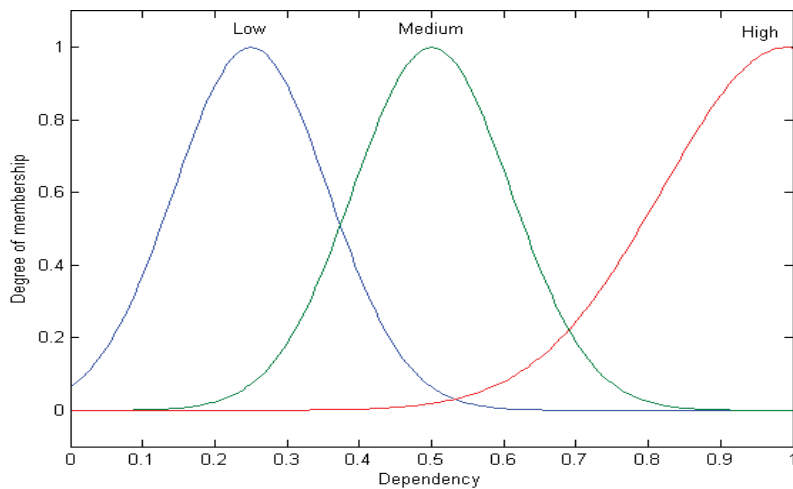


Figure 6: The fuzzy dependency output.

4.2 Fuzzy Rules

According to fuzzy systems theory, we can put the above mentioned attributes together using fuzzy rules to get the desired fuzzy output. The rules combine the

attributes by *if- then-* statements based on the linguistic declarations about the attributes. Rules can be given weights depending on the importance of a rule over others.

The overview of the fuzzy rules rules used in this work is summarized in Table 1. If

Table 1: Rules Overview.

Event	$\xrightarrow{\text{Exploits}}$	Vulnerabilities
Asset	$\xrightarrow{\text{Has}}$	Vulnerabilities
Asset	$\xrightarrow{\text{Depends On}}$	Other Assets
Asset	$\xrightarrow{\text{Has}}$	Risk Value

the set of vulnerabilities belonging to the the event e is V_e , then the total risk value on an asset due to the event is

$$R_e = \sum_{i \in V_e} r(v_i) \quad (5)$$

where $r(v_i)$ is the risk value on an asset associated with vulnerability v_i .

In earlier work [17], we showed how the value for $r(v_i)$ was obtained through the use of fuzzy linguistic declarations about the attributes and the rules combining these attributes. In that work, we showed that, for an asset of value c , the risk value was derived from a fuzzy impact (damage value) value \tilde{t} and the likelihood of attack \tilde{p} for a given vulnerability as

$$r(v) = c \times t \times p \quad (6)$$

where t and p are the defuzzified values of \tilde{t} and \tilde{p} respectively. In that work, we showed the rules used to come up with the values of \tilde{t} and \tilde{p} . We will use these rules in this work and derive new ones for the extra attributes and extensions of this work. These rules are listed in Annex B.

It should be noted that, similar to our earlier work [17], the risk values given in Equation 6 have no significance outside this work. They cannot be used in conjunction with other methods. They are a stepping stone towards meeting the original objective of this work—that is ranking the different computer security events. However, the resultant ranking can be used with other approaches to recommend courses of action.

4.3 Model Implementation

Our model is implemented in two stages. The first stage involves the association of each event with a vulnerability. Each vulnerability’s attributes are identified as in [17]. The second stage assigns all asset dependencies and attribute values to each

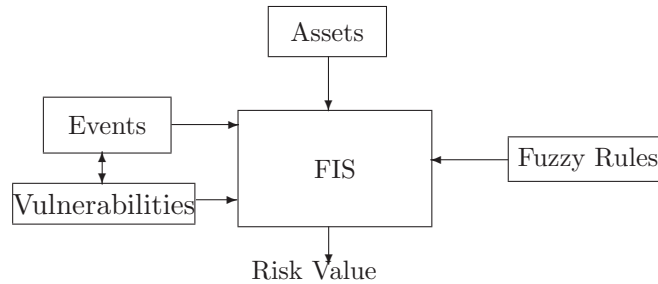


Figure 7: Experimental implementation.

asset and its dependencies. These stages are then applied to the fuzzy inference system (FIS). A layout of the implementation is shown in Figure 7. The output is a risk value for each event, which we use for our comparison.

4.3.1 Dependency Factor

In the second stage, we developed a fuzzy dependency factor which reflects the dependency of one asset on another. We combine the fuzzy attributes illustrated in Figures 4–6 using fuzzy rules obtained through the linguistic declarations about the assets, vulnerabilities and events. Examples of these rules are as follows:

1. If Asset X of *Low* Importance depends on Asset Y which is an *Application* of *Low* Importance, then the Dependency Factor is *Low*
2. If Asset X of *High* Importance depends on Asset Y which is of type *Data* of *Low* Importance, then the Dependency Factor is *Medium*

During the implementation in MATLAB, the rest of the dependency rules are formulated as shown in Table 2. In MATLAB, the number (1) at the end of each rule represents the weight for each rules; since we gave each rule the same weighting, these are all equal to 1.

Table 2: *Dependency rules.*

1.	If (Asset X is Low) and (Asset Y is Low) and (AssetType is Data) then (DF is Low) (1)
2.	If (Asset X is Low) and (Asset Y is Medium) and (AssetType is Data) then (DF is Medium) (1)
3.	If (Asset X is Low) and (Asset Y is High) and (AssetType is Data) then (DF is Medium) (1)
4.	If (Asset X is Low) and (Asset Y is Very High) and (AssetType is Data) then (DF is High) (1)
5.	If (Asset X is Medium) and (Asset Y is Low) and (AssetType is Data) then (DF is Low) (1)
6.	If (Asset X is Medium) and (Asset Y is Medium) and (AssetType is Data) then (DF is Medium) (1)
7.	If (Asset X is Medium) and (Asset Y is High) and (AssetType is Data) then (DF is Medium) (1)
8.	If (Asset X is Medium) and (Asset Y is Very High) and (AssetType is Data) then (DF is High) (1)
9.	If (Asset X is High) and (Asset Y is Low) and (AssetType is Data) then (DF is Low) (1)
10.	If (Asset X is High) and (Asset Y is Medium) and (AssetType is Data) then (DF is Medium) (1)
11.	If (Asset X is High) and (Asset Y is High) and (AssetType is Data) then (DF is High) (1)
12.	If (Asset X is High) and (Asset Y is Very High) and (AssetType is Data) then (DF is High) (1)
13.	If (Asset X is Very High) and (Asset Y is Low) and (AssetType is Data) then (DF is Low) (1)
14.	If (Asset X is Very High) and (Asset Y is Medium) and (AssetType is Data) then (DF is Medium) (1)
15.	If (Asset X is Very High) and (Asset Y is High) and (AssetType is Data) then (DF is High) (1)
16.	If (Asset X is Very High) and (Asset Y is Very High) and (AssetType is Data) then (DF is High) (1)
17.	If (Asset X is Low) and (Asset Y is Low) and (AssetType is Application) then (DF is Low) (1)
18.	If (Asset X is Low) and (Asset Y is Medium) and (AssetType is Application) then (DF is Medium) (1)
19.	If (Asset X is Low) and (Asset Y is High) and (AssetType is Application) then (DF is High) (1)
20.	If (Asset X is Low) and (Asset Y is Very High) and (AssetType is Application) then (DF is High) (1)
21.	If (Asset X is Medium) and (Asset Y is Low) and (AssetType is Application) then (DF is Medium) (1)
22.	If (Asset X is Medium) and (Asset Y is Medium) and (AssetType is Application) then (DF is Medium) (1)
23.	If (Asset X is Medium) and (Asset Y is High) and (AssetType is Application) then (DF is High) (1)
24.	If (Asset X is Medium) and (Asset Y is Very High) and (AssetType is Application) then (DF is High) (1)
25.	If (Asset X is High) and (Asset Y is Low) and (AssetType is Application) then (DF is Medium) (1)
26.	If (Asset X is High) and (Asset Y is Medium) and (AssetType is Application) then (DF is Medium) (1)
27.	If (Asset X is High) and (Asset Y is High) and (AssetType is Application) then (DF is High) (1)
28.	If (Asset X is High) and (Asset Y is Very High) and (AssetType is Application) then (DF is High) (1)
29.	If (Asset X is Very High) and (Asset Y is Low) and (AssetType is Application) then (DF is Medium) (1)
30.	If (Asset X is Very High) and (Asset Y is Medium) and (AssetType is Application) then (DF is Medium) (1)
31.	If (Asset X is Very High) and (Asset Y is High) and (AssetType is Application) then (DF is High) (1)
32.	If (Asset X is Very High) and (Asset Y is Very High) and (AssetType is Application) then (DF is High) (1)

DF == Dependency Factor

The surface plots illustrating the implementation of these rules on all the fuzzy attributes are shown in Figure 8. In this case, we assume Asset X depends on As-

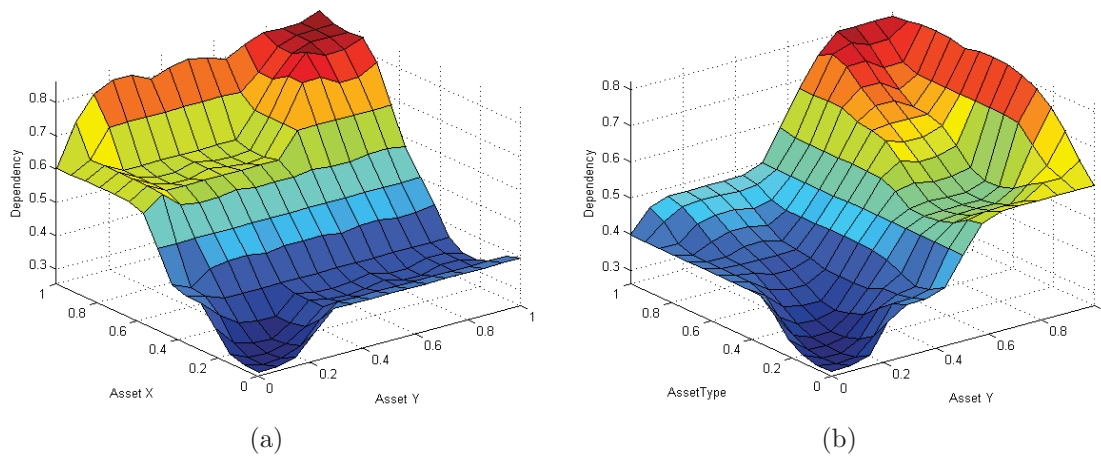


Figure 8: The fuzzy dependency factor.

set Y of type *AssetType*. Since there are three input variables to this fuzzy inference system (FIS), Figure 8 shows two surface plots for two of the three variables for each case. For this work, we limited the maximum number of dependencies to three; that is, if Asset A depends on Asset B which depends on Asset C which depends on Asset D by the respective factors of α , β , and γ , then the total risk on Asset A , R_A is as follows:

$$R_A = r_A + \alpha \left(r_B + \beta \left(\sum_j r_{C_j} + \sum_i \gamma r_{D_i} \right) \right) \quad (7)$$

where r_A , r_B , r_{C_j} , and r_{D_i} are the risk values associated with assets A , B , C , and D respectively. i represents all dependencies at the same level as D , and j represents all dependencies at the same level as C . From the surface plots in Figure 8, the dependency factor varies from 0.1 to 0.9. These values do not mean much on their own; they are relative terms with which we use to compare the risk values on different assets which depend on each other.

Entering the fuzzy variables into the fuzzy inference system (FIS) yields the fuzzy dependency values. To revert back to the crisp domain (which is preferred for making comparisons), we defuzzified the fuzzy dependency factor to a numerical value (α , β , or γ) which can be used in the dependency calculations in Equation 7.

4.3.2 Simulation Data

For this work, we used real vulnerability data obtained from the National Vulnerability Database (NVD). We will demonstrate our approach through a typical 10-host

network as shown in Figure 9. This is a corporate network with 10 hosts and three

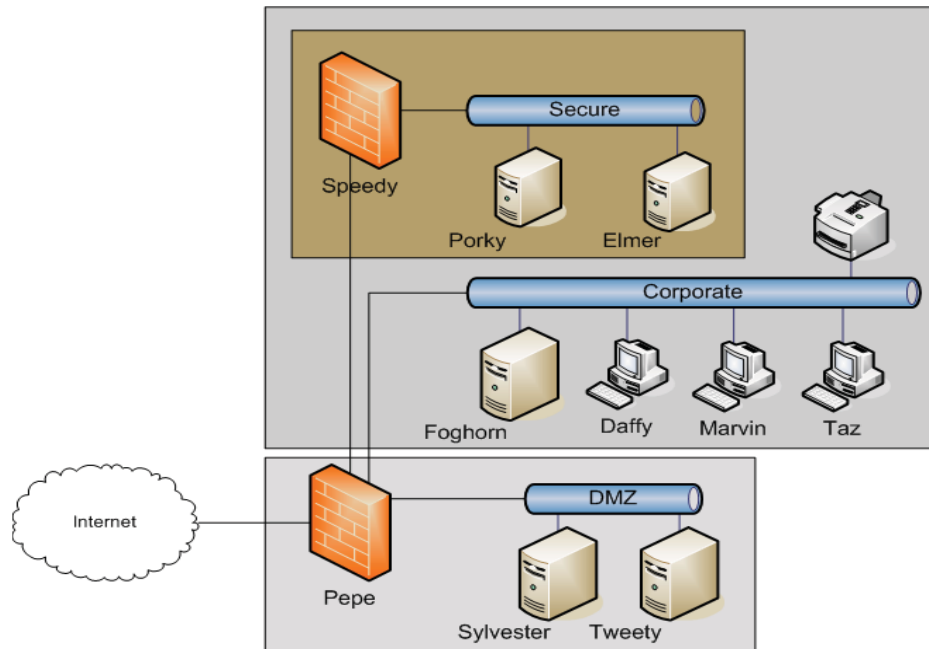


Figure 9: A Typical 10 host network.

networks. It is connected to the internet through one link. Individual assets on each node are represented by numbers. These ID assignments are shown on Table 3. The same type of asset has a different ID on a different node. That enables us to assign different asset values to each asset depending on its role in the organization or the mission at hand.

We did not have historical or recent events to go by and test our model with. However, with the vulnerabilities and assets as connected, we set up some hypothetical events that could be triggered by exploiting the real vulnerabilities associated with this network. A list of these vulnerabilities is shown in Table 4. As will be explained in Section 4.4, the same table is shown in Figure 10 with the vulnerabilities ranked according to the approach in [17]. The majority of these vulnerabilities have Common Vulnerabilities and Exposures (CVE) numbers associated with them. The few that do not have were either entered into the database before a number was available to them or this is an observed vulnerability which is just local to the organization, a capability that we wanted to be able to demonstrate as well.

Table 3: Asset identification for 10-host network (see Figure 9).

Object Type	Number	Name	Object Type	Number	Name
Network (Nt)	1	DMZ	Network (Nt)	2	Corporate
Network (Nt)	2	Secure	Node (Nd)	11	Sylvester
Node (Nd)	12	Tweety	Node (Nd)	21	Foghorn
Node (Nd)	22	Daffy	Node (Nd)	23	Marvin
Node (Nd)	24	Taz	Node (Nd)	31	Porky
Node (Nd)	32	Elmer	Asset (Ast)	112	DNS server
Asset (Ast)	111	Linux Red Hat	Asset (Ast)	113	Mail Server
Asset (Ast)	121	Linux Red Hat	Asset (Ast)	122	Web Server
Asset (Ast)	123	DNS server	Asset (Ast)	211	Windows Server 2003
Asset (Ast)	212	Mail Server	Asset (Ast)	213	Antivirus Software
Asset (Ast)	221	Windows Vista	Asset (Ast)	222	MS Office
Asset (Ast)	223	Internet Explorer	Asset (Ast)	231	MS Office
Asset (Ast)	233	Internet Explorer	Asset (Ast)	232	MS Visual studio
Asset (Ast)	234	Firefox	Asset (Ast)	235	Filezilla FTP Client
Asset (Ast)	236	Cygwin	Asset (Ast)	230	Windows XP Professional
Asset (Ast)	241	MS Office	Asset (Ast)	243	Internet Explorer
Asset (Ast)	242	MS Visual studio	Asset (Ast)	244	Firefox
Asset (Ast)	245	Filezilla FTP Client	Asset (Ast)	246	Cygwin
Asset (Ast)	240	Windows XP Professional	Asset (Ast)	311	Windows Server 2003
Asset (Ast)	312	Citrix Server	Asset (Ast)	313	Quickbooks Pro
Asset (Ast)	321	Fedora	Asset (Ast)	322	Source Repository
Asset (Ast)	323	Samba	Asset (Ast)	324	Mercurial

Table 4: Network vulnerabilities.

Asset	Vulnerabilities	
	ID	Description
112	493	Allows remote attackers to cause a denial of service(CVE-2007-0493)
111	792	Multiple integer overflows/missing upper-bounds checks in iclib CVE-2009-0792)
121	792	Multiple integer overflows/missing upper-bounds checks in iclib CVE-2009-0792)
122	6420	Cross-site request forgery (CSRF)-CVE-2007-6420
123	493	Allows remote attackers to cause a denial of service(CVE-2007-0493)
211	42	Remote attackers to access configuration files (CVE-2007-0042)
221	1084	Allows local users to execute arbitrary code (CVE-2008-1084)
221	42	Remote attackers to access configuration files (CVE-2007-0042)
222	113	MS Office Excel Viewer allows remote attackers to execute arbitrary code
222	2463	Snapshot Viewer ActiveX control in snapview allows remote download of arbitrary files
222	224	Memory Corruption Vulnerability
223	75	Allows remote attackers to execute arbitrary code (CVE-2009-0075) thru crafted HTML
223	554	Arbitrary code execution (CVE-2009-0554)
231	113	MS Office Excel Viewer allows remote attackers to execute arbitrary code
231	2463	Snapshot Viewer ActiveX control in snapview allows remote download of arbitrary files
231	224	Memory Corruption Vulnerability
233	75	Allows remote attackers to execute arbitrary code (CVE-2009-0075) thru crafted HTML
233	554	Arbitrary code execution (CVE-2009-0554)
232	3012	attackers to execute arbitrary code via a malformed EMF image (CVE-2008-3012)
234	1313	Allows remote attackers to cause a denial of service (CVE-2009-1313)
234	5502	Allows remote attackers to cause a denial of service (CVE-2008-5502)
230	1084	allows local users to execute arbitrary code (CVE-2008-1084)
230	42	Remote attackers to access configuration files (CVE-2007-0042)
241	113	MS Office Excel Viewer allows remote attackers to execute arbitrary code
241	2463	Snapshot Viewer ActiveX control in snapview allows remote download of arbitrary files
241	224	Memory Corruption Vulnerability
243	75	Allows remote attackers to execute arbitrary code (CVE-2009-0075) thru crafted HTML
243	554	Arbitrary code execution (CVE-2009-0554)
242	3012	attackers to execute arbitrary code via a malformed EMF image (CVE-2008-3012)
244	244	Allows remote attackers to cause a denial of service (CVE-2009-1313)
244	5502	Allows remote attackers to cause a denial of service (CVE-2008-5502)
240	1084	Allows local users to execute arbitrary code (CVE-2008-1084)
240	42	Remote attackers to access configuration files (CVE-2007-0042)
311	42	Remote attackers to access configuration files (CVE-2007-0042)
321	792	Multiple integer overflows/missing upper-bounds checks in iclib CVE-2009-0792)
323	1105	Heap-based buffer overflow in the receive_smb_raw function in util/sock.c
323	2444	Logic error in the SID/Name translation functionality in smbd
323	2446	Multiple heap-based buffer overflows in the NDR parsing in smbd
323	2447	Allows remote attackers to execute arbitrary commands via shell metacharacters
324	4297	Versions less than 1.0.2 allows remote attackers to read repository files

To demonstrate our approach, we combine these vulnerabilities to simulate a potential event. The simulated events are listed in Table 5. Each event is associated with one or more vulnerabilities, and one or more assets that have these vulnerabilities. In practice, an event is often associated with one vulnerability. In addition we have

Table 5: Simulated events for the 10-Host network.

Event	Vulnerabilities Exploited	Assets affected
1121	493	112
2431	75	122
2401	1084	230, 221, 240
2411	224	241, 222, 231
2431	75	223, 233, 234, 243, 244
2441	5502	234, 244
2442	1313	234, 244
3111	42	221, 230, 240
24111	113	222, 231, 241
3111	42	311
3211	792	321
3231	1105	323
3232	2444	323
3241	4297	324

associated dependencies on a selected number of assets as shown in Table 6. Although not stated here, each application depends on its Operating System (OS).

Table 6: Asset dependencies for the 10-Host network.

Asset	Dependency
234, 244, 243, 233, 223	112, 122
212	122
122	112

4.4 Results

The vulnerabilities in Table 4 are also listed in the table in Figure 10 showing their relative potential risk values as determined in [17]. They are listed in order, starting with the one with the highest risk value to the lowest. An analyst can therefore

ID	Name	Status	Risk Level
1084	Allows local users to execute arbitrary code (CVE-2008-1084)	Open	6.515316
792	Multiple integer overflows/ missing upper-bounds checks in iclib (CVE-2009-0792)	Open	5.742774
554	Arbitrary code execution (CVE-2009-0554)	Open	5.486445
2444	Logic error in the SID/Name translation functionality in smbd	Open	5.212251
2446	Multiple heap-based buffer overflows in the NDR parsing in smbd	Open	5.166816
113	MS Office Excel Viewer allows remote attackers to execute arbitrary code	Open	5.166816
75	Allows remote attackers to execute arbitrary code (CVE-2009-0075) thru crafted HTML	Open	3.947838
3012	attackers to execute arbitrary code via a malformed EMF image (CVE-2008-3012)	Open	2.5633065
224	Memory Corruption Vulnerability	Open	2.1739425
493	Allows remote attackers to cause a denial of service(CVE-2007-0493)	Open	2.15284
244	Allows remote attackers to cause a denial of service (CVE-2009-1313)	Open	1.4599737
1313	Allows remote attackers to cause a denial of service (CVE-2009-1313)	Open	1.4599737
2463	Snapshot Viewer ActiveX control in snapview allows remote download of arbitrary files	Open	1.4158566
1105	Heap-based buffer overflow in the receive_smb_raw function in util/sock.c	Open	1.4158566
42	Remote attackers to access configuration files (CVE-2007-0042)	Open	0.873369
2447	Allows remote attackers to execute arbitrary commands via shell metacharacters	Open	0.625638
494	Allows remote attackers to cause a denial of service(CVE-2007-0493)	Open	0.5899403
5502	Allows remote attackers to cause a denial of service (CVE-2008-5502)	Open	0.2949702
6420	Cross-site request forgery (CSRF)-CVE-2007-6420	Open	0.1487323
4297	Versions <1.0.2 allows remote attackers to read repository files	Open	0.1189859

Confidentiality : 0.118
Integrity : 0
Availability : 0

Figure 10: All the vulnerabilities affecting the 10 node network

prioritize remedial action accordingly. Similarly, the summarized relative risk values associated with each network are shown in Figure 11.

Network ID	Name	Risk Level
2	Corporate	86.2486191
3	Secure	17.8795175
1	DMZ	17.1198409

Figure 11: The affected networks in order of priority.

To achieve comparable results, we used attribute data that is comparable to the

Common Vulnerability Scoring System (CVSS) model. The approach in [17] made comparisons with CVSS output for the ranges of data that CVSS handles. The authors concluded that with data beyond that range, the results could not be verified or compared with other sources. It is with this in mind that we limit our attribute base to match that of CVSS so that we could make some comparisons.

Our approach first lists all the events in the affected networks in order of priority, based on their potential risk values. The events for our networks are shown in Figure 12. The events are ranked in order starting with the one with the highest risk

ID	Description	Risk Level
2401	Arbitrary code execution attempt	24.560847
2431	Remote Attack using crafted HTML	24.1135917
24111	Remote code Execution on MS Office Excel	20.471634
2442	Remote Attack Dos	17.056848
2411	DoS due to Memory corruption	9.355491
3232	Translation Functionality	5.212251
3211	Code Execution attempt due to Integer overflow	4.485786
3111	Remote access of Config files	4.165716
2441	Remote Attack Dos	2.8419663
1121	DNS poisoning	2.15284
3231	DoS due to heap-based buffer overflow	1.4158566
3241	Illegal File Access	0.1189859

Figure 12: All the events on the 10 node network

value. The different colour codes represent the ratio of the calculated risk value to the asset value; the higher the ratio, the closer the colour code is to red—otherwise it will be green. By cross-referencing with the vulnerabilities associated with each event, the vulnerability patching process or other courses of action may be prioritized accordingly. The risk values themselves do not have any significance outside this work. They can only be used for these types of comparisons within a corporation; they cannot be used for comparisons with data obtained in other ways.

We are also able to look at the events for each individual network. In Figure 13, we show the event–ranking for each individual network. The DMZ network has the least number of events and the Corporate network the highest.

It is also important to note that some events like 2431, appear in more than one network; this is because the event affects assets in both networks. In Network 1, this event is associated with a risk value of 2.1, while in Network 2, it has a value of 22.0. From Table 5, we note that this event exploits vulnerability 75. This vulnerability is affects one asset on network 1 and five assets on network 2. That is why Network 2 has a higher risk value associated with event 2431 than Network 1.

ID	Description	Risk Level
1121	DNS poisoning	2.153
2431	Remote Attack using crafted HTML	2.100

(a) Network 1 (DMZ)

ID	Description	Risk Level
2401	Arbitrary code execution attempt	24.561
2431	Remote Attack using crafted HTML	22.013
24111	Remote code Execution on MS Office Excel	20.472
2442	Remote Attack Dos	17.057
2411	DoS due to Memory corruption	9.355
3111	Remote access of Config files	3.292
2441	Remote Attack Dos	2.842

(b) Network 2 (Corporate)

ID	Description	Risk Level
3232	Translation Functionality	5.212
3211	Code Execution attempt due to Integer overflow	4.486
3231	DoS due to heap-based buffer overflow	1.416
3111	Remote access of Config files	0.873
3241	Illegal File Access	0.119

(c) Network 3 (Secure)

Figure 13: The output events from the three networks.

Each network has nodes and we have determined and ranked the events for each one of the nodes as shown in Figure 14. In this case, we selected nodes Daffy, Elmer and

ID	Description	Risk Level
2401	Arbitrary code execution attempt	8.702
24111	Remote code Execution on MS Office Excel	7.335
2431	Remote Attack using crafted HTML	6.035
2411	DoS due to Memory corruption	3.344
3111	Remote access of Config files	1.167

(a) Node 22 (Daffy)

ID	Description	Risk Level
3232	Translation Functionality	5.212
3211	Code Execution attempt due to Integer overflow	4.486
3231	DoS due to heap-based buffer overflow	1.416
3241	Illegal File Access	0.119

(b) Node 32 (Elmer)

ID	Description	Risk Level
2442	Remote Attack Dos	9.540
2401	Arbitrary code execution attempt	8.258
24111	Remote code Execution on MS Office Excel	6.895
2431	Remote Attack using crafted HTML	6.501
2411	DoS due to Memory corruption	3.144
2441	Remote Attack Dos	3.070
3111	Remote access of Config files	1.107

(c) Node 24 (Taz) (Secure)

Figure 14: The output events from three selected nodes.

Taz for this demonstration. In Figure 15, we show the list of assets on node 24 (Taz),

Asset ID	Name	Risk Level
244	Firefox	16.253
241	MS Office	11.928
243	Internet Explorer	10.339
240	Windows XP Professional	9.365
242	MS Visual studio	4.278

Figure 15: The affected assets on node 24.

with the relative risk values associated with each network. Although asset 241 has a higher asset value (10) than asset 244 (5), the risk value associated with asset 241 makes it the top priority for any course of action.

Finally, we selected two assets, Windows XP Professional and Firefox v3.0.2 (both on Taz). The prioritization shows the list of all events on these two assets on Taz.

ID	Description	Risk Level
2401	Arbitrary code execution attempt	8.258
3111	Remote access of Config files	1.107

(a) Asset 240 (Windows XP Professional)

ID	Description	Risk Level
2442	Remote Attack Dos	9.540
2431	Remote Attack using crafted HTML	3.644
2441	Remote Attack Dos	3.070

(b) Asset 244 (Firefox v3.0.2)

Figure 16: The output events from two selected assets.

Based on these two assets, it would be recommended that a course of action be found for event 2401. A cross-reference with the vulnerabilities table shows that this refers to a vulnerability 1084 (CVE-2008-1084).

These results show that we can translate the linguistic declarations about vulnerabilities, assets, events and how they are interconnected, into numerical figures which we can use to prioritize courses of action. The results translates the analyst’s “gut-feel” into a broader networks-wide approach to rank events. The modeling of dependency factors also ensures that remedial action can be taken on all assets that contribute to an event; or that at least the analyst is made aware of the presence of assets whose dependency may contribute to an event.

5 Conclusion

We have developed a fuzzy risk-based approach to prioritize events in a computer network. We achieved this by translating analysts's experiential declarations about vulnerability attributes into a fuzzy mathematical model. Our approach is based on an earlier model which prioritizes vulnerabilities based on the potential risk they pose to the computer network and its assets [17]. The success of that approach in comparison with the Common Vulnerability Scoring System (CVSS) approach for comparable vulnerability attributes enabled us to extend the approach to event prioritization.

In this work, we associated each event with a set of vulnerabilities. Using a fuzzy model, we calculated the risk associated with each event, including dependencies. We then ranked the events based on this risk value, with the highest risk value taking the highest rank. We achieved this at network, node and asset levels. These rankings would help the analyst to prioritize their mitigation processes by attending to the highest ranked events first.

For attributes that lie outside our definition range, it is unpredictable what the ranking produced would be, although the initial objective of this work was to cover all attribute ranges as would be seen by a competent analyst. Significant attribute changes that are not covered in this work reflect a failure of the underlying assumptions of this work that every vulnerability, and therefore event, can be characterized by an all-inclusive set of attributes. As in the vulnerability prioritization, the analysis on events defined outside the attribute range presented in this work, may only be considered promising since there is no way of comparison at this time, except using the analysts' "gut-feel" which we have already captured in this work.

Future work will include the development of a portable demonstrator to be integrated into the ongoing work on a defensive posture demonstrator. At that point, we will be able to make side-by-side comparisons of our relative rankings with those produced by other approaches. To make a complete assessment of this approach for possible client use, we intend to set up experiments with client analysts, and make comparisons of what this method produces and what they would expect as an output. We would also complement the analysts' output expectations with what other tools recommend.

References

- [1] Anderson, K. E. (1998), Intelligence-Based Threat Assessments for Information Networks and Infrastructures: A White Paper, *Global Technology Research, Inc.*
- [2] Pfleeger, C. P. (1997), Security in Computing, 2 ed, Upper Saddle River, NJ: Prentice Hall PTR.
- [3] Mosleh, A., Hilton, E. R., and Browne, P. S. (1985), Bayesian probabilistic risk analysis, *ACM SIGMETRICS-Performance Evaluation Review*, 13(1), 5–12.
- [4] Antón, P. S., Anderson, R. H., Mesic, R., and Scheiern, M. (2003), Finding and fixing vulnerabilities in information systems: The Vulnerability Assessment & Mitigation Methodology, (Technical Report MR-1601-DARPA) RAND National Defence Research Institute, Santa Monica, CA. Available on-line at <http://www.rand.org/publications/MR/MR1601/MR1601.pdf>.
- [5] Anderson, R. H., Feldmman, P. M., Gerwehr, S., Houghton, B., Mesic, R., Pinder, J., Rothenberg, J., and Chiesa, J. R. (1999), Securing the U.S. Defense Information Infrastructure: A Proposed Approach, (Technical Report MR-993-OSD/NSA/DARPA) RAND Corporation, Santa Monica, CA. Available on-line at <http://www.rand.org/publications/MR/MR993/>.
- [6] Relex, Relex Fault Tree/Event Tree (online), <http://www.relexsoftware.com/products/ftaeta.asp> (Access Date: 17 Nov 2005).
- [7] IEE, Quantified Risk Assessment Techniques (online), <http://www.iee.org/Policy/Areas/Health/hsb26c.cfm> (Access Date: 17 Nov 2005).
- [8] Isograph (2005), FaultTree+ - Event Tree Analysis (online), <http://www.isograph-software.com/ftpovereta.htm> (Access Date: 17 Nov 2005).
- [9] Chen, S. and Chen, s. (2003), Fuzzy Risk Analysis Based on Similarity Measures of Generalized Fuzzy Numbers, *IEEE Transactions on Fuzzy Systems*, 11(1), 45–56.
- [10] Chen, S. (1996), New methods for subjective mental workload assessment and fuzzy risk analysis, *Cybernetics and Systems*, 27(5), 449–472.
- [11] Shah, S., Measuring Operational Risk Using Fuzzy Logic Modeling (online), IRMI.com, <http://www.irmi.com/Expert/Articles/2003/Shah09.aspx> (Access Date: 6 Aug 2005).

- [12] Buckshaw, D. L., Parnell, G. S., Unkenholz, W. L., Parks, D. L., Wallner, J. M., and Saydjari, O. S. (2005), Mission Oriented Risk and design Analysis of Critical Information Systems, *Military Operations Research*, V10(N2), 19–38.
- [13] Kangari, R. and Riggs, L. (1989), Construction Risk Assessment by Linguistics, *IEEE Transactions on Engineering Management*, 36(2), 126–131.
- [14] H-J. Zimmerman (1987), Fuzzy Sets, Decision Making and Expert Systems, Kluwer Academic Publishers.
- [15] Symantec Enterprise Security (2006), Symantec Internet Security Threat Report: Trends for July 05-December 05, *Symantec Enterprise Security*, IX, 1–106.
- [16] Prats, J. P. (2003), Development and testing Matlab based fuzzy systems applications, Technical Report The University of Warwick.
- [17] Dondo, M. (2008), A Vulnerability Prioritization System Using A Fuzzy Risk Analysis Approach, In *Proceedings of The Ifip Tc 11 23rd International Information Security Conference*, Vol. 278, pp. 525–540.
- [18] Schiffman, M., The common Vulnerability Scoring System (CVSS) (online), FiRST, <http://www.first.org/cvss/cvss-guide.html> (Access Date: 16 Sep 2005).
- [19] Schiffman, M., The Common Vulnerability Scoring System (CVSS) (online), FiRST, <http://www.packetfactory.net/papers/CVSS/cvss-ppt.pdf> (Access Date: 16 Sep 2005).
- [20] NVD, National Vulnerability Database (online), NVD, <http://nvd.nist.gov> (Access Date: 4 Sept 2006).
- [21] Chambers, J. T. and Thompson, J. W., Common Vulnerability Scoring System (online), NIAC, <http://www.first.org/cvss/cvss-dhs-12-02-04.pdf> (Access Date: 19 Sep 2005).
- [22] Chambers, J. T. and Thompson, J. W., Vulnearbility Disclosure Framework (online), NIAC, <http://www.dhs.gov/interweb/assetlibrary/vdwgreport.pdf> (Access Date: 19 Sep 2005).
- [23] Naraine, R., Cisco’s Free Threat-Alerts Service Uses CVSS (online), eWEEK.com, <http://www.eweek.com/article2/0,1759,1821377,00.asp> (Access Date: 19 Sep 2005).

Acronyms and Abbreviations

AI	availability impact
AC	access complexity
CI	confidentiality impact
CD	collateral damage
CFNOC	Canadian Forces Network Operations Centre
CIA	confidentiality, integrity or availability
COA	courses of action
CVE	Common Vulnerabilities and Exposures
CVSS	Common Vulnerability Scoring System
DOS	denial of service
ETA	Event Trees Analysis
FIS	fuzzy inference system
FL	fuzzy logic
FN	fuzzy number
FTA	Fault Trees Analysis
II	integrity impact
IB	impact bias
KRI	key risk indicator
MF	membership function
NIAC	National Infrastructure Advisory Council
NVD	National Vulnerability Database
OS	Operating System
RC	report confidence
RL	remediation level

RTO	real-time operating system
TD	target distribution
VAM	Vulnerability Assessment and Mitigation
VPA	Vulnerability Prioritization Assistant

This page intentionally left blank.

Annex A: Numerical Significance of Results

In this section, we show the consistency of our numerical results. We do this by showing the significance of the results produced and the rationale behind ranking through risk value comparisons. We split the attributes into two broad categories, namely those that contribute to an increase in risk and those that contribute to its reduction. In both cases, we will show that, with the risk function's monotonicity, it is justifiable to make comparisons of the risk values produced without prejudicing the underlying fuzzy theory or computer security theories; thus giving some justification for the relative rankings produced in this paper.

From Equation 6, the variables that determine the direction and therefore magnitude

$$r(v) = c \times t \times p$$

of this equation are t and p . These variables are in turn affected by the fuzzy attributes introduced in this work. Multiple attributes define the “security posture” of an asset, or set of assets for a given vulnerability. In this work, our aim is to show that the risk exposure on an asset worsens (i.e. its value goes up) as the “security posture” described by the vulnerability attributes worsens. When the posture improves, the risk exposure improves (i.e. risk value decreases). Our objective was that, as the “security posture” worsens, the risk exposure worsens monotonically, i.e. $r(v)$ would have a monotonic increasing behavior. In a similar way, when the “security posture” improves, the risk exposure improves monotonically, i.e. $r(v)$ would have a monotonic decreasing behavior.

Based on this monotonicity in $r(v)$, we propose that we should be able to compare different vulnerabilities based on the value of the risk they expose assets to. Thus, we came up with this ranking system in which provide a ranking based on the value of $r(v)$. In the following sections, we will go on to show which attributes contribute to this monotonicity in $r(v)$.

A.1 Monotonic Increasing Risk Function

To demonstrate the monotonicity (increasing behaviour) of the risk function when the security posture worsens, we chose an arbitrary starting point. We took all attributes that are non-defensive (see Section 2.2). This covers all attributes except *safeguards* and *remediation level* (which we set at the least effective level). We calculated the risk value for the attribute combinations that produces the lowest risk value; this comes from the combination of the lowest fuzzy numbers in each attribute. For example, we would take *Access Complexity = High*, *Access Vector = Local*, *Authentication = Required*, etc. Starting at *Time = 0*, we worsened the security posture of the asset for a given hypothetical vulnerability, and calculated the risk value at each step. We

plotted the risk values at each step as the security posture is worsened. This is shown in Figure A.1.

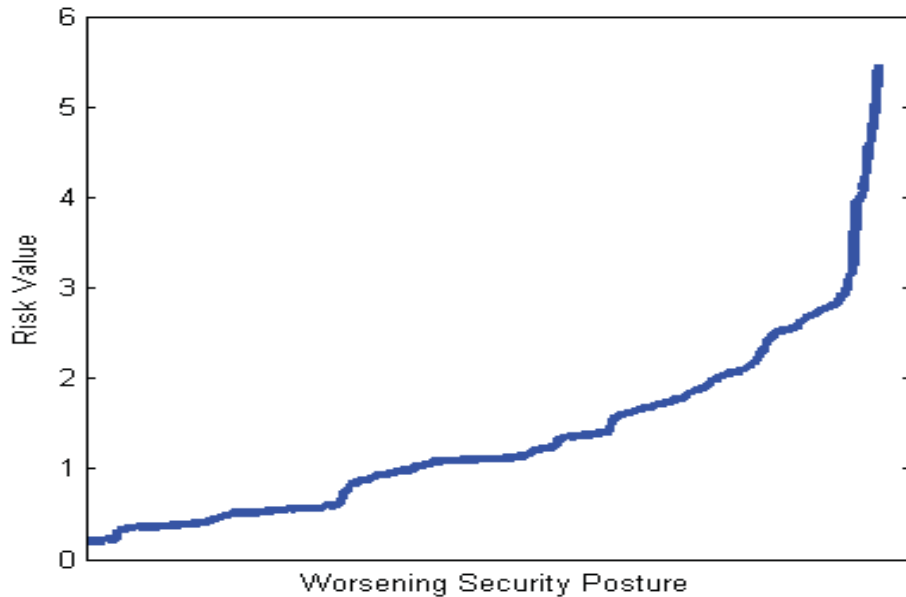


Figure A.1: The risk value as the security posture of an asset deteriorates.

As expected, the risk value increases monotonically as the security posture worsens. Each vulnerability or event can be in one of the states illustrated in the figure. Those higher up the curve reflect a higher risk value and our work recommends that they should get preference in any courses of action (COA) proposed. This is the basis of our ranking approach.

A.2 Monotonic Decreasing Risk function

In a similar way, we took the remaining *risk-reducing* attributes, and changed them to improve the security posture. We calculated the risk value at each step We plotted the risk values at each step as the security posture is worsened. This is shown in Figure A.2. It should be noted that this curve has less security points than the previous curve because we have fewer attributes. The *Time* attribute remained at the *Very High* for this part of the demonstration. This demonstration, of course, assumes that this course of action (ie to apply remediation) was only taken at the that particular time (*Very High*). However, it should be emphasized that our model can handle courses of action at any time

Again, the risk value decreases monotonically as the security posture improves. It is

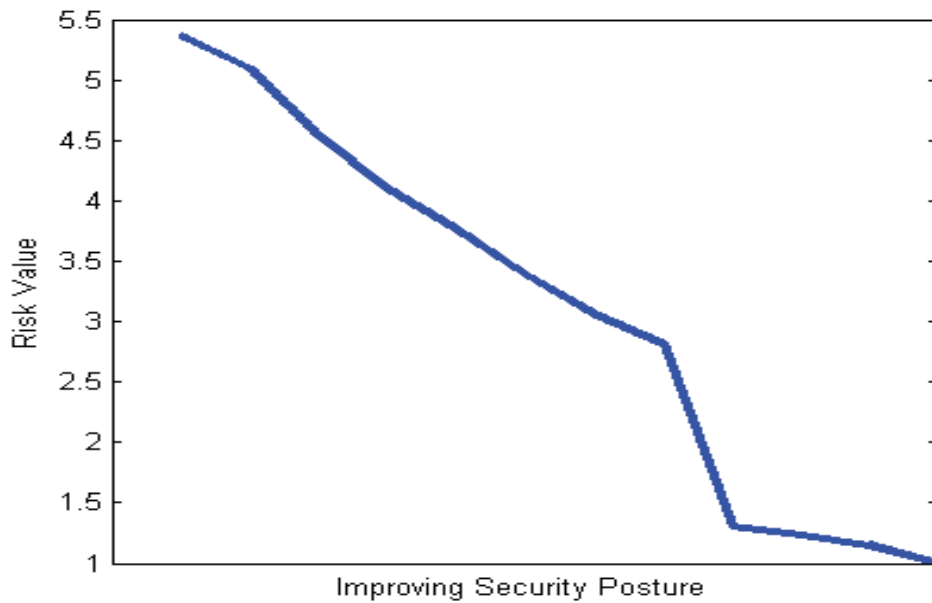


Figure A.2: The risk value as the security posture of an asset improves.

therefore our conclusion that we should be able to compare two or more vulnerabilities whose attributes result in changes to risk values as indicated in Figures A.1 and A.2.

A.3 Time Variability

Another claim that we make in our work is that of the time variability of the risk value. We assume that the longer it takes to patch a vulnerability the higher the risk value. We model this into our fuzzy functions and show that. We went on to claim that we should be able to extend this principle and compare different risk values for events (and vulnerabilities) spanning two or more time periods.

Without the remediation attributes described above, we set the time attribute to the 5 different levels used in this work. We calculated the risk values as we varied the attributes to reflect a worsening security posture. We then plotted the risk values at each step for each time period. The results are shown in Figure A.3.

Each time period is shown to have a monotonic increasing variability as the security posture worsens. Except for very few instances at very low risk values, the general trend shows that for a given instance, the risk value is higher as time increases from *Very Low* to *Very High*. It should be noted that in our earlier work [17], we were able to show that within a given time period (any given curve in Figure A.3), our ranking approach matched that of CVSS.

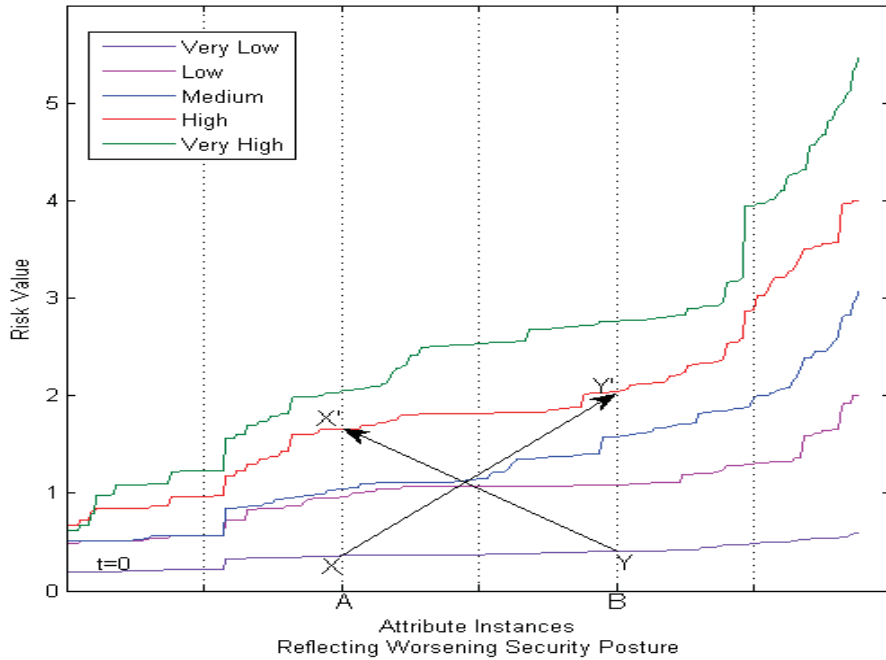


Figure A.3:

Now, consider two vulnerabilities v_A and v_B whose attribute instances are respectively represented by A and B in Figure A.3. The vertical dotted lines in Figure A.3 represent a fixed instance of vulnerability attributes. Consider the *Very Low* time curve to represent a time $t = 0$. If we keep the attributes of v_A constant and change the time to *High* ($t = t_1$), the risk value will change from a value at X to X' . A similar change to the attributes of v_B will see its risk value changing from Y to Y' . So, at time t_1 , v_B is ranked higher than v_A . This is the same relative ranking as at time $t = 0$, but the difference margin is now wider.

However, if at time t_1 the attributes of v_A change to those represented by instance B in Figure A.3 and at the same time the attributes of v_B change to those represented by instance A , then the risk value associated with v_B is now represented by X' , while that for v_A is represented by Y' . This time, v_A is ranked higher than v_B . This is different from the way it was at time $t = 0$.

The curves in Figure A.3 also show that there are many instances of events, on a higher time curve, with risk values higher than events on a lower time curve. The reverse is true. We therefore proposed in this work that, given this scenario, we should be able to compare events in different time zones as well.

Annex B: Fuzzy Rules

For rule clarity, we abbreviated ExploitabilityValue to EV.

B.1 FIS BaseValue Rules

1. If (AV is Local) and (AC is High) and (Auth is Required) and (Impact is None) then (BaseValue is VeryLow) (1)
2. If (AV is Local) and (AC is High) and (Auth is Required) and (Impact is Partial) then (BaseValue is Low) (1)
3. If (AV is Local) and (AC is High) and (Auth is Required) and (Impact is Complete) then (BaseValue is Medium) (1)
4. If (AV is Local) and (AC is High) and (Auth is NotRequired) and (Impact is None) then (BaseValue is VeryLow) (1)
5. If (AV is Local) and (AC is High) and (Auth is NotRequired) and (Impact is Partial) then (BaseValue is Medium) (1)
6. If (AV is Local) and (AC is High) and (Auth is NotRequired) and (Impact is Complete) then (BaseValue is High) (1)
7. If (AV is Local) and (AC is Low) and (Auth is Required) and (Impact is None) then (BaseValue is VeryLow) (1)
8. If (AV is Local) and (AC is Low) and (Auth is Required) and (Impact is Partial) then (BaseValue is Low) (1)
9. If (AV is Local) and (AC is Low) and (Auth is Required) and (Impact is Complete) then (BaseValue is Medium) (1)
10. If (AV is Local) and (AC is Low) and (Auth is NotRequired) and (Impact is None) then (BaseValue is VeryLow) (1)
11. If (AV is Local) and (AC is Low) and (Auth is NotRequired) and (Impact is Partial) then (BaseValue is Medium) (1)
12. If (AV is Local) and (AC is Low) and (Auth is NotRequired) and (Impact is Complete) then (BaseValue is VeryHigh) (1)
13. If (AV is Remote) and (AC is High) and (Auth is Required) and (Impact is None) then (BaseValue is VeryLow) (1)
14. If (AV is Remote) and (AC is High) and (Auth is Required) and (Impact is Partial) then (BaseValue is Medium) (1)
15. If (AV is Remote) and (AC is High) and (Auth is Required) and (Impact is Complete) then (BaseValue is Medium) (1)
16. If (AV is Remote) and (AC is High) and (Auth is NotRequired) and (Impact is None) then (BaseValue is VeryLow) (1)
17. If (AV is Remote) and (AC is High) and (Auth is NotRequired) and (Impact is Partial) then (BaseValue is High) (1)
18. If (AV is Remote) and (AC is High) and (Auth is NotRequired) and (Impact is Complete) then (BaseValue is VeryHigh) (1)
19. If (AV is Remote) and (AC is Low) and (Auth is Required) and (Impact is None) then (BaseValue is VeryLow) (1)
20. If (AV is Remote) and (AC is Low) and (Auth is Required) and (Impact is Partial) then (BaseValue is Medium) (1)
21. If (AV is Remote) and (AC is Low) and (Auth is Required) and (Impact is Complete) then (BaseValue is High) (1)
22. If (AV is Remote) and (AC is Low) and (Auth is NotRequired) and (Impact is None) then (BaseValue is VeryLow) (1)
23. If (AV is Remote) and (AC is Low) and (Auth is NotRequired) and (Impact is Partial) then (BaseValue is High) (1)
24. If (AV is Remote) and (AC is Low) and (Auth is NotRequired) and (Impact is Complete) then (BaseValue is VeryHigh) (1)

B.2 FIS EV Rules

1. If (Exploitability is Unproven) and (RL is OfficialFix) and (RC is Unconfirmed) then (EV is VeryLow) (1)
2. If (Exploitability is Unproven) and (RL is OfficialFix) and (RC is Uncorroborated) then (EV is VeryLow) (1)
3. If (Exploitability is Unproven) and (RL is OfficialFix) and (RC is Confirmed) then (EV is Low) (1)
4. If (Exploitability is Unproven) and (RL is TemporaryFix) and (RC is Unconfirmed) then (EV is VeryLow) (1)
5. If (Exploitability is Unproven) and (RL is TemporaryFix) and (RC is Uncorroborated) then (EV is VeryLow) (1)
6. If (Exploitability is Unproven) and (RL is TemporaryFix) and (RC is Confirmed) then (EV is Low) (1)
7. If (Exploitability is Unproven) and (RL is Unavailable) and (RC is Unconfirmed) then (EV is VeryLow) (1)
8. If (Exploitability is Unproven) and (RL is Unavailable) and (RC is Uncorroborated) then (EV is Low) (1)
9. If (Exploitability is Unproven) and (RL is Unavailable) and (RC is Confirmed) then (EV is Medium) (1)
10. If (Exploitability is Unproven) and (RL is Workaround) and (RC is Unconfirmed) then (EV is Low) (1)
11. If (Exploitability is Unproven) and (RL is Workaround) and (RC is Uncorroborated) then (EV is Medium) (1)
12. If (Exploitability is Unproven) and (RL is Workaround) and (RC is Confirmed) then (EV is Medium) (1)
13. If (Exploitability is ProofOfConcept) and (RL is OfficialFix) and (RC is Unconfirmed) then (EV is VeryLow) (1)
14. If (Exploitability is ProofOfConcept) and (RL is OfficialFix) and (RC is Uncorroborated) then (EV is Low) (1)
15. If (Exploitability is ProofOfConcept) and (RL is OfficialFix) and (RC is Confirmed) then (EV is Low) (1)
16. If (Exploitability is ProofOfConcept) and (RL is TemporaryFix) and (RC is Unconfirmed) then (EV is VeryLow) (1)
17. If (Exploitability is ProofOfConcept) and (RL is TemporaryFix) and (RC is Uncorroborated) then (EV is Low) (1)
18. If (Exploitability is ProofOfConcept) and (RL is TemporaryFix) and (RC is Confirmed) then (EV is Medium) (1)
19. If (Exploitability is ProofOfConcept) and (RL is Unavailable) and (RC is Unconfirmed) then (EV is Low) (1)
20. If (Exploitability is ProofOfConcept) and (RL is Unavailable) and (RC is Uncorroborated) then (EV is Medium) (1)
21. If (Exploitability is ProofOfConcept) and (RL is Unavailable) and (RC is Confirmed) then (EV is Medium) (1)
22. If (Exploitability is ProofOfConcept) and (RL is Workaround) and (RC is Unconfirmed) then (EV is Medium) (1)
23. If (Exploitability is ProofOfConcept) and (RL is Workaround) and (RC is Uncorroborated) then (EV is Medium) (1)
24. If (Exploitability is ProofOfConcept) and (RL is Workaround) and (RC is Confirmed) then (EV is High) (1)
25. If (Exploitability is Functional) and (RL is OfficialFix) and (RC is Unconfirmed) then (EV is Low) (1)
26. If (Exploitability is Functional) and (RL is OfficialFix) and (RC is Uncorroborated) then (EV is Low) (1)
27. If (Exploitability is Functional) and (RL is OfficialFix) and (RC is Confirmed) then (EV is Medium) (1)
28. If (Exploitability is Functional) and (RL is TemporaryFix) and (RC is Unconfirmed) then (EV is Low) (1)
29. If (Exploitability is Functional) and (RL is TemporaryFix) and (RC is Uncorroborated) then (EV is Medium) (1)
30. If (Exploitability is Functional) and (RL is TemporaryFix) and (RC is Confirmed) then (EV is Medium) (1)
31. If (Exploitability is Functional) and (RL is Unavailable) and (RC is Unconfirmed) then (EV is Medium) (1)
32. If (Exploitability is Functional) and (RL is Unavailable) and (RC is Uncorroborated) then (EV is Medium) (1)
33. If (Exploitability is Functional) and (RL is Unavailable) and (RC is Confirmed) then (EV is High) (1)
34. If (Exploitability is Functional) and (RL is Workaround) and (RC is Unconfirmed) then (EV is Medium) (1)
35. If (Exploitability is Functional) and (RL is Workaround) and (RC is Uncorroborated) then (EV is High) (1)
36. If (Exploitability is Functional) and (RL is Workaround) and (RC is Confirmed) then (EV is VeryHigh) (1)
37. If (Exploitability is High) and (RL is OfficialFix) and (RC is Unconfirmed) then (EV is Low) (1)
38. If (Exploitability is High) and (RL is OfficialFix) and (RC is Uncorroborated) then (EV is Medium) (1)
39. If (Exploitability is High) and (RL is OfficialFix) and (RC is Confirmed) then (EV is Medium) (1)
40. If (Exploitability is High) and (RL is TemporaryFix) and (RC is Unconfirmed) then (EV is Medium) (1)

- 214. If (BaseValue is High) and (Safeguards is High) and (EV is Medium) and (Time is High) then (Likelihood is Medium) (1)
- 215. If (BaseValue is High) and (Safeguards is High) and (EV is Medium) and (Time is VeryHigh) then (Likelihood is High) (1)
- 216. If (BaseValue is High) and (Safeguards is High) and (EV is High) and (Time is VeryLow) then (Likelihood is VeryLow) (1)
- 217. If (BaseValue is High) and (Safeguards is High) and (EV is High) and (Time is Low) then (Likelihood is Low) (1)
- 218. If (BaseValue is High) and (Safeguards is High) and (EV is High) and (Time is Medium) then (Likelihood is Medium) (1)
- 219. If (BaseValue is High) and (Safeguards is High) and (EV is High) and (Time is High) then (Likelihood is Medium) (1)
- 220. If (BaseValue is High) and (Safeguards is High) and (EV is High) and (Time is VeryHigh) then (Likelihood is VeryHigh) (1)
- 221. If (BaseValue is High) and (Safeguards is High) and (EV is VeryHigh) and (Time is VeryLow) then (Likelihood is VeryLow) (1)
- 222. If (BaseValue is High) and (Safeguards is High) and (EV is VeryHigh) and (Time is Low) then (Likelihood is Medium) (1)
- 223. If (BaseValue is High) and (Safeguards is High) and (EV is VeryHigh) and (Time is Medium) then (Likelihood is Medium) (1)
- 224. If (BaseValue is High) and (Safeguards is High) and (EV is VeryHigh) and (Time is High) then (Likelihood is High) (1)
- 225. If (BaseValue is High) and (Safeguards is High) and (EV is VeryHigh) and (Time is VeryHigh) then (Likelihood is VeryHigh) (1)

This page intentionally left blank.

Annex C: Common Vulnerability Scoring System (CVSS)

CVSS is a relatively new approach used to quantitatively analyse vulnerabilities [18–20]. It is a product of the National Infrastructure Advisory Council (NIAC) effort to introduce an open standard for vulnerability scoring [21, 22]. There have been significant efforts to encourage all vulnerability assessment tool vendors to use CVSS, and some of the larger vendors are heeding the call [23].

The CVSS approach is based on three basic metric groups. Each metric is a characteristic or a group of characteristics of a vulnerability that can be measured quantitatively or qualitatively. They are defined as follows:

1. Base metric group *b*: This defines the characteristics of some aspects of a vulnerability that do not change with time, nor in different target environments. These characteristics are as follows:
 - (a) *Confidentiality impact (CI)* metric b_{ci} measures the impact on confidentiality of a successful exploit of the vulnerability on the target asset. The possible scores for this metric are as follows:
 - i. *none*: No impact
 - ii. *partial*: There is significant informational disclosure
 - iii. *complete*: A total compromise of critical system information
 - (b) *Integrity impact (II)* metric b_{ii} measures the impact on integrity a successful exploit of a vulnerability will have on the target asset. The possible scores for this metric are as follows:
 - i. *none*: No impact
 - ii. *partial*: Significant breach of integrity
 - iii. *complete*: A total compromise of system integrity
 - (c) *Availability impact (AI)* metric b_{ai} measures the impact on availability a successful exploit of the vulnerability will have on the target asset. The possible scores for this metric are as follows:
 - i. *none*: No impact
 - ii. *partial*: There is significant resource interruption
 - iii. *complete*: A total shutdown of the resource
 - (d) *Impact bias (IB)* metric b_{ib} gives a stronger weighting to one of the impact metrics over the other two. This allows for distinctions to be made on the importance of CIA functionalities and services on the asset. The corresponding CIA bias terms are b_{cib} , b_{iib} , and b_{aib} . The possible scores for this metric are as follows:

- i. *Normal*: Weights on “Impact scores” for CIA are all equal
 - ii. *Confidentiality*: confidentiality impact (CI) is assigned greatest weight
 - iii. *Integrity*: integrity impact (II) is assigned greatest weight
 - iv. *Availability*: availability impact (AI) is assigned greatest weight
- (e) *Access complexity (AC)* metric b_{ac} measures the complexity of attack required to exploit the vulnerability once an attacker has access to the target system. The possible scores for this metric are as follows:
- i. *High*: Specialised access conditions exist
 - ii. *Low*: System always exploitable
- (f) *Authentication (Au)* metric b_{au} measures whether or not an attacker needs to be authenticated to the target system in order to exploit the vulnerability. The possible scores for this metric are as follows:
- i. *Required*: Authentication required to exploit the vulnerability
 - ii. *Not Required*: Authentication not required to exploit the vulnerability
- (g) *Access vector (AV)* metric b_{av} measures whether or not the vulnerability is locally or remotely exploitable. The possible scores for this metric are as follows:
- i. *Local*: For local exploitation
 - ii. *Remote*: For remote exploitation
- (h) CVSS quantifies all the above properties and calculates the Base Score as follows:

$$b = 10b_{av}b_{ac}b_{au}((b_{ci}b_{cib}) + (b_{ii}b_{iib}) + (b_{ai}b_{aib})) \quad (\text{C.1})$$

2. Temporal metric group z : These are metrics which give an indication of events that may occur which affect the urgency of the threat posed by the vulnerability. These metrics are as follows:

- (a) *Exploitability* metric z_{ex} “attempts” to measure the current state of exploit technique or code availability and suggests a likelihood of exploitation. This assumes that there are more unskilled attackers than there are attackers who are skilled enough to research vulnerabilities and then create their own version of exploit code. The possible scores for this metric are as follows:
- i. *Unproven*: No exploit code available yet
 - ii. *Proof of Concept*: The code or technique is not functional in all situations and may require substantial hand tuning by a skilled attacker
 - iii. *Functional*: Functional exploit code available
 - iv. *High*: The code works in every situation where the vulnerability is exploitable

- (b) *Remediation Level (RL)* metric z_{rm} gives an indication of the effectiveness of the safeguards put in place. The possible scores for this metric are as follows:
- i. *Official Fix*: A complete vendor solution is available
 - ii. *Temporary Fix*: An temporary official fix is available
 - iii. *Workaround*: An unofficial, non-vendor solution available
 - iv. *Unavailable*: No solution available or the solution is impossible to apply
- (c) *Report Confidence (RC)* metric z_{rc} measures the degree of confidence in the existence of the reported vulnerability and the credibility of the known technical details. The possible scores for this metric are as follows:
- i. *Unconfirmed*: There is little confidence in the validity of the report, e.g. rumours.
 - ii. *Uncorroborated*: Multiple, non-official sources. There may be conflicting reports.
 - iii. *Confirmed*: Vendor of the affected technology has acknowledged that the vulnerability exists.
- (d) This gives the Temporal Score, given by:

$$z = bz_{ex}z_{rm}z_{rc} \tag{C.2}$$

3. The environmental metric group e : The metrics in this group give an indication of the risk posed to different operational environments by a vulnerability. The metrics are as follows:
- (a) *Collateral Damage potential (CD)* metric e_{cd} measures the potential for a loss in physical equipment, property damage or loss of life or limb. The possible scores for this metric are as follows:
- i. *None*: There is no potential for property or physical damage
 - ii. *Low*: There is light property or physical damage if the vulnerability is exploited
 - iii. *Medium*: There is significant property or physical damage if the vulnerability is exploited
 - iv. *High*: There is catastrophic property or physical damage if the vulnerability is exploited
- (b) *Target distribution (TD)* metric e_{td} measures the relative size of the field of target systems susceptible to the vulnerability. The possible scores for this metric are as follows:
- i. *None*: No target systems exist
 - ii. *Low*: Between 1% – 15% of the total environment is at risk.

- iii. *Medium*: Between 16% – 49% of the total environment is at risk.
- iv. *High*: Over 50% of the environment is at risk

(c) The environmental score is given by:

$$e = z + ((10 - z)e_{cd})e_{td} \tag{C.3}$$

CVSS is an empirical approach whose focus is on simplicity. Industry is gradually adopting it [23]. This approach has the advantage that it takes into consideration vulnerability attributes, and uses them to calculate a score for relative comparison. However, **CVSS**'s rough estimates of the number of assets affected by a vulnerability (through the **TD** metric), its course-grained inclusion of asset values and the limited variability of its temporal metrics makes its vulnerability prioritisation less accurate than what we propose in this work.

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when document is classified)

1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's report, or tasking agency, are entered in section 8.) Defence R&D Canada – Ottawa 3701 Carling Avenue, Ottawa ON K1A 0Z4, Canada		2. SECURITY CLASSIFICATION (Overall security classification of the document including special warning terms if applicable.) UNCLASSIFIED	
3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title.) Event prioritisation using a fuzzy risk analysis approach			
4. AUTHORS (Last name, followed by initials – ranks, titles, etc. not to be used.) Dondo, M.			
5. DATE OF PUBLICATION (Month and year of publication of document.) March 2010		6a. NO. OF PAGES (Total containing information. Include Annexes, Appendices, etc.) 64	6b. NO. OF REFS (Total cited in document.) 23
7. DESCRIPTIVE NOTES (The category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.) Technical Memorandum			
8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.) Defence R&D Canada – Ottawa 3701 Carling Avenue, Ottawa ON K1A 0Z4, Canada			
9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.) 15bo02		9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)	
10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.) DRDC Ottawa TM 2009-287		10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)	
11. DOCUMENT AVAILABILITY (Any limitations on further dissemination of the document, other than those imposed by security classification.) (X) Unlimited distribution () Defence departments and defence contractors; further distribution only as approved () Defence departments and Canadian defence contractors; further distribution only as approved () Government departments and agencies; further distribution only as approved () Defence departments; further distribution only as approved () Other (please specify):			
12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11)) is possible, a wider announcement audience may be selected.)			

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

Analysts handle multitudes of computer network security events on a daily basis. They must make an assessment on the potential impact these events have on their organization's assets. As the number of events increases, it becomes increasingly difficult for the analyst to make an assessment as to which events to handle first. This can be resolved by calculating a potential risk metric associated with each event, and then prioritizing the events based on the calculated risk values. Most risk analysis approaches available are based on models which require historical data. In many cases, numerical data related to uncertainty factors about the risk calculations is not available, but the experiential expertise of analysts is. This experiential expertise can be modeled as linguistic variables and functions about an event, and be used to model the risk value associated with each event. In this paper, we present an approach to determine the potential risk value associated with each computer security event by modeling the experiential expertise of analysts through fuzzy linguistic declarations about an event. We then rank these events based on the relative calculated risk values for each. We test our approach on a prototype network using real vulnerability data.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Risk, Event, Vulnerability
Threat, Security, Impact
Safeguards, Fuzzy, Prioritisation

Defence R&D Canada

Canada's leader in Defence
and National Security
Science and Technology

R & D pour la défense Canada

Chef de file au Canada en matière
de science et de technologie pour
la défense et la sécurité nationale



www.drdc-rddc.gc.ca