IMPROVING THE RELEVANCE OF CYBER INCIDENT NOTIFICATION FOR MISSION ASSURANCE

THESIS

Stephen M. Woskov, Captain, USAF

AFIT/GIR/ENV/11-M06

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GIR/ENV/11-M06

IMPROVING THE RELEVANCE OF CYBER INCIDENT NOTIFICATION FOR MISSION ASSURANCE

THESIS

Presented to the Faculty

Department of Systems and Engineering Management

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Information Resource Management

Stephen M. Woskov, BS

Captain, USAF

March 2011

AFIT/GIR/ENV/11-M06

IMPROVING THE RELEVANCE OF CYBER INCIDENT NOTIFCATION FOR MISSION ASSURANCE

Stephen M. Woskov, BS
Captain, USAF

Approved:

_____	_____
Michael R. Grimaila, PhD, CISM, CISSP (Chairman)	date
_____	_____
Robert F. Mills, PhD (Member)	date
_____	_____
Michael W. Haas, PhD (Member)	date

Abstract

Military organizations have embedded Information and Communication Technology (ICT) into their core mission processes as a means to increase operational efficiency, improve decision making quality, and shorten the kill chain. This dependence can place the mission at risk when the loss, corruption, or degradation of the confidentiality, integrity, and/or availability of a critical information resource occurs. Since the accuracy, conciseness, and timeliness of the information used in decision making processes dramatically impacts the quality of command decisions, and hence, the operational mission outcome; the recognition, quantification, and documentation of critical mission-information resource dependencies is essential for the organization to gain a true appreciation of its operational risk. This knowledge provides utility to commanders both during the mission planning phase, as a means to rationally mitigate mission risks, and during the mission execution phase, by providing rapid situational awareness and understanding following a cyber incident.

The objective of this research is to identify existing decision support technologies and evaluate their capabilities as a means for capturing, maintaining and communicating mission-to-information resource dependency information in a timely and relevant manner to assure mission operations. This thesis intends to answer the following research question: Which decision support technology is the best candidate for use in a cyber incident notification system to overcome limitations identified in the existing United States Air Force cyber incident notification process?

Acknowledgements

I would first like to thank my thesis advisor, Dr. Michael Grimaila, for his constant support and valued input during this research endeavor. His mentorship was critical to the success of this work. I would also like to thank my thesis committee, Dr. Robert Mills and Dr. Michael Haas, for their contributions in guiding my thoughts in this research. Furthermore, I would like thank the AFIT faculty that instructed me throughout the Information Resource Management program. They supplied me with the right tools needed to accomplish this thesis.

Additionally, I would like to thank my family and friends for their encouragement throughout this effort. Last but not least, I would like to thank my wonderful wife, whom I had the joy of marrying during my time AFIT. Her unconditional love and support were nothing short of amazing.

Stephen M. Woskov

Table of Contents

Page

Abstract iv

Acknowledgements v

Table of Contents vi

List of Figures viii

List of Tables ix

I. Introduction 1

1.1 Background 1
1.2 Problem Statement 3
1.3 Research Goals 4
1.4 Scope and Assumptions 4
1.5 Cyber Incident Mission Impact Assessment Research 5
1.6 Methodology 5
1.7 Preview 6

II. Literature Review 7

2.1 Importance of Information and Notification 7
2.2 Modeling a Complex Domain 8
2.3 What is Relevance? 12
2.4 Decision Support Systems 15
2.4.1 Rules-Based Systems 15
2.4.2 Case-Based Reasoning 17
2.4.3 Bayesian Networks 20
2.4.4 Neural Networks 23
2.4.5 Hybrid Systems 26

III. Methodology 28

3.1 Methodology Strategy 28
3.2 Developing an Evaluation Table 29
3.3 Content Analysis 33
3.3.1 Sampling 33
3.3.2 Coding 34
3.3.3 Intercoder Reliability 35

IV. Analysis and Results 37

4.1 Introduction 37

Page

4.2 Publications Selected 37
4.3 Content Analysis Discussion 40
4.3.1 Adaptable to Environment 40
4.3.2 Functions with Uncertainty 42
4.3.3 Facilitates Knowledge Acquisition 43
4.3.4 Low Maintainability 46
4.3.5 Provides Information Depth 46
4.3.6 Presents Information Clearly 48
4.3.7 Provides Tangible Information 49
4.4 Content Analysis Results 50
4.5 Design Considerations for Applying CBR 52
4.5.1 Case Representation 52
4.5.2 Case Indexing 58
4.5.3 Knowledge Acquisition 60
4.5.4 Usability 61

V. Conclusions and Recommendations 64

5.1 Conclusions 64
5.2 Limitations 64
5.3 Recommendations for Future Research 65

Appendix A: Coder Training Handout 67

Appendix B: Supporting Text Extracts 73

Appendix C: Coding Results from Second Coder 80

Bibliography 81

List of Figures

Page

Figure 1. Model of Situational Awareness (Endsley, 1995) 9

Figure 2. Approaches to Modeling Cyber Incident Impact (Grimaila, Fortson, & Mills, 2009). 10

Figure 3. Knowledge Engineering (Buchanan, et al., 1983) 16

Figure 4. Example Rule from the MYCIN RBS (Alty, 1985; Shortliffe, 1976) 17

Figure 5. The CBR Cycle (Aamodt & Plaza, 1994) 18

Figure 6. BN for Lung Cancer Example (Korb & Nicholson, 2004) 21

Figure 7. BN Reasoning Tasks (Korb & Nicholson, 2004) 23

Figure 8. Diagram of Artificial Neuron (adopted from Haykin, 1994 and Tarassenko, 1998) ... 24

Figure 9. Single-Layer Feed Forward Network (Schocken & Ariav, 1994) 25

Figure 10. Multi-Layer Feed Forward Network (Drew & Monson, 2000) 25

Figure 11. Multi-Layer Recurrent Network (Pearlmutter, 1989) 26

Figure 12. Example C4 NOTAM 55

Figure 13. Example Case for Cyber Incident Notification 57

Figure 14. Accuracy of RBSs vs. CBR Systems (Chan, 2005) 62

List of Tables

Page

Table 1. User-Defined Relevance Criteria (Barry & Schamber, 1998) 14

Table 2. Nodes and Values for Lung Cancer Example (Korb & Nicholson, 2004) 21

Table 3. Evaluation Table for Ranking DSSs 32

Table 4. Content Analysis Coding Scheme 34

Table 5. Books Used in Content Analysis 38

Table 6. Articles Used for Content Analysis 39

Table 7. Content Analysis Results 50

Table 8. Highest Ranked DSS by Characteristic 51

Table 9. Reliability Results for Two Coders 52

IMPROVING THE RELEVANCE OF CYBER INCIDENT NOTIFICATION FOR MISSION ASSURANCE:

I. Introduction

1.1 Background

Military organizations continue to embed Information and Communication Technology (ICT) into their core mission processes as a means to increase their operational efficiency, exploit automation, reduce response times, improve decision making quality, minimize costs, maximize profit, and shorten the kill chain. This dependence can place mission operations at risk when the loss, corruption, or degradation of the confidentiality, integrity, and/or availability of a critical information resource, system, or infrastructure device occurs. Despite developing a robust security capability, it is inevitable that an organization will experience an information incident. Information incidents can occur for any number of reasons including external attacks, malicious insiders, natural disaster, accidents, and/or equipment failure and can occur within, or external to, the organizational boundary. Regardless of where the incident occurs, it is desirable to notify all organizations whose mission is critically dependent upon the impacted information resource in a timely and relevant manner so they can take appropriate contingency measures to assure their mission operations. Unfortunately, this task can be quite difficult when the dependent information resource is external to the organization, the affected information passes through multiple organizations between the information provider and information consumer, or the information resource is classified (Grimaila, Fortson, & Sutton, 2009; Grimaila, Schechtman, & Mills, 2009).

Since the accuracy, conciseness, and timeliness of the information used in decision making processes dramatically impacts the quality of command decisions, and hence, the

operational mission outcome; the recognition, quantification, and documentation of critical mission-information resource dependencies is essential for the organization to gain a true appreciation of its operational risk. This knowledge provides utility to commanders both during the mission planning phase, as a means to rationally mitigate mission risks, and during the mission execution phase, by providing rapid situational awareness and understanding following a cyber incident. In military contexts, the failure to understand and appreciate the relationship between mission objectives and the underlying ICT resources can have dire consequences including physical destruction and the loss of life when a cyber incident occurs. To reduce the likelihood of this outcome, personnel must maintain real-time awareness of how resources that are critical to their mission's success are affected by a cyber incident. When an information incident occurs, it is important to notify and inform decision makers within organizations whose mission is critically dependent upon the affected information in a timely and relevant manner so they can take appropriate contingency measures (Grimaila, Fortson, & Mills, 2009; Grimaila, Fortson, & Sutton, 2009).

Unfortunately, the existing incident notification process within the United States Air Force (USAF) has several limitations which severely limit the usefulness of incident notification. Specifically, it was determined that 1) notifications may fail to identify all affected parties due to a focus on physical systems rather than information; 2) notifications may fail to reach the organization's decision makers who can take the proper contingency actions; 3) the incident response process lacks automated delivery methods, which creates a delay during dissemination; 4) notifications may become irrelevant as an organization's mission or resource dependencies change over time; and 5) the process does not allow organizations to communicate the criticality

of their resource dependencies to external entities (Grimaila, Schechtman, et al., 2009). These deficiencies place organizational missions at risk and motivate the need for this research.

1.2 Problem Statement

Military organizations are increasingly dependent on information, information systems, and the cyber infrastructure to achieve their mission goals and objectives. Research has shown that existing cyber incident notification processes within the USAF do not provide the capability to supply actionable information to contingency decision makers' need for mission assurance. There is a critical need to develop a decentralized, scalable incident notification system that enables personnel to capture and maintain knowledge of the potential mission impacts resulting from the loss or degradation of an information resource. The system would provide the ability for personnel to document scenarios which may place the organizational mission at risk. The collection of scenarios would serve to document the organization's understanding of mission risk as a function of the underlying cyber resources; provide mission planners with a view of cyber resources in terms of their mission criticality; and provide a framework to communicating potential mission impacts to the users in a timely and relevant manner following a cyber incident. There are several decision support technologies in existence that have been used in knowledge retention and retrieval tasks across a broad spectrum of application environments. This thesis intends to answer the following research question: Which decision support technology is the best candidate for use in a cyber incident notification system to overcome limitations identified in the existing USAF cyber incident notification process?

1.3 Research Goals

The goals for this research are:

- Identify characteristics that are desirable to improve the timeliness and relevance of a cyber incident notification system.
- Survey decision support technologies and identify potential candidates for use in a cyber incident notification system.
- Evaluate potential decision support technologies using the desired characteristics to identify the best technology to use in a cyber incident notification system.
- Establish initial design considerations and demonstrate the feasibility of the selected decision support technology using a fictional example scenario based upon a real world military unit.

1.4 Scope and Assumptions

This thesis represents the initial research toward applying decision support technology to improve cyber incident notification within the USAF, and is therefore an exploratory study. Before an incident notification system can be built, a preliminary direction must be established as to what reasoning method is appropriate for the domain. This research intends to explore that initial direction and, as a result, set the stage for future progress in engineering and implementing a system for operational use.

Additionally, this research focuses on effectively communicating the potential mission impact resulting from a cyber attack to affected mission operators in a relevant manner. However, the proposed decision support system may also provide the ability to inform decision makers about non-cyber losses, such as equipment and personnel, when dependency status

information is available. While the focus will be upon cyber dependencies, the ability to account for other mission dependencies will be considered.

Finally, this thesis centers on identifying, collecting, and organizing mission-information resource dependencies solely at low-level organizations within the USAF. As a result, an "organization" is defined as a typical USAF base-level squadron, which is assumed to have approximately 100-300 personnel, and rely on a multiple internal and external information systems to achieve its mission objectives.

1.5 Cyber Incident Mission Impact Assessment Research

The primary objective of this research is to contribute to the goals and objectives of the Cyber Incident Mission Impact Assessment (CIMIA) program by investigating technologies that can be applied to improve the timeliness and relevance of cyber incident notification (Grimaila, Fortson, & Mills, 2009; Grimaila, Fortson, & Sutton, 2009; Grimaila, Schechtman, et al., 2009). However, it is important to distinguish the focus of this thesis from other research within the CIMIA project. Specifically, this work places an emphasis on relevance. A separate project is developing an architecture to improve the timeliness limitation. Additionally, there is a third study performing an experiment to determine whether a trial cyber incident notification system can improve decision making quality when compared to current processes. This thesis extends upon that research by helping shape the notification system for future testing and implementation.

1.6 Methodology

This exploratory research provides a feasibility study of, and initial support for, the development of a decision support system to improve the relevance of cyber incident notifications to assure organization mission operations. To achieve the research goals, this thesis

draws upon existing literature to determine: 1) ideal features that a cyber incident notification system should possess, and 2) potential technologies that could improve the existing process. The ideal system features are used to evaluate the technologies and determine the most promising option via a content analysis. Once the most suitable technology has been selected, initial design considerations are established for its application to the domain of cyber incident notification. Finally, an example scenario demonstrates how the design features are exploited to provide relevant notification to organization personnel.

1.7 Preview

This thesis is organized into five chapters. This chapter presented background material and an introduction to the research goals. Chapter II provides a literature review on fundamental aspects about the military domain and the concept of relevance in order to establish ideal criteria for an incident notification system. Also, this chapter contains a review of, and background for, decision support systems. Chapter III presents a methodology for isolating the most suitable technology for improving relevant notification. In Chapter IV, the most ideal technology is selected and design considerations are established to mold it to the cyber domain. Also, this chapter contains a fictional example scenario that shows the feasibility of the selected technology. Finally, Chapter V presents a conclusion and a discussion of future research areas.

II. Literature Review

2.1 Importance of Information and Notification

Information, and subsequently knowledge, has become an important asset within modern organizations (Davenport & Prusak, 1998; Denning, 1999; Pipkin, 2000). This information age has brought increased reliance upon ICT to increase operational efficiency, exploit automation, reduce response times, improve decision making quality, minimize costs, maximize profit, and shorten the kill chain. Unfortunately, this dependence can place an organization's mission at risk if the confidentiality, integrity, or availability of the information needed from these systems, or the cyber resources used to store, process, transport, or disseminate the information, have been lost or degraded. This concern generates the need for operations personnel to be aware of how cyber incidents affect their organization's mission (Grimaila, Fortson, & Sutton, 2009; Grimaila, Schechtman, et al., 2009).

Currently, the USAF uses Time Compliance Network Orders (TCNOs) and Command, Control, Communications, and Computers Notices to Airmen (C4 NOTAMs) as the primary methods to notify organizations of cyber incidents (Department of the Air Force, 2005). However, the present process has several limitations: 1) notifications may fail to identify all affected parties due to a focus on physical systems rather than information; 2) notifications may fail to reach the organization's decision makers who can take the proper contingency actions; 3) the process lacks automated delivery methods, which creates a delay during dissemination; 4) notifications may become irrelevant as an organization's mission or resource dependencies change over time; and 5) the process does not allow organizations to communicate the criticality of their resource dependencies to external entities (Grimaila, Schechtman, et al., 2009). From these limitations, two main areas have been identified as crucial to the improvement of the

current cyber incident notification process: timeliness and relevance. Timeliness refers to reducing the time between an incident and the notification of the appropriate decision maker, while relevance focuses on enhancing the usefulness of a notification (Grimaila, Fortson, & Sutton, 2009). This thesis will primarily focus upon improving the relevance of incident notification.

2.2 Modeling a Complex Domain

Before examining the concept of relevance, it is essential to have an understanding about the domain of interest: the military environment. There are some fundamental distinctions between military operations and non-military operations. One of these differences lies in the criticality of decision making. While most organizations experience loss in terms of dollars, poor decision making in the military domain can also result in physical destruction and loss of life (Grimaila & Fortson, 2008; Grimaila, Fortson, & Sutton, 2009). These severe consequences demand that a cyber incident notification system take into account some key attributes that are intrinsic to military operations.

First, the military environment is dynamic in nature (Department of Defense, 2006; Department of the Air Force, 2003; Leonhard, 1998). This aspect creates the need to continually update resource dependencies to reflect current operational objectives (Grimaila, Fortson, & Sutton, 2009). Having accurate knowledge about resource dependencies is fundamental for maintaining situational awareness. Endsley (1988) defines situational awareness (SA) as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future" (p. 97). This definition reduces to "knowing what is going on around you" (Endsley & Garland, 2000, p. 5). Endsley (1995) identifies SA as a precursor to decision making, and ultimately the

performance of actions (see Figure 1). Therefore, a cyber incident notification system must have the ability to adapt to the changing military environment to enhance SA.

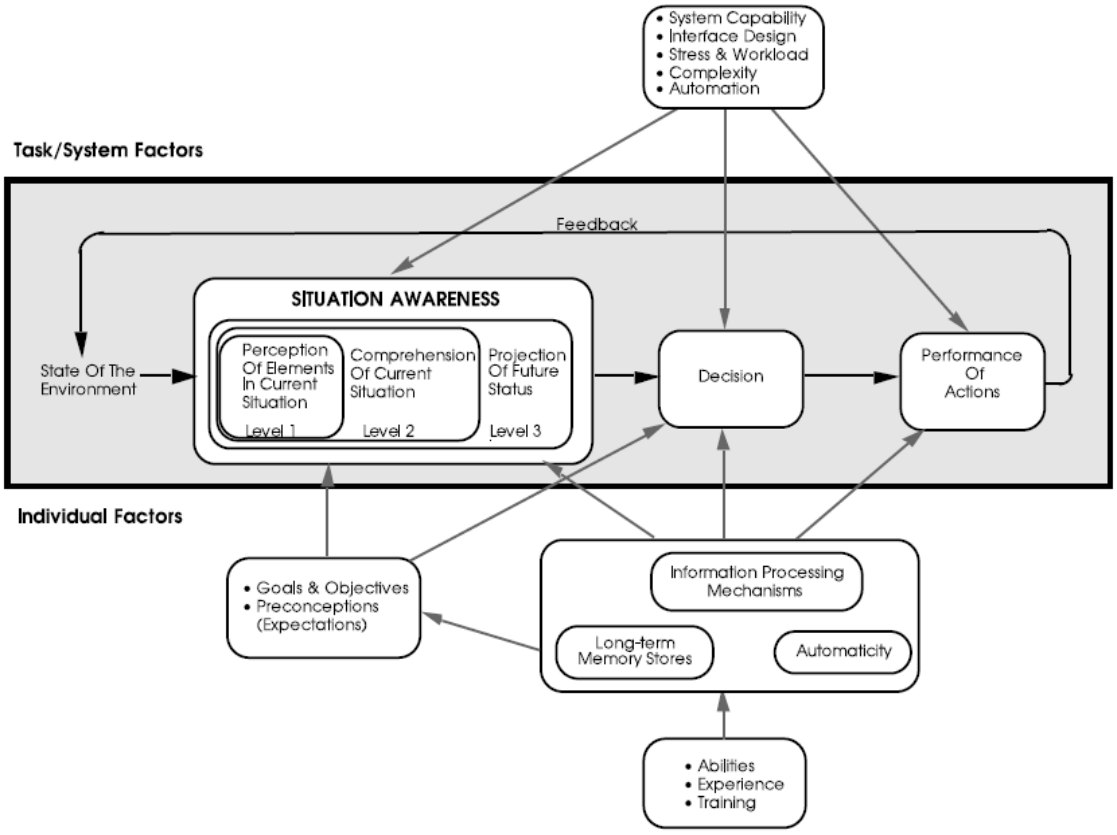


Figure 1. Model of Situational Awareness (Endsley, 1995)

Second, warfare is inherently uncertain and unpredictable. This aspect is sometimes called the "fog of war" (Alberts, Garstka, Hayes, & Signori, 2001; Clausewitz, 1976; Department of the Air Force, 2003). As a result, military commanders are often limited to what information they can use for decision making. *Joint Publication 3-13* explains that "decisions are made based on the information available at the time" (p. I-8). Consequently, it is important that a cyber incident notification system can provide benefit to a decision maker even when information is missing, incomplete, or uncertain.

The dynamic and uncertainty aspects of military operations have led to some ideas on how the mission impact resulting from cyber incidents can be modeled. Figure 2 shown below demonstrates two approaches for comparison (Grimaila, Fortson, & Mills, 2009). The approach on the left is only suitable for static environments because it uses a rigid modeling technique, such as enterprise architecture (EA) (Department of Defense, 2009; Wong-Jiru, Colombi, Suzuki, & Mills, 2007). In contrast, the approach on the right seeks to take advantage of more adaptive methods via a decision support system that can be populated with a variety of information collected from subject matter experts (SMEs), historical mission impacts, and explicit mission models to identify and estimate the value of critical mission-information dependencies. The approaches differ in the effort required to construct and maintain the knowledge base, the accuracy of the assessment, the ability to adapt to change, and the ability to account for uncertainty. This thesis focuses on the latter approach and aims to advance it by investigating the design of the decision support system.

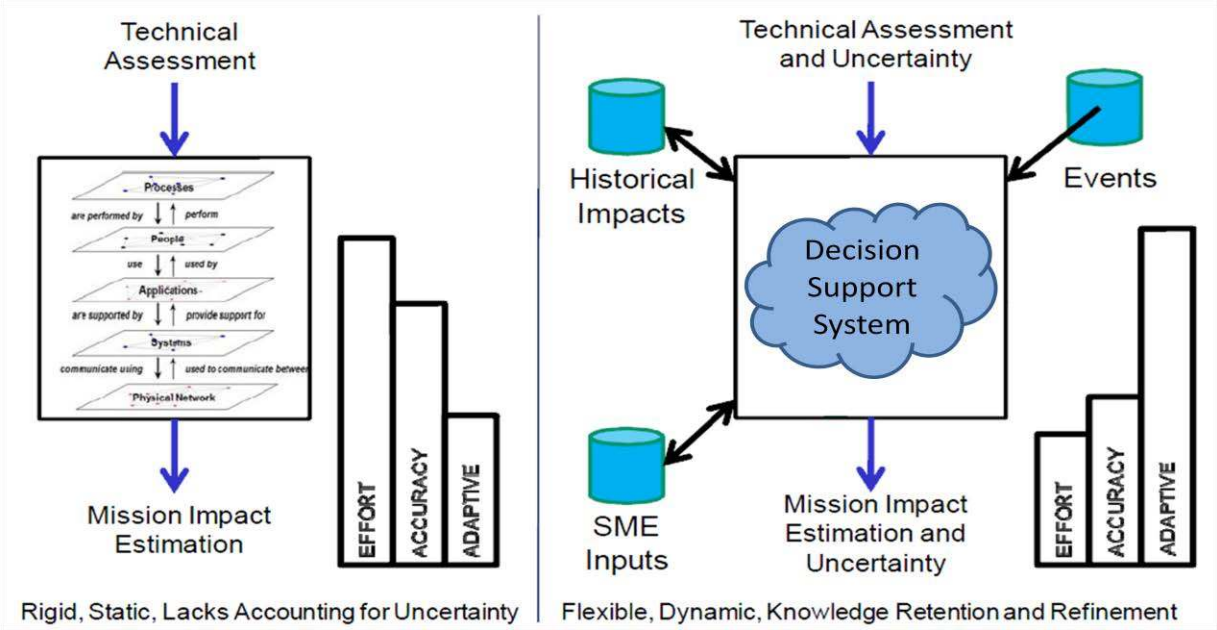


Figure 2. Approaches to Modeling Cyber Incident Impact (Grimaila, Fortson, & Mills, 2009)

Figure 2 above also helps point out another fundamental reason for choosing the right side approach, which has to do with how problems are solved. In the left side approach, the domain is modeled as accurately as possible, primarily using quantitative methods, with the intent of supplying a user with exact solutions. This task is tricky in the military environment where there are many resources to account for. The difficulty associated with modeling large domains is referred to as the "state space explosion" problem (Demri, Laroussinie, & Schnoebelen, 2006; Gemikonakli, Ever, & Kocyigit, 2009). Instead, the right side approach focuses on providing lower solution accuracy, primarily using qualitative methods, but requires less effort to build and is more adaptive. As the environment becomes better understood, quantitative metrics can be added to improve accuracy. Overall, this method appears to fit the complexity of the domain more appropriately (Grimaila, Fortson, & Mills, 2009).

Inexact solutions, but ones that are helpful for problem solving, are known in the artificial intelligence (AI) literature as *heuristics*. Heuristics have been defined in many ways; however, Romanycia and Pelletier (1985) attempt to synthesize all these views into one broad meaning: "A heuristic in AI is any device, be it a program, rule, piece of knowledge, etc., which one is not entirely confident will be useful in providing a **practical** solution, but which one has reason to believe will be useful, and which is added to a problem-solving system in expectation that on average the performance will improve" (p. 57). Similarly, the use of stories, or narratives, play an important role in communication and decision making (Gargiulo, 2006; Pennington & Hastie, 1988). In the context of a court room, Pennington and Hastie (1988) showed that "evidence, in the form of stories, play a causal role in determining verdict decisions" (p. 531). As a result, this thesis accepts heuristics and stories as an effective method for collecting and communicating knowledge about cyber incidents when quantitative metrics are unavailable.

2.3 What is Relevance?

With an understanding of the domain established, the concept of relevance must be explored. The study of relevance is most prominent within the field of information science, specifically information retrieval (IR) (Borlund, 2003; Brouard & Nie, 2004; Cosijn & Ingwersen, 2000; Harter, 1992; Park, 1993; Saracevic, 2007; Schamber, Eisenberg, & Nilan, 1990; Swanson, 1986). Borlund (2003) states that the objective of IR is the "retrieval of *relevant* information" (p. 913). However, there is no universally accepted definition of relevance and a number of perspectives exist (Mizzaro, 1997; Schamber, et al., 1990). Perhaps a good fundamental definition is best articulated by Saracevic (1975): "Relevance is considered as a measure of the effectiveness of a contact between a source and a destination in a communication process" (p. 321). In a military context, a more specific meaning arises. Bass and Baldwin (2007) define relevance as "a measure of applicability to a purpose or a customer" (p. 105). By combining these views, the goal of this thesis can be thought of as improving the *effectiveness* and *applicability* of cyber incident notifications.

In general, all definitions of relevance can be grouped into two main categories: objective and subjective relevance (Borlund, 2003; Harter, 1992; Saracevic, 1975; Swanson, 1986). Objective relevance primarily deals with how well a topic search returns results that deal with that topic. In this view, relevance is dependent upon a query and the search algorithm of the information system being used. Consequently, this concept is also referred to as system-oriented relevance because the role of the user is neglected (Barry, 1994; Schamber, et al., 1990). However, Cuadra et al. (1967) explain that relevance is not a meaningful concept "as long as it is construed and used only as a relation between strings of written words independent of a judging process" (p. 23). This "judging process" is the main factor in subjective relevance, which

focuses on how a user perceives the effectiveness of information. Thus, this concept is also referred to as user-oriented relevance (Schamber, et al., 1990). This view has gained more interest due to the realization that end users are the ones who decide whether retrieved information is useful (Barry, 1994). This thesis takes the position that subjective relevance is the most important type for achieving effective cyber incident notification.

Because subjective relevance is dependent on the user's perspective, it is much harder to determine appropriate measures. However, there has been some progress toward establishing a core set of determinants. One such study was performed by Schamber (1991), who interviewed users of weather information in three different fields: aviation, electric power utilities, and construction. The respondents were asked about a particular situation when weather information was needed in their job. For these situations, the subjects explained which information source they used and what features of the source were useful. As a result, Schamber formed a list of user-defined relevance criteria from the interview data.

In another subjective relevance study, Barry (1994) elicited participants in a academic environment with a stated research purpose (e.g. class assignments, masters theses, etc.) to view a set of document representations, and some full-text versions, related to their individual needs. For each document, the subjects circled anything that would incline them to pursue or not pursue that reference. They were then asked about their choices in an open-ended interview. This study also produced a set of user-defined relevance criteria.

Barry and Schamber (1998) combined their results from the two studies mentioned above to produce a list of 10 criteria that were common between their findings (see Table 1). The significance of this research lies in the result that certain relevance criteria overlapped even though the individual experiments were performed in different environments. This conclusion

provides support for the existence of relevance criteria that are important in any context. Therefore, in addition to the military domain aspects mentioned in previous section, Table 1 provides a good foundation for establishing a set of desired characteristics for a cyber incident notification system as proposed by the first research goal.

Table 1. User-Defined Relevance Criteria (Barry & Schamber, 1998)

Depth/Scope/Specificity	The extent to which information is in-depth or focused; is specific to the user's needs; has sufficient detail or depth; provides a summary, interpretation, or explanation; provides a sufficient variety or volume
Accuracy/Validity	The extent to which information is accurate, correct or valid
Clarity	The extent to which information is presented in a clear and well-organized manner
Currency	The extent to which information is current, recent, timely, up-to-date
Tangibility	The extent to which information relates to real, tangible issues; definite, proven information is provided; hard data or actual numbers are provided
Quality of Sources	The extent to which general standards of quality or specific qualities can be assumed based on the source providing the information; source is reputable, trusted, expert
Accessibility	The extent to which some effort is required to obtain information; some cost is required to obtain information
Availability of Information/Sources of Information	The extent to which information or sources of information are available
Verification	The extent to which information is consistent with or supported by other information within the field; the extent to which the user agrees with information presented or the information presented supports the user's point of view
Affectiveness	The extent to which the user exhibits an affective or emotional response to information or sources of information; information or sources of information provide the user with pleasure, enjoyment or entertainment.

2.4 Decision Support Systems

A central part of this thesis involves identifying available technologies that have the potential to improve cyber incident notification. The technologies considered in this research are categorized as decision support systems (DSSs). A DSS is defined by Carter et al. (1992) as "an interactive IT-based system that helps decisionmakers utilize data and models in making their decisions" (p. 3). Another closely related field is expert systems (ESs). ESs are computers that embody human expertise and allow users to call upon that stored knowledge for advice, conclusions, and/or explanations about a specific domain (Shu-Hsien, 2005). Due to their similarities, the terms DSS and ES can be used interchangeably.

This section presents a review of, and background for, four types of DSSs with the potential to help improve relevant notification: rules-based systems (RBSs), case-based reasoning (CBR) systems, Bayesian networks (BNs), and neural networks (NNs). These four technologies were selected among others based on their prominence in literature dealing with decision support. By reviewing these DSSs, the second research goal was achieved.

2.4.1 Rules-Based Systems

RBSs were a popular technology for decision support in the 1970's and 80's; however, they are still showing potential in more current applications (Hayes-Roth, Waterman, & Lenat, 1983; Michie, 1982; Shu-Hsien, 2005). These systems use rules in the form of "IF-THEN" statements to reason, where "IF" is a condition and "THEN" is an action. There may be multiple conditions for an action to take place, for example, "*if* X and Y and Z, *then* deduce A" (Bramer, 1982, p. 5). Reasoning is accomplished using one of two methods, forward- or backward-chaining. In forward-chaining, all rule conditions are assessed first to determine whether an

action is triggered. In contrast, backward-chaining assumes all rule actions are initially true and identifies the applicable action by determining which conditions have been met (Alty, 1985).

In general, a RBS is made up of two components, a knowledge base and an inference engine. The knowledge base is where the rules are stored. The inference engine is the mechanism for interpreting the rules and producing results based on user input. There is an important distinction between these two elements that is highlighted by Quinlan (1982): "The power of the system does not come principally from this knowledge application mechanism (the *inference engine*) but from the richness, pertinence and redundancy of the knowledge itself" (p. 34). This characteristic emphasizes an important theme that must be embraced when building any type of ES. The process of obtaining and transferring knowledge (also called knowledge acquisition, knowledge engineering, or knowledge elicitation) from a human to a system is critical to the usefulness of that system (see Figure 3) (Buchanan, et al., 1983; Cooke & McDonald, 1986).

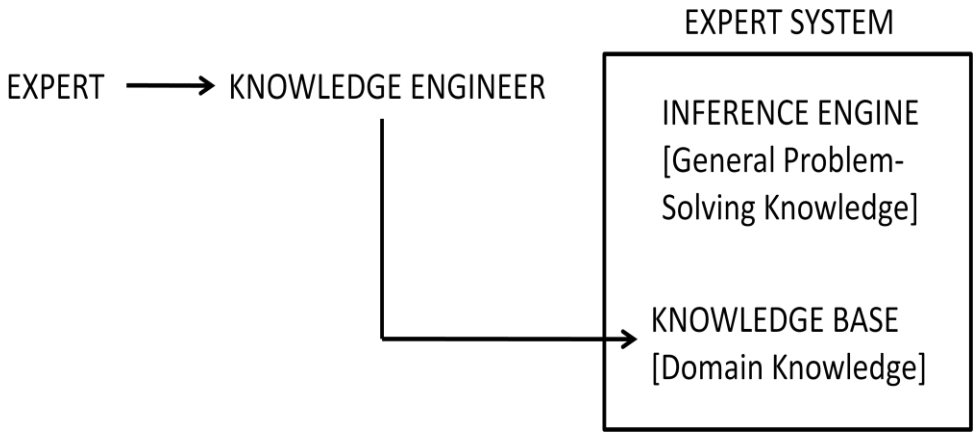


Figure 3. Knowledge Engineering (Buchanan, et al., 1983)

One of the most prominent early RBSs was MYCIN, which was used to aid physicians in diagnosing and treating bacterial infections (Alty, 1985). An example rule from MYCIN is

shown in Figure 4. In addition to medical diagnosis, other early applications of RBSs included helping geologists assess mineral sites (i.e. PROSPECTOR) and assisting technicians with configuring computer systems (i.e. R1) (Bramer, 1982; McDermott, 1982). Current applications include fault diagnosis, management fraud assessment, and tutoring (Shu-Hsien, 2005).

IF the infection is primary-bacteremia
AND the site of the culture is one of the sterile sites
AND the suspected portal of entry is the gastro-intestinal tract
THEN there is suggestive evidence (.07) that the identity of organism is bacteriodes

Figure 4. Example Rule from the MYCIN RBS (Alty, 1985; Shortliffe, 1976)

2.4.2 Case-Based Reasoning

The concept behind CBR is summarized by Riesbeck and Schank (1989): "A case-based reasoner solves new problems by adapting solutions that were used to solve old problems" (p. 25). This logic is founded on three underlying assumptions listed by Watson (2003): 1) CBR assumes that the world is regular; what holds true today will most likely be true tomorrow, 2) CBR anticipates that events will repeat because it is the sole reason they are remembered, and finally 3) similar problems have similar solutions. Overall, CBR shares similarities with RBSs. Kolodner (1993) explains that "we can think of case-based reasoning as a type of rule-based reasoning in which the rules are very large, the antecedents need to be only partially matched, and the consequents need to be adapted before they are applied" (p. 93).

A unique aspect about CBR is that it relies on specific knowledge from past events, instead of generalized relationships about a specific domain. Additionally, CBR is an approach that allows incremental learning. Once a new case is added to its library, it can be retrieved in the future (Aamodt & Plaza, 1994). While commonly labeled as a technology, CBR is actually a methodology for problem solving (Watson, 1999). Researchers suggest that people cognitively use CBR on a daily basis (Kolodner, 1992).

A few models have been developed to explain the CBR process; however, the most popular one was established by Aamodt & Plaza (1994) (see Figure 5). In their model, the CBR processes are described by the four REs: RETRIEVE, REUSE, REVISE, and RETAIN. A problem is solved by *retrieving* a past case, *reusing* the previous case in some way, *revising* the solution after using it, and finally *retaining* the new experience in the case-base (i.e. the knowledge repository) by either adding the new case or updating existing cases.

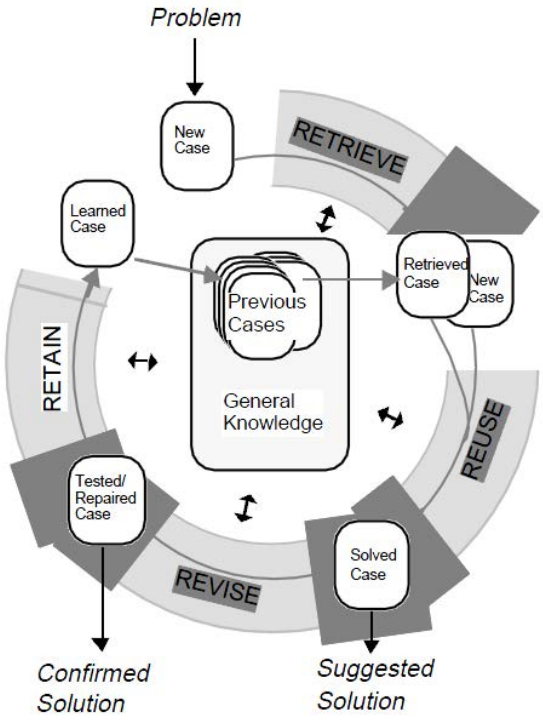


Figure 5. The CBR Cycle (Aamodt & Plaza, 1994)

This concept is best served with an example. Suppose a banker has to make a decision on whether to grant a loan to a client. The banker does not want to lend money to a person unable to pay it back, but a client should not be turned down pointlessly; the bank will make money on the loan's interest. In order to make the appropriate decision, the banker could use his or her knowledge from previous loans and apply it to the new situation. If the new client meets criteria similar to others with successful repayment in the past, the banker would grant the loan. However, if the client meets criteria similar to others who have failed to repay, the banker would not grant the loan (Watson, 1997).

CBR has been successfully implemented in a number of different areas including law, management, health sciences, planning, and technical support (Ashley, 1991; Hammond, 1986; Koton, 1988; Watson, 1997, 2003). The potential benefits of using CBR in a military context have already been recognized. One of the earliest studies in this area was performed by Goodman (1989), who developed a decision support aid for battle planning. By taking advantage of an existing database containing historical land battles, this CBR system retrieved past conflicts most similar to a present operation based on user input. In more current research, Jakobson et al. (2004) discuss CBR's potential to aid in battlespace management. Their work particularly focuses on the usefulness of CBR within the dynamic environment of military operations. Finally, Weber and Aha (2002) used CBR as a framework to design a Lessons Learned System (LLS). Lessons learned are past successes or failures that are pertinent to tasks within an organization. In their work, they combine an LLS with a DSS used for military mission planning. While using the DSS, the LLS automatically notifies a user when there is a lesson applicable to the part of the plan that he or she is working on. This design allows lessons to be delivered *when* and *where* they are needed.

2.4.3 Bayesian Networks

Bayesian networks use probability theory to reason about problems in uncertain environments (Jensen & Nielsen, 2007; Korb & Nicholson, 2004; Pearl, 1988; Russell & Norvig, 2003). The term Bayesian network (BN) will be used throughout this thesis; however, other names include belief network, probabilistic network, casual network, and knowledge map. BNs are defined by four characteristics (Russell & Norvig, 2003):

1. A BN is made up of nodes that represent random variables within a domain. These variables can either be discrete or continuous.
2. The nodes are connected to each other by links, called arcs, which signify direct dependencies. An arc from node X to node Y signifies that X is the *parent* of Y. This relationship also means that Y is the *child* of X.
3. Each node in the network has a conditional probability distribution in the form of $P(X_i|Parents(X_i))$, which means that the probability of node X_i is dependent on the parents of X_i.
4. One cannot return to a node by following arcs (i.e. a directed cycle). This property categorizes a BN as a directed acyclic graph, or DAG.

An example BN will help illustrate these characteristics and how they are used for reasoning. First consider the nodes and their possible values shown in Table 2. This information is paired with Figure 6 to show how a doctor may determine whether a patient has lung cancer.

Table 2. Nodes and Values for Lung Cancer Example (Korb & Nicholson, 2004)

Cancer	{*Tru*
Dyspnoea	{*Tru*
X-Ray	{*Pos*

Figure 6. BN for Lung Cancer Example (Korb & Nicholson, 2004)

Suppose a patient's chance of having *Cancer* is due to two factors: *Pollution* and *Smoker*. Also, the presence or non-presence of *Cancer* will in turn have an effect on the outcome of the patient's *XRay* result, as well as their chance of having shortness of breath (i.e. *Dyspnoea*). The nodes *Pollution* and *Smoker* are called the root nodes because they do not have any parents. Additionally, the nodes *XRay* and *Dyspnoea* are called the leaf nodes because they do not have

any children. Each node is associated with a conditional probability table (CPT), which displays all of the possible probabilities for that node given the state of its parents. For example, the CPT for *Cancer* shows four probabilities a patient has cancer given his or her level of pollution exposure and smoking status (Korb & Nicholson, 2004).

This illustration helps highlight another important concept related to BNs. First, using BNs necessitates the assumption of the Markov property, which means that no dependencies are present other than the ones already established using arcs. In the example, smoking cannot directly cause *Dyspnoea*. *Dyspnoea* is *only* influenced by the presence or non-presence of *Cancer*. As an extension of this property, one can determine whether nodes are conditionally independent. In the example, *Pollution* and *Smoker* are said to be conditionally independent from *XRay* and *Dysponea*. This determination means that if *Cancer* is present, the knowledge of a patient's pollution exposure or smoking status will have no impact on his or her *XRay* result or chance of having *Dsypnoea* (Korb & Nicholson, 2004).

BNs support reasoning tasks such as diagnosis, prediction, intercausal (i.e. "explaining away"), and a combination of these (see Figure 7). The first two types are different only in their direction of inference. Diagnostic reasoning starts with evidence of symptoms and infers their cause. In contrast, predictive reasoning uses evidence from causes to infer possible outcomes. Intercausal reasoning, or "explaining away", can occur because of conditional independence. In the model, *Cancer* is only caused by *Pollution* and *Smoker*. If cancer is detected and the patient is also known to be a smoker, this information lowers the probability that the patient has been exposed to high levels of pollution. In other words, *Pollution* has been "explained away" as a possible cause (Korb & Nicholson, 2004).

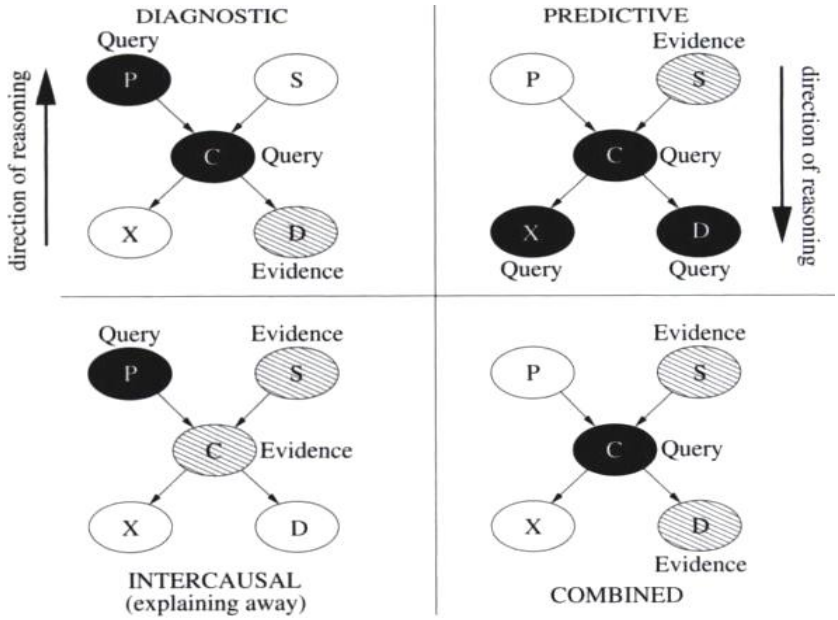


Figure 7. BN Reasoning Tasks (Korb & Nicholson, 2004)

The application of BNs to risk analysis has been increasing in recent years (Weber et al., 2010). Hudson et al. (2001) use a BN to develop software called Site Profiler, which aids military planners in performing antiterrorism risk management. Falzon (2006) developed the Centre of Gravity Network Effects Tool (COGNET). COGNET uses a BN to help the military determine the impact of an enemy's center of gravity if certain vulnerabilities are compromised.

2.4.4 Neural Networks

The study of the human brain has motivated the development of neural networks (NNs), also commonly referred to as artificial neural networks (ANNs) or connectionist networks (Drew & Monson, 2000; Gallant, 1993; Haykin, 1994; Russell & Norvig, 2003; Schalkoff, 1997; Tarassenko, 1998). Haykin (1994) defines a NN as "a machine that is designed to model the way in which the brain performs a particular task or function of interest…" (p. 2).

Neurons are cells in the brain that produce electrical signals and are thought to allow information-processing capability. Thus, ANNs aim to model this behavior (Russell & Norvig,

2003). A neuron in an ANN receives an input and computes an output using an activation function. In addition to the input, there is a synaptic weight that influences the output of the neuron (Gallant, 1993). An artificial neuron is shown in Figure 8.

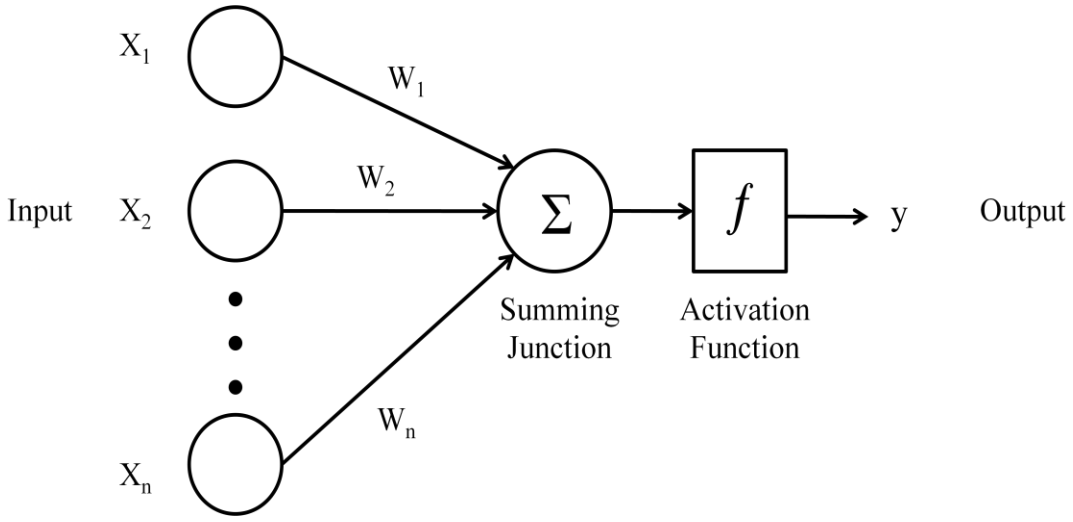


Figure 8. Diagram of Artificial Neuron (adopted from Haykin, 1994 and Tarassenko, 1998)

The output is determined by the following equation:

where y is the neuron output, f is the activation function, w_i is the synaptic weight associated with the links, and x_i is the input signal (Tarassenko, 1998).

While there are different types of NNs, the most common is a feed-forward network. In this structure, the neuron output only goes in one direction (see Figure 9 and Figure 10). In contrast, recurrent networks distinguish themselves by having feedback loops, which means that neuron outputs can cycle back as inputs to previous neurons (see Figure 11). Additionally, these networks can consist of multiple layers. In a single layer network, the inputs are connected directly to the outputs. In a multi-layer network, the input neurons feed into hidden layers before

reaching the output. Adding layers increases the reasoning ability; however, selecting the right amount of hidden neurons is not fully understood (Russell & Norvig, 2003).

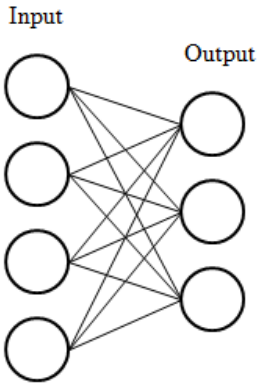


Figure 9. Single-Layer Feed Forward Network (Schocken & Ariav, 1994)

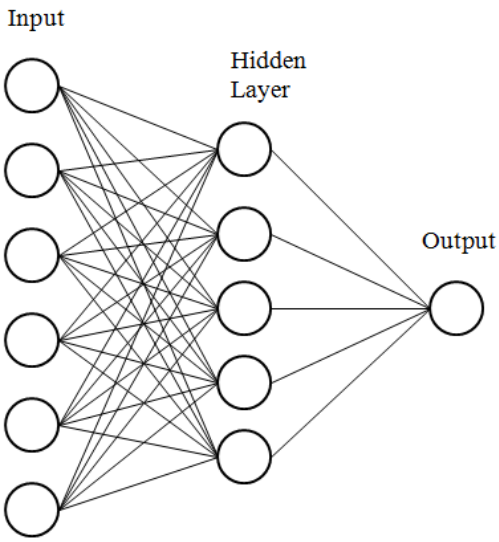


Figure 10. Multi-Layer Feed Forward Network (Drew & Monson, 2000)

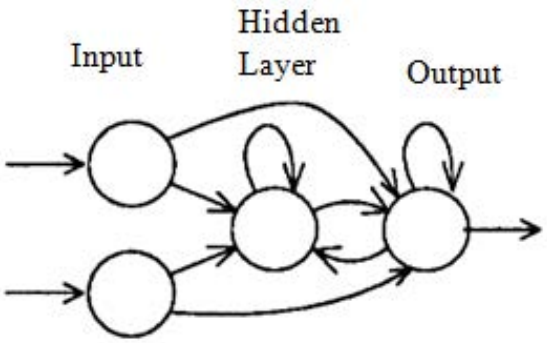


Figure 11. Multi-Layer Recurrent Network (Pearlmutter, 1989)

NNs have several key qualities that are identified by Tarassenko (1998):

1. NNs can learn from experience, which makes them appropriate for solving problems that are not fully understood. However, a lot of training examples are required for learning.
2. By generalizing from training examples, NNs can accurately solve problems that have not been encountered before.
3. NNs can be built faster due to less reliance on domain expert input. Although, expertise is important for constructing an ideal NN design.
4. NNs can be computationally fast and efficient because of parallel processing.
5. NNs can provide non-linear solutions, which again makes them suitable for solving complex problems.

NNs have been used in a number of applications which include classification, image processing, pattern recognition, risk assessment, and medicine and engineering diagnostics (Drew & Monson, 2000; Russell & Norvig, 2003; Schalkoff, 1997).

2.4.5 Hybrid Systems

The DSSs above have been reviewed independently; however, applications usually involve a combination of reasoning methods to enhance problem-solving capability. There are

many different ways to merge these techniques. One common approach is the fusion of RBSs and CBR systems. For example, Kumar, Singh, and Sanyal (2009) combine these two methods to enhance decision support in an intensive care unit. Additionally, Tung et al. (2010) present a rules-based CBR system to enhance case retrieval speed.

There have been many examples of other hybrid systems as well. Hajmeer and Basheer (2003) combine NNs and BNs to model bacteria growth. Yang, Han, and Kim (2004) integrate CBR and NNs to improve fault diagnosis. Finally, Hatzilygerdoudis and Prentzas (2004) propose the combination of three methods: rules, NNs, and CBR.

While hybrid systems are a common approach for decision support, this thesis focuses on selecting a single method as the most suitable framework for incident notification. It is inevitable that a hybrid system will ultimately be the end result; however, it is essential to determine one key reasoning technique as the foundation for the development of a cyber incident notification system.

III. Methodology

3.1 Methodology Strategy

This thesis represents the initial research toward applying decision support technology to improve cyber incident notification within the USAF. As such, this thesis is primarily an exploratory study. Exploratory research is defined by Neuman (2006) as "research in which the primary purpose is to examine a little understood issue or phenomenon to develop preliminary ideas...." (p. 33). An important first step in the pursuit of building and implementing a DSS for cyber incident notification requires some direction on which technologies are appropriate for the domain. This research intends to explore that initial direction.

To achieve this objective, each of the technologies outlined in the previous chapter were ranked based upon a list of desired characteristics. The first part of this methodology discusses the characteristics that were used for the ranking, and ultimately results in the creation of an evaluation table. Once established, the second part of the methodology describes *how* the table was coded. This task was completed using a content analysis method. Content analysis is described by Krippendorff (2004) as "… a research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use" (p. 18). As this definition states, a context must be established. The context for the analysis, as well as the development of the evaluation table, is stated by the following scenario: *A cyber incident has just occurred at base X that could potentially affect flying operations. The operations squadron commander needs to know that a cyber incident has occurred, how it affects his or her organization's mission, and what courses of action can be taken to avoid detrimental consequences. What tool could perform these actions and ultimately allow the commander to make the best decisions about the execution of operations in the post-incident environment?*

3.2 Developing an Evaluation Table

To determine which DSS is most suitable for incident notification, there must be a means to compare them. Because this thesis is primarily an exploratory study, there were no pre-defined metrics to use. Instead, the literature review in the previous chapter served as the foundation for establishing the criteria. This section will describe which metrics were picked and why. As a result, an evaluation table was created to rank and compare the technologies (see Ladd, Datta, & Sarker, 2010).

One could easily provide a long list of criteria for evaluating DSSs in the cyber incident notification domain. However, the objective of this section was to produce a concise group of the most important factors. First, the beginning of the literature review highlighted some key aspects about the military environment. It was established that military operations are dynamic. Therefore it is important that a technology is adaptable. Next, it was discussed that commanders may need to make decisions based on missing or uncertain data. Therefore, a DSS should be able to function and provide benefit even during uncertainty. Finally, it was identified that a DSS is only as good as the knowledge that it contains. Consequently, a technology applied to incident notification should facilitate knowledge acquisition. This characteristic means that it should be easy for anyone, not just domain experts, to enter new knowledge into a DSS. Additionally, it is just as important that a technology makes it easy for users to maintain the knowledge base over time. Thus, four initial characteristics were deemed as desirable in a notification system: *Adaptable to Environment*, *Functions with Uncertainty*, *Facilitates Knowledge Acquisition*, and *Low Maintainability*.

After acknowledging the domain aspects, the 10 user-defined relevance criteria identified by Barry and Schamber (1998) in Table 1 were considered. However, it is important to note that

when viewed as a subjective concept, relevance is not inherent to information. Only the users of the information can decide whether a particular notification is relevant or not. Therefore, there is no method to objectively determine whether a message is more "relevant" than another without asking a specific individual in a specific information need situation. As an alternative, this research intends to find a technology that has the most *capacity* for relevant notification to take place. From this view point, Barry and Schamber's study can still be used as a foundation.

An important difference between the relevance literature and this research is the end users' motivation. In the relevance literature, it is assumed that the users are actively looking for information to satisfy a need (i.e. weather information needed to perform a job, or documents needed for research). In this research, the users (i.e. a squadron commander) are *not* actively looking for notifications. Instead, they are automatically warned about incidents while performing their normal duties. Due to this distinction, it was determined that some of categories from Table 1 were not applicable. First, the *Accessibility* category was not considered because it is assumed that all individuals within an organization will be able to view any notification that is sent to them without difficulty or cost. Next, the *Availability of information/Sources of information* criterion was not considered because it is assumed that the end users will be present at the DSS when an incident occurs and ready to view a notification. The ability to provide notifications while personnel are away from the DSS is outside the scope of this thesis. Next, the *Verification* criterion was not included because it is assumed that all information entered into a DSS will be consistent with experience gathered from an organization over time. Also, failure to verify a notification does not mean a user will find it irrelevant. On the contrary, a user may learn something new upon reviewing a notification. Finally, the *Affectiveness* criterion was not considered because pleasure or entertainment should not be

experienced from reading notifications. If affection existed, then some notifications may be favored over others, which is not ideal. After excluding these metrics, six were left from Table 1 as potential desired characteristics for a cyber incident notification system: *Depth/Scope/Specificity*, *Accuracy/Validity*, *Clarity, Currency*, *Tangibility*, and *Quality of Sources*.

After reviewing all the potential characteristics, it was recognized that some overlapped with each other. First, it was determined that the *Accuracy/Validity* and *Currency* criteria were unneeded because they were believed to be functions of *Adaptable to Environment* and *Facilitates Knowledge Acquisition*. A technology that is adaptable will provide accurate and current information because it will adjust to the environment. Additionally, a technology that makes it easy for users to enter knowledge will help ensure that the information in the repository is accurately entered and up-to-date. Next, the *Quality of Sources* criterion was also viewed as a function of knowledge acquisition. If it is easy to enter knowledge, all levels of military personnel (i.e. airmen to colonels) will be able to contribute even if they are not domain experts.

As a result, the desired characteristics were finalized: *Adaptable to Environment*, *Functions with Uncertainty*, *Facilitates Knowledge Acquisition*, *Low Maintainability*, *Provides Information Depth* (changed from *Depth/Scope/Specificity*), *Presents Information Clearly* (changed from *Clarity*), and *Provides Tangible Information* (changed from *Tangibility*). By establishing these characteristics, the first research goal was achieved. The final evaluation table is shown in Table 3. This table was used as the means to rank and compare the technologies. The ranking, or coding process, was accomplished by performing a content analysis and is explained in the following section.

Table 3. Evaluation Table for Ranking DSSs

Desired Characteristics	**Definition**	**Decision Support System**			
		RBS	**CBR**	**BN**	**NN**
Adaptable to Environment	Ability of the system to continually provide accurate information over time; flexible to change				
Functions with Uncertainty	Ability of the system to provide benefit when decision making information is uncertain or missing				
Facilitates Knowledge Acquisition	Ease at which the system allows any user (i.e. domain experts to novices) to enter new knowledge into its repository				
Low Maintainability	Ease at which the system allows users to maintain the knowledge base				
Provides Information Depth	Ability of the system to provide sufficient and focused information to a decision maker (i.e. problem, solutions, additional context)				
Presents Information Clearly	Ability of the system to display information in a way that is easy to understand				
Provides Tangible Information	Ability of the system to provide definite proven information (i.e. scenarios or hard data)				

3.3 Content Analysis

With the evaluation table established, the next task was to rank each DSS against the desired characteristics. This objective was completed by performing a content analysis. As defined in the beginning of this chapter, a content analysis is a process in which text is examined to answer a research question. Because this technique analyzes documents, it is a qualitative method (Patton, 2002).

3.3.1 Sampling

The first step in conducting the content analysis was to determine which publications would be used for coding Table 3. There are a number of different ways to obtain an appropriate sample, which include both random and non-random sampling techniques (Krippendorff, 2004; Neuendorf, 2002). While random sampling with a large sample size is preferred for generalizing results, this technique is not typically used in qualitative research. Instead, qualitative research focuses on in-depth analysis, which necessitates the need to select sources which can best answer the research question. This type of sampling is called purposeful sampling, and was the method used in this thesis (Patton, 2002). Patton (2002) states: "The logic and power of purposeful sampling lie in selecting *information-rich cases* for study in depth. Information-rich cases are those from which one can learn a great deal about the issues of central importance to the purpose of the inquiry, thus the term *purposeful* sampling" (p. 230).

To find the most information-rich publications, online databases as well as library resources were used at the Air Force Institute of Technology (AFIT). Journal articles and books on RBSs, CBR systems, BNs, and NNs, were searched by using key words that included, but not limited to: "rules-based systems", "rules-based reasoning", "case-based reasoning", "expert systems", "Bayesian networks", and "Neural Networks". The sampling process was completed

once "theoretical saturation" was achieved. Theoretical saturation is defined by Mack et al. 2005 as "the point in data collection when new data no longer bring additional insights to the research question" (p. 5). This term is also referred to as "theoretical sufficiency" (Andrade, 2009).

3.3.2 Coding

After the resources were identified, the next step was to establish a coding scheme. According to Krippendorff (2004), ordinal scales consisting of 3, 5, or 7 metrics are a natural choice for ranking data. For this thesis, a 5-point ordinal scale was created (see Table 4). For each desired characteristic in Table 3, the DSSs were ranked by how well they embodied that characteristic using the scale. For example, the first characteristic is *Adaptable to Environment*. If RBSs were found to be extremely adaptable, then they would be coded as a "5" for that variable. During the process, the documents from the sample served as support for each coding decision. Once all of the fields were coded, each characteristic (i.e. row) was analyzed independently to determine which DSS was ranked the best. The DSS that was ranked the best in a majority of the categories was selected as the most suitable for consideration in an incident notification system.

Table 4. Content Analysis Coding Scheme

Code	Definition
1	DSS does not support this characteristic
2	DSS scarcely supports this characteristic
3	DSS moderately supports this characteristic
4	DSS greatly supports this characteristic
5	DSS fully supports this characteristic

3.3.3 Intercoder Reliability

To provide meaningful conclusions about the DSSs, reliability in the coding process must be established. In a content analysis, reliability is obtained when there is an acceptable level of agreement between two or more coders. This concept is generally referred to as intercoder reliability (Krippendorff, 2004; Neuendorf, 2002). The simplest form of intercoder reliability is percent agreement, in which the number of agreements between coders is divided by the total number of coding decisions. However, this calculation fails to take into account agreement by *chance*. To cope with this problem, several coefficients have been developed: Scott's pi (π), Cohen's kappa (κ), and Krippendorff's alpha (α). For this research, Krippendorff's alpha was used because it is the only coefficient that can adjust for ordinal coding data. In contrast, Scott's pi and Cohen's kappa assume the use of nominal level metrics (Neuendorf, 2002).

As mentioned above, an acceptable level of agreement must be established between coders. However, there are differing opinions on what an "acceptable" level is (Neuendorf, 2002). Because Krippendorff's alpha was used, the rule of thumb set by its creator was applied. Krippendorff (2004) suggests that variables with an alpha above .8 can be considered reliable, while variables with an alpha between .667 and .8 can be used for making cautious conclusions.

For this content analysis, Table 3 was coded by the author of this thesis as well as a primary researcher on the CIMIA team. The second coder was supplied with a training handout that is located in Appendix A. This handout included a description on the purpose of the analysis, directions for coding, and short descriptions of each DSS along with text extracts. Because the second coder was a subject matter expert on all of the DSSs, the training process was smooth and minimal. After Table 3 was coded independently by both members, the results

were compared and a Krippendorff's alpha was computed for each desired characteristic. To compute each alpha, a free online tool called ReCal was used (Freelon, 2010, 2011).

IV. Analysis and Results

4.1 Introduction

This chapter presents and discusses the results of the content analysis. First, the selected publications are identified. Next, there is a discussion on how the DSSs were ranked with respect to each of the desired characteristics from the author's perspective. Following this examination, a coded Table 3 is displayed as the final product. After analyzing the table, initial conclusions are made about the most suitable DSS for incident notification. Then, the results from the second coder are used to determine whether reliability in the conclusion exists. Once the most suitable DSS is selected, initial design considerations are proposed. Finally, this chapter shows the feasibility of the DSS by presenting a fictional example scenario based upon a real world military unit.

4.2 Publications Selected

By using the library and online database resources at AFIT, there were many publications that described the fundamental aspects of RBSs, CBR systems, BNs, and NNs. As explained in the methodology chapter, articles were chosen based on their ability to provide insight on ranking the DSSs according to the desired characteristics listed in Table 3. Because a qualitative sampling method was used, there was no target sample size. Overall, there were 23 publications chosen for the study, which included both books and journal articles. The books and articles that were selected are displayed in Table 5 and Table 6, respectively.

Table 5. Books Used in Content Analysis

Title	Author(s)	Year Published
*Building Expert Systems[1]	Frederick Hayes-Roth Donald A. Waterman Douglas B. Lenat (editors)	1983
*Introductory Readings in Expert Systems[2]	Donald Michie (editor)	1982
*Case-Based Reasoning: Experiences, Lessons, & Future Directions[3]	David B. Leake (editor)	1996
Case-Based Reasoning	Janet Kolodner	1993
Applying Case-Based Reasoning: Techniques for Enterprise Systems	Ian Watson	1997
Bayesian Artificial Intelligence	Kevin B. Korb Ann E. Nicholson	2004
Bayesian Networks and Decision Graphs	Finn V. Jensen Thomas D. Nielson	2007
A Guide to Neural Computing Applications	Lionel Tarassenko	1998
Neural Networks: A Comprehensive Foundation	Simon Haykin	1994
Artificial Neural Networks	Robert J. Schalkoff	1997
Neural Network Learning	Stephen I. Gallant	1993
Artificial Intelligence: A Modern Approach	Stuart Russell Peter Norvig	2003

* Denotes an edited book

[1] The following chapter was used from this book: *An Overview of Expert Systems*, by Hayes-Roth, Waterman, and Lenat

[2] The following chapter was used from this book: *A survey and critical review of expert systems research*, by Bramer

[3] The following chapters were used from this book: 1) *CBR in Context: The Present and Future*, by Leake, and 2) *A Tutorial Introduction to Case-Based Reasoning*, by Kolodner and Leake

Table 6. Articles Used for Content Analysis

Title	Journal	Author(s)	Year Published
Rule-Based Systems	Communications of the ACM	Frederick Hayes-Roth	1985
An Approach to Verifying Completeness and Consistency in a Rule-Based Expert System	AI Magazine	Motoi Suwa A. Carlisle Scott Edward H. Shortliffe	1982
Applying Case-Based Reasoning to Autoclave Loading	IEEE Expert	Daniel Hennessy David Hinckle	1992
Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches	AI Communications	Agnar Aamodt Enric Plaza	1994
Integrating Case- and Rule-Based Reasoning	International Journal of Approximate Reasoning	Soumitra Dutta Piero P. Bonissone	1993
Application of a hybrid case-based reasoning approach in electroplating industry	Expert Systems with Applications	Felix T.S. Chan	2005
A Bayesian Belief Network for IT implementation decision support	Decision Support Systems	Eitel J.M. Lauría Peter J. Duchessi	2006
Using Bayesian network analysis to support centre of gravity analysis in military planning	European Journal of Operational Research	Lucia Falzon	2006
An Application of Bayesian Networks to Antiterrorism Risk Management for Military Planners	Technical report by Digital Sandbox, Inc.	Linwood D. Hudson Bryan S. Ware Suzanne M. Mahoney Kathryn B. Laskey	2002
Neural networks for decision support: Problems and opportunities	Decision Support Systems	Shimon Schocken Gad Ariav	1994
Artificial neural networks	Surgery	Philip J. Drew John R. T. Monson	2000

4.3 Content Analysis Discussion

This section presents a discussion on how the DSSs were coded with respect to each desired characteristic from the author's perspective. The rankings are supported by the literature listed in Table 5 and Table 6 above. The text extracts in this discussion, as well as additional findings, are located in Appendix B.

4.3.1 Adaptable to Environment

The *Adaptable to Environment* characteristic evaluates DSSs on their ability to provide accurate notifications in a dynamic environment. Because resource dependencies can change during military operations, a DSS must be flexible. It was first determined that BNs were the least adaptable. BNs rely on causal events and probability theory to reason, which means that the domain of interest must be modeled up front. Falzon (2004) uses BNs as a way for military planners to model centers of gravity, or COGs. She states, "since the structure of a COG model is so important for accurate analysis users are encouraged to check it carefully before moving on to populate the model [with conditional probabilities]" (p. 637). Also, Lauría and Duchessi (2006) use BNs to provide advice for companies implementing information technology. They determined 13 factors that influenced implementation, which were used as the nodes in their BN. In their summary, they mention that "the implementation factors that appear here may change over time… necessitating the development of another data set and BBN [Bayesian Belief Network] model" (p. 1586). Due to this inflexibility, BNs were coded as a 1.

NNs were ranked higher than BNs because they allow for learning to take place. Haykin (1994) explains that "neural networks have a built-in capability to *adapt* their synaptic weights to changes in the surrounding environment" (p. 4). However, this adaptability is still bound by the quality of training data. Tarassenko (1998) states: "During development, the neural network was

trained on data collected from the operating environment, probably over a limited period of time. For some applications, the operating environment will change over time, possibly leading to a reduction of performance" (p. 48). Overall, NNs were coded as a 3.

RBSs also have some potential for adaptability. Watson (1997) explains that "rules can encapsulate small chunks of knowledge that collectively can model a complex problem" (p. 7). Similarly, Hayes-Roth (1985) says that the skill of a RBS "increases at a rate proportional to the enlargement of their knowledge bases" (p. 921). These findings suggest that RBSs can adapt as the environment changes by adding new rules, unlike the more rigid models that must be established in BNs or NNs. However, rules must conform to a certain format, i.e. IF-THEN statements. As a result, the environment may not be modeled correctly. Dutta and Bonissone (1993) say that rules require a strict match, which is "very restrictive as real-world situations are often fuzzy and do not match exactly with rule premises and conclusions" (p. 166). Overall, RBSs were coded as a 3.

CBR systems were found to be the most adaptable. Like RBSs, knowledge can be continually added to improve the system as the environment changes. Aamodt and Plaza (1994) say that CBR supports "incremental, sustained learning, since a new experience is retained each time a problem has been solved, making it immediately available for future problems" (p. 2). Also, adaptation is a fundamental step in the CBR process. Kolodner (1993) states: "Because new situations rarely match old ones exactly, however, old solutions must be fixed to fit new situations. In this step, called adaptation, the ballpark solution is adapted to fit the new situation" (p. 21). However, adaptation can be a difficult aspect to include and few systems actually use it (Watson, 1997). Regardless, CBR was coded as a 4 due to its potential.

4.3.2 Functions with Uncertainty

The *Functions with Uncertainty* criterion focuses on how much value a DSS can provide when information is uncertain, incomplete, or missing. A commander makes decisions based on the information that can be gathered from the environment. But, this information may not always be certain or complete. A DSS should be able to function and provide helpful feedback even when inputs are degraded.

While BNs were ranked as the least adaptable, they were found to be the best during uncertainty. This rationale stems from the same explanation of why BNs are inflexible: BNs model the domain completely before reasoning begins. Korb and Nicholson (2004) state that "Bayesian networks provide full representations of probability distributions over their variables. That implies that they can be conditioned upon any subset of their variables, supporting any direction of reasoning" (p. 34). Even if a particular piece of evidence in a BN is missing, a user could still use the network to infer about the domain. It is important to note that this capability assumes the model accurately represents the environment. However, since that aspect was addressed in the last characteristic (i.e. adaptability), model accuracy is independent of this category. As a result, BNs were coded as a 5.

NNs were also ranked highly. Instead of a complete domain model, NNs define the inputs and outputs of the environment and learn with training data. A high amount of input variables reduces the importance of any one piece of information. Gallant (1993) explains that "most connectionist models [i.e. NNs] naturally extend to cases where some inputs are unknown…. Because cells [i.e. neurons] can easily examine large numbers of inputs, they naturally tend to be less sensitive to noise; the greater number of correct input variables can outvote the fewer number of incorrect input values" (p. 10). This capability is also supported by

Haykin (1994), who explains that "owing to the distributed nature of information in the network, the damage [to neurons] has to be extensive before the overall response of the network is degraded seriously" (p. 5). Therefore, missing information may have no affect on the overall outcome of the NN. For this reason, NNs were coded as a 4.

RBSs and CBR systems were not as favorable because missing information could mean poor or no solutions. Hennessy and Hinkle (1992) explain that "because a rule-based system rigidly matches rules to a problem description, a missing rule halts the reasoning process" (p. 25). Similarly Chan (2005) explains that "when the CBR system is initially applied for a particular problem, only a few cases will be stored in the database. This leads to a problem of 'openness'.… The system will fail to generate [a] solution, or generate very unreasonable solutions" (p. 125). However, more cases could yield better solutions. It was ultimately determined that CBR systems were better with missing information because the retrieval algorithms can still provide the nearest solution even if it may not be the most accurate. Kolodner (1993) explains that "cases are retrieved that match the input *partially*" (p. 94). On the other hand, missing rules may stop the entire reasoning process as noted above. With this in mind, CBR systems were coded as a 3 while RBSs were coded as a 2.

4.3.3 Facilitates Knowledge Acquisition

The *Facilitates Knowledge Acquisition* characteristic evaluates the ease at which a DSS allows users to enter new knowledge. As discussed in the literature review, a DSS is only as valuable as the information contained in it. Therefore, it is crucial that a DSS makes it easy for users to provide input.

It was found that NNs were the most complicated for facilitating knowledge acquisition because of their "black box" approach. Schalkoff (1997) explains that there are "no clear rules or

design guidelines for arbitrary application" (p. 10). Also, adding hidden neuron layers can be troublesome. Russell and Norvig (2003) state: "The problem of choosing the right number of hidden units in advance is still not well understood" (p. 744). If there are no procedures for building a network, then adding to it will be just as complicated. Also, it was discussed earlier that training must be accomplished when a new network is constructed. However, Tarassenko (1998) points out that "a sufficient number of training examples is required to ensure that the neural network is trained to recognize and respond to the full range of conditions" (p. 69). This requirement makes it difficult to update the network. Therefore, NNs were coded as a 2.

Similarly, BNs scored low in this category because of their formal structures. This difficulty applies to both entering new nodes and determining conditional probability distributions. First, the nodes are arranged by causality. Any new node must fit into the network at the right spot or the model could become inaccurate. Jensen and Nielson (2007) explain the difficulties with determining causality: "First, causal relations are not always obvious.… Furthermore, causality is not a well understood concept. Is a causal relation a property of the real world or rather, is it a concept in our minds helping us to organize our perception of the world?" (p. 60). For this reason, it is hard for users to enter new nodes unless they have an extremely good grasp of the domain. Second, creating conditional probabilities can be troublesome. Hudson et al. (2002) explain that in their BN they "populated their conditional probability tables with 'rough guess' values based on information we had obtained from domain experts and literature" (p. 5). This finding suggests that it takes time to build the appropriate knowledge before updates can be made. Overall, BNs were coded as a 2.

Both RBSs and CBR systems showed more promise with facilitating knowledge acquisition because the addition of new knowledge does not necessitate the need to understand

the structure of a model. RBSs store knowledge in small chunks that can be accumulated over time to form an accurate picture of a domain. Because rules are small pieces of information, they are easy for users to associate with. Watson (1997) explains that "rules are a part of everyday life, and so again people can relate to them" (p. 10). Also, rules are independent of each other (Watson, 1997). Therefore, users do not need to understand causality before entering new rules. However, because rules are chained together, there is the possibility of contradicting previous rules or adding redundant rules (Suwa, Scott, & Shortliffe, 1982). Also, since knowledge must be transformed into a specific format, building a rule base usually involves a third party. Kolodner (1993) states: "In rule-based reasoning, knowledge is extracted from experts and encoded into rules. This is often difficult to do" (p. 94). So, while rules are easy to understand, the typical user may not be able to accurately write them. An ideal DSS should allow users to enter information without having to think too much about how to condense their own knowledge. Overall, RBSs were coded as a 4.

CBR is more focused on capturing whole experiences. Leake (1996) states: "Because case-based reasoners reason from complete specific episodes, CBR makes it unnecessary to decompose experiences and generalize their parts into rules" (p. 5). He also explains that "experts who are resistant to attempts to distill a set of domain rules are often eager to tell their 'war stories' - the cases they have encountered" (p. 6). This finding shows that it may be more natural for users to enter cases as opposed to rules. Ultimately this notion could mean the elimination of system experts, allowing many military personnel the ability to enter knowledge with little expertise or training. Finally, Kolodner (1993) states that "several recent studies point to the relative ease with which case-based reasoners can be built as compared to building the same rule-based systems" (p. 94). From these findings, CBR was coded as a 5.

4.3.4 Low Maintainability

The *Low Maintainability* category evaluates the ease at which a DSS allows users to maintain the knowledge base. After knowledge is acquired it must be maintained over time to remain accurate. Previous findings were used to code this category. From the discussion above, it was shown that BNs and NNs were the least adaptable to the environment. They also require a significant understanding of the underlying theory to be used, which makes it complicated for average users to add knowledge. Because of these difficulties, it was determined that significant effort would be needed to maintain the knowledge base of a BN or NN. Therefore, they were both ranked as a 2.

In contrast RBSs and CBR systems are more natural in their reasoning processes. From this finding, it was determined that less expertise is needed to maintain these systems over time. Because CBR systems draw on whole experiences that do not have to conform to any particular standard, they were ranked as a 4. However, RBSs were ranked as a 3 because some expertise is required for knowledge to be transformed into rules.

4.3.5 Provides Information Depth

The *Provides Information Depth* category evaluates a DSSs ability to provide sufficient and focused information to a decision maker. A commander must have the appropriate amount of information to make the right decision. Providing situation context and courses of action can aid him or her in this process. NNs were not highly ranked in this category because they do not provide an explanation to the user. Drew and Monson (2002) explain that "clinicians remain wary of computer-aided diagnosis - and ANNs in particular - because many of them believe they require an insight into the system's behavior to assess the relevance of a computer-aided diagnosis decision to a particular patient. They thus resent the 'black box' nature of ANNs" (p.

8). This view is also expressed by Schocken and Ariav (1994) who sate: "Neural computing is extremely convoluted, and therefore it is difficult to explain or defend the system's 'rationale' (unlike expert systems, where one can trace reasoning chains or invoke some sort of belief calculus)" (p. 402). As a result, NNs were coded as a 1.

BNs primarily deal with providing quantitative information, i.e. probabilities. Lauría and Duchessi (2006) state that a "BBN makes probabilistic assertions as to the outcome of certain actions" (p. 1582). However, BNs also present a picture of the domain that can be useful to a decision maker. Falzon (2006) explains that BNs "provide a visual representation that facilitates reasoning and enhances shared understanding of complex situations" (p. 632). While this aspect is positive, there is still some qualitative information that is missing from BNs . Overall, BNs were coded as a 3.

RBSs can offer the qualitative information that BNs lack. Hayes-Roth et al. (1983) explain that RBSs can "provide explanations or justifications for conclusions reached" (p.5). The program TEIRESIAS, which worked in conjunction with the MYCIN RBS, exemplifies this concept. While using TEIRESIAS, Bramer (1981) explains that "the user (expert) can ask WHY (to query the significance of a request by MYCIN for information) or HOW (to ask how deductions so far considered as established by MYCIN were arrived at)" (p. 14). This type of information allows users to gain a deeper understanding behind the reasoning process. However, since knowledge is constricted to rule format, some information could be lost. Overall, RBSs were coded as a 4.

CBR systems offer similar benefits when compared to RBSs. However, it was found that CBR is more configurable than RBSs because there is no restriction to the type of information that can be presented to a user. This freedom to design cases is a fundamental aspect

of CBR systems. Kolodner (1993) explains: "In case-based reasoning… the majority of intellectual emphasis has been on content issues: What kinds of content should cases have?" (p. 94). Watson (1997) explains that cases can contain "names, product identifiers, values like cost or temperature, and textual notes. An increasing number of CBR tools also support multimedia features, such as photographs, sound, and video" (p. 19). Because the cases in CBR systems are not forced to conform to any specific format, this DSS offers the greatest ability to provide any and all information that a consumer needs when making a decision. This extra flexibility allowed CBR to be coded as a 5.

4.3.6 Presents Information Clearly

The *Presents Information Clearly* criterion evaluates a DSSs ability to display information in a manner that is easy to understand. A user must be able to quickly comprehend cyber notifications so that contingency actions are not delayed. This was a difficult category to code because all DSSs have the ability to present information that is clear. Part of this concept has to do with how the user interface is designed, which is outside the scope of this thesis. However, findings from the previous characteristics provide insight on which systems have more capability.

Because CBR is the most configurable with respect to knowledge representation, as discussed in the previous section, it was coded as a 4. BNs and RBSs were close in this category because they both offer advantages. BNs have the ability to show a pictorial model. In contrast, RBSs can display rules to users allowing them to understand the reasoning process. Because these systems provide equally valuable information in different ways, they were both coded as 3's. NNs were ranked the lowest because of their "black box" approach, as discussed earlier, and were coded as a 2.

4.3.7 Provides Tangible Information

The *Provides Tangible Information* characteristic evaluates a DSSs ability to provide definite information, which includes hard data or scenarios. Tangible information helps the decision maker gain a more concrete understanding of a given situation. As with the previous characteristic, this category was difficult to rank because all of the DSSs in consideration have the ability to provide tangible information. Findings from the previous characteristics were used to code this section.

As discussed in section 4.3.3, CBR focuses on capturing experiences. Also, section 4.3.5 showed that CBR is able to present knowledge in many ways. Due to these advantages, CBR was coded as a 5. RBSs offer similar advantages with the exception that knowledge is constricted to rule format. Therefore, RBSs were coded as a 4.

The ranking between BNs and NNs was difficult because they are both primarily quantitative. However, BNs can present a decision maker with the probability of an event occurring as well as a picture of the domain, which is useful for decision making. As a result BNs were coded as a 3. Like BNs, NNs are typically mathematical in their output, which can provide concrete results to a decision maker. However, because the output of an NN can be hard to understand, they were coded as a 2.

4.4 Content Analysis Results

The preceding discussion explained the reasoning behind each of the coding decisions in the content analysis. Table 7 formally presents these results.

Table 7. Content Analysis Results

Desired Characteristics	Definition	Decision Support System			
		RBS	CBR	BN	NN
Adaptable to Environment	Ability of the system to continually provide accurate information over time; flexible to change	**3**	**4**	**1**	**3**
Functions with Uncertainty	Ability of the system to provide benefit when decision making information is uncertain or missing	**2**	**3**	**5**	**4**
Facilitates Knowledge Acquisition	Ease at which the system allows any user (i.e. domain experts to novices) to enter new knowledge into its repository	**4**	**5**	**2**	**2**
Low Maintainability	Ease at which the system allows users to maintain the knowledge base	**3**	**4**	**2**	**2**
Provides Information Depth	Ability of the system to provide sufficient and focused information to a decision maker (i.e. problem, solutions, additional context)	**4**	**5**	**3**	**1**
Presents Information Clearly	Ability of the system to display information in a way that is easy to understand	**3**	**4**	**3**	**2**
Provides Tangible Information	Ability of the system to provide definite proven information (i.e. scenarios or hard data)	**4**	**5**	**3**	**2**

After examining the table above, it was observed that CBR was ranked the highest in six out of the seven desired characteristics (see Table 8). The exception was found in the *Functions with Uncertainty* characteristic, where BNs were ranked the highest.

Table 8. Highest Ranked DSS by Characteristic

Desired Characteristics	**Highest Ranked DSS**
Adaptable to Environment	CBR
Functions with Uncertainty	BN
Facilitates Knowledge Acquisition	CBR
Low Maintainability	CBR
Provides Information Depth	CBR
Presents Information Clearly	CBR
Provides Tangible Information	CBR

From this analysis, it appeared that CBR systems were the most suitable for consideration in a cyber incident notification system. However, intercoder reliability was required to ensure that this conclusion was accurate. The coding results obtained from the second coder are located in Appendix C. These results were used to calculate a Krippendorff's alpha for each characteristic and are displayed in Table 9. This table shows that high reliability (i.e. an alpha above .8) was achieved for *Adaptable to Environment*, *Functions with Uncertainty*, *Facilitates Knowledge Acquisition*, *Provides Information depth*, and *Provides Tangible Information*. Considerable reliability (i.e. an alpha above .667) was achieved for *Low Maintainability* and *Presents Information Clearly*. From these results, it was solidified that CBR was the most suitable framework for improving incident notification.

Table 9. Reliability Results for Two Coders

Desired Characteristics	Krippendorff α
Adaptable to Environment	0.91
Functions with Uncertainty	0.98
Facilitates Knowledge Acquisition	0.81
Low Maintainability	0.71
Provides Information Depth	0.81
Presents Information Clearly	0.79
Provides Tangible Information	0.89

4.5 Design Considerations for Applying CBR

This section presents some initial design considerations for applying CBR to the cyber incident notification domain. Most importantly, a case representation format will be determined. The feasibility of this case design will be explored using a fictional real-world example. Other design considerations examined are case indexing, knowledge acquisition, and usability.

4.5.1 Case Representation

Before knowledge can be acquired for use within a CBR system, a case representation format should first be determined (Althoff & Weber, 2006). Kolodner and Leake (1996) define a case as a "contextualized piece of knowledge representing an experience that teaches a lesson fundamental to achieving the goals of the reasoner" (p. 36). The three major parts of a case typically include a problem, a solution, and an outcome. The problem shows what is wrong, the solution provides a possible answer, and the outcome describes the end result of carrying out the proposed solution (Kolodner, 1993). Using these basic fields is a good pace to start in developing a case representation for cyber incident notification.

The added focus on relevance in this research provides new insight to how cases can be designed. Most case representations only include the use of text. However, adding unit specific mission representation in the form of diagrams or pictures can further enhance recognition of the relevance of a case upon retrieval. Additionally, the information represented within a case is limited to one perspective. While this may be suitable for some domains, the concept of relevance highlights the issue of subjectivity. Therefore, multiple views and solutions with respect to a specific case may provide benefit to a wider audience since each user has a different view of the world. These ideas can be incorporated into the problem, solution, and outcome fields described in the previous paragraph.

First, a *problem* statement is needed to explain why a case is being displayed. However, this element could be enhanced by adding a picture. Glenberg and Langston (1992) showed that text accompanied by pictures helped individuals build mental models. Thus, a pictorial problem representation may allow end users to have a better understanding about how their mission is affected in a short amount of time. Business process models are one method that can be used for this idea. A business process model is a way to represent the sequence of activities and their relationships that lead to a specific goal. In other words, these diagrams externalize knowledge about organizational processes which lead to mission success (Kalpic & Bernus, 2006).

Next, the *solution* field provides a possible course of action for solving the problem. However, as described above, several solutions may be useful. Multiple stories can provide a richer knowledge base when attempting to make the best decision. Therefore, consumers should be able to contribute to the solution pool, much like Web 2.0 web sites. However, users must not be overwhelmed by an abundance of information. For this reason, only one solution should be listed along with the option to show more if desired. To provide the one "best" solution up front,

users should be allowed rank solutions based on their experience. The solution with the highest ranking could be displayed as the best option. But again, the user should be able expand this section to reveal more solutions submitted by other users. This idea is not unlike how reviews are made online for travel web sites, like www.tripadvsior.com (Trip Advisor, 2011). When searching for a hotel in a certain city, the highest ranked hotel is listed first.

Finally, the *outcome* field shows the result from implementing a particular solution. Like the solutions field, multiple outcomes should be displayed with respect to a specific case. Again, these will be determined by input provided from experienced organizational members. One possible idea would be to present a "best case", "worst case", and "historical case". The "best case" would present the most positive outcome that could result from implementing a solution. In contrast, the "worst case" would be the most negative outcome. The "historical case" would present the most likely outcome based on previous occurrences. These different views allow a decision maker to be more informed about a situation.

To present this case representation design, an example scenario will be used. In particular, this example focuses on the Maintenance Operations Center (MOC) located at Wright-Patterson Air Force Base, Ohio. The MOC is an organization that operates within a maintenance squadron on a USAF installation. They are ultimately responsible for helping the squadron ensure aircraft readiness, which is the main mission objective for any maintenance unit (Department of the Air Force, 2010). As part of their responsibility, the MOC must enter the readiness status of aircraft into a system called GO81. This system then automatically updates another system called the Global Decision Support System (GDSS). The GDSS is used by higher echelons of command in determining which aircraft are available to perform specific missions. If the link between GO81and GDSS is lost or compromised, then commanders will

not have accurate information about the aircraft that can be tasked. This issue creates the need for MOC personnel to be aware that their updates are not being observed via the GDSS. With the current USAF C4 NOTAM process, the information provided to an end user would be limited. Figure 12 shows what this notification may look like.

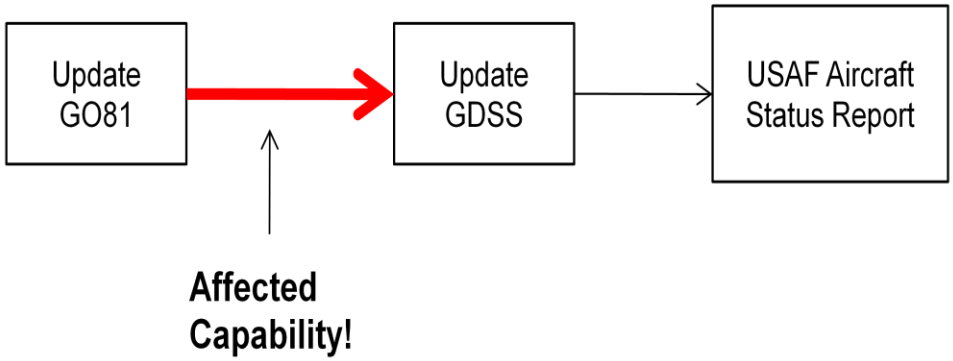
```
UNCLASSIFIED//FOR OFFICIAL USE ONLY

ACKNOWLEDGEMENT DATE: 9 DEC 2010

INTITAL RELEASE TIME:  09 0245Z DEC 10

TRACKING NUMBER:  1

ORIGINATING AGENCY:  HIGHER HEADQUARTERS

TYPE:  INFORMATIVE

CATEGORY:  NOTAM

PRIORITY:  SERIOUS

SUBJECT:  VULNERABILITY IN GDSS

MISSION IMPACT:  LOSS OF SYSTEM AVAILABILITY

EXECUTIVE SUMMARY:
G081 IS NOT AUTOMATICALLY UPDATING GDSS.  AIRCRAFT READINESS STATUS MAY NOT BE
CURRENT, WHICH COULD ULTIMATELY AFFECT MISSION OPERATIONS.

SYSTEM(S) AFFECTED:
GO81 AND GDSS USERS

ACTION:
SYSTEM ADMINSTRATORS SHOULD REFER TO GDSS POINT OF CONTACT TO DETERMINE
APPROPRIATE FIX ACTIONS.

REPORTING REQUIREMENTS:
REPORT COMPLIANCE.

REMARKS:
Please contact HELP DESK - if you have questions and/or concerns
```

Figure 12. Example C4 NOTAM

Overall, the C4 NOTAM in the example above provides some important information; however, there is a lack of detail in the problem description and course of action. This issue stems from the fact that these notifications are typically written by personnel outside of the MOC. Contribution from members inside the MOC organization could tailor notifications to be more relevant. Additionally, the lack of a pictorial representation of the problem may prevent users from understanding the criticality of the message. The flexibility of CBR systems to represent knowledge in many ways can help improve upon these limitations. Figure 13 shows what this same scenario might look like to a user equipped with a CBR notification system.

Problem: GO81 is Not Automatically Updating the Global Decision Support System (GDSS)

USAF aircraft tasking decisions could be affected



Solutions: Communicate with personnel at the GDSS location to explain the problem. Continue to make aircraft readiness updates in GO81, but use other means to relay those updates for manual entry into GDSS (i.e. over the phone or by email). View other solutions

Outcomes: *Best Case* – The automatic update problem is fixed quickly and no flying operations are affected.

Worst Case – The problem may take days to be fixed and manual updates to GDSS take too long to keep up with operations. Missions may be degraded.

Historical Case – The problem is resolved within 24 hours and no operations are greatly affected.

Figure 13. Example Case for Cyber Incident Notification

In the example case representation, the *problem* field explains to the user that the automatic update between GO81 and GDSS is not occurring. It is further stated that this issue could possibly affect aircraft tasking decisions. A business process model is displayed to help the user understand the how the activities relate to each other. The *solutions* field shows the highest rated course of action, which states that the personnel associated with the GDSS should be notified and updates to the GDSS must still be made using other means. At the bottom of this description, the user can click the "View other solutions" link to present more options. Finally, the *outcomes* field states possible results associated with this scenario. The historical case shows that the problem will most likely be resolved without operations being affected. But the worst case shows the potential for degraded operations, which allows the decision maker to prepare for further action.

4.5.2 Case Indexing

Indices are features that represent cases so they can be quickly retrieved (Watson, 1997). Determining the right indices for cases is commonly referred to as the indexing problem (Kolodner, 1993; Riesbeck & Schank, 1989). Kolodner (1993) presents four guidelines for creating indices: they should 1) be predicative, 2) address the purpose of the case, 3) be abstract enough for use in multiple situations, and 4) be recognizable.

The user requires an incident notification only when there is a problem with the confidentiality, integrity, or availability (C-I-A) of the information needed from a resource. From this notion, the resources can become the indices which trigger a case. When all information resources are unaffected by C-I-A, no cases are retrieved. Once a problem arises with a resource, the CBR system will search the library for any cases that match the current

situation. This means that the CBR engine must continually check the current state of information resources. It will only act once there is a deviation from the steady state.

The level of index abstraction is crucial within the military environment. As information systems and their connections may change frequently, it is important that indices remain dynamic. For example, consider a domain name server (DNS). At a detailed level, a DNS could be located via its internet protocol (IP) address. However, using the IP address as an index is not appropriate as the address is subject to change. Instead, indices must be stored in a more abstracted representation. In the example, an index labeled "DNS" would be a better choice because it is suitable for *any* IP address. As a result, this means there must be a separate storage location linking the specific IP address to the DNS.

Due to the complex operating environment, there are upstream and downstream dependencies between organizations. To make this concept concrete, an organization can be thought of as a Lego block. A Lego block has two sides: one side allows it to connect with other pieces, while the reverse side allows other pieces to connect to it. When many blocks are placed together to form a structure, each individual block is dependent on its other connections and vice versa. Similarly, in the operating environment organizations rely on other organizations for mission assurance. Due to this dependence, organizations should publish the status of their resources so that others can maintain SA. Thus, indices should include resources that are supplied by other parties.

Finally, while this thesis focuses on providing relevant notification about cyber incidents, non-cyber resources could also be incorporated as indices. While information resources are critical to any mission, personnel and equipment are also required for success. As a result, the proposed CBR framework could capture resources that have the ability to be *monitored* through

cyber means. For example, a USAF maintenance squadron may be required to have a certain amount of aircraft fuel on hand to support flying operations. If the fuel runs out, missions could be halted. But, an electronic sensor could be implemented to warn personnel when the fuel is low. This sensor input could be integrated into the CBR system as an index to provide notification that more fuel is needed.

4.5.3 Knowledge Acquisition

The process of acquiring knowledge has been identified as a bottleneck for developing DSSs. As discussed earlier, the knowledge acquisition processes may involve knowledge engineers, who are people dedicated to collecting knowledge from experts (Cooke & McDonald, 1986). However, the ability of CBR to collect experiences can make the knowledge collection process easier and less expensive. Puppe and Gappa (1992) explain that direct knowledge acquisition is the best approach in terms of cost. The direct method allows experts themselves to formalize their own knowledge and transfer it into a system.

However, this method must be addressed carefully. Aha (1998) explains that the difficulty of writing cases is a major reason that CBR systems fail in organizations. Therefore, it is essential that users do not feel overwhelmed when creating cases. To avoid this problem, an incremental case acquisition strategy could be implemented. In this approach, cases would be slowly pieced together over time.

For this strategy to be successful, the use of automated mapping agents and tutoring techniques could reduce the workload placed on the user. Automated mapping agents can determine the most frequently used resources. For example, a user may connect daily to a specific server outside his or her unit. When a threshold is reached, the agent would assume that

this connection has an important meaning to the user. Now, tutoring principles can be used as the actual method to elicit knowledge about this connection.

Kim and Gil (2007) discuss the benefit of including tutoring methods into knowledge acquisition systems. They describe 15 tutoring principles that should be considered. From this research, two principles carry over to this thesis: 1) generate educated guesses, and 2) indicate a lack of understanding. For the former principle, an agent could ask a user the following: "You [the user] seem to connect to resource A frequently, are you performing an important task?" This dialogue could then be followed up using the latter principle: "How important is this connection to you?" A scale could be included with this message to allow the user to rate the importance level. By using these principles, knowledge about resource dependencies and their criticality to mission objectives can be established over time.

4.5.4 Usability

When new CBR systems are implemented, they may not contain many cases in the case-base. This issue could result in receiving unreasonable solutions from the system, which ultimately prevents users from trusting it. Chan (2005) explains that adding rules in the early stages of implementation can help fix this problem. Rule-based reasoning (RBR) systems (i.e. RBSs) have higher initial solution accuracy than CBR systems, as shown in Figure 14.

To increase usability, some rules should be populated into the CBR system during the early implementation stages. Once enough cases are added to the case-base, the system will no longer need to rely on rules to provide accurate solutions. To determine an appropriate set of rules, organization members should meet and enumerate the most obvious problems from C-I-A incidents. For example, one possible rule might be "IF the internet is unavailable, THEN GO81 cannot be used for aircraft status reporting."

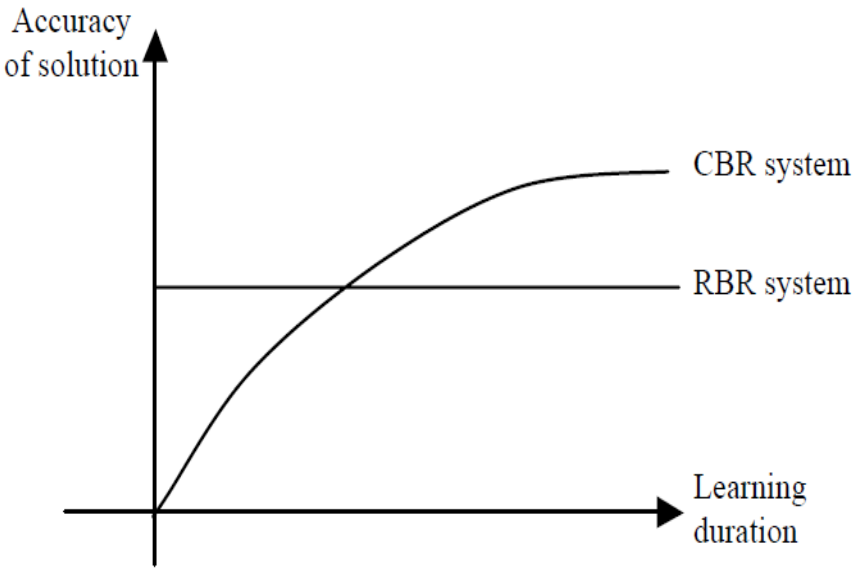


Figure 14. Accuracy of RBSs vs. CBR Systems (Chan, 2005)

Many CBR systems have failed due to the lack of user participation, which ceases case library development. In an effort to improve this issue, research by He et al. (2009) proposes that the integration of Web 2.0 technology and CBR systems will help encourage users to become more involved with the CBR process. Because problem solving is a social endeavor, they suggest that one reason CBR systems are not broadly accepted is due to the lack of a social environment in current systems (He, Xu, et al., 2009). Similar research on the topic of usability is investigating how the design of CBR interfaces can be optimized to encourage acceptance by more users. The users' mental model about how a CBR system searches for information is also important for success. Therefore, a good interface should provide training to help users understand the system (He, Wang, Means, & Xu, 2009).

Based on the success of the web page Wikipedia, its architecture is an ideal framework for aiding usability. Cases can be accumulated in a library that has a similar design as Wikipedia. Using this structure would take advantage of consistency, which is one of the key

human computer interaction design principles (Shneiderman & Plaisant, 2004). By maintaining a format that is familiar, potential users of the proposed CBR notification system may feel more comfortable using it.

V. Conclusions and Recommendations

5.1 Conclusions

The primary goal of this research was to determine which DSS is the best candidate for use in a notification system to overcome existing limitations in the USAF cyber incident notification process. Additionally, an emphasis was placed on enhancing notification *relevance*. To achieve this objective, a set of criteria were established as the basis for evaluating four types of potential systems: RBSs, CBR systems, BNs, and NNs. Overall, CBR was found to be the most suitable framework. However, a hybrid system should not be discounted. Instead, CBR appears to be a good foundation on which other methods can be added.

Once CBR was selected as a suitable DSS, some initial design considerations were proposed. These included case representation, case indexing, knowledge acquisition, and usability. The feasibility of the case representation design was demonstrated using an example scenario, which showed how CBR could improve the existing notification process. Among the many positive aspects, CBR's ability to represent knowledge in different ways showed its potential for providing relevant notifications to an end user.

During the process of conducting this research, another important product emerged: DSS evaluation criteria. Table 3 can be used as a tool by future researchers and practitioners to evaluate DSSs in other domains where incident notification is needed. While this thesis was specifically focused on the military domain, many organizations in the business industry operate in complex and dynamic environments that make the criteria appropriate.

5.2 Limitations

There are a few limitations from this study that must be pointed out. First, the content analysis methodology as whole has some short comings. In general, a content analysis takes

information that is qualitative and coverts it into quantitative data, i.e. the coding process. Unfortunately, knowledge richness can be lost in this exchange. Therefore, other methodologies that take a pure qualitative approach could provide a different view of this problem. Other appropriate methods may include open ended interviews with subject matter experts or a case study analysis.

Second, the evaluation criteria were limited in scope and partially biased by the author's perspective. There are many desirable features that an incident notification system could have; however, a parsimonious list was developed on purpose in an effort to make the evaluation process concise. Additionally, the characteristics were created based on the author's understanding of the problem and review of the literature. While the criteria were verified by the advisor of this research, collaboration with more subject matter experts may expose other factors that were not considered.

Finally, there were only four types of DSSs selected for this study. While other reasoning methods exist, the ones picked were thought to be the most promising with respect to improving incident notification. However, it is possible that some potential systems were overlooked and not included for consideration. Performing the content analysis with a broader group of systems would make this research stronger.

5.3 Recommendations for Future Research

This thesis represents the initial research toward applying CBR within the cyber incident notification domain. As a result, there are several areas that require future work:

Engineering. This thesis takes a management level perspective on how a CBR system could provide relevant notification. However, an engineering view is needed to determine how the technology would actually be applied. Some important areas in CBR that require

consideration include retrieval and adaptation. Retrieval focuses on finding cases that most appropriately match the current situation. A reliable retrieval method is needed for a CBR system to present accurate cases to a user. Adaptation is a process through which cases are altered to help fit a problem more precisely. Including adaptation can enhance the usefulness of a retrieved solution. Engineering research is required to determine how a CBR system could provide these functions from an algorithmic standpoint.

Interface Design. Providing relevant notification is partly dependent on the how notifications are presented to the user. CBR was found to have a lot of flexibility when it comes to knowledge representation. While this thesis has offered some possible ideas on how cases could look, a more in-depth examination of human computer interaction principles should be conducted to aid in the refinement of a case representation design. This research should also work in conjunction with the engineering efforts outlined above to determine the technical feasibility of applying specific designs within a CBR system.

Implementing/Testing. A recent experiment conducted by a member of the CIMIA research team found that a cyber incident notification system can provide benefit to a decision maker when compared to the current USAF notification process. However, the specifics about the system were not a focus of that research. This thesis complements that study by supporting CBR as a way to help shape the incident notification system. Therefore, a new experiment can be created to implement CBR reasoning methods. Open source CBR software scripts are available and could be used for initial testing. If proved useful, these scripts could be modified over time, with the help of engineering input, to ultimately create a system that is appropriately molded to the cyber incident notification context.

Appendix A: Coder Training Handout

Improving the Relevance of Cyber Incident Notification for Mission Assurance: A Request for Subject Matter Expert Input

Introduction:

Military organizations continue to embed Information and Communication Technology (ICT) into their core mission processes as a means to increase their operational efficiency, exploit automation, reduce response times, improve decision quality, minimize costs, maximize profit, and shorten the kill chain. This dependence can place mission operations at risk when the loss, corruption, or degradation of the confidentiality, integrity, and/or availability of a critical information resource, system, or infrastructure device occurs. It is desirable to notify all organizations whose mission is critically dependent upon the impacted information resource so they can take appropriate contingency measures to assure their mission operations. Unfortunately, the existing incident notification process within the United States Air Force has several limitations which severely limit the usefulness of incident notification. Research is being conducted on how decision support technologies can help improve these deficiencies. Specifically, the focus is on improving the relevance of notification.

Four decision support systems (DSSs) have been identified as having the potential to provide benefit: rules-based systems (RBSs), case-based reasoning (CBR) systems, Bayesian networks (BNs), and neural networks (NNs). Additionally, seven desired characteristics have been identified as essential to improving relevant notification. These characteristics along with the four DSSs have been combined into a table that is presented on the next page. Subject matter experts are needed to code this table according the directions listed below. The data collected from this survey will be used to help advance the initial design considerations for a cyber incident notification system.

Directions:

The four pages after the coding table contain general information about each of the four DSSs as well as text extracts from various documents. Extracts with a "+" after them represent positive findings that while extracts with a "-" after them represent negative findings. Please read this information before coding to gain an understanding about the different systems. After reviewing the information, feel free to use your own experience to supplement what has been provided. Finally, please code the table on the following page. For each characteristic, rate each DSS on its ability to support that characteristic using a 5 point ordinal scale shown below*:

Code	Definition
1	DSS does not support this characteristic
2	DSS scarcely supports this characteristic
3	DSS moderately supports this characteristic
4	DSS greatly supports this characteristic
5	DSS fully supports this characteristic

* DSSs are able to have the same score for any given characteristic. For example, a BN and NN could both be coded as a 2 for the *Adaptable to Environment* metric.

Coding Table

Desired Characteristics	Definition	Decision Support System			
		RBS	CBR	BN	NN
Adaptable to Environment	Ability of the system to continually provide accurate information over time; flexible to change				
Functions with Uncertainty	Ability of the system to provide benefit when decision making information is uncertain or missing				
Facilitates Knowledge Acquisition	Ease at which the system allows any user (i.e. domain experts to novices) to enter new knowledge into its repository				
Low Maintainability	Ease at which the system allows users to maintain the knowledge base				
Provides Information Depth	Ability of the system to provide sufficient and focused information to a decision maker (i.e. problem, solutions, additional context)				
Presents Information Clearly	Ability of the system to display information in a way that is easy to understand				
Provides Tangible Information	Ability of the system to provide definite proven information (i.e. scenarios or hard data)				

Decision Support System Summaries

Rules-Based Systems (RBSs)

These systems use rules in the form of "IF-THEN" statements to reason, where "IF" is a condition and "THEN" is an action. In general, a RBS is made up of two components, a knowledge base and an inference engine. The knowledge base is where the rules about a certain domain are stored. The inference engine is the mechanism for interpreting the rules and producing results from user input (Bramer, 1982; Hayes-Roth, 1985).

Text Extracts:

"Rules can encapsulate small chunks of knowledge that collectively can model a complex problem" (Watson, 1997, p. 7). (+)

The skill of RBSs "increases at a rate proportional to the enlargement of their knowledge bases" (Hayes-Roth, 1985, p. 921). (+)

"Rules are a part of everyday life, and so again people can relate to them" (Watson, 1997, p. 10). (+)

"Rules can be placed in any order in a program" (Watson, 1997, p. 7). (+)

"The user (expert) can ask WHY (to query the significance of a request by MYCIN [an RBS for medical diagnosis] for information) or HOW (to ask how deductions so far considered as established by MYCIN were arrived at)" (Bramer, 1981, p. 14). (+)

"Real-world situations are often fuzzy and do not match exactly with rule premises and conclusions" (Dutta & Bonissone, p. 166). (-)

"Because a rule-based system rigidly matches rules to a problem description, a missing rule halts the reasoning process" (Hennessy & Hinkle, 1992, p. 25). (-)

"In rule-based reasoning, knowledge is extracted from experts and encoded into rules. This is often difficult to do" (Kolodner, 1993, p. 94). (-)

Case-Based Reasoning (CBR) Systems

The concept behind CBR is summarized by Riesbeck and Schank (1989): “A case-based reasoner solves new problems by adapting solutions that were used to solve old problems” (Pg. 25). CBR works well in domains that are not fully understood, or rules cannot be formed. A unique aspect about CBR is that it relies on specific knowledge from past events, instead of generalized relationships about a specific domain. Additionally, CBR is an approach that allows incremental learning. Once a new case is added to its library, it can be retrieved in the future (Aamodt & Plaza, 1994).

Text Extracts:

CBR supports "incremental, sustained learning, since a new experience is retained each time a problem has been solved, making it immediately available for future problems" (Aamodt & Plaza, 1994, p. 65). (+)

"Because new situations rarely match old ones exactly, however, old solutions must be fixed to fit new situations. In this step, called adaptation, the ballpark solution is adapted to fit the new situation" (Kolodner, 1993, p. 21). (+)

"Cases are retrieved that match the input *partially*" (Kolodner, 1993, p. 94). (+)

"Because case-based reasoners reason from complete specific episodes, CBR makes it unnecessary to decompose experiences and generalize their parts into rules" (Leake, 1996, p. 6). (+)

"Several recent studies point to the relative ease with which case-based reasoners can be built as compared to building the same rule-based systems" (Kolodner, 1993, p. 94). (+)

"Cases can come in many different shapes and sizes…" (Kolodner & Leake, 1996, p. 38). (+)

"CBR researchers believe there is one great challenge facing them-namely, adaptation. Relatively few commercial systems adapt cases..." (Watson, 1997, p. 210). (-)

"When the CBR system is initially applied for a particular problem, only a few cases will be stored in the database. This leads to a problem of 'openness'.… The system will fail to generate [a] solution, or generate very unreasonable solutions" (Chan, 2005, p. 125). (-)

Bayesian Networks (BNs)

BNs use probability theory to reason about problems in uncertain environments. Typically, the domain is modeled up front using nodes to represent variables. These nodes are also placed in a causal order. Next, each variable is assigned a conditional probability based on subject matter input or experience. These probabilities can be used to support reasoning tasks such as diagnosis and prediction (Jensen & Nielsen, 2007; Korb & Nicholson, 2004). Information provided from these networks is usually quantitative in nature (i.e. probability of an even occurring).

Text Extracts:

"Bayesian networks provide full representations of probability distributions over their variables. That implies that they can be conditioned upon any subset of their variables, supporting any direction of reasoning" (Korb & Nicholson, 2004, p. 34). (+)

"A BBN can answer queries, or 'what-if' questions, about the variables that appear in the network" (Lauría & Duchessi, 2006, p. 1575). (+)

BNs "provide a visual representation that facilitates reasoning and enhances shared understanding of complex situations" (Falzon, 2006, p. 632). (+)

"The implementation factors [nodes] that appear here may change over time… necessitating the development of another data set and BBN [Bayesian Belief Network] model" (Lauría & Duchessi, 2006, p. 1586). (-)

"First, causal relations are not always obvious.… Furthermore, causality is not a well understood concept. Is a causal relation a property of the real world or rather, is it a concept in our minds helping us to organize our perception of the world?" (Jensen & Nielson, 2007, p. 60). (-)

Neural Networks (NNs)

Neurons are cells in the brain that produce electrical signals and are thought to allow information-processing capability. Thus, ANNs aim to model this behavior. Networks can consist of multiple layers. Adding layers increases the reasoning ability; however, selecting the right amount of hidden neurons is not well understood (Russell & Norvig, 2003). NNs work well in domains that are not well understood. Once network inputs and outputs are established, the network is "trained" using existing data. These networks can adapt using the training data.

Text Extracts:

"Neural networks have a built-in capability to *adapt* their synaptic weights to changes in the surrounding environment" (Haykin, 1994, p. 4). (+)

"Owing to the distributed nature of information in the network, the damage [to neurons] has to be extensive before the overall response of the network is degraded seriously" (Haykin, 1994, p. 5). (+)

"Decision support applications are typically hampered by low structurability and noisy or missing data. In contrast to traditional DSS resources, neural networks are quite oblivious to both limitations" (Schocken & Ariav, 1994, p. 412). (+)

"During development, the neural network was trained on data collected from the operating environment, probably over a limited period of time. For some applications, the operating environment will change over time, possibly leading to a reduction of performance" (Tarassenko, 1998, p. 48). (-)

"No clear rules or design guidelines for arbitrary application" (Schalkoff, 1997, p. 10). (-)

"A sufficient number of training examples is required to ensure that the neural network is trained to recognize and respond to the full range of conditions" (Tarassenko, 1998, p. 69). (-)

"Neural computing is extremely convoluted, and therefore it is difficult to explain or defend the system's 'rationale' (unlike expert systems, where one can trace reasoning chains or invoke some sort of belief calculus)" (Schocken & Ariav, 1994, p. 402). (-)

Appendix B: Supporting Text Extracts

Extracts with a "+" after them represent findings that positively support a characteristic while extracts with a "-" after them negatively support a characteristic.

Desired Characteristic: ***Adaptable to Environment***

DSS	Code	Text Extracts
RBS	3	"Rules can encapsulate small chunks of knowledge that collectively can model a complex problem" (Watson, 1997, p. 7). (+) The skill of RBSs "increases at a rate proportional to the enlargement of their knowledge bases" (Hayes-Roth, p. 921). (+) "Real-world situations are often fuzzy and do not match exactly with rule premises and conclusions" (Dutta & Bonissone, p. 166). (-)
CBR	4	CBR supports "incremental, sustained learning, since a new experience is retained each time a problem has been solved, making it immediately available for future problems" (Aamodt and Plaza, 1994, p. 65). (+) "Because new situations rarely match old ones exactly, however, old solutions must be fixed to fit new situations. In this step, called adaptation, the ballpark solution is adapted to fit the new situation" (Kolodner, 1993, p. 21). (+) "CBR researchers believe there is one great challenge facing them-namely, adaptation. Relatively few commercial systems adapt cases..." (Watson, 1997, p. 210). (-)
BN	1	"since the structure of a COG model is so important for accurate analysis users are encouraged to check it carefully before moving on to populate the model [with conditional probabilities]" (Falzon, 2004, p. 637). (-) "the implementation factors [nodes] that appear here may change over time… necessitating the development of another data set and BBN [Bayesian Belief Network] model" (Lauría & Duchessi, 2006, p. 1586). (-)
NN	3	"neural networks have a built-in capability to *adapt* their synaptic weights to changes in the surrounding environment" (Haykin, 1994, p. 4). (+) "During development, the neural network was trained on data collected from the operating environment, probably over a limited period of time. For some applications, the operating environment will change over time, possibly leading to a reduction of performance" (Tarassenko, 1998, p. 48). (-)

Desired Characteristic: ***Functions with Uncertainty***

DSS	Code	Text Extracts
RBS	2	"Rules are retrieved that match the input *exactly*" (Kolodner, 1993, p. 94). (-) "Because a rule-based system rigidly matches rules to a problem description, a missing rule halts the reasoning process" (Hennessy & Hinkle, 1992, p. 25). (-)
CBR	3	"Cases are retrieved that match the input *partially*" (Kolodner, 1993, p. 94). (+) "When the CBR system is initially applied for a particular problem, only a few cases will be stored in the database. This leads to a problem of 'openness'…. The system will fail to generate [a] solution, or generate very unreasonable solutions" (Chan, 2005, p. 125). (-)
BN	5	"Bayesian networks provide full representations of probability distributions over their variables. That implies that they can be conditioned upon any subset of their variables, supporting any direction of reasoning" (Korb & Nicholson, 2004, p. 34). (+) "A BBN can answer queries, or 'what-if' questions, about the variables that appear in the network" (Lauría & Duchessi, 2006, p. 1575). (+)
NN	4	"Most connectionist models naturally extend to cases where some inputs are unknown…. Because cells [i.e. neurons] can easily examine large numbers of inputs, they naturally tend to be less sensitive to noise; the greater number of correct input variables can outvote the fewer number of incorrect input values" (Gallant, 1993, p. 10). (+) "Owing to the distributed nature of information in the network, the damage [to neurons] has to be extensive before the overall response of the network is degraded seriously" (Haykin, 1994, p. 5). (+) "Decision support applications are typically hampered by low structurability and noisy or missing data. In contrast to traditional DSS resources, neural networks are quite oblivious to both limitations" (Schocken & Ariav, 1994, p. 412). (+)

Desired Characteristic: ***Facilitates Knowledge Acquisition***

DSS	Code	Text Extracts
RBS	4	"Rules are a part of everyday life, and so again people can relate to them" (Watson, 1997, p. 10). (+) "Rules can be placed in any order in a program" (Watson, 1997, p. 7). (+) "In rule-based reasoning, knowledge is extracted from experts and encoded into rules. This is often difficult to do" (Kolodner, 1993, p. 94). (-)
CBR	5	"Because case-based reasoners reason from complete specific episodes, CBR makes it unnecessary to decompose experiences and generalize their parts into rules" (Leake, 1996, p. 6). (+) "Experts who are resistant to attempts to distill a set of domain rules are often eager to tell their 'war stories' - the cases they have encountered" (Leake, 1996, p. 6). (+) "Several recent studies point to the relative ease with which case-based reasoners can be built as compared to building the same rule-based systems" (Kolodner, 1993, p. 94). (+)
BN	2	"First, causal relations are not always obvious.… Furthermore, causality is not a well understood concept. Is a causal relation a property of the real world or rather, is it a concept in our minds helping us to organize our perception of the world?" (Jensen & Nielson, 2007, p. 60). (-) "… populated their conditional probability tables with 'rough guess' values based on information we had obtained from domain experts and literature" (Hudson et al., 2002, p. 5). (-)
NN	2	"No clear rules or design guidelines for arbitrary application" (Schalkoff, 1997, p. 10). (-) "A sufficient number of training examples is required to ensure that the neural network is trained to recognize and respond to the full range of conditions" (Tarassenko, 1998, p. 69). (-)

Desired Characteristic: ***Low Maintainability***

DSS	Code	Text Extracts
RBS	3	"Rules are a part of everyday life, and so again people can relate to them" (Watson, 1997, p. 10). (+) "In rule-based reasoning, knowledge is extracted from experts and encoded into rules. This is often difficult to do" (Kolodner, 1993, p. 94). (-)
CBR	4	"Because case-based reasoners reason from complete specific episodes, CBR makes it unnecessary to decompose experiences and generalize their parts into rules" (Leake, 1996, p. 6). (+) "Experts who are resistant to attempts to distill a set of domain rules are often eager to tell their 'war stories' - the cases they have encountered" (Leake, 1996, p. 6). (+)
BN	2	"The implementation factors [nodes] that appear here may change over time… necessitating the development of another data set and BBN [Bayesian Belief Network] model" (Lauría & Duchessi, 2006, p. 1586). (-) "… populated their conditional probability tables with 'rough guess' values based on information we had obtained from domain experts and literature" (Hudson et al., 2002, p. 5). (-)
NN	2	"No clear rules or design guidelines for arbitrary application" (Schalkoff, 1997, p. 10). (-) "During development, the neural network was trained on data collected from the operating environment, probably over a limited period of time. For some applications, the operating environment will change over time, possibly leading to a reduction of performance" (Tarassenko, 1998, p. 48). (-)

Desired Characteristic: ***Provides Information Depth***

DSS	Code	Text Extracts
RBS	4	"…provide[s] explanations or justifications for conclusions reached" (Hayes-Roth et al., 1983, p. 5). (+) With the RBS TEIRESIAS "the user (expert) can ask WHY (to query the significance of a request by MYCIN for information) or HOW (to ask how deductions so far considered as established by MYCIN were arrived at)" (Bramer, 1982, p. 14). (+)
CBR	5	"In case-based reasoning… the majority of intellectual emphasis has been on content issues: What kinds of content should cases have?" (Kolodner, 1993, p. 94). (+) Cases can store "names, product identifiers, values like cost or temperature, and textual notes. An increasing number of CBR tools also support multimedia features, such as photographs, sound, and video" (Watson, 1997, p. 19). (+)
BN	3	BNs "provide a visual representation that facilitates reasoning and enhances shared understanding of complex situations" (Falzon, 2006, p. 632). (+) "The BBN makes probabilistic assertions as to the outcome of certain actions" (Lauría & Duchessi, 2006, p. 1582). (-)
NN	1	"Clinicians remain wary of computer-aided diagnosis - and ANNs in particular - because many of them believe they require an insight into the system's behavior to assess the relevance of a computer-aided diagnosis decisions to a particular patient. They thus resent the 'black box' nature of ANNs" (Drew & Monson, 2000, p. 8). (-) "Neural computing is extremely convoluted, and therefore it is difficult to explain or defend the system's 'rationale' (unlike expert systems, where one can trace reasoning chains or invoke some sort of belief calculus)" (Schocken & Ariav, 1994, p. 402). (-)

Desired Characteristic: ***Presents Information Clearly***

DSS	Code	Text Extracts
RBS	3	"Rules are a part of everyday life, and so again people can relate to them" (Watson, 1997, p. 10). (+) "The control structure is relatively simple and can be understood by people other than computer scientists" (Watson, 1997, p. 10). (+) "Real-world situations are often fuzzy and do not match exactly with rule premises and conclusions" (Dutta & Bonissone, p. 166). (-)
CBR	4	"Cases can come in many different shapes and sizes, covering large or small time slices, associating solutions with problems, outcomes with situations, or both" (Kolodner & Leake, 1996, p. 38). (+) "Cases have been represented using a variety of notations" (Kolodner, 1993, p. 165). (+)
BN	3	BNs "provide a visual representation that facilitates reasoning and enhances shared understanding of complex situations" (Falzon, 2006, p. 632). (+) "This visual arrangement [of a BN] provides a convenient knowledge representation" (Lauría & Duchessi, 2006, p. 1575). (+)
NN	2	"Neural computing is extremely convoluted, and therefore it is difficult to explain or defend the system's 'rationale' (unlike expert systems, where one can trace reasoning chains or invoke some sort of belief calculus)" (Schocken & Ariav, 1994, p. 402). (-) "No general way to assess the internal operation of the network" (Schalkoff, 1997, p. 10). (-)

Desired Characteristic: ***Provides Tangible Information***

DSS	Code	Text Extracts
RBS	4	With the RBS TEIRESIAS "the user (expert) can ask WHY (to query the significance of a request by MYCIN for information) or HOW (to ask how deductions so far considered as established by MYCIN were arrived at)" (Bramer, 1981, p. 14). (+)
CBR	5	"Cases can come in many different shapes and sizes, covering large or small time slices, associating solutions with problems, outcomes with situations, or both" (Kolodner & Leake, 1996, p. 38). (+) Cases can store "names, product identifiers, values like cost or temperature, and textual notes. An increasing number of CBR tools also support multimedia features, such as photographs, sound, and video" (Watson, 1997, p. 19). (+)
BN	3	BNs "provide a visual representation that facilitates reasoning and enhances shared understanding of complex situations" (Falzon, 2006, p. 632). (+) "The BBN makes probabilistic assertions as to the outcome of certain actions" (Lauría & Duchessi, 2006, p. 1582). (-)
NN	2	"Neural computing is extremely convoluted, and therefore it is difficult to explain or defend the system's 'rationale' (unlike expert systems, where one can trace reasoning chains or invoke some sort of belief calculus)" (Schocken & Ariav, 1994, p. 402). (-) "The subject of knowledge representation inside an artificial neural network is, however, very complicated" (Haykin, 1994, p. 24). (-)

Appendix C: Coding Results from Second Coder

Desired Characteristics	Definition	Decision Support System			
		RBS	CBR	BN	NN
Adaptable to Environment	Ability of the system to continually provide accurate information over time; flexible to change	**3**	**4**	**1**	**2**
Functions with Uncertainty	Ability of the system to provide benefit when decision making information is uncertain or missing	**1**	**3**	**5**	**4**
Facilitates Knowledge Acquisition	Ease at which the system allows any user (i.e. domain experts to novices) to enter new knowledge into its repository	**3**	**4**	**2**	**1**
Low Maintainability	Ease at which the system allows users to maintain the knowledge base	**4**	**4**	**2**	**3**
Provides Information Depth	Ability of the system to provide sufficient and focused information to a decision maker (i.e. problem, solutions, additional context)	**3**	**5**	**4**	**2**
Presents Information Clearly	Ability of the system to display information in a way that is easy to understand	**3**	**5**	**4**	**1**
Provides Tangible Information	Ability of the system to provide definite proven information (i.e. scenarios or hard data)	**4**	**5**	**4**	**1**

Bibliography

Aamodt, A., & Plaza, E. (1994). Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. *AI Communications, 7*(1), 39-59.

Aha, D. W. (1998). The omnipresence of case-based reasoning in science and application. *Knowledge-Based Systems, 11*(5-6), 261-273.

Alberts, D., Garstka, J., Hayes, R., & Signori, D. (2001). *Understanding Information Age Warfare*. Washington D.C.: Command and Control Research Program.

Althoff, K., & Weber, R. (2006). Knowledge management in case-based reasoning. *The Knowledge Engineering Review, 20*(03), 305-310.

Alty, J. L. (1985). Use of expert systems. *Computer-Aided Engineering Journal, 2*(1), 2-9.

Andrade, A. D. (2009). Interpretive Research Aiming at Theory Building: Adopting and Adapting the Case Study Design. *The Qualitative Report, 14*(1), 42-60.

Ashley, K. D. (1991). Reasoning with cases and hypotheticals in HYPO. *International Journal of Man-Machine Studies, 34*(6), 753-796.

Barry, C. L. (1994). User-Defined Relevance Criteria: An Exploratory Study. *Journal of the American Society for Information Science, 45*(3), 149-159.

Barry, C. L., & Schamber, L. (1998). Users' criteria for relevance evaluation: A cross-situational comparison. *Information Processing and Management, 34*(2-3), 219-236.

Bass, S. D., & Baldwin, R. O. (2007). A Model for Managing Decision-Making Information in the GIG-Enabled Battlespace. *Air & Space Power Journal, 21*(2), 100-108.

Borlund, P. (2003). The Concept of Relevance in IR. *Journal of the American Society for Information Science and Technology, 54*(10), 913-925.

Bramer, M. A. (1982). A survey and critical review of expert systems research. In D. Michie (Ed.), *Introductory readings in expert systems*. New York: Gordon and Breach Science Publishers.

Brouard, C., & Nie, J. (2004). Relevance as resonance: a new theoretical perspective and a practical utilization in information filtering. *Information Processing and Management, 40*(1), 1-19.

Buchanan, B. G., Barstow, D., Bechtal, R., Bennett, J., Clancey, W., Kulikowski, C., et al. (1983). Constructing an expert system. In F. Hayes-Roth, D. Waterman & D. Lenat (Eds.), *Building expert systems* (pp. 127-167). Reading, MA: Addison Wesley Publishing Company.

Carter, G. M., Murray, M. P., Walker, R. G., & Walker, W. E. (1992). *Building Organizational Decision Support Systems*. San Diego, CA: Academic Press Inc.

Chan, F. T. S. (2005). Application of a hybrid case-based reasoning approach in electroplating industry. *Expert Systems with Applications, 29*(1), 121-130.

Clausewitz, C. V. (1976). *On War*. Princeton, NJ: Princeton University Press.

Cooke, N. M., & McDonald, J. E. (1986). A Formal Methodology for Acquiring and Representing Expert Knowledge. *Proceedings of the IEEE, 74*(10), 1422-1430.

Cosijn, E., & Ingwersen, P. (2000). Dimensions of relevance. *Information Processing and Management, 36*(4), 533-550.

Cuadra, C. A., Katter, R. V., Holmes, E. H., & Wallace, E. M. (1967). *Experimental Studies of Relevance Judgments: Final Report. Volume 1: Project Summary*. System Development Corp., Santa Monica, CA.

Davenport, T., & Prusak, L. (1998). *Working Knowledge: How Organizations Manage What They Know*. Boston: Harvard Business School Press.

Demri, S., Laroussinie, F., & Schnoebelen, P. (2006). A parametric analysis of the state-explosion problem in model checking. *Journal of Computer and System Sciences, 72*(4), 547-575.

Denning, D. (1999). *Information Warfare and Security*. Upper Saddle River, NJ: Pearson.

Department of Defense. (2006). *Information Operations*. JP 3-13. Washington: United States Department of Defense, Joint Chiefs of Staff, 13 February 2006.

Department of Defense. (2009). *DoD Architecture Framework Version 2.0* Volume 1: Introduction, Overview, and Concepts. Washington: United States Department of Defense, 28 May 2009.

Department of the Air Force. (2003). *Air Force Basic Doctrine*. AFDD 1. Washington: HQ USAF, 17 November 2003.

Department of the Air Force. (2005). *Enterprise Network Operations and Tracking*. AFI 33-138. Washington: HQ USAF, 28 November 2005.

Department of the Air Force. (2010). *Aircraft and Equipment Maintenance Management*. AFI21-101. Washington: HQ USAF, 26 July 2010.

Drew, P. J., & Monson, J. R. T. (2000). Artificial neural networks. *Surgery, 127*(1), 3-11.

Dutta, S., & Bonissone, P. P. (1993). Integrating Case- and Rule-Based Reasoning. *International Journal of Approximate Reasoning, 8*(3), 163-203.

Endsley, M. (1988). *Design and evaluation for situation awareness enhancement.* Paper presented at the Human Factors Society 32nd Annual Meeting, Santa Monica, CA.

Endsley, M. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors, 37*(1), 32-64.

Endsley, M., & Garland, D. (Eds.). (2000). *Situation Awareness: Analysis and Measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.

Falzon, L. (2006). Using Bayesian network analysis to support centre of gravity analysis in military planning. *European Journal of Operational Research, 170*(2), 629-643.

Freelon, D. G. (2010). ReCal: Interocder Reliability Calculation as a Web Service. *International Journal of Internet Science, 5*(1), 20-33.

Freelon, D. G. (2011). ReCal for Ordinal, Interval, and Ratio Data (OIR). Retrieved February 16, 2011, from http://dfreelon.org/utils/recalfront/recal-oir/

Gallant, S. I. (1993). *Neural network learning and expert systems*. Cambridge, MA: The MIT Press.

Gargiulo, T. L. (2006). Power of stories. *The Journal for Quality & Participation, 29*(1), 4-8.

Gemikonakli, O., Ever, E., & Kocyigit, A. (2009). Approximate solution for two stage open networks with Markov-modulated queues minimizing the state space explosion problem. *Journal of Computational and Applied Mathematics, 223*(1), 519-533.

Glenberg, A. M., & Langston, W. E. (1992). Comprehension of Illustrated Text: Pictures Help to Build Mental Models. *Journal of Memory and Language, 31*(2), 129-151.

Goodman, M. (1989). *CBR in Battle Planning.* Paper presented at the Proceedings of the Second Workshop on Case-Based Reasoning, Pensacola Beach, FL, US.

Grimaila, M. R., & Fortson, L. W. (2008). Improving the Cyber Incident Damage and Mission Impact Assessment. *IAnewsletter, 11*(1), 10-15.

Grimaila, M. R., Fortson, L. W., & Mills, R. F. (2009). *Developing Methods for Timely and Relevant Mission Impact Estimation.* Paper presented at the Proceedings of the 2009 SPIE Defense, Security and Sensing Conference Orlando, FL.

Grimaila, M. R., Fortson, L. W., & Sutton, J. L. (2009). *Design Considerations for a Cyber Incident Mission Impact Assessment (CIMIA) Process.* Paper presented at the Proceedings of the 2009 International Conference on Security and Management, Las Vegas, NV.

Grimaila, M. R., Schechtman, G., & Mills, R. F. (2009). *Improving Cyber Incident Notification in Military Operations.* Paper presented at the Proceedings of the 2009 Industrial Engineering Research Conference, Miami, FL.

Hajmeer, M. N., & Basheer, I. A. (2003). A hybrid Bayesian-neural network approach for probabilistic modeling of bacterial growth/no-growth interface. *International Journal of Food Microbiology, 82*(3), 233-243.

Hammond, K. (1986). CHEF: A Model of Case-based Planning. *Proceedings of AAAI-86*.

Harter, S. (1992). Psychological Relevance and Information Science. *Journal of the American Society for Information Science, 43*(9), 602-615.

Hatzilygeroudis, I., & Prentzas, J. (2004). Integrating (rules, neural networks) and cases for knowledge representation and reasoning in expert systems. *Expert Systems with Applications, 27*(1), 63-75.

Hayes-Roth, F. (1985). Rule-Based Systems. *Communications of the ACM, 28*(9), 921-932.

Hayes-Roth, F., Waterman, D. A., & Lenat, D. B. (Eds.). (1983). *Building expert systems*. Reading, MA: Addison-Wesley.

Haykin, S. (1994). *Neural Networks: A Comprehensive Foundation*. New York: Macmillan College Publishing Company.

He, W., Wang, F.-K., Means, T., & Xu, L. D. (2009). Insight into interface design of web-based case-based reasoning retrieval systems. *Expert Systems with Applications, 36*(3, Part 2), 7280-7287.

He, W., Xu, L. D., Means, T., & Wang, P. (2009). Integrating Web 2.0 with the Case-Based Reasoning Cycle: A systems Approach. *Systems Research and Behavioral Science, 26*(6), 717-728.

Hennessy, D., & Hinkle, D. (1992). Applying Case-Based Reasoning to Autoclave Loading. *IEEE Expert, 7*(5), 21-26.

Hudson, L., Ware, B., Mahoney, S., & Laskey, K. (2002). *An Application of Bayesian Networks to Antiterrorism Risk Management for Military Planners*. Technical Report, Digital Sandbox, Inc.

Jakobson, G., Lewis, L., Buford, C., & Sherman, C. (2004). *Battlespace Situation Analysis: The Dynamic CBR Approach*. Paper presented at the MILCOM 2004 - 2004 IEEE Military Communications Conference.

Jensen, F., & Nielsen, T. (2007). *Bayesian Networks and Decision Graphs*. New York: Springer

Kalpic, B., & Bernus, P. (2006). Business process modeling through the knowledge management perspective. *Journal of Knowledge Management, 10*(3), 40-56.

Kim, J., & Gil, Y. (2007). Incorporating tutoring principles into interactive knowledge acquisition. *International Journal of Human-Computer Studies, 65*(10), 852-872.

Kolodner, J. (1992). An Introduction to Case-Based Reasoning. *Artificial Intelligence Review, 6*(1), 3-34.

Kolodner, J. (1993). *Case-Based Reasoning*. San Mateo, CA: Morgan Kaufmann.

Kolodner, J., & Leake, D. B. (1996). A Tutorial Introduction to Case-Based Reasoning. In D. B. Leake (Ed.), *Case-Based Reasoning: Experiences, Lessons, and Future Directions*. Menlo Park, CA: AAAI Press/MIT Press.

Korb, K., & Nicholson, A. (2004). *Bayesian Artificial Intelligence*. Boca Raton: Chapman & Hall/CRC.

Koton, P. (1988). A medical reasoning program that improves with experience. *Computer Methods and Programs in Biomedicine, 30*(2-3), 177-184.

Krippendorff, K. (2004). *Content analysis: an introduction to its methodology*. Thousand Oaks, CA: Sage Publications, Inc.

Kumar, K. A., Singh, Y., & Sanyal, S. (2009). Hybrid approach using case-based reasoning and rule-based reasoning for domain independent clinical decision support in ICU. *Expert Systems with Applications, 36*(1), 65-71.

Ladd, D. A., Datta, A., & Sarker, S. (2010). *Trying to Outrun a Speeding Environment: Developing "High-velocity" Strategic DSS Evaluation Criteria.* Paper presented at the Proceedings of the Sixteenth Americas Conference on Information Systems, Lima, Peru.

Lauría, E. J. M., & Duchessi, P. J. (2006). A Bayesian Belief Network for IT implementation decision support. *Decision Support Systems, 42*(3), 1573-1588.

Leake, D. B. (1996). CBR in Context: The Present and Future. In D. B. Leake (Ed.), *Case-Based Reasoning: Experiences, Lessons, and Future Directions*. Menlo Park, CA: AAAI Press/MIT Press.

Leonhard, R. R. (1998). *The Principles of War for the Information Age*. New York: The Ballantine Publishing Group.

Mack, N., Woodsong, C., MacQueen, K. M., Guest, G., & Namey, E. (2005). *Qualitative Research Methods: A Data Collector's Field Guide*. Research Triangle Park, NC: Family Health International.

McDermott, J. (1982). R 1: A rule-based configurer of computer systems. *ARTIFICIAL INTELLIG., 19*(1), 39-88.

Michie, D. (Ed.). (1982). *Introductory readings in expert systems*. New York: Gordon and Breach Science Publishers.

Mizzaro, S. (1997). Relevance: The whole History. *Journal of the American Society for Information Science, 48*(9), 810-832.

Neuendorf, K. A. (2002). *The Content Analysis Guidebook*. Thousand Oaks, CA: Sage Publications.

Neuman, L. W. (2006). *Social research methods: qualitative and quantitative approaches*. Boston, MA: Pearson Education, Inc.

Park, T. (1993). The Nature of Relevance in Information Retrieval: An Empirical Study. *The Library Quarterly, 63*(3), 318-351.

Patton, M. Q. (2002). *Qualitative Research and Evaluation Methods*. Thousand Oaks, CA: Sage Publications.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. San Francisco, CA: Morgan Kaufmann.

Pearlmutter, B. (1989). Learning State Space Trajectories in Recurrent Neural Networks. *Neural Computation, 1*, 263-269.

Pennington, N., & Hastie, R. (1988). Explanation-Based Decision Making: Effects of Memory Structure on Judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 521-533.

Pipkin, D. L. (2000). *Information Security: Protecting the Global Enterprise*. Upper Saddle River, NJ: Prentice-Hall, Inc. .

Puppe, F., & Gappa, U. (1992). Towards Knowledge Acquisition by Experts. In F. Belli & F. Radermacher (Eds.), *Industrial and Engineering Applications of Artificial Intelligence and Expert Systems* (Vol. 604, pp. 546-555): Springer Berlin / Heidelberg.

Quinlan, J. R. (1982). Fundamentals of the knowledge engineering problem. In D. Michie (Ed.), *Introductory readings in expert systems*. New York: Gordon and Breach Science Publishers.

Riesbeck, C., & Schank, R. (1989). *Inside Case-Based Reasoning*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Romanycia, M. H. J., & Pelletier, F. J. (1985). What is a heuristic? *Computational Intelligence, 1*(1), 47-58.

Russell, S., & Norvig, P. (2003). *Artificial Intelligence: A Modern Approach*. Upper Saddle River, NJ: Pearson Education

Saracevic, T. (1975). Relevance: A review of and a framework for the thinking on the notion in information science. *Journal of the American Society for Information Science, 26*(6), 321-343.

Saracevic, T. (2007). Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science. Part II: Nature and Manifestations of Relevance. *Journal of the American Society for Information Science and Technology, 58*(13), 1915-1933.

Schalkoff, R. J. (1997). *Artificial neural networks*. New York: McGraw-Hill

Schamber, L. (1991). *Users' criteria for evaluation in a multimedia environment.* Paper presented at the Proceedings of the 54th Annual Meeting of the American Society for Information Science.

Schamber, L., Eisenberg, M. B., & Nilan, M. S. (1990). A re-examination of relevance: toward a dynamic, situational definition. *Information processing & management, 26*(6), 755-776.

Schocken, S., & Ariav, G. (1994). Neural networks for decision support: Problems and opportunities. *Decision Support Systems, 11*(5), 393-414.

Shneiderman, B., & Plaisant, C. (2004). *Designing the user interface: strategies for effective human-computer interaction*. Boston: Pearson/Addison Wesley.

Shortliffe, E. H. (1976). *Computer-based medical consultations: MYCIN*. New York: Elsevier

Shu-Hsien, L. (2005). Expert system methodologies and applications--a decade review from 1995 to 2004. *Expert Systems with Applications, 28*(1), 93-103.

Suwa, M., Scott, A. C., & Shortliffe, E. H. (1982). An Approach to Verifying Completeness and Consistency in a Rule-Based Expert System. *AI Magazine, 3*(4), 16-21.

Swanson, D. R. (1986). Subjective versus objective relevance in bibliographic retrieval systems. *Library Quarterly, 56*(4), 389-398.

Tarassenko, L. (1998). *A Guide to Neural Computing Applications*. London: Arnold Publishers.

Trip Advisor. (2011). Retrieved February 20, 2011, from http://www.tripadvisor.com/

Tung, Y.-H., Tseng, S.-S., Weng, J.-F., Lee, T.-P., Liao, A. Y. H., & Tsai, W.-N. (2010). A rule-based CBR approach for expert finding and problem diagnosis. *Expert Systems with Applications, 37*(3), 2427-2438.

Watson, I. (1997). *Applying Case-Based Reasoning: Techniques for Enterprise Systems*. San Francisco, CA: Morgan Kaufmann

Watson, I. (1999). Case-based reasoning is a methodology not a technology. *Knowledge-Based Systems, 12*(5-6), 303-308.

Watson, I. (2003). *Applying Knowledge Management: Techniques for Building Corporate Memories*. San Francisco, CA: Morgan Kaufmann.

Weber, R. O., & Aha, D. W. (2002). Intelligent delivery of military lessons learned. *Decision Support Systems, 34*(3), 287-304.

Wong-Jiru, A., Colombi, J., Suzuki, L., & Mills, R. F. (2007). *Graph Theoretical Analysis of Network Centric Operations Using Multi-Layer Models*. Paper presented at the 5th Annual Conference on Systems Engineering Research.

Yang, B. S., Han, T., & Kim, Y. S. (2004). Integration of ART-Kohonen neural network and case-based reasoning for intelligent fault diagnosis. *Expert Systems with Applications, 26*(3), 387-395.

REPORT DOCUMENTATION PAGE

Form Approved

OMB No. 074-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From – To)*
24-03-2011	**Master's Thesis**	April 2010 – March 2011

4. TITLE AND SUBTITLE

Improving the Relevance of Cyber Incident Notification for Mission Assurance

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Woskov, Stephen M., Capt, USAF

5d. PROJECT NUMBER
10ENV297

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)
Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB OH 45433-776

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT/GIR/ENV/11-M06

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Douglas Kelly, PhD, Cyber Team Lead
Air Force Research Laboratory
711th Human Performance Wing
Sense-making and Organizational Effectiveness Branch (RHXS)
2698 G Street, Bldg 190
Wright-Patterson AFB OH 45433-7604
Comm: (937) 656-4391

10. SPONSOR/MONITOR'S ACRONYM(S)

AFRL/HPW/RHXS

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

13. SUPPLEMENTARY NOTES
This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

14. ABSTRACT

Military organizations have embedded Information and Communication Technology (ICT) into their core mission processes as a means to increase operational efficiency, improve decision making quality, and shorten the kill chain. This dependence can place the mission at risk when the loss, corruption, or degradation of the confidentiality, integrity, and/or availability of a critical information resource occurs. Since the accuracy, conciseness, and timeliness of the information used in decision making processes dramatically impacts the quality of command decisions, and hence, the operational mission outcome; the recognition, quantification, and documentation of critical mission-information resource dependencies is essential for the organization to gain a true appreciation of its operational risk. This research identifies existing decision support systems and evaluates their capabilities as a means for capturing, maintaining and communicating mission-to-information resource dependency information in a timely and relevant manner to assure mission operations. This thesis answers the following research question: Which decision support technology is the best candidate for use in a cyber incident notification system to overcome limitations identified in the existing United States Air Force cyber incident notification process?

15. SUBJECT TERMS
Cyber Incident, Notification, Relevance, Decision Support Systems, Case-Based Reasoning

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
REPORT	**ABSTRACT**	**c. THIS PAGE**			**Michael R. Grimaila, PhD, CISSP, CISM**
U	**U**	**U**	**UU**	100	**19b. TELEPHONE NUMBER** *(Include area code)* (937) 255-3636 X4800, Michael.grimaila@afit.edu

Standard Form 298 (Rev: 8-98)
Prescribed by ANSI Std. Z39-18