HAL
archives-ouvertes.fr

Cyber Situation Awareness: Modeling the Security Analyst in a Cyber-Attack Scenario through Instance-Based Learning

Varun Dutt, Young-Suk Ahn, Cleotilde Gonzalez

▶ **To cite this version:**

Varun Dutt, Young-Suk Ahn, Cleotilde Gonzalez. Cyber Situation Awareness: Modeling the Security Analyst in a Cyber-Attack Scenario through Instance-Based Learning. Yingjiu Li. 23th Data and Applications Security (DBSec), Jul 2011, Richmond, VA, United States. Springer, Lecture Notes in Computer Science, LNCS-6818, pp.280-292, 2011, Data and Applications Security and Privacy XXV. <10.1007/978-3-642-22348-8_24>. <hal-01586581>

HAL Id: hal-01586581
https://hal.inria.fr/hal-01586581

Submitted on 13 Sep 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution| 4.0 International License

Cyber Situation Awareness: Modeling the Security Analyst in a cyber-attack scenario through Instance-based Learning

Varun Dutt[1,1], Young-Suk Ahn[1], Cleotilde Gonzalez[1],

[1] Dynamic Decision Making Laboratory, 4609 Winthrop Street, Pittsburgh, PA, 15213, USA
varundutt@cmu.edu, ysahn@altenia.com, coty@cmu.edu

Abstract. In a corporate network, the situation awareness (SA) of a security analyst is of particular interest. A security analyst is in charge of observing the online operations of a corporate network (e.g., an online retail company with an external webserver and an internal fileserver) from threats of random or organized cyber-attacks. The current work describes a cognitive Instance-based Learning (IBL) model of the recognition and comprehension processes of a security analyst in a simple cyber-attack scenario. The IBL model first recognizes cyber-events (e.g., execution of a file on a server) in the network based upon events' situation attributes and the similarity of events' attributes to past experiences (instances) stored in analyst's memory. Then, the model reasons about a sequence of observed events being a cyber-attack or not, based upon instances retrieved from memory and the risk-tolerance of a simulated analyst. The execution of the IBL model generates predictions of the recognition and comprehension processes of security analyst in a cyber-attack. An analyst's decisions are evaluated in the model based upon two cyber SA metrics of accuracy and timeliness of analyst's decision actions. Future work in this area will focus on collecting human data to validate the predictions made by the model.

Keywords: cyber-situation awareness; cyber-attack; dynamic decision-making; instance-based learning theory; intrusion-detection system; security analyst; threat event.

1 Introduction

Recently, President Barack Obama declared that the "cyber threat is one of the most serious economic and national security challenges we face as a nation" [1]. According to his office, the nation's cyber-security strategy is twofold: (1) improve our resilience to cyber incidents; and, (2) reduce the cyber threat [1]. Similarly, in the United Kingdom, organizers of the London 2012 Olympic Games believe that there is an

[1] Please address all correspondence about this paper to: Varun Dutt, Dynamic Decision Making Laboratory, 4609 Winthrop Street, Pittsburgh, PA, 15213. Phone: +1-412-628-1379; Fax: +1-412-268-6938; Email: varundutt@cmu.edu

increased danger of cyber-attacks that could fatally undermine the technical network that supports everything from recording world records to relaying results to commentators at the Games [2]. At the lowest level, meeting both the objectives of the Whitehouse and those of the organizers of the Olympic Games in a corporate network requires cyber situation-awareness (SA), a three stage process which includes: recognition (or the awareness of the current situation in the network); comprehension (or the awareness of malicious behavior in the current situation in the network); and, projection (assessment of possible future courses of action resulting from the current situation in the network) [3, 4].

In cyber SA, the ability of a security system to protect itself from a cyber-attack without any interventions from a human decision-maker is still a distant dream [5]. Thus, the role of human decision-makers in security systems is one that is crucial and indispensible [6, 7]. A key role in the cyber-security process is that of a security analyst: a decision-maker who is in charge of observing the online operations of a corporate network (e.g., an online retail company with an external webserver and an internal fileserver) from threats of random or organized cyber-attacks. The purpose of this paper will be to describe a cognitive model of the recognition and comprehension processes of a security analyst, where the model is based on the Instance-Based Learning Theory (IBLT, hereafter, IBL model) [8]. Furthermore, we evaluate the IBL model of the security analyst using two cyber SA measures: accuracy and timeliness [5]. The IBLT is well suited to modeling the decisions of a security analyst as the theory provides a generic decision-making process that starts by recognizing and generating experiences through interaction with a changing decision environment, and closes with the reinforcement of experiences that led to good decision outcomes through feedback from the decision environment.

2 A Simple Scenario of a Cyber Attack

The cyber-infrastructure in a corporate network typically consists of a webserver and a fileserver [9, 10]. The webserver handles customer interactions on a company's webpage. However, the fileserver handles the working of many workstations that are internal to the company and that allow company employees to do their daily operations. A bidirectional firewall (firewall 1 in Figure 1) protects the path between the webserver and the company's website on the Internet. Thus, firewall 1 allows both the incoming "request" traffic and the outgoing "response" traffic between the company's website and the webserver. Another firewall (firewall 2 in Figure 1) protects the path between the webserver and the fileserver. Firewall 2 is a much stronger firewall than the firewall 1 as it only allows a very limited Network File System (NFS) access of the fileserver from the webserver, but an easy access of the webserver from the fileserver (this latter access allows company employees to make changes on the webserver that would later show-up on the company's website). For this cyber-infrastructure, attackers follow a sequence of an "island-hopping" attack[5, pg. 30], where the webserver is compromised first, and then the webserver is used to originate attacks on the fileserver and other company workstations (the workstations are directly connected to the fileserver).

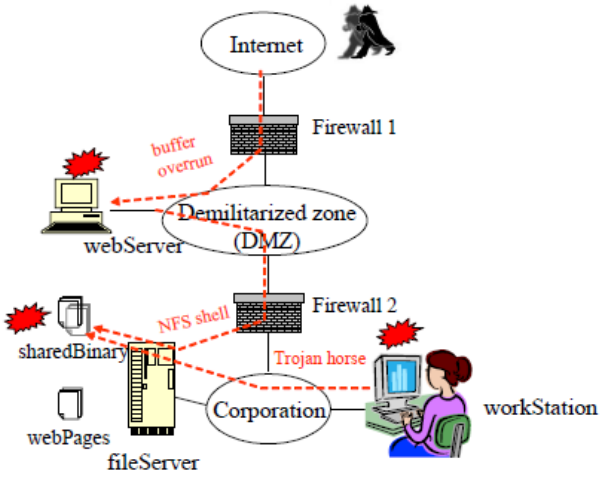

Fig. 1. A simple scenario of a cyber-attack. The attacker (shown as a black person) tries to gain access of a company's fileserver indirectly through the company's webserver. Source: [10].

Ou et al. [9] and Xie et al. [10] defined a simple scenario of an island-hopping cyber-attack within the cyber-infrastructure discussed above (see Figure 1). In the simple scenario, a security analyst is expose to a sequence of 25 network events (consisting of both threat and non-threat events) whose nature (threat or non-threat) is not precisely known to a security analyst. Out of the total of 25 events, there are 8 predefined threat events in the sequence that are initiated by an attacker. The attacker, through some of these 8 events, first compromises the webserver by remotely exploiting vulnerability on the webserver and getting a local access to the webserver. If the cyber-attack remains undetected by the security analyst by the 8^{th} event out of a total of 25 events, then the attacker gains full access of the webserver. Since typically in a corporate network and in the simple scenario, a webserver is allowed to access the fileserver through only a NFS event, the attacker then modifies data on the fileserver through the vulnerability in the NFS event. If the cyber-attack remains undetected by the security analyst by the 11^{th} event out of a total of 25, then the attacker gains full access of the file server. Once the attacker gets an access to modify files on the fileserver, he then installs a Trojan-horse program (i.e., a virus) in the executable binaries on fileserver that is used by different workstations (event 19^{th} out of 25). The attacker can now wait for an innocent user on a workstation to execute the virus program and obtain control of user's workstation (event 21^{st} out of 25).

During the course of the simple scenario, a security analyst is able to observe all the 25 events corresponding to file executions and packets of information transmitted on and between the webserver, fileserver, and different workstations. He is also able to observe alerts that correspond to some network events using an intrusion-detection system (IDS) [5]. The IDS raises an alert for a suspicious file execution or a packet transmission event that is generated on the corporate network. However, among the alerts generated by the IDS in the simple scenario, there is both a false-positive and a false-negative alert and one alert that correspond to the 8^{th} event, but which is

received by the analyst after the 13th event in the sequence (i.e., a time delayed alert). Most importantly, due to the absence of a precise alert corresponding to a potential threat event, the analyst does not have precise information on whether a network event and its corresponding alert (from the IDS) are initiated by an attacker or by an innocent company employee. Even through the analyst lacks this precise information, he needs to decide, at the earliest possible and most accurately, whether the sequence of events in the simple scenario constitutes a cyber-attack. The earliest possible or proportion of timeliness is determined by subtracting the percentage of events seen by the analyst before he makes a decision about cyber-attack in the simple scenario to the total number of events (25) in the scenario from 100%. The accuracy of the analyst is determined by whether the analyst's decision was to ignore the sequence of events or declare a cyber-attack based upon the sequence of observed network events.

3 Motivation

Prior literature has shown that the SA of a security analyst is a function of the a priori experiences and knowledge level of the analyst about a cyber-attack scenario [5], and the willingness of the analyst to take risks, i.e., analyst's risk-tolerance [11, 12]. Prior research in judgment and decision making (JDM) has also discussed how our prior experiences of events in the environment shape our decision choices [13, 14]. Typically, having a greater number of bad experiences in memory about an activity makes a decision-maker avoid the activity; whereas, good experiences with an activity boost the likelihood of a decision-maker to undertake the same activity [13, 14]. Although there is abundant literature that discusses the role of prior experiences in general and the relevance of risk-tolerance in network security, there exists lack of a study that empirically investigates the role of both these factors together on the SA of a security analyst.

We believe that an analyst's correct and timely classification of a sequence of network events in the simple scenario as a cyber-attack or not, is based upon the following two factors:

1. The knowledge level of the analyst in terms of the mix of experiences stored in analyst's memory, and,

2. The analyst's risk-tolerance level, i.e., the willingness of an analyst to classify a sequence of events as a cyber-attack.

The above two factors as well as many other cognitive factors that may limit on enhance the cyber-SA of an analyst can be studied through computational cognitive modeling. In this paper, we use IBLT to develop a model of the security analyst and we assess the effects on the two factors on the accuracy and timeliness of the analyst to detect a cyber-attack in the simple scenario.

4 Instance-based Learning Theory (IBLT) and IBL model of Security Analyst

IBLT is a theory of how people make decisions from experience in a dynamic task [8]. In the past, computational models based on IBLT have proven to be accurate in generating predictions of human behavior in many dynamic-decision making situations like those faced by the security analyst [15, 16].

IBLT proposes that people represent every decision making situation as *instances* that are stored in memory. For each decision-making situation, an instance is retrieved from memory and reused depending on the similarity of the current situation's attributes to the attributes stored in instances in memory. An instance in IBLT is composed of three parts: situation (S) (the knowledge of situation attributes in a situation event), decision (D) (the course of action to take for a situation event), and utility (U) (i.e., a measure of the goodness of a decision made or the course of action taken for a situation event).

In the case of the decision situations faced by the security analyst, these attributes are those that characterize potential threat events in a corporate network and that needs to be investigated continuously by the analyst. The situation attributes that characterize potential threat events in the simple scenario are the *IP* address of the location (webserver, fileserver, or workstation) where the event took place, the *directory* location in which the event took place, whether the IDS raised an *alert* corresponding to the event, and whether the *operation* carried out as part of the event (e.g., a file execution) by a user of the network succeeded or failed.

In the IBL model of the security analyst, an instance's S slots refers to the situation attributes defined above; the D slot refers to the decision, i.e., whether to classify a sequence of events as constituting a cyber-attack or not; and, the U slot refers to the accuracy of the classification of an situation as a threat. IBLT proposes five mental phases in a closed-loop decision making process: recognition, judgment, choice, execution, and feedback (see Figure 2). The five decision phases of IBLT represent a complete learning cycle where the theory explains how knowledge is acquired, reused, and learnt by human decision-makers. Because the focus of this study is on the recognition and comprehension process in the SA of a security analyst, we will only focus on and discuss the recognition, judgment, choice, and execution phases in the IBLT (for details on the feedback phase, refer to [8, 15]). In addition to the IBLT's decision-making process, IBLT borrowed some of the proposed statistical-learning mechanisms from a popular cognitive architecture called ACT-R [17, 18]. Thus, most of the previous cognitive models that have used IBLT were developed within the ACT-R architecture.

The IBLT's process starts in the recognition phase in search for alternatives and classification of the current situation as *typical* or *atypical*. The current situation is typical if there are memories of similar situations (i.e., instances of previous trials that are similar enough to the current situation). If the situation is typical, then in the judgment phase, the most similar instance is retrieved from memory and is used in determining the value of the expected utility of the situation being evaluated. In the IBL model of the security analyst, the decision alternatives refer to whether a sequence of events constitutes a cyber-attack or not. For the model, the determination of the utility in the judgment phase means whether to comprehend a potential network

event as a threat to the network or not. The actual determination of the utility is based upon the value in the utility slot of an instance retrieved from memory. The decision to retrieve an instance from memory for a situation event is determined based upon a comparison of the instance's memory strength, called *activation*. Thus, an instance is retrieved from memory if the instance has the highest activation among all instances in memory.

However, if the situation event in the network is atypical (i.e., no instance similar to the situation event is found in memory), then a judgment heuristic rule is applied to determine the value of the utility of a new instance corresponding to a decision alternative. In the IBL model of the security analyst, we prepopulate the memory of a simulated analyst with certain instances to start with. These are assumed to be pre-stored experiences of past situations in the analyst's memory, and thus all situation events are treated by the model as typical.

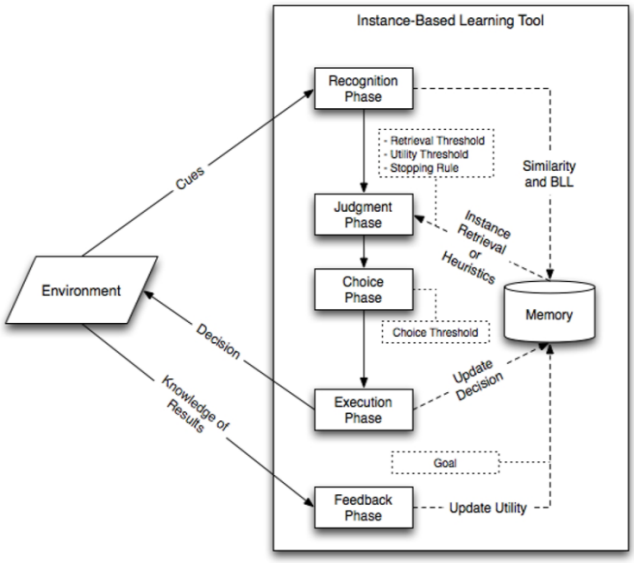

Fig. 2. The five phases of IBL theory (right) and an environment, i.e., a decision task with which a model developed according to the IBLT interacts (left).

Next, in the choice phase, a decision alternative is selected based upon the utility determined in the judgment phase (above). Thus, the choice phase in the IBL model of the security analyst consists of whether to classify a set of network events seen up to the current event in the scenario as constituting a cyber-attack, or whether to accumulate more evidence by further observing incoming situation events before such a classification could be made. According to IBLT, this decision is determined in the "necessity level" which represents a satisficing mechanism to stop search of the environment and be "satisfied" with the current evidence (e.g., the *satisficing strategy*) [19]. In the IBL model of the security analyst, we will call this parameter the "risk-tolerance level" (a free parameter) to represent the number of events the model has to classify as threats in the simple scenario before it classifies the scenario to constitute a cyber-attack. For the risk-tolerance level, each time the model

classifies a situation event in the network as a threat (based upon retrieval of an instance from memory), a counter increments and signifies an accumulation of evidence in favor of a cyber-attack. If the value of the accumulated evidence (represented by the counter) becomes equal to the analyst's risk-tolerance level, the analyst will classify the scenario as a cyber-attack based upon the sequence of already observed network events; otherwise, the model will decide to continue obtaining more information from the environment, and observe the next situation event in the network. We manipulate the risk-tolerance parameter in this study at different number of events 2, 4, or 6 (more details ahead). Regardless, the main outcome of the choice phase in the model is whether to classify a set of network events as a cyber-attack or not.

The choice phase in the model is also based upon a property of analyst to exhibit "inertia," i.e., simply not to decide to classify a sequence of observed network events as a cyber-attack due to lack of attention and continue to wait for the next situation event. The inertia in the model is governed by a free parameter called *probability of inertia* (Pinertia) [15, 20]. If the value of a random number derived from a uniform distribution between 0 and 1 is less than Pinertia, the model will choose to observe another event in the scenario and will not classify the sequence of already observed events as a cyber-attack; otherwise, the model will make a decision to classify the sequence of already observed events based upon the set risk-tolerance level. We assumed a default value of Pinertia at 0.3 (or 30%).

The choice phase is followed by the execution of the best decision alternative. The execution phase for the IBL model of the security analyst means either to classify a sequence of observed events as a cyber-attack and stop online operations in the company, or *not* to classify the sequence of events as a cyber-attack and to let the online operations of the company continue undisrupted.

In IBLT, the activation of an instance i in memory is defined using the ACT-R architecture's activation equation:

$$A_i = B_i + \sum_{l=1}^{k} P_l \times M_{li} + \varepsilon_i \quad \textbf{(1)}$$

where, B_i is the base-level learning parameter and reflects the recency and frequency of the use of the ith instance since the time it was created, which is given by:

$$B_i = ln\left(\sum_{t_i \in \{1,\ldots,t-1\}} (t - t_i)^{-d}\right) \quad \textbf{(2)}$$

.

The frequency effect is provided by $t-1$, the number of retrieval of the i^{th} instance from memory in the past. The recency effect is provided by $t - t_i$, i.e., the event since the t^{th} past retrieval of the ith instance (in equation 2, t denotes the current event number in the scenario). The d is the decay parameter and has a default value of 0.5 in the ACT-R architecture and it is the value we assume for the IBL model of the security analyst.

The $\sum_{l=1}^{k} P_l \times M_{li}$ summation is the similarity component and represents the mismatch between a situation event's attributes and the situation (S) slots of an instance *i* in memory. The *k* is the total number of attributes of a situation event that are used to retrieve the instance *i* from memory. In the IBL model of the security analyst, the value of $k = 4$, as in the simple scenario, there are 4 attributes that characterize a situation event in the network and that are also used to retrieve instances from memory. As mentioned above, these attributes are *IP*, *directory*, *alert*, and *operation* in an event. The match scale (P_l) reflects the amount of weighting given to the similarity between an instance *i*'s situation slot *l* and the corresponding situation event's attribute. P_l is generally a negative integer with a common value of -1.0 for all situation slots *k* of an instance *i*. We assume a value of -1.0 for the P_l in the IBL model of the security analyst. The M_{li} or match similarities represents the similarity between the value *l* of a situation event's attribute that is used to retrieve instances from memory and the value in the corresponding situation slots of an instance *i* in memory. Typically, M_{li} is defined using a squared distance between the situation event's attributes and corresponding situation slots in instances in memory [21]. Thus, in the IBL model of the security analyst, M_{li} is equal to sum of squared differences between a situation event's attributes and the corresponding situation slots of an instance. In order to find the sum of these squared differences, the situation events' attributes and the values in the corresponding slots of instances in memory were coded using numeric codes. Table 1 shows the codes assigned to the SDU slots of instances in memory and the situation events' attributes in the simple scenario.

Table 1. The coded values in the slots of an instance in memory and attributes of a situation event.

Attributes	Values	Codes
IP (S)	Webserver	1
	Fileserver	2
	Workstation	3
Directory (S)	Missing value	-100
	File X	1
Alert (S)	Present	1
	Absent	0
Operation (S)	Successful	1
	Unsuccessful	0
Decision (D)	Cyber-attack	1
	No Cyber-attack	0
Threat (U)	Yes	1
	No	0

Due to the $\sum_{l=1}^{k} P_l \times M_{li}$ specification, instances that encode a similar situation to the current situation event's attributes, receive a less negative activation (in equation 1). In contrast, instances that encode a dissimilar situation to the current situation event's attributes receive a more negative activation.

Furthermore, ε_i is the noise value that is computed and added to an instance *i*'s activation at the time of its retrieval attempt from memory. The noise value is characterized by a parameter *s*. The noise is defined as,

$$\varepsilon_i = s \times ln\left(\frac{1-\eta_i}{\eta_i}\right). \quad (3)$$

where, η_i is a random draw from a uniform distribution bounded in [0, 1] for an instance *i* in memory. We set the parameter *s* in an IBL model to make it a part of the activation equation (equation 1). The *s* parameter has a default value of 0.25 in the ACT-R architecture and we assume the default value of s in the IBL model of the security analyst.

5 Implementation and execution of the IBL model

The IBL model of the security analyst was created using Matlab software (the Matlab representation has already been evaluated to work similarly to the ACT-R representation, see [22]). The IBL model of the security analyst goes over a sequence of 25 network events in the simple scenario (Figure 1). We pre-populated the memory of a simulated analyst in the model with instances encoding all possible sequences of network events based upon values of events' attributes. Some of these instances in memory contained a threat value as the utility and some which do not (more information below). Unbeknownst to the model (but known to the modeler), out of the 25 events in the scenario (mentioned above), there are 8 pre-defined threat events that are executed by an attacker outside the company [9-10]. For each event in the scenario, the IBL model uses equation 1 – 3, to retrieve an instance that is most similar to the encountered event. Based upon the value of the utility slot of a retrieved instance, the situation event is classified as a threat or not a threat. Depending upon the inertia mechanism and the risk-tolerance level of a simulated analyst in the model, a decision is made to classify a sequence of observed events as a cyber-attack and stop company's online operations, or to let the company continue its online operations (no cyber-attack).

The IBL model was executed for a set of 500 repeated simulated trials of the same scenario where each simulated trial made the model to process 25 situation events in the network. For each set of 500 simulated trials, we manipulated the mix of threat and non-threat instances in memory of a simulated analyst, i.e., experience of the analyst, and the risk-tolerance level of the analyst.

The mix of threat and non-threat instances in the IBL model's memory could be one of the following three kinds: ambivalent analyst (Ambi): 50% of threat instances and 50% non-threat instances for each situation event in the scenario; an extra-careful analyst (Extra): 75% of threat instances and 25% of non-threat instances for each situation event in the scenario; and, a less-careful analyst (Less): 25% of threat instances and 75% of non-threat instances for each situation event in the scenario. The risk-tolerance level of analyst was manipulated as the following three levels: low (2 events out of a possible 25 event need to be classified as threats before the analyst classifies a sequence of observed events as cyber-attack); medium (4 events out of a

possible 25 event to be classified as threats before the analyst classifies a sequence of observed events as cyber-attack); and, High (6 events out of a possible 25 event to be classified as threats before the analyst classifies a sequence of observed events as cyber-attack).

We wanted to derive predictions of the effect of the above manipulations in the IBL model upon the performance of the analyst. The performance of a simulated analyst was measured using the accuracy and timeliness of the analyst. The accuracy was evaluated using two different cyber-SA metrics, recall and precision, and the timeliness was evaluated in the model using a single timeliness cyber-SA metric [5]. Recall is the percent of events correctly detected as threats out of the total number of known threat events observed by the model before the model stopped (Recall is the same as hit rate in Signal Detection Theory). Precision is percentage of events correctly detected as threats out of the total number of threat events detected by the model before the model stopped. Timeliness is 100%-percentage of events, out of a total 25, after which the model stops and classifies the scenario to be a cyber-attack (the timeliness could be defined as the number of events out of 25, but defining it as a percentage allows us to compare it to other two SA measures). We expected best performance for the IBL model representing an extra-careful analyst with a low risk-tolerance, and the worst performance for the IBL model representing a less-careful analyst with a high risk-tolerance. This is because an extra-careful analyst with a low risk-tolerance will be classifying network events more cautiously as constituting a cyber-attack compared to a less-careful analyst with a high risk-tolerance.

6 Results

Figure 3 shows the predictions from the IBL model of the effects of manipulating the memory and the risk-tolerance upon the performance of the security analyst. Generally, as per our expectation, an extra-careful analyst with a low risk-tolerance did better on all three performance measures compared to a less-careful analyst with a high risk-tolerance. Furthermore, risk-tolerance and memory seem to impact all three performance measures; however, the Precision is always smaller than the Recall and Timeliness measures and this observation is to do with the fact that a model that is able to retrieve more threat instances from memory not necessarily retrieves them correctly, i.e., retrieve a threat instance from memory that *always* corresponds to a network threat event. Also, the effect of memory appears to be more impacting that the risk-tolerance of the analyst.

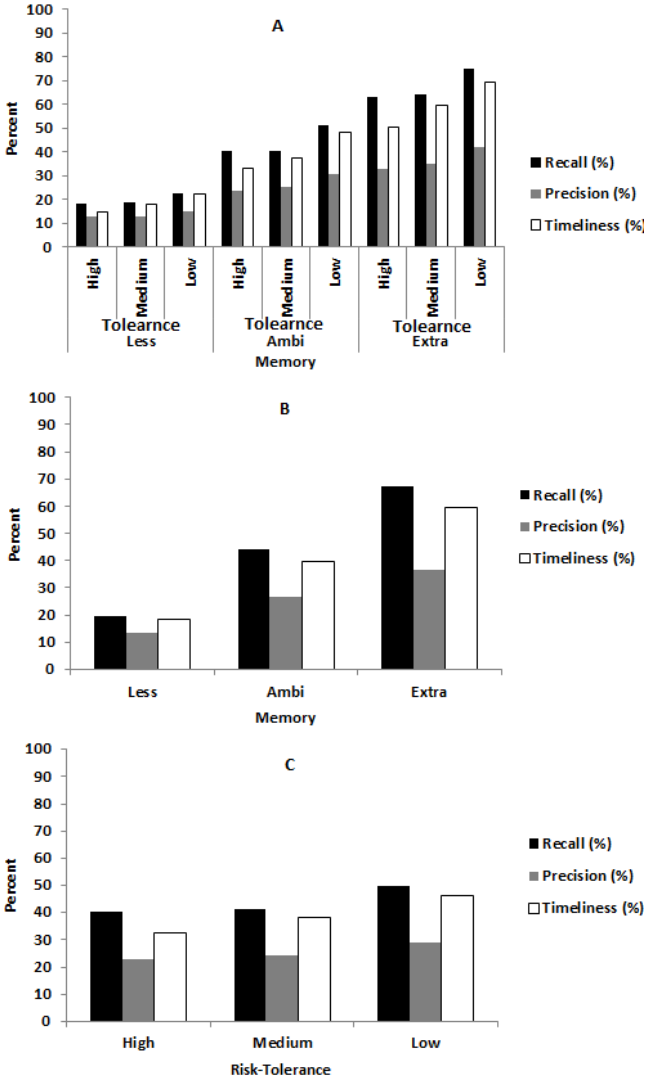

Fig. 3. (A) The interaction effect of memory and risk-tolerance on cyber SA of an analyst. (B) The effect of memory alone on cyber SA of an analyst. (C) The effect of risk-tolerance alone on cyber SA of an analyst. A greater percentage on all three cyber SA measures is more desirable as it makes the simulated analyst more efficient.

7 Discussion

In this paper, we have proposed that computational models based on the IBLT can be used to make predictions of the SA of a security analyst. Particularly, the model can

make concrete predictions of the level of recall, precision, and timeliness of the security analyst given some level of experience (in memory) and risk-tolerance.

We created an IBL model of the analyst for a simple scenario of a typical island-hopping cyber-attack. Then, using the scenario, we evaluated the performance of a simulated analyst on three commonly used measures of cyber-SA. These measures are based upon accuracy of analyst (Precision and Recall) and the timeliness of the analyst (Timeliness). Our results reveal that both the risk-tolerance level of an analyst and the mix of threat and non-threat instances in analyst's memory affect the analyst's cyber SA with the effect of the analyst's experiences (in memory) slightly more impacting compared to analyst's risk-tolerance. The less impact of the risk-tolerance factor compared to memory could be due to the nature of IBL models that are strongly dependent upon retrieval of instances from memory to make choice decisions.

When the simulated analyst is less-careful, then for any situation event the model has only a 25% chance of retrieving threat instances and 75% chance of it retrieving non-threat instances. As a consequence, the model has a lesser chance to classify actual threats in the simple scenario as threats and it takes more time for the model to accumulate evidence that is more than the risk-tolerance level (decreasing the Timeliness). However, when the simulated analyst is more-careful, then for any situation event there is a 75% chance of the model retrieving threat instances and 25% chance of it retrieving non-threat instances. As a consequence, the model has a greater chance to classify actual threats in the simple scenario as threats and it takes less time for the model to accumulate evidence that is equal to the risk-tolerance level (increasing the Timeliness).

The important aspect of the model is the fact that although the Recall and Timeliness increase as a direct function of the ability of the model to retrieve threat instances from the memory and its risk-tolerance, there is not a substantial increase in model's Precision when either of the two factors is favorable (Figure 3 A, B, and C). The slow increase in Precision is expected because a model that is able to retrieve more threat instances from memory and is less risk-tolerant, might not necessarily be more precise in its actions. However, there is still an increase in Precision with a manipulation of both factors and this suggests that making a security analyst less risk-tolerant as well as extra-careful might help increase the job-efficiency of the analyst. These are only some of the many predictions that the IBL model is able to make regarding the Cyber-SA of human analysts.

Although the current model is able to make precise predictions, these need to be validated with human data, i.e., observed behavior from a human security analyst operating in the simple scenario. We plan to run laboratory studies in the near future to assess human behavior in this simple scenario. An experimental approach will allow us to validate our model predictions and improve the relevance of the model and assumptions made in it on its free parameters. In these experimental studies, we believe that some of the interesting factors to manipulate would include the experiences of the human analyst (stored in memory). One method we have thought currently is to make the analyst read examples of more and less threat scenarios before the analyst participates in the act of detecting cyber-attacks in the simple scenario. Also, we plan to record the risk-taking and risk-averse behavior of the analyst in the study to control for the risk-tolerance factor. Thus, our next goal will be to validate the predictions from the IBL model.

If our model is able to represent the Cyber-SA of human analysts accurately, this model would have significant potential to contribute towards the design of training and decision support tools for security analysts.

8 Acknowledgements

This research was a part of a Multidisciplinary University Research Initiative Award (MURI; # W911NF-09-1-0525) from Army Research Office for a research project on Cyber Situation Awareness. We would like to thank Hau-yu Wong, Dynamic Decision Making Laboratory, for help with editorial work in the paper.

References

1. Cybersecurity, http://www.whitehouse.gov/the_press_office/Remarks-by-the-President-on-Securing-Our-Nations-Cyber-Infrastructure/
2. London 2012 Olympics faces increased cyber attack threat, http://www.guardian.co.uk/uk/2011/jan/19/london-2012-olympics-cyber-attack
3. Endsley, M.: Toward a Theory of Situation Awareness in Dynamic Systems. Hum. Fact. 37, , 32--64 (1995)
4. Tadda, G., Salerno, J. J., Boulware, D., Hinman, M., Gorton, S.: Realizing Situation Awareness Within a Cyber Environment. In: Proceedings of SPIE Vol. 6242, pp. 624204. SPIE, Kissimmee, FL (2006)
5. Jajodia, S., Liu, P., Swarup, V., Wang, C. (Eds.): Cyber Situational Awareness. Springer, New York, (2001)
6. Gardner, H.: The Mind's New Science: A History of the Cognitive Revolution. Basic Books, New York (1987)
7. Johnson-Laird, P.: How We Reason. Oxford University Press, London (2006)
8. Gonzalez, C., Lerch, J. F., Lebiere, C.: Instance-Based Learning in Dynamic Decision Making. Cog. Sci. 27(4), 591--635 (2003)
9. Ou, X., Boyer, W. F., McQueen, M. A.: A Scalable Approach to Attack Graph Generation. In: Proceedings of the 13th ACM Conference on Computer and Communications Security, pp. 336--345. ACM, Alexandria (2006)
10. Xie, P., Li, J. H., Ou X., Liu, P., Levy, R.: Using Bayesian Networks for Cyber Security Analysis. In: Proceedings of the 2010 IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), pp. 211--220. IEEE Press, Hong Kong (2010)
11. McCumber, J.: Assessing and Managing Security Risk in IT Systems: A Structured Methodology. Auerbach Publications, Boca Raton, FL (2004)
12. Salter, C., Saydjari, O., Schneier, B., Wallner, J.: Toward a Secure System Engineering Methodology. In: Proceedings of New Security Paradigms Workshop, pp. 2--10. ACM, Charlottesville (1998).
13. Lejarraga, T., Dutt, V., Gonzalez, C.: Instance-Based Learning: A General Model of Decisions from Experience in Repeated Binary Choice. J. Behav. Dec. Mak. (in press)
14. Hertwig, R., Barron, G., Weber, E. U., Erev, I.: Decisions from Experience and the Effect of Rare Events in Risky Choice. Psych. Sci. 15, 534--539 (2004)
15. Gonzalez, C., Dutt, V.: Instance-Based Learning: Integrating Decisions from Experience in Sampling and Repeated Choice Paradigms. Manuscript submitted for publication (2010)
16. Dutt, V., Cassenti, D. N., Gonzalez, C. Modeling a Robotics Operator Manager in a Tactical Battlefield. Manuscript submitted for publication (2010)
17. Anderson, J. R., Lebiere, C.: The Atomic Components of Thought. Lawrence Erlbaum Associates, Hillsdale (1998)
18. Anderson, J. R., Lebiere, C.: The Newell Test for a Theory of Mind. Behav. Brain Sci. 26, 587--639 (2003)
19. Simon, H. A., March, J. G.: Organizations. Wiley, New York (1958)

20. Gonzalez, C., Dutt, V, Lejarraja, T.: How Did an IBL Model Become the Runners-up in the Market Entry Competition? Unpublished manuscript in preparation (2011)
21. Shepard, R. N.: The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function. Psychometrika 2, 125--140 (1962)
22. Gonzalez, C., Dutt, V., Lebiere, C.: Building a New Instance-Based Learning Modeling Tool. Unpublished manuscript in preparation (2011)