Global Information Assurance Certification Paper

Predicting Control Attributes With Bayesian Networks

GIAC GCIA Gold Certification

Author: Dan Lyon, danlyon@mac.com
Advisor: Rich Graves

Accepted: Oct 1, 2013

Abstract

Attack trees have been used to formally measure system security with attributes such as attacker cost and skill. Existing work has focused on the mathematics behind calculating optimal paths through the system for an attacker; however the actual attribute values are not discussed. To use attacker attributes as a measure of system security, a valid source of data is required. The research presented by this paper was designed to help fill the gap. The research study implemented a survey to collect valid estimates of attacker skill, time and cost for a small set of information security controls. Finally, a Bayesian Network was created to illustrate a predictive model for overcoming security controls.

1. Introduction

Attack trees have been used as a mechanism to formalize security analysis of a system for over a decade (Amoroso, 1994; Schneier, 1999), and have gone through various adaptations including Defense Trees, Attack Response Trees and Attack Countermeasure Trees. Measurements applied to quantify the security of a system in conjunction with attack trees have included attacker time, skill and cost (Roy, 2010).

Software packages such as Secure*IT*ree and AttackTree+ have implemented attack trees to perform threat analysis. Attacker attributes included with the software packages are difficulty and resources required for a successful attack. While the tools provide the analysis capability, the assumptions and values for each control must be provided or validated by the security engineer.

The research performed was designed to address the difficulties in validating these assumptions by obtaining expert opinion. A subset of security controls at five different attack surfaces was identified and an online survey was used to gather expert opinion on the attacker skill, time and cost required to overcome each control. Survey participants were selected through several email distribution lists comprised of security professionals. Those responding to the survey were assessed for level of expertise and the results included in Appendix A.

In addition to published results from the survey, it is important to show how the results could be used to predict a control's attribute at design time. To show this, a predictive model was created with a Bayesian Network. A brief overview of Bayesian Networks is provided and the predictive model is described.

The information acquired and published by this research allows engineers the ability to derive or validate assumptions about control attributes. Additionally a more formalized approach to secure design is shown through the use of a Bayesian model.

2. Attack Tree History

Attack trees were derived from fault trees as a formal tree structure to analyze the security of a system. Attack trees are created from the attacker perspective, where the

Dan Lyon, danlyon@mac.com

goal of the attacker is the root node and connecting nodes represent the multiple ways the goal can be realized (Schneier, 1999).

Bistarelli, Peretti and Trubitsyna (2006) extended attack trees to create Defense Trees. Defense trees differed by placing controls at the leaf nodes. These defense nodes allowed evaluation of security investments through example metrics such as Annualized Loss Expectancy, Return on Investment and Return on Attack. While the examples showed theoretical cost to overcome controls, including a security door or video camera, no attempt was made to place a realistic value on common information security controls such as overcoming a network protocol.

Another extension of attack trees was created to focus on attack consequences by including defender response, and termed an Attack Response Tree (Zonouz, Khurana, Sanders & Yardely, 2009). Attack Response Trees added defender response to attack trees and instead of attack scenarios were focused on attack consequences. The goal of Zonouz et al's proposed Response and Recovery Engine (RRE) is to model the interaction between an attacker and the system's countermeasures. The underlying theory of the RRE model presented is that both the attacker and the defender try to maximize their benefit. For this theory to be useful in the business world, some measure of attacker expenditure must be quantified.

Two final examples of attack trees, Attack Countermeasure Trees (Roy, 2010) and Protection Trees (Edge, 2007), note that various metrics can be applied and the system optimized for these metrics. Both of these focus on the optimizations for the system and no formal research performed of what it would actually cost an attacker to overcome controls.

For any of the formal methods where a control attribute is used as a predictive metric, a valid source of data needs to be provided.

3. Research Method and Design

The basic research method was simple: identify a set of cyber security controls and ask experts their opinion on how much skill, time, and money it would take an

attacker to overcome each control. Several factors drove how the survey was designed and are described below.

The first idea that influenced the research design is the concept of Voice of the Customer (VoC) used in Six Sigma manufacturing techniques. Part of Six Sigma is identifying all the stakeholders of a system, and including them in the requirements gathering and definition process (Maass & McNair, 2009). Incorporating the malicious user as a stakeholder can be accomplished by taking the view that they should not be successful in their goals. Using this concept, the survey was created to obtain a measure of what would cause a malicious user to fail.

A second concept that influenced the research design is that overcoming a security control is a tradeoff between attacker skill, time and cost. A control that requires a low-skill individual to spend a month to overcome may only take a higher skill individual a couple days. Similarly, if an attacker has expensive hardware and software available that can automate overcoming a control, then the time required is reduced. Because the variables of cost, skill and time are related it was important to allow the participants the ability balance appropriately.

A third inspiration for the research came from the Defense Science Board's task force report Resilient Military Systems and the Advanced Cyber Threat (2013). In the report, attackers were classified in a six-tiered hierarchy. The report defined Tiers I through VI. Tier I included attackers that use known exploits with tens of dollars available in resources. Tier VI was defined as attackers that invest large sums (billions) of money and time to accomplish their goals. The scale in resources ranged from tens of dollars for Tiers I and II, to millions of dollars available for Tiers III and IV, and finally billions of dollars available for Tiers V and VI. The defined resource scale in the Defense Science Board's report quickly escalated from tens to millions between Tier I and Tier III. The sharp increase may have suited the intended audience of the defense board, but businesses may need more granular data than tens and millions.

3.1. Scenarios and Controls

Controls were divided into five subgroups based on common attack surfaces. For each control subgroup, a scenario was presented with an attacker goal that defined some

assumptions. Scenarios covered attacks on the following surfaces: hardware, software executable, local system interfaces, wireless protocols, and network authentication. Scenarios are described below in Table 1 Attack Scenarios. The attack surface and associated controls are defined in the Table 2 Attack Surface, Controls and Identifier.

Attack Surface	Scenario and Assumptions
Local System Interface	An attacker wants to acquire access to an end user system by executing malicious software or gaining shell access. The only interfaces on the target machine are a RS232 serial port, USB port and an Ethernet port.
Network Authentication	An attacker wants to compromise a secure communications channel between a user's computer and a well-managed network infrastructure. Assume the attacker does not have physical access to either endpoint. Assume the attacker is not on the local subnet. Assume all communications are using TCP/IP over the internet. Assume the certificate authority is well-managed.
Wireless	An attacker wants to obtain network access to a target network through the wireless access points. Assume the attacker does not have physical access to the target network.
Software	An attacker wants to understand a software binary on a target machine. Assume the attacker has physical access. Assume all software is running on a hardened Linux platform.
Hardware	An attacker wants to access data on a target machine. Assume the attacker has physical access. Assume the Operating System uses an encrypted file system with a symmetric encryption key.

Table 1 Attack Scenarios

Attack Surface	Control
Hardware	The encryption key is stored in a Trusted Platform Module (TPM chip) and the encrypted data stored in removable memory. Bus traffic is not encrypted.
Hardware	The encryption key is stored in a Trusted Platform Module (TPM chip) and the encrypted data stored in Non-removable memory. Bus traffic is not encrypted.
Hardware	The encryption key is stored in on-board memory and the encrypted data stored in NON-removable memory. Bus traffic is not encrypted, on-board memory is encrypted. Bus traffic protected from access via physical layer (epoxy).

Hardware	The encryption key is stored in on-board memory and the encrypted data stored in NON-removable memory. Bus traffic is not encrypted, on-board memory is not encrypted.
Hardware	The encryption key is stored in on-board memory and the encrypted data stored in removable memory. Bus traffic is not encrypted, on-board memory is not encrypted.
Software	The software is a release build compiled C program that has anti-tampering compiled into it.
Software	The software is a release build compiled C program that has complex obfuscation.
Software	A program that is a release build compiled C program, without the symbol table present.
Software	A Java program that has had simple obfuscation applied.
Software	A Java program that has not been obfuscated.
Software	Software binaries that are not encrypted or obfuscated.
Wireless	WPA-Enterprise 802.1X
Wireless	WPA-PSK(TKIP)
Wireless	Bluetooth 2.1+ EDR SSP
Wireless	Cellular-GSM
Wireless	Cellular-UMTS
Wireless	Cellular-CDMA
Wireless	WPA2-PSK (AES)
Local Interface	RS-232 Serial Port that provides access to a diagnostic debugger capability.
Local Interface	Ethernet port is available with a local operating system firewall blocking unsolicited inbound communications
Local Interface	USB port restricted to Mass Storage USB devices that have signed content
Network Authentication	Single SSL (client authenticates server certificate)
Network Authentication	Mutual SSL (client authenticates server certificate and server authenticates client certificate)
Network Authentication	Basic Authentication (server authenticates client UID/Password over HTTP)

Dan Lyon, danlyon@mac.com

Network Authentication	Digest Authentication (server authenticates client UID/Password over HTTP)
Network Authentication	Single SSL followed by Basic Authentication in SSL session (client authenticates server certificate, then server authenticates client UID/Password)
Network Authentication	Mutual SSL followed by Basic Authentication in SSL session (client authenticates server certificate and server authenticates client certificate, then server authenticates client UID/Password)

Table 2 Attack Surface, Controls and Identifier

3.2. Attacker Attribute Definitions

After controls were identified, the possible answers were created for skill, time and cost. The goal of each answer set was to go from almost no resources required to resource levels available to a nation state. Attacker skill, time and cost values are defined in Table 3, Table 4, and Table 5 respectively. In addition to the values and a definition, a numeric value used in results analysis is provided.

Attacker Skill Value	Definition	Numeric Value
Lucky, Accidental	Attack occurs by accident without any effort or skills required	1
Basic Computer Skills	Attacker has basic computer skills, average office worker	2
Hobbyist	Attacker has moderate scripting skills, uses attacks discovered and built by others; script kiddie	3
Amateur	Dedicated, attacks for fun, works alone, some technical knowledge beyond basic tool and script use	4
Apprentice	Limited collaboration with peer group, in-depth technical knowledge in some areas	5
Professional	Limited-to-full access to diverse skills in personal network, broader technical expertise, uses professional tools, limited-to-full access to lab environments	6
International Professional	Internationally recognized expert, full access to diverse skills, full access to lab environments	7

Organizational Team	Teams comprised of dedicated and aligned members with full spectrum of skills, professional and custom tools	8
Limited By Math	An Organizational Team that is limited by provable mathematical limitations, (example: must guess 256-bit number)	9

Table 3 Attacker Skill Definitions

Attacker Time Value	Numeric Value
0-1 days	1
2-7 days	2
1-2 weeks	3
3-4 weeks	4
3-4 months	5
6-12 months	6
Greater than 1 year	7

Table 4 Attacker Time Definitions

Attacker Cost Value (US $)	Numeric Value
1	1
10	2
100	3
1000	4
5000	5
10000	6
50000	7
100000	8
500000	9
1000000	10
10000000	11

Table 5 Attacker Cost Definitions

3.3. Survey Process

For each attack surface subgroup, participants were required to answer if they had experience attacking the area under assessment. Respondents rated attacker skill, attack duration, and attack cost for each control. The survey design allowed skill, duration and cost to be traded off. Participants were allowed to enter freeform text for each scenario, and were also asked to identify related demographic information including all active certifications, years of experience and education level. Survey completion was encouraged through use of guilt (in the case of fellow master's students), humor, and a single $25.00 Amazon gift card for those wishing to enter an email address.

3.4. Participants

The online survey was administered to four email distribution lists: the SANS Advisory Board, SANS Technology Institute master's program, Pauldotcom and SecurityMetrics. The qualifications for the SANS Advisory Board list are to obtain 90% or better on a GIAC proctored exam. The qualifications for the STI master's program list are to be faculty, enrolled in or graduated from the institution. Both SANS mailing lists are considered to contain individuals with expertise in information security. The SecurityMetrics mailing list (www.securitymetrics.org) is controlled with effort made at validating individuals that join, but has no technical requirements to join. Pauldotcom (www.pauldotcom.com) is an industry specific mailing list popular among professionals but also has no technical requirements to join. Due to the specific nature of the Pauldotcom and SecurityMetrics mailing lists, each was determined to contain industry professionals with some level of expertise.

4. Results

The total number of responses for the survey was 19 complete and 78 partial responses for one or more control. Not all responses contained useful data, and the number of responses that contained information on one or more controls was 41. It is unknown exactly how many individuals the survey link reached from the various mailing lists, but is estimated between two and three thousand.

Dan Lyon, danlyon@mac.com

Answers were converted from the textual representation used in the survey to their numerical equivalent, and the median value and the variance were calculated from the data. The median was chosen over the mean to get the data point that was in the middle of all responses. All median and variance values are reported in Appendix A.

The median and variance were calculated because a normal distribution was desired, not because the normal curve was statistically significant. The choice for the distribution was deemed appropriate for the model because the median represents the most likely value and the variance represents the uncertainty.

The reported demographic data was analyzed to obtain a measure of confidence for each individual response. Education, experience, and certifications were used in conjunction with the attack experience. Numerical values were assigned to each indicator and then summed for each respondent. This allowed stratification of the data such that those with more experience and education could be considered high confidence data that should be weighted heavier than those with less experience or education.

The definition for high confidence data was an individual who had the following qualifications: a bachelor's degree, experience attacking the control, four to six years of information security industry experience, and an active certification. Those participants matching the qualifications were considered high confidence data. All other participants were considered low confidence data. For example, some respondents answered the questions but provided no information other than the required attack experience. Without the demographic data, individuals were considered low confidence. The results are reported in Appendix A.

The goal of the survey was to gather raw data and show how the data could be used, not to infer any kind of relationship between the data. Therefore no further statistical analysis of the data was performed. The next section is provided to illustrate the potential use of the numbers obtained in the survey.

Dan Lyon, danlyon@mac.com

5. Discussion

The model presented used a Bayesian network to take the survey numbers and generate a probability curve for attack skill. Before the attribute model is discussed, a brief introduction to Bayesian networks is provided.

One problem with the data acquired is that it is subjective and dependent upon subject matter expert opinion. While the data is useful, it is not completely accurate and is only expert opinion on the attacker attributes, not the actual values themselves. There is some level of uncertainty in the results. The next step in analyzing the values was to use a Bayesian Network to predict a probability curve.

The following section provides a brief overview of Bayesian networks showing how the information learned from this survey is used to model the actual attacker.

5.1. Bayesian Network Introduction

According to Fenton and Neil's excellent text *Risk Assessment and Decision Analysis with Bayesian Networks*, a Bayesian Network (BN) is a graphical representation of events and relationships (2012). In the simplest form, a BN is represented with two nodes and an arc between them that defines the node relationship. The Bayesian network examples were created using a software program called AgenaRisk, which has a free version available with purchase of Fenton and Neil's book.

Each node must be defined to have different states and a probability value for each state. A mathematical relationship between two nodes must be defined in terms of probabilities, and is termed the *Node Probability Table*. Thus probability values are calculated from one node to the next. An illustration of a simple BN is provided using an example of a student writing a research paper.

Assume a student is finishing a research paper and has a day off of work coming up. The student needs to complete the paper soon, and so how the student will spend the upcoming day off is influenced by whether or not they have completed the paper. The Bayesian network for this is shown in Figure 1.

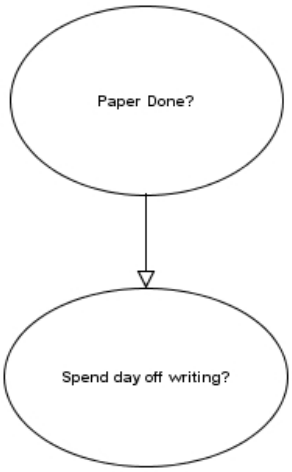

Figure 1 Paper Done BN

If Paper Done is True, then the student will not be spending the day off writing. However, if Paper Done is False, then the student will certainly work to finish the paper.

Suppose that we have no prior knowledge of the student's progress on the paper, such that whether or not they are done is unknown. This results in a 50 percent probability for each of the true and false values in node Paper Done. The resulting probability for whether or not the student spends the day writing is also evenly distributed. Figure 2 shows the probabilities for the scenario with no prior information.

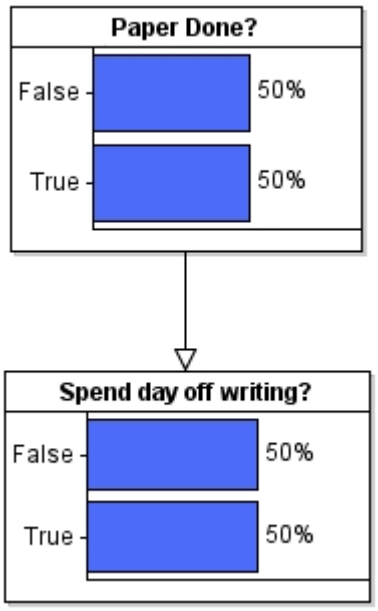

Figure 2 No prior information

Dan Lyon, danlyon@mac.com

Bayesian networks have the ability to represent uncertainty in values and to use evidence. Both of these concepts are important, and clarified by furthering the student paper example.

Assume that the student has a ten percent chance of going golfing even if the paper is not complete. Changing the node probability table to reflect the uncertainty is displayed in Figure 3 and the updated probabilities are shown Figure 4. Note that the student now has only a forty five percent chance of spending their next day off writing.

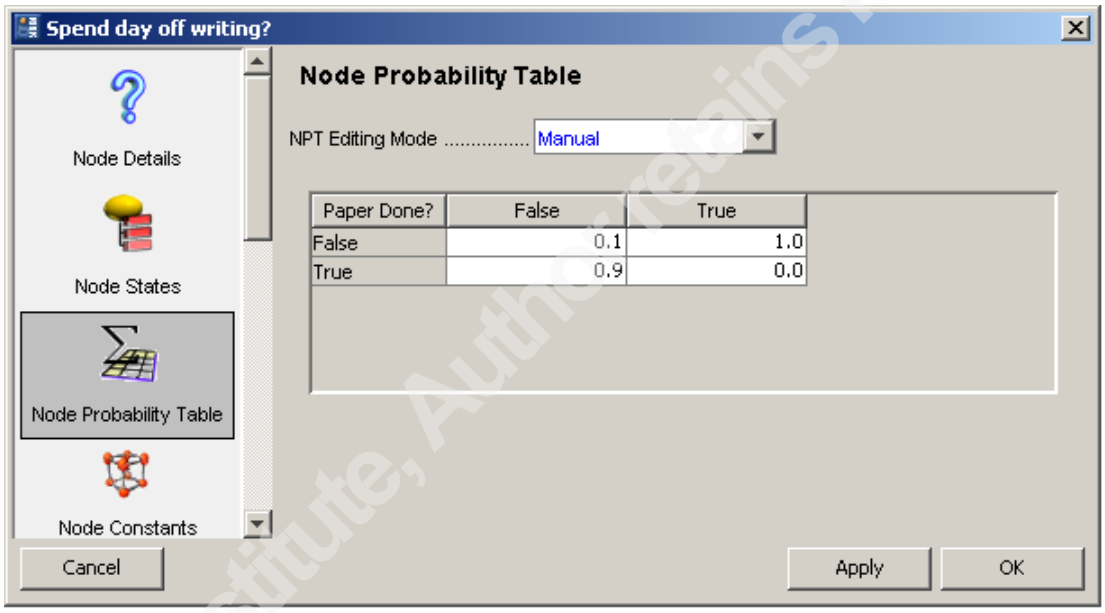

Figure 3 Uncertainty in NPT

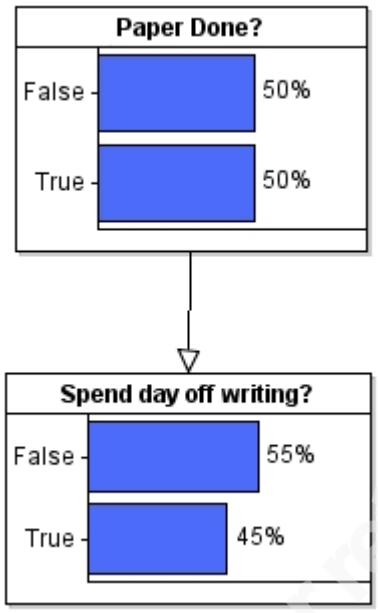

Figure 4 Updated Node with Uncertainty

Another key concept of BN's is using evidence and evidence propagation to learn information, and it is implemented in AgenaRisk using scenario observations. The student paper is expanded to demonstrate this by adding an observation in the form of a scenario. Assume that the student's professor observes the student golfing on the student's day off. Because of the ten percent uncertainty used previously, the professor cannot be certain if the student has completed the paper, even though they are golfing. When the evidence is entered into the Bayesian network (Figure 5) the evidence is propagated back to the Paper Done node. This propagation leaves a nine percent chance that the student has not completed their paper, even though they are golfing.

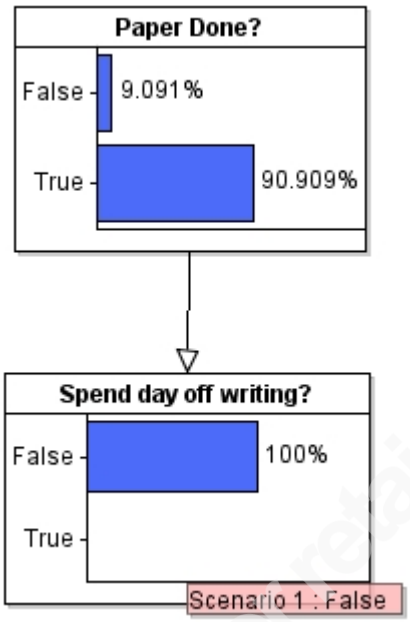

Figure 5 Observed evidence propagation

A full discussion of Bayesian networks, as well as how AgenaRisk can be used is beyond the scope of this paper. For those interested in further understanding, Fenton and Neil's book (which deals largely with accurate risk calculations) is highly recommended.

Probability tables, uncertainty, evidence and evidence propagation were described to give a limited basis for understanding a more complex model.

5.2. Bayesian Network Attacker Skill Example

A Bayesian network was created to use the survey data and provide predictions. A model for attacker skill is described, both the model structure as well as the model with survey data, and a hypothetical situation shown where evidence is collected. While only attacker skill is implemented here, each of the attributes evaluated could be used in a similar model.

The created BN structure is shown in Figure 6. The goal was to predict the required attack skill and the output node of interest is the Predicted Skill node. The prediction node had three identified parents; a Measurement Type, a Weighted Median, and a Weighted Variance. Measurement Type was defined to be either Estimated or Actual, depending on if the data was only an estimate or if it was actual data from a documented breach. The median and variance each were characterized by two parents of

low and high confidence. The weighting was used to appropriately weight the values towards a higher confidence data set.

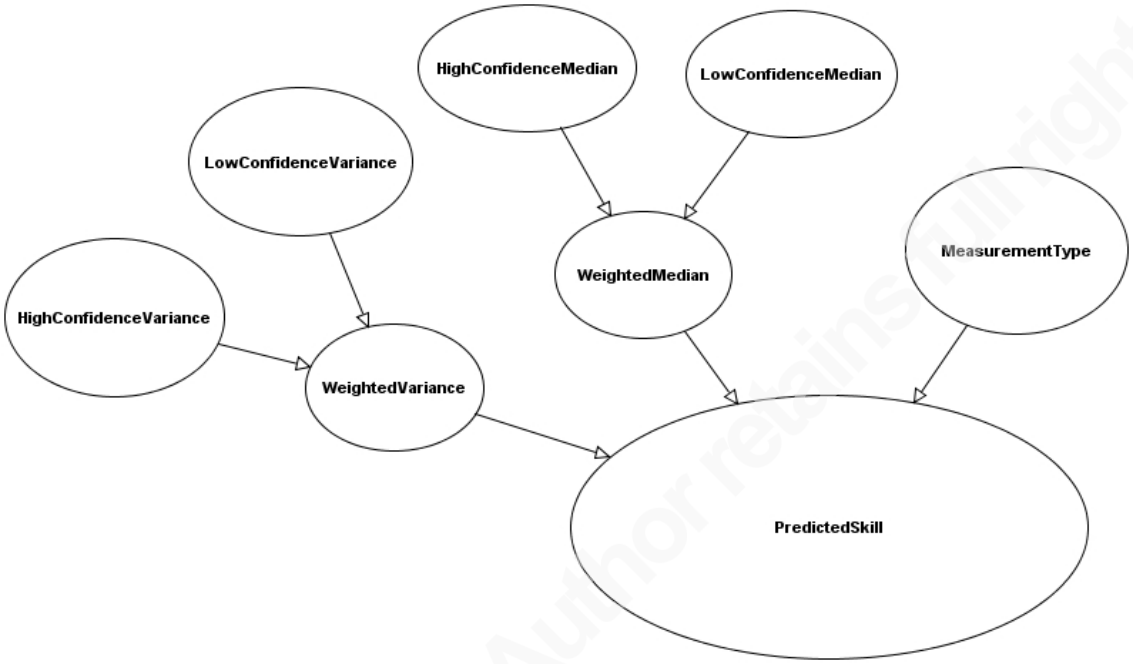

Figure 6 Attacker skill with no observations

The Predicted Skill node is shown in Figure 7 Prior with no observed data. The curve shows the probability of skill required for a security control with no information known. Predicted skill values 1 through 9 were derived from the numerical skill values in the survey (see Table 3 for definitions). The probability distribution for Predicted Skill was calculated from the parent nodes. Some assumptions were defined in the node table for Predicted Skill, but are not explicitly discussed here. The focus was intended to be the Bayesian network as a tool, not a full review of the example model. Many such models may be created, depending on the needs of the organization.

The predicted skill line displayed three peaks due to limited variances, parent limits and assumptions in the node probability table. Logically the curve displayed made sense. The probabilities at the extreme values should be less than the values in the middle, but they should not be zero. The values in the range from 3 to 7 are about equally likely, given no information about the control. For purposes of this example, the probability curve was determined to be adequate.

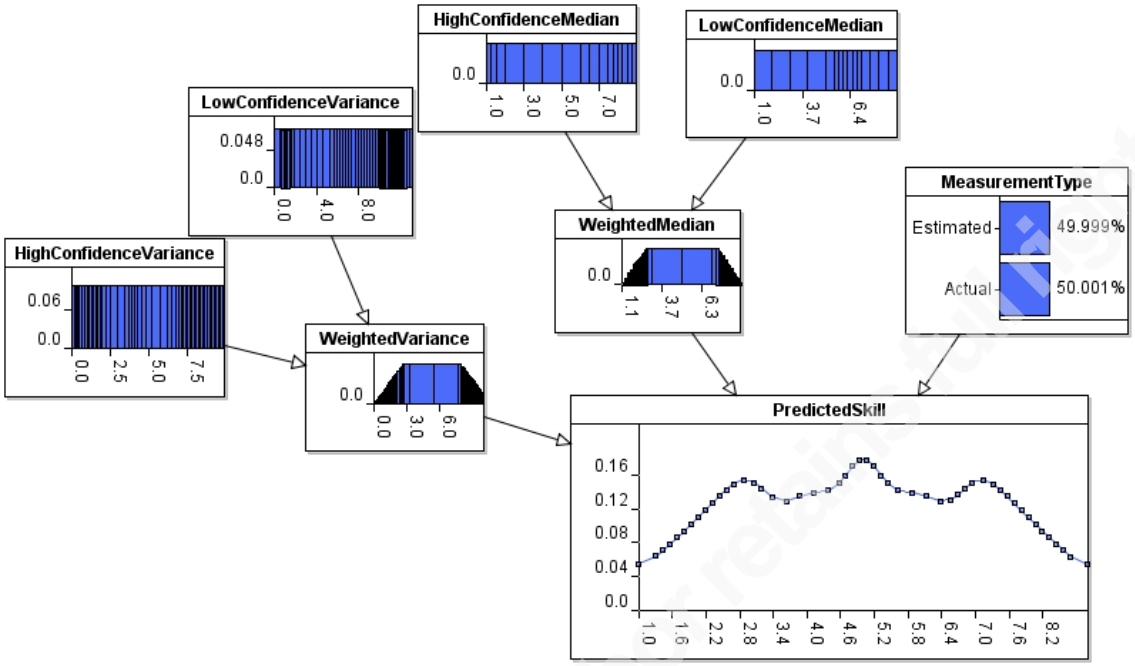

Figure 7 Prior with no observed data

Using the above model, survey data is entered for the evaluated control of digest authentication (NetworkAuth4). The measured values from the survey were a high and low confidence median that attacker skill was a 4. The variances were minimally different, with the high confidence variance at 1.874 and the low confidence variance at 1.667. The data was entered and the impact on the predicted skill probabilities presented by the green line in Figure 8.

Dan Lyon, danlyon@mac.com

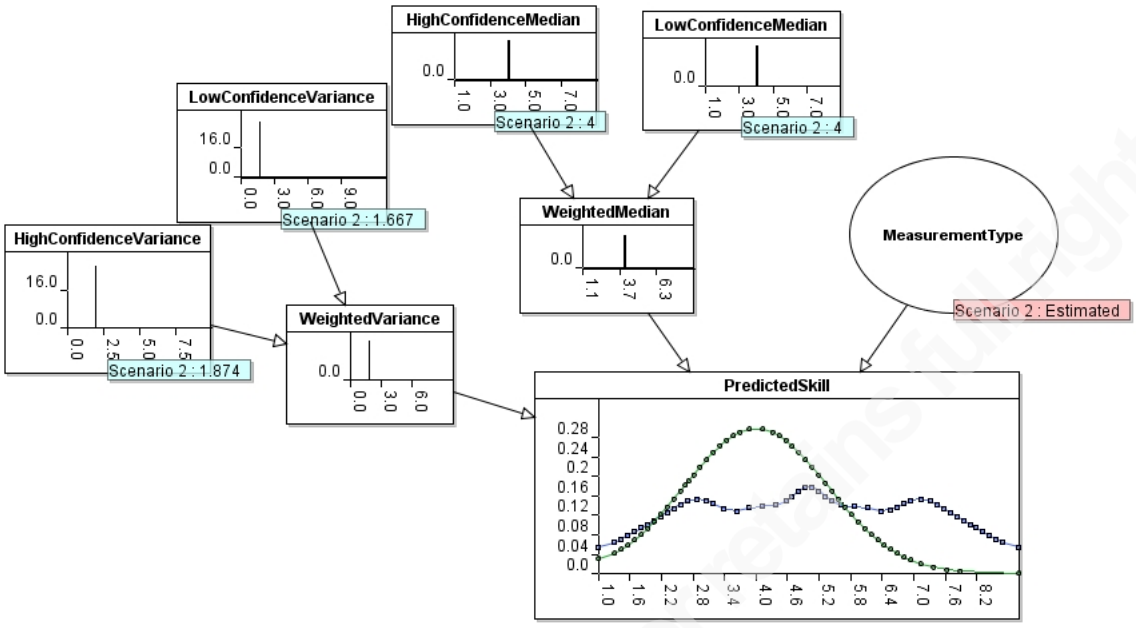

Figure 8 Digest authentication skill

Scenario 3 in Figure 9 is used to evaluate the WPA-PSK TKIP control (Wireless2). The high confidence median and variance were 6 and 1.156 and the low confidence median and variance were 3 and 3, respectively. This example is chosen because the high and low confidence data is different. The model assumptions weighted the high confidence data over the low confidence data, and the curve reflected those assumptions.

Dan Lyon, danlyon@mac.com

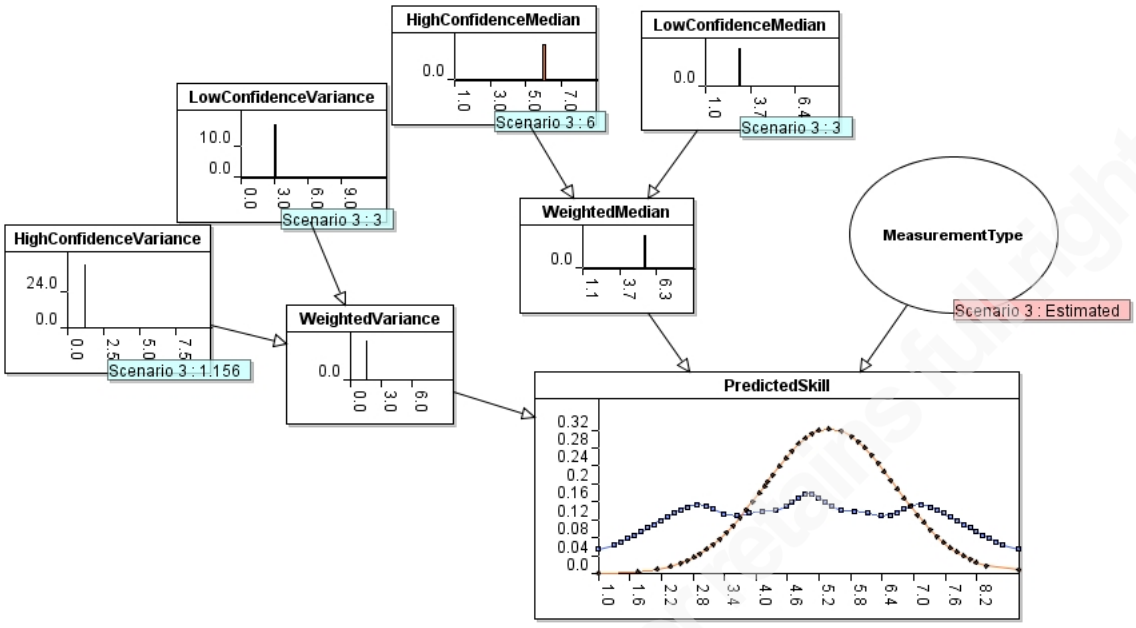

Figure 9 Weighted values

One final benefit of Bayesian networks is highlighted by entering actual evidence into the measurement type node, along with a measured value. Assume that evidence was discovered that a hobbyist compromised a given control. This could have been through testing, or perhaps public disclosure. The skill of a hobbyist was translated to the value of 3 for attacker skill, and the evidence entered in. However there is still some uncertainty around the value, and the Predicted Skill node accounted for this in the node probability table definition. Note that in the example, there is no variance data entered for scenario 4, only the observed value. A small variance value is calculated from the parent variance and represents the uncertainty. The curve is calculated using the entered median and a variance with no assumptions. Given that there is still some uncertainty about the attacker skill required to overcome the control, the graph makes sense. Note that the probability drops off, such that attack skills beyond Professional are not likely.

Dan Lyon, danlyon@mac.com

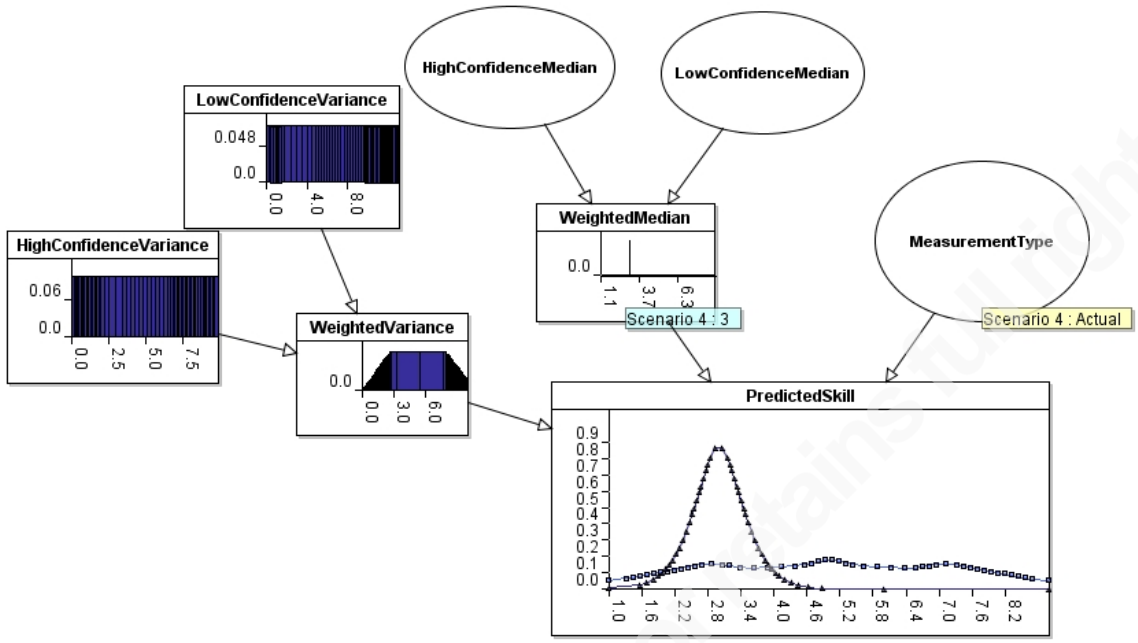

Figure 10 Observed evidence

The Bayesian model was presented as just one way to incorporate the results. Many assumptions were made, and to implement this for a given organization the assumptions would need validation by that organization's experts.

6. Results Interpretation

One of the goals of Six Sigma techniques was to increase predictability of the product design (Ginn, 2004). Analysis of the overall variances revealed one example of a significant difference in predictability with controls of similar median values.

The Cellular-GSM and Cellular-CDMA median skill values were calculated to be the same, however the variances differed for greatly for attacker skill.

Attribute	Cellular GSM	Cellular CDMA
Median Skill	6	6
Skill Variance	3.744	1

Table 6 GSM vs. CDMA Skill

The difference exposed by the survey was the variance, or predictability, of GSM compared to CDMA.

Dan Lyon, danlyon@mac.com

Analysis continued using the BN model, and the cumulative probability distributions plotted in Figure 11 Cumulative Prediction Comparison for GSM and CDMA. The model showed the difference in predictability.

The takeaway from the survey data and the model was that an attacker of amateur skill had roughly a 2% probability to overcome CDMA, while the same amateur had a 17% probability to overcome GSM. A greater degree of confidence at values below the median is provided to CDMA due to the limited variance values.

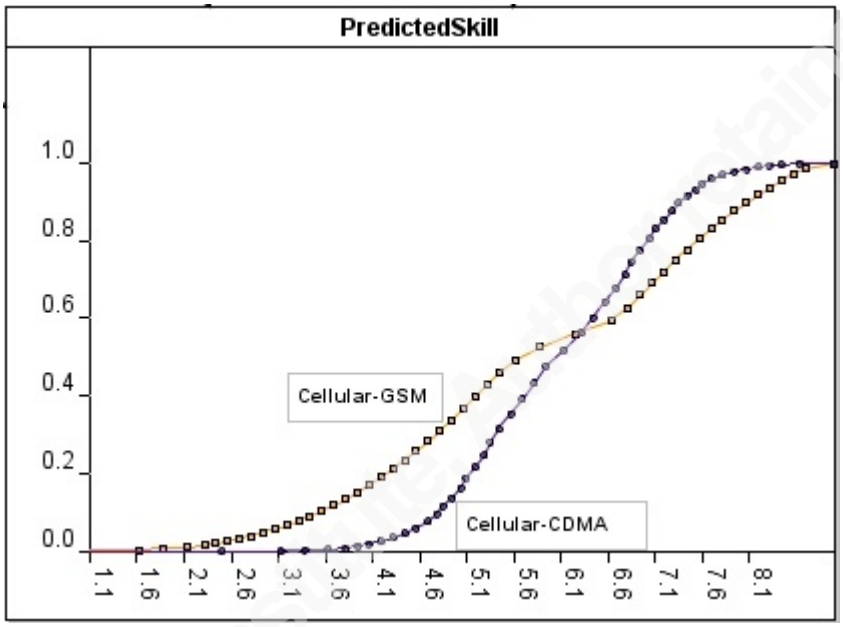

Figure 11 Cumulative Prediction Comparison for GSM and CDMA

An additional observation on the data was that attacker cost was nearly always a higher variance than skill or time. This indicated a reduced measure of predictability when estimating the cost required by an attacker compared to skill and time. This could have resulted from more ambiguity in the survey question, as instructions did not explicitly define when cost was incurred to an attacker. This could be accounted for in future surveys by providing detailed definition on what exactly constitutes a cost.

7. Limitations and Future Work

Potential ambiguity in the cost data was not the only limitation identified, and several others are described in this section.

Dan Lyon, danlyon@mac.com

An assumption was made when analyzing the data that it would fit into a normal distribution because the model was relied on median and variance. Detailed statistical analysis was not performed. Rationale for not performing detailed analysis was because of the limited responses and the intended use of the data. Future studies could incorporate a larger population and more thorough analysis of the data obtained.

Another limitation is that attacker skill, time and cost are dependent variables. In other words, when attacker skill and cost were increased, time could have been reduced. No attempt was made to quantify the relationships between the three variables.

It is also recognized that this survey data was collected at a point in time, and that the information security landscape can shift rapidly. An argument could be made that the survey data is outdated shortly after it is acquired.

Even with the limitations on the survey and the data, valuable information is still available for system designers using this model.

8. Conclusion

Attack trees were described as an analysis tool that can help with system design and optimization. The problem with current literature's lack of published security control attributes was highlighted and a survey was designed that attempted bridge the gap. The survey inspiration from the Voice of the Customer concept used in Six Sigma design was discussed. Following the VoC concept, a link to the current approach described by the Defense Science Board's recent report showed the need for data tailored to a business rather than the military.

The survey design, process and the results were discussed and the results made available to the community. After discussing the results, a brief introduction to Bayesian networks was covered. Finally, it was demonstrated how a predictive model could be applied using a Bayesian Network.

Dan Lyon, danlyon@mac.com

9. References

Amoroso, E. (1994). *Fundamentals of computer security technology.* Upper Saddle River, NJ: Prentice-Hall PTR.

Fenton, N., & Neil, M. (2012). *Risk Assessment and Decision Analysis with Bayesian Networks.* Boca Raton, FL: CRC Press.

Roy, A. (2010). *Attack countermeasure trees: A non-state-space approach towards analyzing security and finding optimal countermeasure sets.* (Master's thesis) Retrieved from http://dukespace.lib.duke.edu/dspace/bitstream/handle/10161/3148/thesis.pdf?sequence=1

Schneier, B. (1999, December). *Attack trees.* Retrieved from http://www.schneier.com/paper-attacktrees-ddj-ft.html

Department of Defense, Defense Science Board. (2013). *Resilient military systems and the advanced cyber threat.* Retrieved from Office of the Under Secretary of Defense for Acquisition website: http://www.acq.osd.mil/dsb/reports/ResilientMilitarySystems.CyberThreat.pdf

Bistarelli, S., Peretti, P., & Trubitsyna, I. (2006). *Defense trees for economic evaluation of security investments. Informally published manuscript, Department of Mathematics and Computer Science, University of Perugia, Perugia, Perugia, Italy.* Retrieved from http://www.dmi.unipg.it/~bista/papers/papers-download/26_Bistarelli_S.pdf

Edge, K. S. (2007*). A framework for analyzing and mitigating the vulnerabilities of complex systems via attack and protection trees*. (Doctoral dissertation) Retrieved from http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA472310

Ginn, D. (2004). *The design for six sigma memory jogger*. Salem, NH: GOAL/QPC.

Maass, E., & McNair, P. (2009). *Applying Design for Six Sigma to Software and Hardware Systems.* Prentice Hall.

Zonouz, S., Khurana, H., Sanders, W., & Yardely, T. (2009). *RRE: A game-theoretic intrusion response and recovery engine*. Manuscript submitted for publication, Performability Engineering Research Group, University of Illinois at Urbana-Champaign, Champaign, Illinois, Retrieved from

https://www.perform.csl.illinois.edu/Papers/USAN_papers/08ZON01.pdf

Dan Lyon, danlyon@mac.com

Appendix A: Summarized Results

Results are summarized in the following sections based on attack surface.

Hardware Control Data Results Summary

ID	Control
HW1	The encryption key is stored in a Trusted Platform Module (TPM chip) and the encrypted data stored in removable memory. Bus traffic is not encrypted.
HW2	The encryption key is stored in a Trusted Platform Module (TPM chip) and the encrypted data stored in Non-removable memory. Bus traffic is not encrypted.
HW3	The encryption key is stored in on-board memory and the encrypted data stored in NON-removable memory. Bus traffic is not encrypted, on-board memory is encrypted. Bus traffic protected from access via physical layer (epoxy).
HW4	The encryption key is stored in on-board memory and the encrypted data stored in NON-removable memory. Bus traffic is not encrypted, on-board memory is not encrypted.
HW5	The encryption key is stored in on-board memory and the encrypted data stored in removable memory. Bus traffic is not encrypted, on-board memory is not encrypted.

Control	Attribute	Median	Variance	High Conf Median	High Conf Var	Low Conf Median	Low Conf Var
HW1	Skill	6	0.700	6.500	0.917	6.000	0.000
HW1	Time	3	0.567	3.500	0.333	2.500	0.500
HW1	Cost	7	1.000	7.000	1.000	7.000	2.000
HW2	Skill	6.5	0.667	6.500	0.917	6.500	0.500
HW2	Time	3.5	0.667	3.500	0.333	3.000	2.000
HW2	Cost	7	2.300	7.000	2.333	7.500	4.500
HW3	Skill	6	1.905	6.000	2.000	7.000	2.000
HW3	Time	3	1.952	3.000	1.800	3.500	4.500
HW3	Cost	6	4.800	6.000	5.333	8.000	8.000
HW4	Skill	6	1.200	6.000	1.000	5.000	0.000
HW4	Time	3	0.800	3.000	0.667	3.000	2.000
HW4	Cost	6	5.200	6.000	6.333	6.000	8.000
HW5	Skill	6	0.967	6.000	1.000	5.500	0.500
HW5	Time	3	0.800	3.500	0.917	2.500	0.500
HW5	Cost	6	2.200	6.000	4.000	6.500	0.500

Dan Lyon, danlyon@mac.com

Software Control Data Results Summary

ID	Control
SW1	The software is a release build compiled C program that has anti-tampering compiled into it.
SW2	The software is a release build compiled C program that has complex obfuscation.
SW3	A program that is a release build compiled C program, without the symbol table present.
SW4	A Java program that has had simple obfuscation applied.
SW5	A Java program that has not been obfuscated.
SW6	Software binaries that are not encrypted or obfuscated.

Control	Attribute	Median	Variance	High Conf Median	High Conf Var	Low Conf Median	Low Conf Var
SW1	Skill	6	1.970	6.500	2.982	6.000	0.250
SW1	Time	3	1.418	3.000	1.619	2.500	0.917
SW1	Cost	5	4.278	5.000	6.967	5.000	0.667
SW2	Skill	7	2.397	7.000	3.122	6.500	0.917
SW2	Time	3	1.364	3.000	1.429	2.500	0.917
SW2	Cost	6	5.833	5.000	7.667	6.500	0.917
SW3	Skill	6	1.692	6.000	2.278	6.000	0.250
SW3	Time	2	1.231	3.000	1.500	2.000	0.000
SW3	Cost	5	4.673	4.000	7.619	5.000	0.250
SW4	Skill	6	1.897	5.500	1.822	6.000	2.250
SW4	Time	2.5	1.174	2.000	1.028	2.500	1.667
SW4	Cost	4	5.291	3.500	5.929	4.000	3.667
SW5	Skill	5	1.114	5.000	1.067	5.000	1.333
SW5	Time	1.5	1.302	2.000	1.250	1.000	1.700
SW5	Cost	3	3.879	3.000	5.952	3.000	1.500
SW6	Skill	5	0.981	4.500	1.122	5.000	0.500
SW6	Time	2	1.412	2.000	1.444	2.000	1.700
SW6	Cost	4	3.231	3.500	4.214	4.000	2.300

Dan Lyon, danlyon@mac.com

Wireless Control Data Results Summary

ID	Control
Wireless1	WPA-Enterprise 802.1X
Wireless2	WPA-PSK(TKIP)
Wireless3	Bluetooth 2.1+ EDR SSP
Wireless4	Cellular-GSM
Wireless5	Cellular-UMTS
Wireless6	Cellular-CDMA
Wireless7	WPA2-PSK (AES)

Control	Attribute	Median	Variance	High Conf Median	High Conf Var	Low Conf Median	Low Conf Var
Wireless1	Skill	6	1.329	6.000	1.656	5.000	0.333
Wireless1	Time	2	1.667	2.500	1.656	1.500	0.500
Wireless1	Cost	4	3.346	4.000	4.456	3.500	0.500
Wireless2	Skill	6	2.000	6.000	1.156	3.000	3.000
Wireless2	Time	2	1.933	2.000	2.722	1.500	0.500
Wireless2	Cost	3	3.810	4.000	4.767	3.000	0.000
Wireless3	Skill	5	2.859	5.000	2.855	4.500	4.500
Wireless3	Time	2	0.568	2.000	0.456	2.000	2.000
Wireless3	Cost	4	2.564	4.000	3.111	3.500	0.500
Wireless4	Skill	6	3.744	6.000	4.400	7.000	1.000
Wireless4	Time	2.5	2.447	2.500	2.678	2.000	2.000
Wireless4	Cost	6	4.964	6.000	6.194	6.000	0.000
Wireless5	Skill	6	2.103	6.000	2.265	7.000	N/A
Wireless5	Time	2	1.964	2.000	2.178	3.000	N/A
Wireless5	Cost	6	5.211	6.000	5.750	5.000	N/A
Wireless6	Skill	6	1.000	6.000	1.167	6.500	0.500
Wireless6	Time	3	2.000	2.500	2.100	4.000	N/A
Wireless6	Cost	6	5.567	6.000	6.194	5.000	N/A
Wireless7	Skill	6	1.838	6.000	2.265	6.000	0.000
Wireless7	Time	3	2.000	3.000	1.970	4.000	2.000
Wireless7	Cost	4	5.192	4.000	5.655	5.500	4.500

Dan Lyon, danlyon@mac.com

Network Authentication Control Data Results Summary

ID	Control
NetworkAuth1	Single SSL (client authenticates server certificate)
NetworkAuth2	Mutual SSL (client authenticates server certificate and server authenticates client certificate)
NetworkAuth3	Basic Authentication (server authenticates client UID/Password over HTTP)
NetworkAuth4	Digest Authentication (server authenticates client UID/Password over HTTP)
NetworkAuth5	Single SSL followed by Basic Authentication in SSL session (client authenticates server certificate, then server authenticates client UID/Password)
NetworkAuth6	Mutual SSL followed by Basic Authentication in SSL session (client authenticates server certificate and server authenticates client certificate, then server authenticates client UID/Password)

Control	Attribute	Median	Variance	High Conf Median	High Conf Var	Low Conf Median	Low Conf Var
NetworkAuth1	Skill	5	2.490	5.000	1.981	4.500	3.500
NetworkAuth1	Time	2	1.390	2.000	1.638	1.500	0.667
NetworkAuth1	Cost	3	3.421	3.500	4.401	3.000	1.367
NetworkAuth2	Skill	6	1.957	6.000	2.124	6.000	1.767
NetworkAuth2	Time	3	1.362	3.000	1.410	2.500	1.467
NetworkAuth2	Cost	5	3.205	5.000	4.256	5.000	1.200
NetworkAuth3	Skill	3	1.690	3.000	1.829	3.000	1.571
NetworkAuth3	Time	1	0.736	1.000	0.952	1.000	0.238
NetworkAuth3	Cost	2	2.162	2.000	2.577	2.000	1.619
NetworkAuth4	Skill	4	1.729	4.000	1.874	4.000	1.667
NetworkAuth4	Time	1	0.748	1.500	0.951	1.000	0.238
NetworkAuth4	Cost	3	2.366	3.000	2.744	3.000	1.905
NetworkAuth5	Skill	6	2.261	6.000	2.571	5.000	1.767
NetworkAuth5	Time	2	1.463	2.000	1.720	1.500	0.967
NetworkAuth5	Cost	4	3.322	4.000	4.667	3.500	0.667
NetworkAuth6	Skill	6	2.490	6.000	1.952	6.500	4.400
NetworkAuth6	Time	3	1.748	3.000	1.210	3.000	3.600
NetworkAuth6	Cost	5	3.232	5.000	3.455	5.500	3.367

Dan Lyon, danlyon@mac.com

Local Interface Control Data Results Summary

ID	Control
LI1	RS-232 Serial Port that provides access to a diagnostic debugger capability.
LI2	Ethernet port is available with a local operating system firewall blocking unsolicited inbound communications
LI3	USB port restricted to Mass Storage USB devices that have signed content

Control	Attribute	Median	Variance	High Conf Median	High Conf Var	Low Conf Median	Low Conf Var
LI1	Skill	5	1.910	4.000	1.250	6.000	1.930
LI1	Time	2	0.457	1.500	0.286	2.000	0.507
LI1	Cost	3	4.660	3.000	0.667	3.000	5.660
LI2	Skill	5	2.220	5.000	1.067	5.000	2.466
LI2	Time	2	0.669	1.500	0.667	2.000	0.689
LI2	Cost	3	3.516	4.000	1.067	3.000	4.099
LI3	Skill	5	3.272	4.500	3.467	5.000	3.339
LI3	Time	2	1.386	1.500	0.300	2.000	1.566
LI3	Cost	4	3.802	3.500	1.467	4.000	4.375

Dan Lyon, danlyon@mac.com