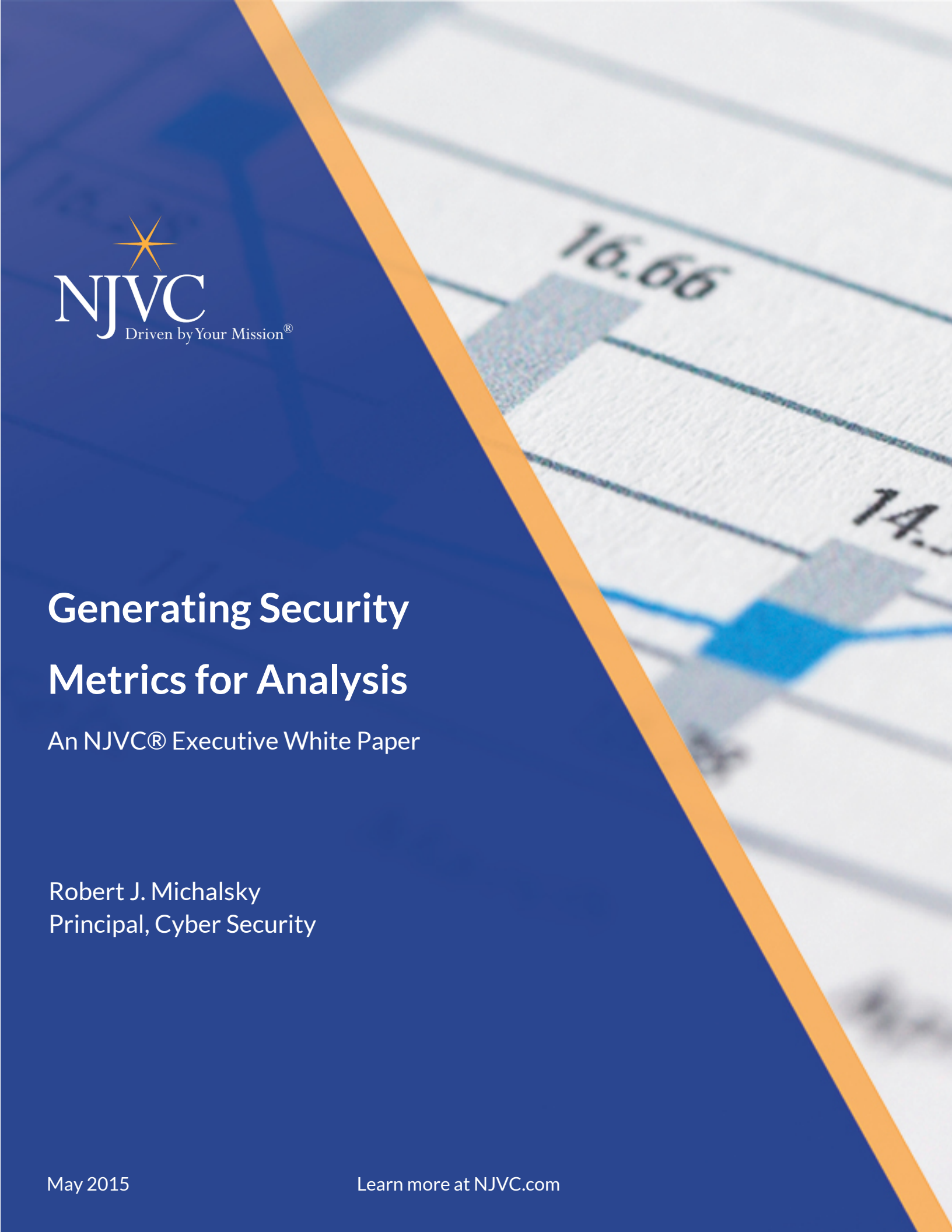
NJVC
Driven by Your Mission®



Generating Security Metrics for Analysis

An NJVC® Executive White Paper

Robert J. Michalsky
Principal, Cyber Security

May 2015

Learn more at NJVC.com

Executive Summary

Detecting and responding to cyber security threats is an extensive challenge for all IT organizations regardless of their industry.

Simply put, protecting all the valuable online assets of an organization is difficult even with the multitude of modern controls available. Physical security concepts have built up and matured over literally thousands of years, whereas the digital realm has only existed since dawn of the computer age in the 1940s and 1950s.

Organizations need to not only protect the assets; the business process workflows that produce the valuable assets must also be protected. For example the intelligence Community needs to protect not only the products of analysis and the products produced, but also the sources of the data – where the raw data originates from. These sources must be protected as a supply chain must be – any injections of inaccurate information in the process can lead to an analysis product that is fundamentally flawed.

Security metrics can be used to determine and monitor the source of information – that it came from reputable or trusted individuals and can also support subsequent manipulation of that data to produce intelligence products. Data marking helps differentiate the value or worth of specific data items and serves to focus protection efforts.

The utilization of purposely built and maintained security metrics provides insight to malicious activities. If baselines defining normal operations can be established, it becomes easier to detect malware or rootkits that may have penetrated network defenses allowing attackers to carry out a range of malicious activities.

Creating this 'fingerprinting' of IT operations requires an extensive array of sensors which produce copious amounts of data which in turn supports algorithms to dissect that data and then automation to provide aggregated alerts for more detailed analysis.

Visualization of these vast quantities of raw data facilitates human analysis of suspicious events. As operational insights mature, these alerts can be continually updated leading to the 'holy grail' of cyber security – Continuous Diagnostics and Mitigation (CDM). The Department of Homeland Security is developing that program for Federal organizations, but all companies can benefit from creating an infrastructure that uses their specific data to operate in a similar manner.

http://www.dhs.gov/cdm

Security Metric Introduction

Organizations of any size – from the smallest start up to the largest global enterprise – all have something of value to protect. This can be some type of intellectual property, a product formulation, a process workflow, a recipe, billing information, financial data, whatever provides that enterprise some type of competitive value. When this data is stored in digital form, cyber security is deployed to protect those assets.

Cyber security in its most primitive form is concerned with the management of risk. For that risk to be assessed, quantified and mitigated, it must be measured. This is where security metrics come into play. It is only through measurements that operational security controls can be brought into play and risk reduced. Cyber security in its purest form is focused on the continual reduction of risk to achieve some acceptable level based on available limited resources such as time and labor and dollars.

Security metrics are measures, and when applied in a contextual setting, allow cyber security teams to quantify various types of risk and seek the most cost effective methods to reduce risk. These metrics provide detailed insight into operational situations and support cyber engineering activities such as detecting malware, preventing command and control beaconing, discovering unauthorized data exfiltration, preventing rootkit installations, uncovering credential tampering and reporting on elevation of privileges across network domains.

Security metrics – along with a consistent and well defined interpretation of their meaning – are the raw material for conducting analysis on operational IT systems that follows scientific principles, is repeatable, and will support any required subsequent forensic analysis should a data breach occur. Given the volume of transactional data this requires, the process of visualization can give rapid insight to ops personnel on areas of concern.

For CIO and CISO with budget constraints – and that would be everyone - measurements also allow a business Return on Investment (ROI) focus to be brought into the situation, by assessing the effectiveness of security metrics to get at the underlying vulnerability producing that risk.

Security controls all have a cost to impose. It is one thing to identify some gap and state which security control to use to plug that gap, but having security metrics at our disposal allows us to select an approach that provides the best combination of low cost and risk reduction.

For example, malware is detected on a system server. Sifting through audit logs points to an unsuspecting user who clicked on an email advertisement and infected the corporate network. How to mitigate this risk in the future? Disable the user email account? Not likely. Block the originating address of the offending message? Absolutely. But what if the email came from a legitimate customer who had their system compromised? Or a legitimate website was serving up malware until it was discovered and disabled? Often, entire domain names cannot – or should not - be blocked.

The idea is to support subsequent analysis by providing data that an investigator can use to discover the origin of the infection, identify how the malware spread through the internal network domains, and to eradicate all traces.

All mitigation activities inevitably become labor intensive but can be invaluable in strengthening defenses and continuously improving the overall organization defensive posture. By having security metrics available, each analysis method used can be evaluated over time allowing for optimization of expensive labor assets (those whip smart cyber engineers you employ).

What are some sources of security metrics?

So where does all this mystical data come from? Virtually any component in an IT infrastructure can be a source. It is helpful to consider them from two source categories, a physical hardware device, and application software (including Operating System) that runs on those devices to enable their functionality.

The hardware metrics are most often quantities or counts of defined IT activities, whereas the application software metrics are more user activity based. Both can provide differing insights when evaluating network and data breaches.

Hardware Metrics

Device	*Sample metrics*	*Units*	*Source*
Servers	CPU utilization Running processes Hidden processes Hard drive statistics CD / DVD / USB usage Transaction logs Processes loaded at start	% over time Quantity Quantity % utilization User ID, quantity User ID, activity Process executable	Operating System

	up Availability Up time Unplanned outages Vulnerability alerts	list % over time Length of time Quantity Unapplied patches	
Routers, Switches, Bridges	Access Control Lists Link utilization Number of hops Packet loss Availability	User ID, roles % Quantity Quantity %	Device routing tables; internal storage
Firewalls	Blocked IP addresses	Country of origin	Device
End user devices	Authorized users Connection attempts	User ID Quantity	Device
Mobile devices	Apps being used Data being accessed Timeline usage profile	Approved Yes/No Data type Data transfers	Device
Network, IDS, IPS	Bandwidth utilization Protocols in use Packet counts	%, # packets / time Protocol listing Quantity	Network sniffer
SAN	Utilization	Data read/write latency	Operating System
Virtual Machines (VM)	Operating Systems emulated Hypervisor instantiations	Listing Quantity	OS of host system
Ports	Authorized ports Application interface	Listing Description	Operating System

Software Metrics

Source	*Sample Metric*	*Units*	*Evaluation*
Application (type) Mission critical apps Internet browser type	# users (total) # users (concurrent) Release date Patch version in use Software scans	Application Patch version Log of all configuration changes Unauthorized software	Unauthorized User activity? Suspicious Activity? New software installed All software formally configuration controlled

Logon tracking – failed logons, # logons etc.	Failed logons Password resets	Cyber COTS tool	Are failed log ins trending up over time?
Databases	Total # records Queries run Unexpected data encryption	Quantities	Does the volume of 'long running' queries exceed normal limits?
Anti-virus	# signatures used	Quantities	Are anti-virus signatures up to date?
User satisfaction	Time to process a log on Time to execute specific data queries	Seconds Seconds	Are users reporting any aberrant system behavior?
Vulnerability alerts	Software to be patched	List of s/w products	Track volume over time –decreasing trend is less risk, increasing trend is more risk incurred
VPN status	Users, packets Access to privileged accounts	VPN software	SPAM or DDOS alerts Role based access control
Malware detection	Rootkit IP addresses in use	Virus / worm	Detect beaconing to external command and control server White list / black list software
Corporate e-mail	Traffic Volume Amount of spam # phishing attacks	Quantities	Track # threats which penetrate network defenses

What makes a good security metric?

Not all security metrics are created equal. The following are some qualities that comprise a good metric:

- Easy to collect – automatically generate via software tools
- Consistent measure – day to day and week to week
- Unambiguous - avoid 'small, medium, large' as an assessment of impact
- Clear definition – do the users of the data understand it?
- Repeatable – collected in the same manner each time
- Able to utilize directly to draw comparisons – compare versus a baseline of normal

- The cost to collect is quantifiable – Does the ROI justify the collection?
- Can be used to generate alerts – indicators of potential security issues
- Create actionable intelligence – support potential forensic activities

All this data does not have to generated and processed in real time. Malicious activities that may take systems off line (such as DDOS attacks) benefit from real time alerts, but the vast majority of data results from scans or aggregating real time data into summary form. It is expected that there be a time delay between data generation and resulting 'actionable intelligence'.

Another point to observe is that even something as seemingly innocent such as a risking disk latency time could indicate a incident if it occurs during off hours when user activity is typically low.

Generation and Storage

Having identified the appropriate security metrics for an organization, the next challenge is to define ***what to store*** and ***how much to store.***

Given the volumes of transactions in modern IP based networks, it is beyond infeasible to store every data item that may hold some type of cyber security value. Generating voluminous logs of data that might prove to be of use some day, is not an optimal way to proceed. Consider the end game first – what data might be required in order to conduct detailed forensic analysis of an attack. Think of the investigation itself and the data that would prove most useful to have. Start with that as an objective as you determine your security data logging and archiving strategy.

Another technique is using data aggregation. This is the process of combining two or more singular metrics into a summary. Data with like-minded attributes can also be grouped together and only then might that data be stored and archived. That data might be grouped by type or department or some other measure that makes particular sense to the organization.

Data should also be standardized based on a consistent time metric. That could be once an hour for certain data types and once a day for other data types. Any numerical values can also have various statistical measures used – min/max/standard deviation, ordinal count, value ranges etc. that can focus analysis on an appropriate level of abstraction.

All these techniques seek to reduce the raw amount of data to be stored without losing any data integrity and yet retaining sufficient details to carry out future forensic studies should they be needed.

Simple data sets can use something like Excel pivot tables which are built explicitly to group and categorize large volumes of data in spreadsheets. More sophisticated analyses can use a formal statistical analysis program such as SAS or SPSS. Data can then be stored in a relational data base structure facilitating ongoing analysis and new analysis of older archived data through standard SQL constructs.

Other mathematic techniques may prove useful – particularly when looking to identify long term trends. For instance time series analysis can be used to determine if a data sequence has violated some threshold and should generate a real-time alert to a cyber analyst.

Probably the single greatest challenge in the security metrics domain is establishing a baseline –what a specific environment defines as 'normal'. Variances against that serve to indicate changes in the network environment or overall architecture infrastructure. This implies creating custom log files with pertinent security data and evaluating them over normal operating conditions.

For instance, individual transaction events against key application software or protected database records may look like the following:

Application Name	User ID	Timestamp	# Logon Attempts	# Records Accessed	# data items Accessed	Protected data Access Flag (Yes/No)

This type of log can be used to create a detailed time ordered sequence of events for an individual logon account which supports analysis of an attacker moving through a network. Storing this in a relational database structure further supports queries such as when searching for individuals trying to access accounts through repeated logon failures.

Going further, different types of aggregated records can be generated. These records can provide more of a macro perspective across Business Units or across network domains.

Example:

Application Name	Max Quantity of Active User ID'	Avg Quantity of Active User ID's	Time Period	# Records Retrieved (summary)	# data items Retrieved (summary)	Quantity of Protected data Access Flags (Yes/No)

Custom log transaction records such as these can be used for both static analysis – generating real time alerts – as well as longer term trend analysis over days, weeks or months. It has been shown that many attackers lie low in systems for months at a time evading detection and exfiltrating data over time. Also, records such as these help detect long term activities that look perfectly normal during normal daily operations.

Analysis of this type can also take advantage of data ***correlation***. This is when multiple records or events are compared against each other or against historical baseline values to determine if further cyber analysis is required. These transaction logs may represent aggregate user activity summaries over extended time periods. Another aspect is to hone in on User ID that may be accessing data across multiple departments – consistent with the authorization of their role, but inconsistent with their normal Use Case. This can help identify the insider threat.

When using all this security metric data, security controls can be utilized to guide the analytical queries. For instance, has the data volume of access requests exceeded normal parameters? This might indicate a violation of a Data Loss Prevention (DLP) control. A User ID accessing data across multiple network domains may exceed their defined Role Based Access controls (RBAC).

Cyber Security Use Cases

Cyber Security while often thought of as the sole domain of the 'security' team, in actuality is a team sport requiring participation from a range of personnel, both inside and outside of the often narrow confines of IT.

Having taken great care in collecting raw data and aggregating it into useable form, the question becomes how to display the data to end users – in this case the cyber analysts responsible to maintain organizational security. A 'dashboard' display is often used to convey specific data to specific classes of user. But how to determine how many unique dashboard displays might be needed in an enterprise environment? One available technique from the software engineering world is that of Use Case modeling.

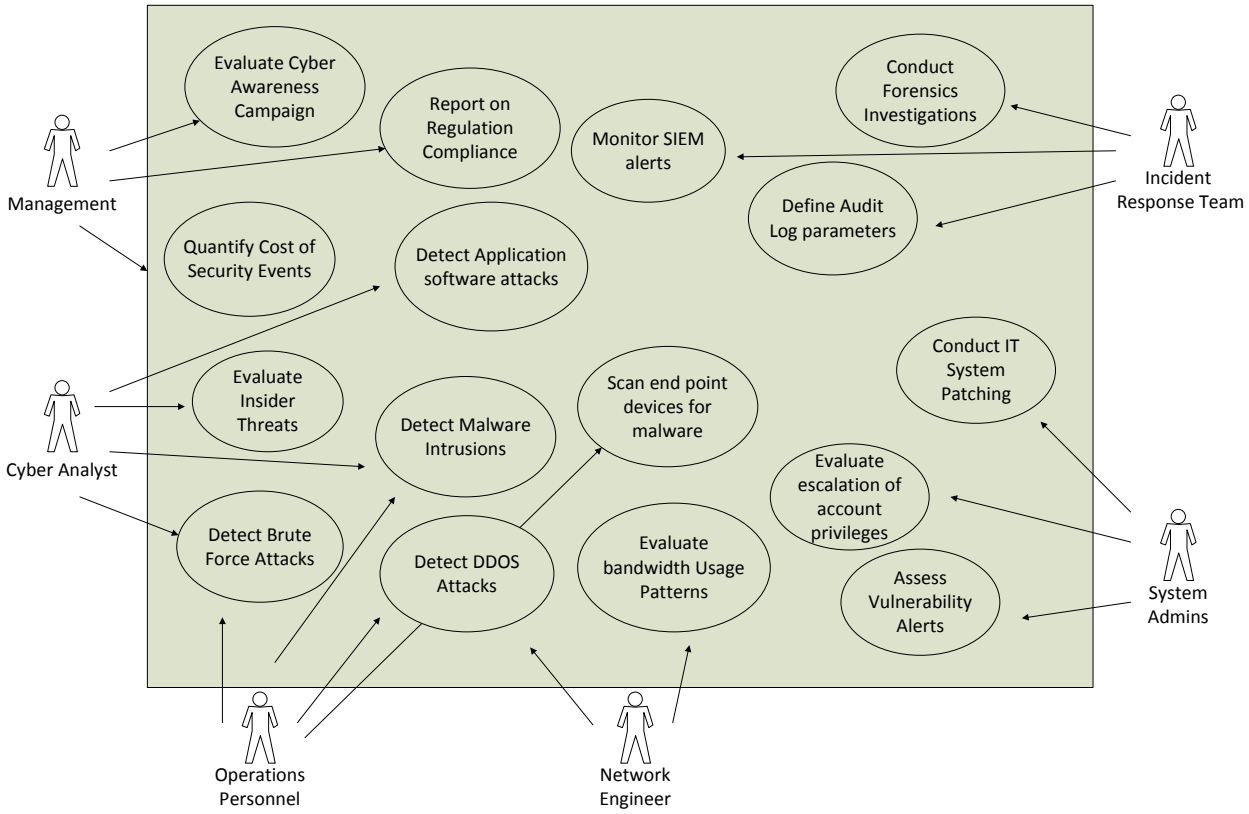
Use Case modeling – within the confines of the Unified Modeling Language (UML) - defines an 'actor' as a 'role' which interacts with a system to achieve an objective. An actor can be a person or a system and as such, can help determine the security metrics that can flow to a person via a visual display – or be accommodated elsewhere in the IT environment by other software or hardware components.

In a security engineering context, all actors are 'stakeholders' in that they need to see only the data relevant to their role in assessing security vulnerabilities and threats. If all security metrics are aggregated into actionable intelligence specific to each unique role, then security incidents can be quickly identified, analyzed and mitigated – hopefully before the impact of a data breach.

Potential consumers of security metrics include the following:

- Organizational Management
- Cyber Analysts
- Security Operations Personnel
- Incident Response Team
- System Administrators
- Network Engineers

The summation of all of these Use Cases constitutes the overall IT infrastructure security profile. This is illustrated in the following diagram. Each role focuses on utilizing the security metrics they require to conduct their required activities. Some overlap may be advantageous, but each role is focused on collecting and utilizing the security metrics required to evaluate each Use Case.



Visualization

Having defined user roles and supporting Use Cases for each, visualization aspects can now be defined and matured. The intent is to assemble the raw security metrics into forms that can enable quick response and adequately address each of the defined Use Cases.

For the cyber analyst charged with detecting malware intrusions – what graphs, diagrams, alerts and dashboard displays are most useful? Those individuals define their data need requirements and data integration development results in custom dashboards that support each role.

Because each role is tasked with solving very different problems and has unique role reporting requirements, the security metrics of interest to each can be quite different and taken together support customization of all data generating vendor products, along with custom built sensor data unique to each IT infrastructure.

The overall intent is to generate visuals in each dashboard that would support both real time alerts that necessitate quick response (such as a network DDOS attack), along with longer term trend data that would support ongoing organization operations (such as monthly status reports).

Summary

It takes an extensive amount of effort to instrument security metrics as described herein – yet things may not be as bleak as they seem at first glance. Commercial Off The Shelf (COTS) tools come out of the box with automated generation of many of these data items. Most often, this data is not then customized into a series of transaction logs that support normal business operations and enhance the cyber security protection of the organization.

Dashboards can provide a readily usable perspective into what can be an overwhelming amount of transactional data. Care must be taken to consider the role of the dashboard user and providing them the specific data which supports their IT system perspective.

Taking the time and effort to perform this level of system integration can result in a hardened defensive posture for an IT infrastructure, as well as a key reduction in the amount of time which transpires between network breach and breach detection. Cyber security is a process and it can – and should - be continually matured over time.

Organizations need evolving, maturing security metrics since enterprise conditions constantly change due to new adversarial actions. Bad actors are always probing for existing and new weaknesses and looking for exploitation methods to leverage and utilize. Cyber security controls must also constantly be updated to reflect this continuing escalation.

It is more important that attackers be uncovered in a timely manner, than trying to fortify perimeter defenses that can often be breached by a simple phishing email at a targeted user.

Not all malware is created equal. By focusing on the most extreme threats and working toward a continuous diagnostics and mitigation operational perspective, an organization is well on the way towards offering a rock solid defensive posture to an increasingly interconnected world

About the Author

Robert J. Michalsky has served government and commercial customers for more than 30 years providing a wide range of IT and analytical services. As NJVC Principal, Cyber Security, he quantifies and pursues new business opportunities in cyber security. Mr. Michalsky is a strong advocate for protecting user data through technology enablement and enhancing business processes through modeling and analysis.