





Advanced Security Analytics

Igor Kotenko

Laboratory of Computer Security Problems

St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS)

and

International Laboratory "Information security of cyberphysical systems", ITMO University

St. Petersburg, Russia

Table of content

- Introduction. Cyber situational awareness, security monitoring and SIEM systems
- Security event correlation
- Technologies of advanced security analytics
- SIEM systems and big data
- State-of-the-art
- Security evaluation framework
- Conclusion



Main aspects of cyber situation awareness

- Be aware of the current situation (situation perception)
- Be aware of the impact of the attack (impact assessment)
- Be aware of how situations evolve (situation tracking)





- Be aware of the quality (and trustworthiness) of the collected information items and the knowledge-intelligence-decisions derived from these information items.
- Assess plausible futures of the current situation (projecting future possible actions/activities of an adversary, understanding of adversary intent, opportunity, and capability as well as understanding own vulnerabilities, possible countermeasures, etc.)

[P. Barford, M. Dacier, T. G. Dietterich et al., 2010]



Well-defined metrics can help answer to the following questions

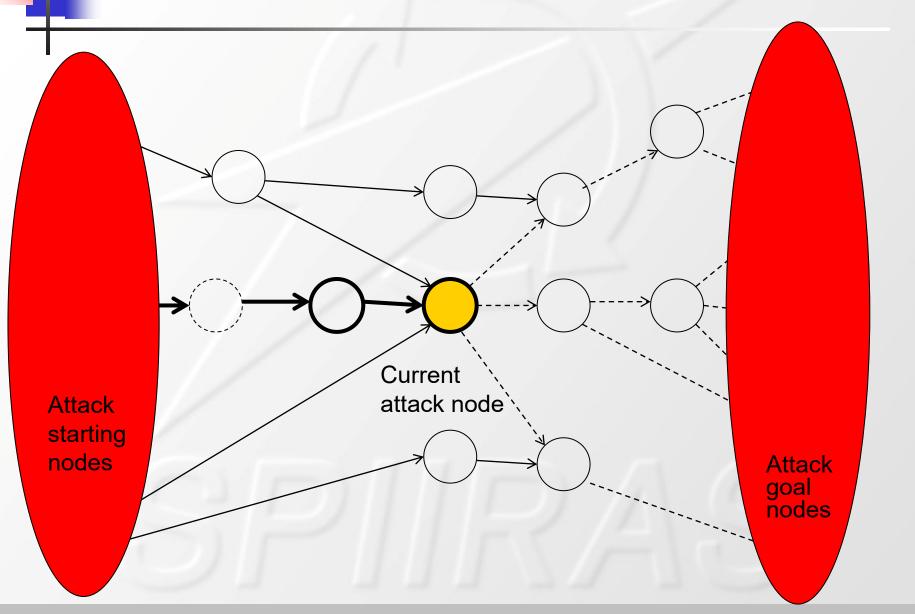
- Are there any vulnerabilities in the system? Which ones are critical? What should be eliminated first?
- Is there (currently) a network attack?
- What component (system / application / service) was and / or would be compromised?
- Who is attacking the system?
- How can you measure (potential) risk?
- What is the most likely target of the attack and the damage from the attack?
- Can we prevent an attack?
- What are the response options?
- What are the rational response options and which one is optimal?
- How many computing resources (memory, bandwidth, etc.) will be lost due to the attack?
- Is the mission / task / operation still performed (or partially)?





- Figure-out possible sequences of attacks, and to preemptively identify the security objectives that are most likely to be targeted by the attacker
- Correlate sequences of alerts as they pertain to specific actions within an attack model
- Identify appropriate sets of countermeasures, that is actions taken by the system to subvert the ongoing sequence of attacker actions
- Dynamically compute the impact of attacks and countermeasures; the former when they violate the normal security policy, and the latter when they modify the system configuration, so it no longer complies with the default policy requirements.

Using attack models for situation awareness



SIEM systems

Security information and event management (SIEM) system – security monitoring and incident management system.

The main purpose of SIEM is to increase security by providing the ability to, in near-real time mode, manipulate security information and implement proactive incident and event management.

"Proactive" means "acting before the situation becomes critical." It is assumed that proactive incident and security event management is based on automatic mechanisms that use information about the "history" of analyzed network events and the forecast of future events, as well as on automatic adjustment of event monitoring parameters to the current state of the protected system

Limitations of SIEM systems and the requirements to the new-generation SIEM systems

Limitations:

- restrictions on the target infrastructure;
- inability of multi-level interpretation of incidents and events (levels physical, network, applications, business processes);
- failure to provide a high degree of reliability and robustness of the event data collection environment;
- low scalability, etc.

Functional requirements:

- use of proactive incident and event management,
- generation of countermeasures in real time;
- intelligence, high scalability, multilevel and multidomain security event handling;
- proactive security management and reliable and robust event data collection.



Extended list of tasks solved by the SIEM system

- collection, processing and analysis of security events entering to the system from a variety of heterogeneous sources;
- real-time detection of attacks and violations of criteria and security policies;
- security assessment of information, telecommunication and other critical resources;
- analysis and management information security risks;
- investigating incidents;
- detecting the divergence of critical resources and business processes with internal security policies and bringing them in line with each other;
- development and implementation of information security solutions;
- formation of reports.



SIEM-system components

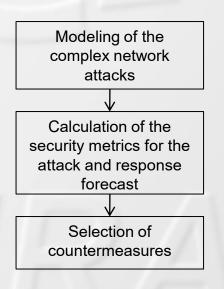
Security information and events

Security evaluation, countermeasure generation and visual analytics

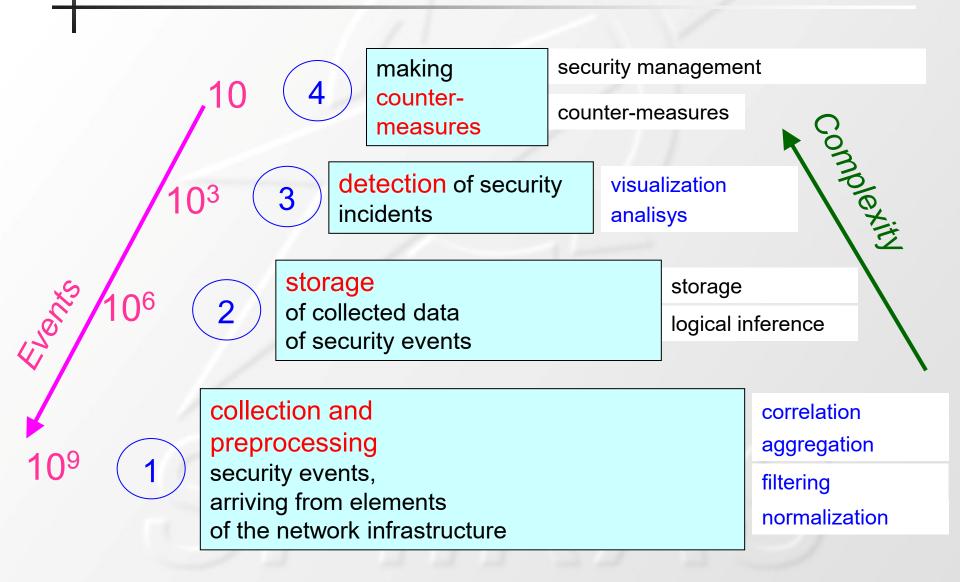
Security metrics and selected countermeasures

Functions (automate process of the information and security events processing):

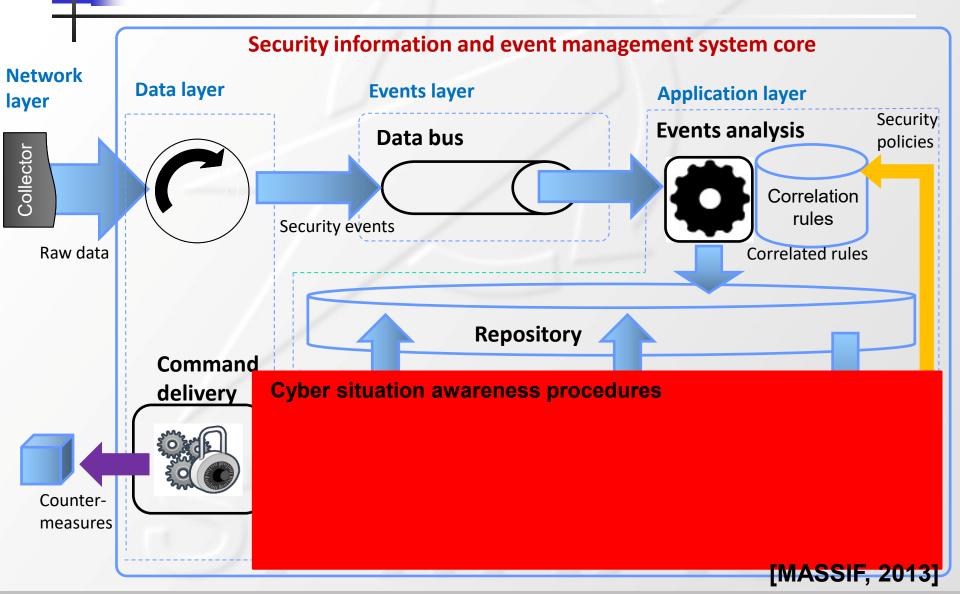
- -Log storage;
- -Event management;
- -Correlation;
- -Threat detection;
- -Vulnerability assessment;
- -Asset discovery



Stages of perspective SIEM operations



Common architecture of the SIEM system, data flows and place of cyber situation awareness



New features of next generation systems for cyber security monitoring and management

- interlevel correlation of security events from various non-uniform sources
- adaptive, scalable event processing to manage large amounts of security data in real time or near real time.
- prognostic security analysis that allows proactive detection and prevention of attacks by taking appropriate countermeasures in a near real time
- high availability and resiliency of collecting data on security events and forcing solutions in conditions of distributed infrastructure and active malicious and / or unintended impacts on communication channels
- real-time countermeasure generation
- possibility of building integrated monitoring and response systems, like SOC, or connecting to "FinCERT" of the Bank of Russia or GosSOPKA (in the case of domestic solutions)

Top SIEM vendors

Vendor/Product	Use Cases	Metrics	Intelligence	Delivery	Pricing
HPE ArcSight	Enterprises	350+ data sources, 75,000 events per second (EPS)	Integrates with machine learning, intelligence platforms	Appliance, software or cloud	Based on data ingested and events per second (EPS)
Splunk Enterprise Security	Highly-regulated industries	Most users ingest several petabytes daily	Integrates with Splunk UBA & machine learning toolkit	Software or cloud	Based on max daily data volume; starts at \$1,800/GB/day
IBM Security QRadar	Enterprises and regulated industries	400+ sources, scales to millions of events per second	UBA, forensics, packet inspection, Watson integration	Cloud or hardware, software or virtual appliance	Cloud starts at \$800/month; on- premises at \$10,400
AlienVault Unified Security Management	Lower-cost option for on-premises or AWS	Up to 15,000 EPS	Global network sharing 1 million threats daily	Cloud or virtual or hardware appliance	Lower-cost open source-based product
LogRhythm	Scales from midrange to enterprise	Highly scalable decentralized architecture	Machine analytics for advanced threats	Appliance, software or virtual instance	Subscription pricing tied to volume consumption
McAfee Enterprise Security Manager	Support for public sector, education and healthcare	50,000+events per second, billions of events stored	Automated task and policy changes	Physical or virtual appliance	Based on EPS capacity, starting at \$39,995
Micro Focus Sentinel Enterprise	MSSPs and distributed enterprises	Event taxonomy comprises more than 200 fields	Integrates with NetIQ technologies	Software or virtual appliance	Based on EPS and per device
Solar Winds Log & Event Manager	Security teams looking for easy, lower-cost solution	Up to 250 million events per day	Thresholds can be set for abnormal behavior	Virtual appliance	Starts at \$4,495 for 30 nodes
Trustwave SIEM Enterprise	Mid-market and enterprise	Millions of daily events	Analytics and threat intelligence from SpiderLabs	Appliance, software or managed service	Subscription or fee- based consulting
RSA NetWitness	Financial, government, energy, telecoms	30,000 EPS, 10Gbps & 100,000 endpoints per scalable system	Streaming analytics, machine learning, automation	On-premises, virtual, cloud and hybrid options	Based on throughput per 50 GB of logs and 1TB of packets

[eSecurityPlanet]

2017 Magic Quadrant for SIEM (Gartner)



Examples of SIEM systems developed in the Russian Federation

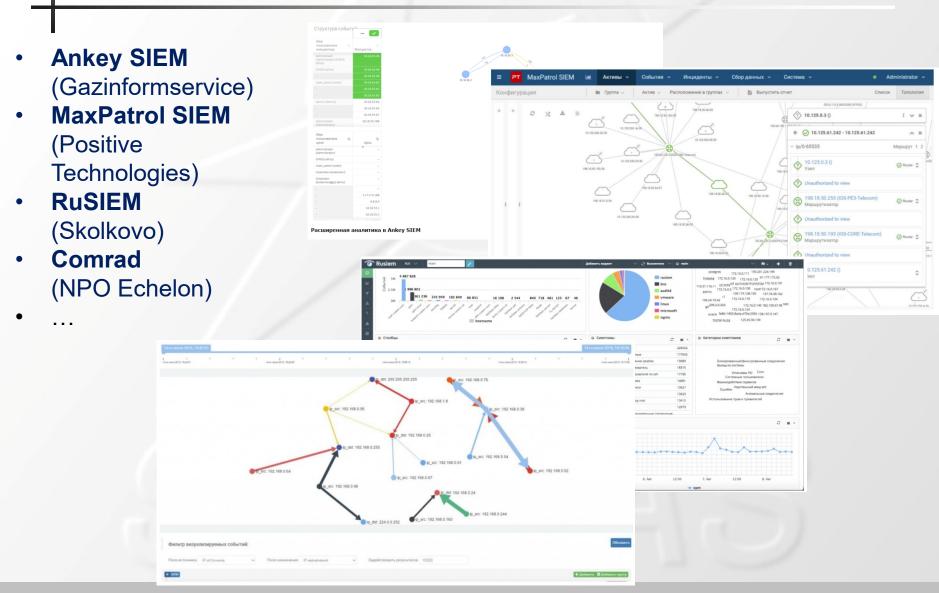


Table of content

- Introduction. Cyber situational awareness, security monitoring and SIEM systems
- Security event correlation
- Technologies of advanced security analytics
- SIEM systems and big data
- State-of-the-art
- Security evaluation framework
- Conclusion

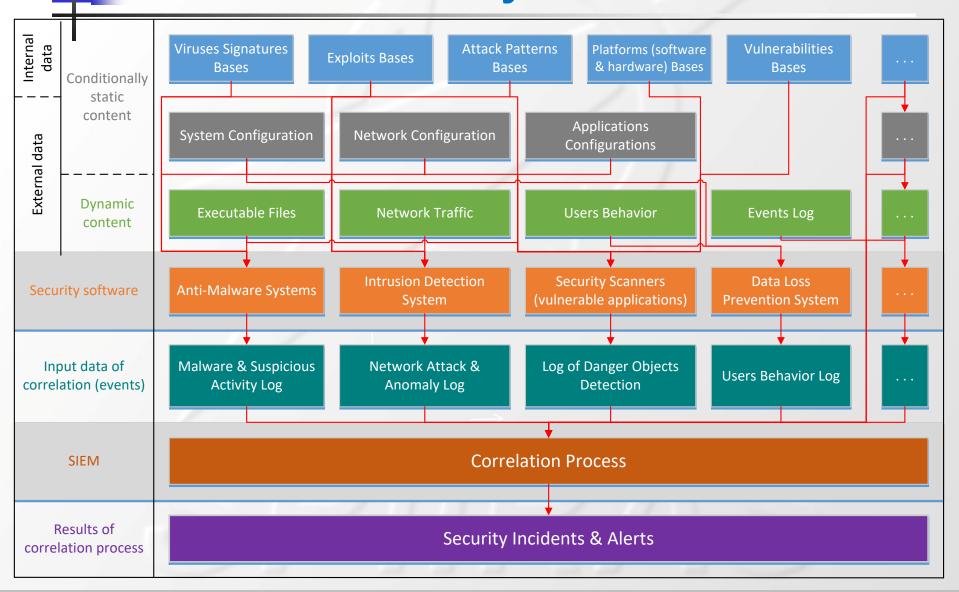




Role of the correlation process

- Identify the links between events
- Group low-level events into higher-level events
- Define the relationships between events and security information
- Range the importance of events and their groups within the security task
- Detect malicious, attacking and abnormal activity
- Detect multi-step attacks, incidents and security alerts
- Determine the source and purpose of the attack

Input data for the correlation process in SIEM systems



Stages of data correlation process

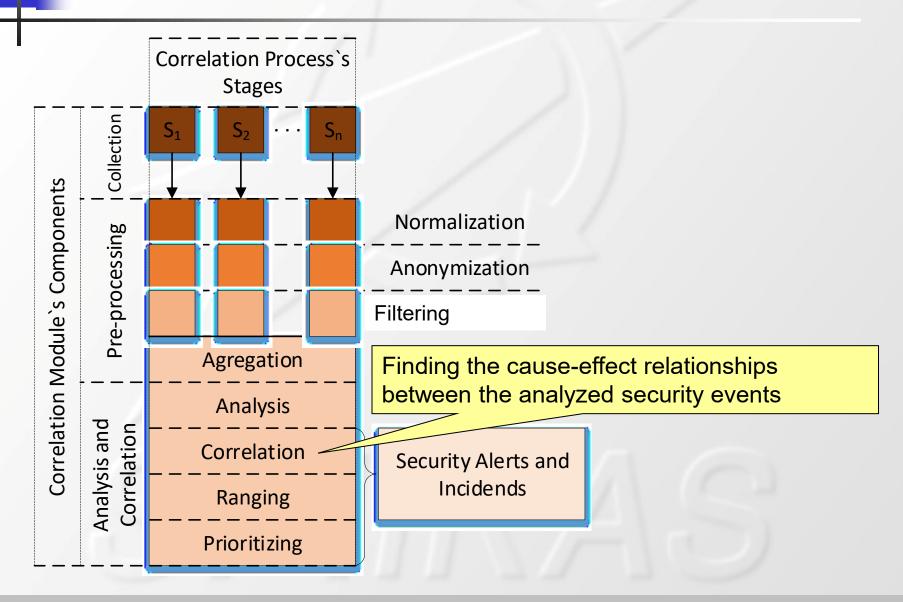
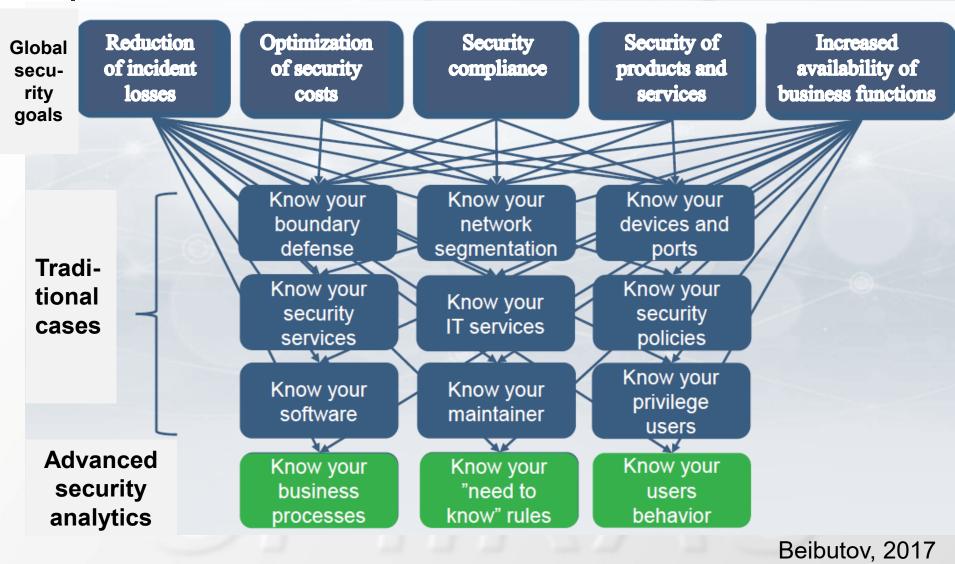


Table of content

- Introduction. Cyber situational awareness, security monitoring and SIEM systems
- Security event correlation
- Technologies of advanced security analytics
- SIEM systems and big data
- State-of-the-art
- Security evaluation framework
- Conclusion



Next-generation SIEM systems and SOC: goals and means



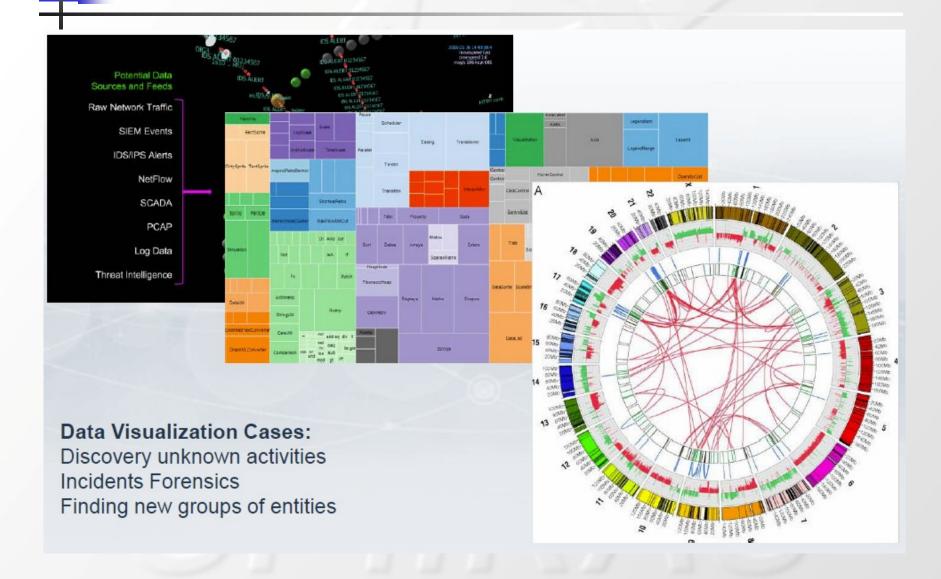








Efficient visualization of security data





SIEM vs Future Machine Learning - User Behavior Analytics (UBA)

	SIEM	UBA
Detection logic	Clear scenarios for incidents, use of basic statistics and thresholds	Machine learning
Focus on threats	A wide range of detected threats	Threats related to user activity
Types of data analyzed	Logs, NetFlow	Data from SIEM + additional contextual directories
Average interval of data analysis	In real time, historical analysis for a short time interval (3- 5 days)	Analysis of historical information for a long period (3-12 months)
Types of detected threats	Known, Known-Unknowns	Unknown, Unknown-unknowns
Additional contexts	Basic data for enrichment are used	Often they themselves become a new context

User Behavior Analytics

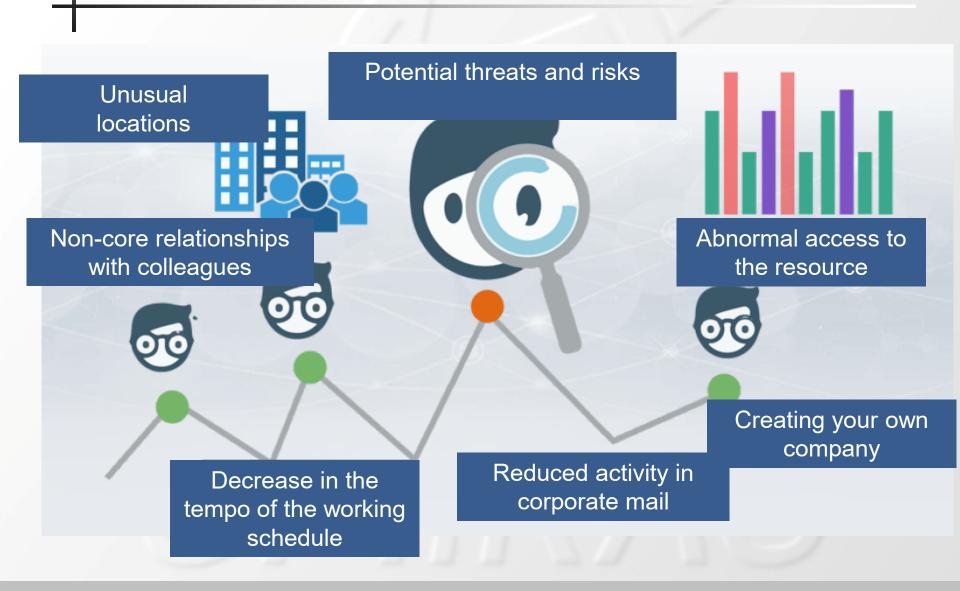
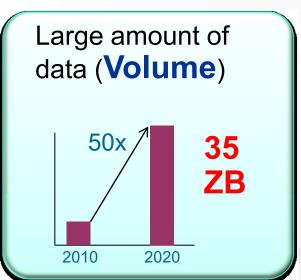


Table of content

- Introduction. Cyber situational awareness, security monitoring and SIEM systems
- Security event correlation
- Technologies of advanced security analytics
- SIEM systems and big data
- State-of-the-art
- Security evaluation framework
- Conclusion



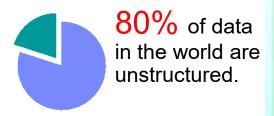
Main characteristics of big data



Big gain and high processing speed (Velocity)



Big data

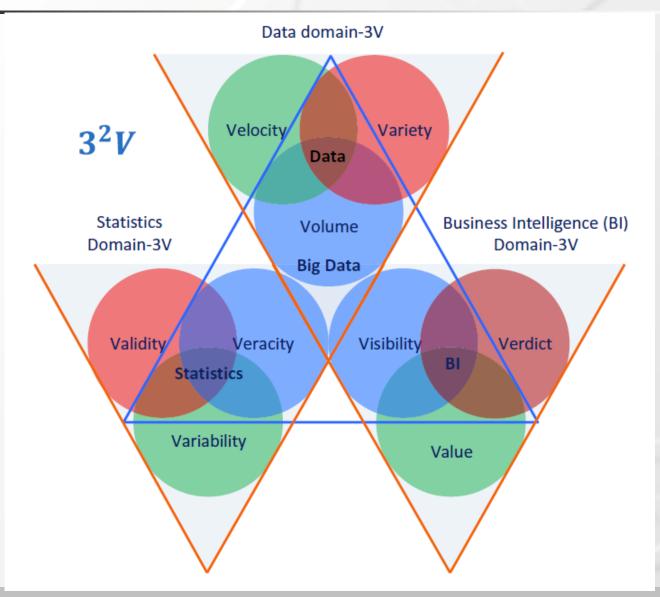


arge ifferences in ata reliability **/eracity**)

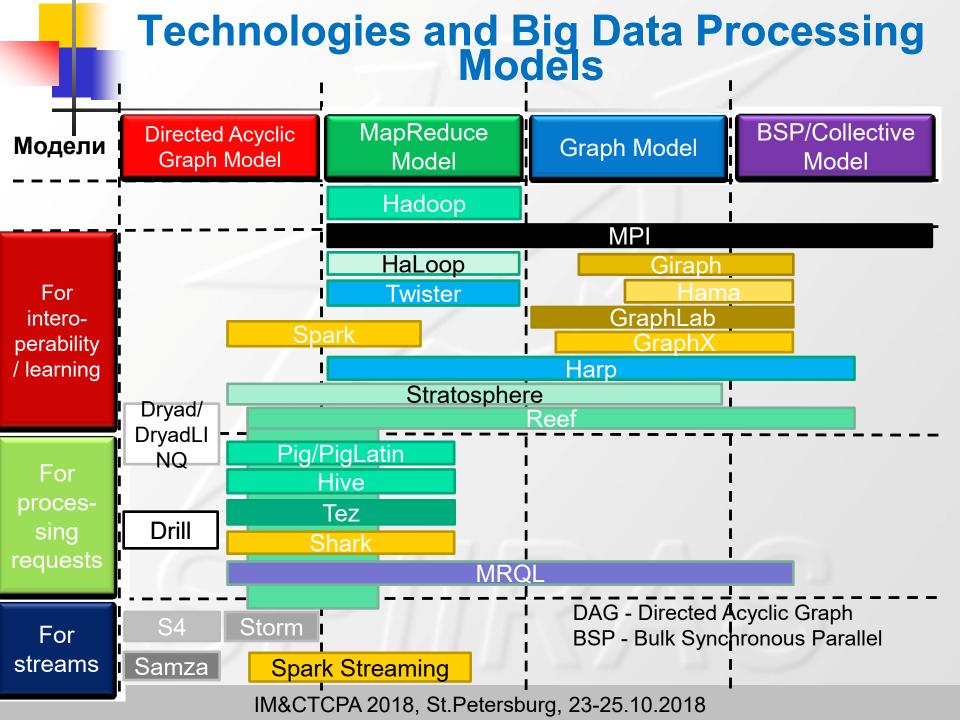
1 of 3 decision makers do not trust the information that is used to make decisions

[IBM]

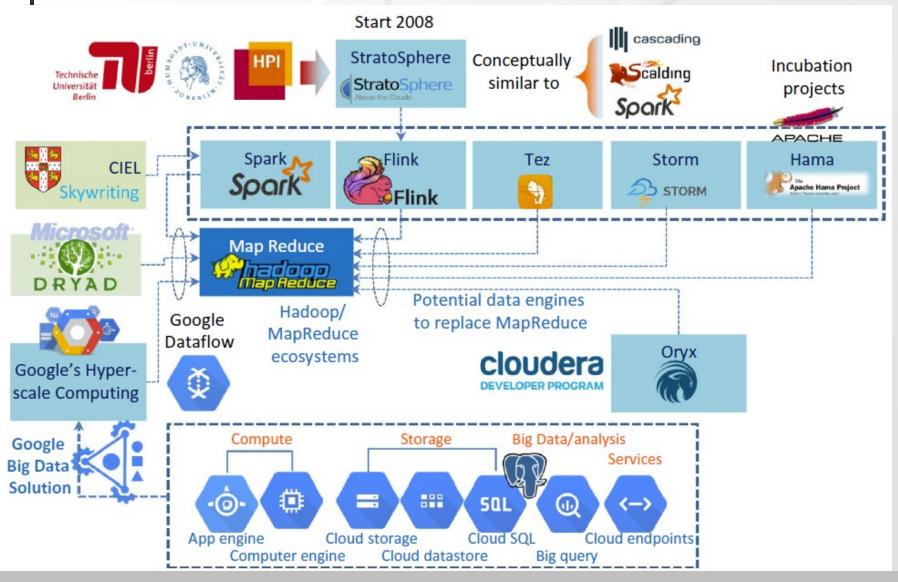
Big Data Models (9V)



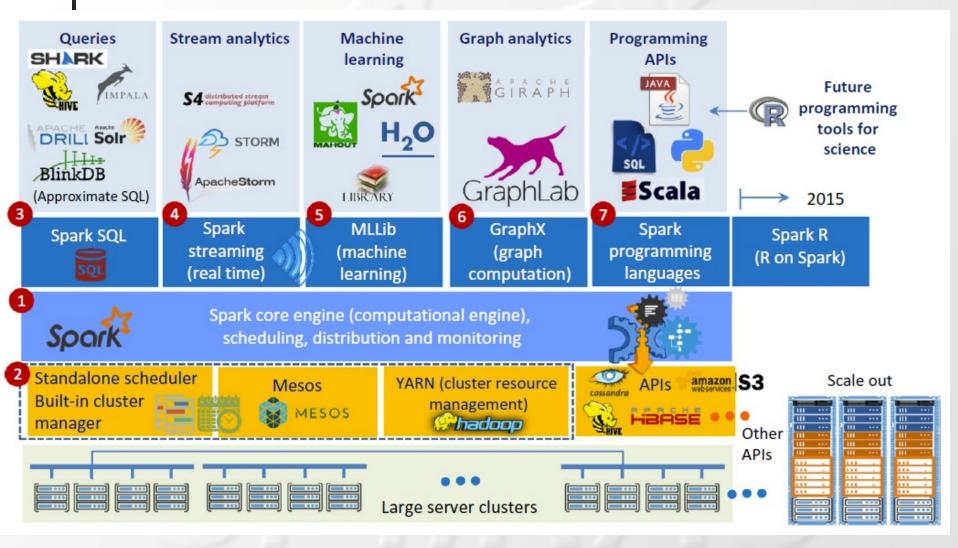
IM&CTCPA 2018, St.Petersburg, 23-25.10.2018



Technologies replacing MapReduce



Spark ecosystem



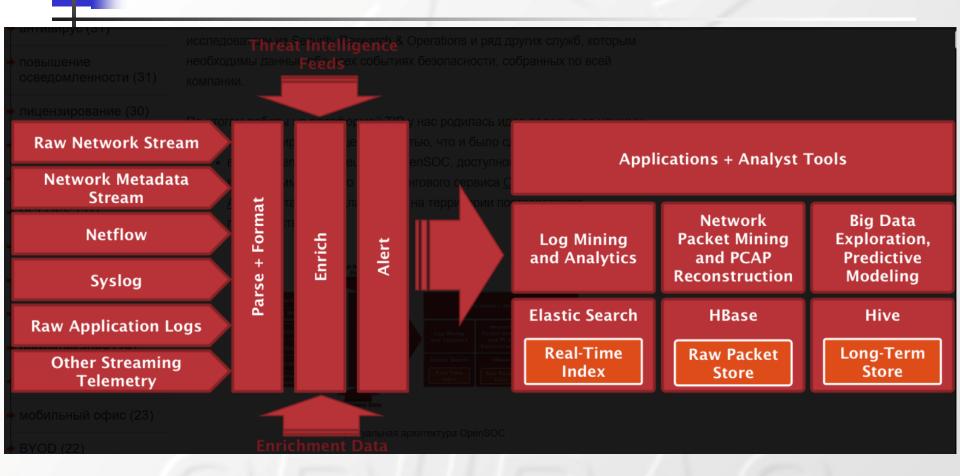
2017 Magic Quadrant for SIEM (Gartner)



Leading SIEM systems have the ability to integrate with big data platforms (native or open source, such as **Hadoo**p). For example, Fortinet (AccelOps) includes **Apache Kafka**, Intel Security, providing McAfee Enterprise Security Manager (ESM), has introduced two-way integration with **Hadoop**.

The use of Elasticsearch,
Logstash and Kibana (Elastic
Stack), OpenSOC, Apache
Metron and other tools using big
data platforms such as Hadoop in
SIEM systems allows one to
provide data collection, incident
management and analytics.

OpenSOC conceptual architecture



Apache Flume 1.4.0 +, Apache Kafka 0.8.1+, Apache Storm 0.9 +, Apache Hadoop 2.x, Apache Hive 12 +, Apache Hbase 0.94+, Elastic Search 1.1 +, MySQL 5.6+

Table of content

- Introduction. Cyber situational awareness, security monitoring and SIEM systems
- Security event correlation
- Technologies of advanced security analytics
- SIEM systems and big data
- State-of-the-art
- Security evaluation framework
- Conclusion



Correlation

Stages of correlation process in Intrusion Detection Systems (IDS):

- (1) normalization; (2) aggregation; (3) filtering; (4) anonymization; (5) prioritization;
- (6) correlation [C.Kruegel et al., 2005]

Signature-based methods of event correlation:

- Rule-based [R. Sadoddin, A. Ghorbani, 2006]
 [A. Hanemann, P. Marcu, 2008; T. Limmer and F. Dressler, 2008]
- Template-based (scenario-based) [R. Sadoddin, A. Ghorbani, 2006]
- Graph-based [A. Muller, 2009],[P. Ning and D. Xu, 2008]
- Based on finite state machines [A. Muller, 2009; A.A. Ghorbani et al., 2010]
- Based on similarity [M. A. Hasan, 1999; U. Zurutuza, R. Uribeetxeberria, 2004]
- and others.

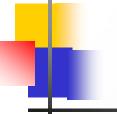
Self-learning methods of event correlation:

- Bayesian networks [R. Sadoddin, A. Ghorbani, 2006; A. Muller, 2009],
 [D.W. Guerer et al., 1996]
- Immune networks [A. Muller, 2009; D.W. Guerer et al., 1996]
- Artificial neural networks [A.Muller, 2009; D.W.Guerer et al., 1996; H.T. Elshoush and I.M. Osman, 2001]
- and others



Attack modelling (1/2) Formalisms

- Colored Petri nets [Kumar S., Spafford E.H., 1994; ...]: Each intrusion signature is defined as a pattern that represents relation between events and their context
- Model checking [C.Ramakrishnan and R.Sekar; R.Ritchey and P.Ammann; O.Sheyner; S.Jha and J.Wing; Giannakopoulou, 2011; SMV, NuSMV, SPIN, ...]: Hypothesis (system state) should be defined to check its violation with model checking technique
- Expert systems [M.Danforth Java Expert System Shell; Gamal et al., 2011; ...]: Rules are implementation of attack actions, facts are system states. Attacks have preconditions/post conditions.
- Logical approach [X.Ou, W.Boyer, M.McQueen, 2009 Datalog language; ...]: Graph consists of the input vertexes and fact vertexes.
 Network model is a set of Datalog statements, attacks are Datalog rules
- Attack graphs [Ortalo et al., 1999; Ritchey&Ammann, 2000; Sheyner et al., 2002; Rieke, 2004; Ingols, 2009; ...]: Vertexes are system states, arcs are transitions, etc.



Attack modelling (2/2) Important research directions

- Representing attack scenarios and malefactors [Schneier, 1999; Dawkins et al., 2002;[Shepard et al., 2005; ...]
- Specification of platforms, vulnerabilities, vulnerability scorings, attacks, weaknesses and configurations [NVD; OSVDB; CVE; CVSS; CPE; CCE; CWE; CAPEC; ...]
- Combining service dependency graphs with attack graphs [Kheir et al., 2009; Kheir et al., 2010; ...]
- Representing zero day attacks [Ingols et al., 2009; Wang et al., 2010; ...]



Protocols for Specification of platforms, vulnerabilities, vulnerability scorings, attacks, weaknesses and configurations (1/2)

Security Content Automation Protocol (SCAP)

- Common Vulnerabilities and Exposures (CVE)
- Common Configuration Enumeration (CCE)
- Common Platform Enumeration (CPE)
- eXtensible Checklist Configuration Description Format (XCCDF)
- Open Vulnerability Assessment Language (OVAL)
- Common Vulnerability Scoring System (CVSS)

Threat Analysis Automation Protocol (TAAP)

- Malware Attribute Enumeration & Characterization (MAEC)
- Common Attack Pattern Enumeration & Classification (CAPEC)
- Common Platform Enumeration (CPE)
- Common Weakness Enumeration (CWE)
- Open Vulnerability and Assessment Language (OVAL)
- Common Configuration Enumeration (CCE)
- Common Vulnerabilities and Exposures (CVE).





Protocols for Specification of platforms, vulnerabilities, vulnerability scorings, attacks, weaknesses and configurations (2/2)

- Event Management Automation Protocol (EMAP)
 - Common Event Expression (CEE)
 - Malware Attribute Enumeration & Characterization (MAEC)
 - Common Attack Pattern Enumeration & Classification (CAPEC).
- Incident Tracking and Assessment Protocol (ITAP)
 - Open Vulnerability and Assessment Language (OVAL)
 - Common Platform Enumeration (CPE)
 - Common Configuration Enumeration (CCE)
 - Common Vulnerabilities and Exposures (CVE)
 - Common Vulnerability Scoring System (CVSS)
 - Malware Attribute Enumeration & Characterization (MAEC)
 - Common Attack Pattern Enumeration & Classification (CAPEC)
 - Common Weakness Enumeration (CWE)
 - Common Event Expression (CEE)
 - Incident Object Description Exchange Format (IODEF)
 - National Information Exchange Model (NIEM)
 - Cybersecurity Information Exchange Format (CYBEX).

Complex event processing

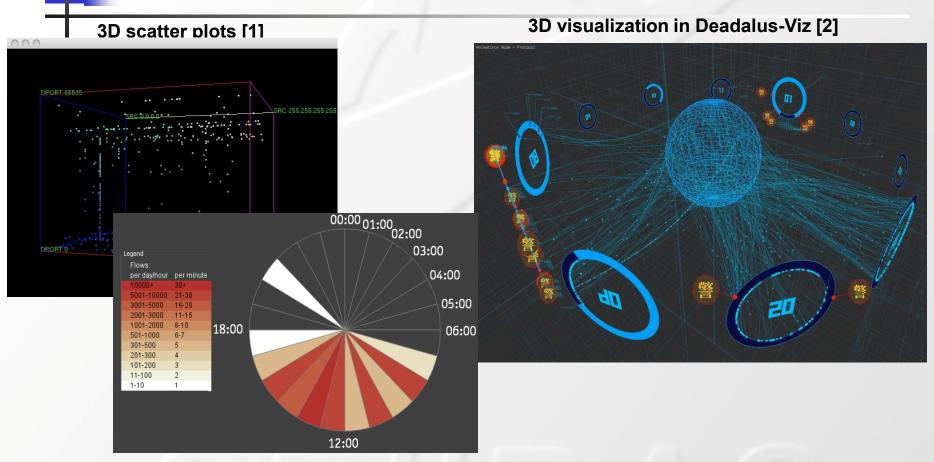
- [Gyllstrom et al.] considered the system intended for collection, cleaning, and processing of RFID data. However, the parallel stream processing is not analyzed.
- [Liu et al., 2010] suggested the frameworks allowing to process web data by means of CEP engines. However, parallel computing is not considered.
- [Wang et al., 2011] suggested to use active rules within the CEP engine. However, extension of these results on parallel event processing in IoT is not justified.
- [Gulisano et al., 2010 and 2012] considered a highly scalable data streaming infrastructure for CEP. In our work, we are guided by these results, but try to develop them further.

Visual analytics

- (1) Models for Network Perimeter Monitoring
 - 3D scatter plots
 - 3D visualization in Deadalus-Viz
 - ClockView of network traffic
 - **...**
- (2) Models for Policy Assessment
 - Matrices
 - Graphs
 - Treemaps
 - SpiralView
 - Starburst
 - ...

- (3) Vulnerability Assessment
 - Treemaps
 - Histograms
 - ...
- (4) Models for Attack Graph Analysis
 - Matrices
 - Graphs
 - Treemaps
 - •
- (5) Combination of security metrics
 - Maps
 - Dial-based
 - Gauge clusters
 - · ...

(1) Models for Network Perimeter Monitoring

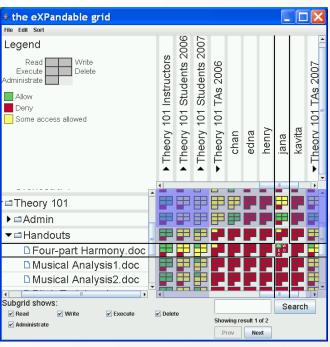


ClockView of network traffic [3]

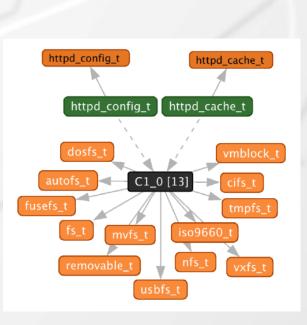
- [1] Lau S. The spinning cube of potential doom. In Communications of the ACM, vol. 47(6), 2004. P.24-26.
- [2] Inoue D., Eto M., Suzuki K., Suzuki M., Nakao K. DAEDALUS-VIZ: Novel Real-time 3D Visualization for Darknet Monitoring-based Alert System". Proc. VizSec '12, October 15, Seattle, WA, USA (2012)
- [3] C. Kintzel, J. Fuchs and F. Mansmann. "Monitoring Large IP Spaces with ClockView". In Proc. of Int. Symp. on Visualization for Cyber Security (VizSec), 2011.



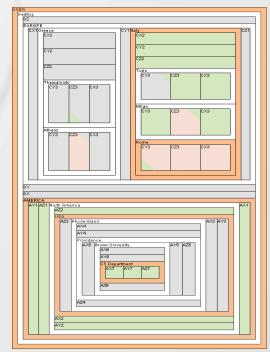
Matrix-based visualization of access rules [1]



Graph-based visualization of access rules [2]



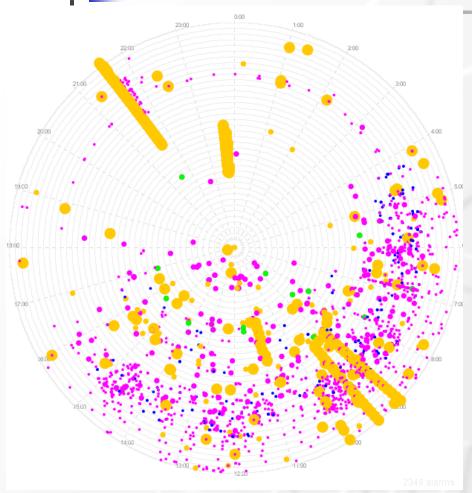
Treemap-based visualization of access rules [3]



- [1] R. W. Reeder, L. Bauer, L. F.Cranor, M. K. Reiter, K. Bacon, K.How, and Heather Strong. 2008. Expandable grids for visualizing and authoring computer security policies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '08).ACM, New York, NY, USA, 1473-1482.
- [2] S. Marouf, M. Shehab. SEGrapher: Visualization-based SELinux PolicyAnalysis // Proc. of 4th Symposium on Configuration Analytics and Automation (SAFECONFIG), 2011 P. 1 8.
- [3] Heitzmann, A., Palazzi, B., Papamanthou, C., Tamassia, R.: Effective Visualisation of File System Access-Control. Proc. of the 5th international workshop on Visualisation for Computer Security (VizSec'08), LNCS, Vol.5210, Springer-Verlag, Berlin, Heidelberg, pp.18-25 (2008).

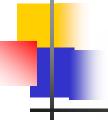


SpiralView for policy assessment [Bertini et al., 2007]

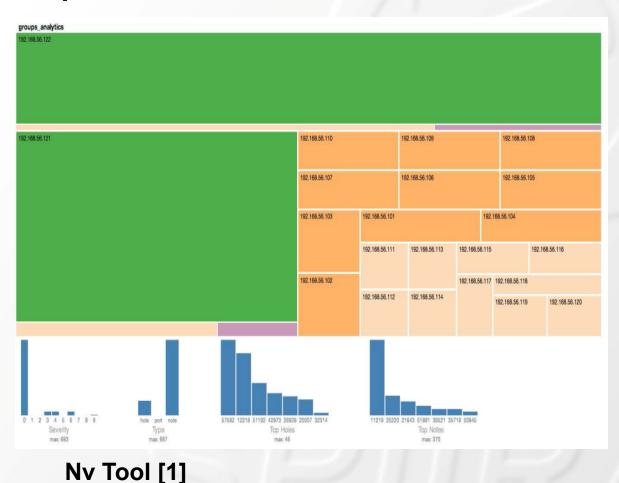


Bertini E., Hertzog P., Lalanne D. SpiralView: Towards Security Policies Assessment through Visual Correlation of Network Resources with Evolution of Alarms. In Proceedings of the IEEE Symposium on Visual Analytics Science and Technology (VAST) 2007. pp.139-146.

- Spiral axises represent time-based structures on which alarms are positioned using their time of appearance in the network.
- All the alarms generated in the system in the last k months are displayed on the k rings, starting from the older in the center up to the newer alarms in the outer ring.
- The choice of the spiral shape :
- (1) it can represent data sequentially;
 (2) it exposes periodic behaviour through radial alignments of objects;
 (3) it assigns more space to recent alarms
 - The colour of alarms represents alarm type (User, Application Behaviour, Scan and Propagation, etc.). Their size is mapped to alarm severity.



(3) Models for Vulnerability Assessment [Harison et al., 2012]



Nv tool uses treemaps and linked histograms to allow security analysts and systems administrators to analyze vulnerabilities detected by the Nessus vulnerability scanner.

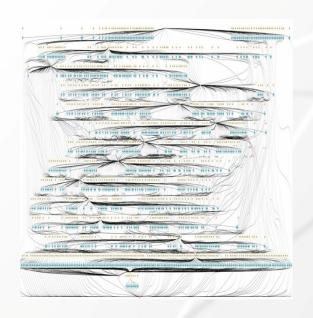
Nv tool uses a semantic based color scheme where, for example, different colors are used for fixed vulnerabilities, new ones, and open vulnerabilities

[1] Harrison, L., Spahn, R., Iannacone, M., Downing, E., Goodall, J.R.: NV: Nessus Vulnerability Visualisation for the Web. Proc. of the VizSec'12, October 15 2012, Seattle, WA, USA (2012)

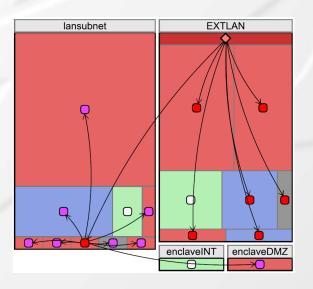


(4) Models for Attack Graph Analysis

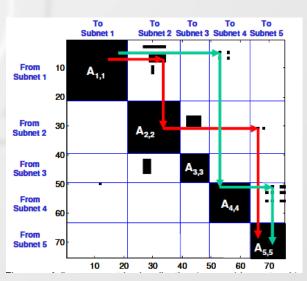
Graph-based visualization of access rules [1]



Treemap-based visualization of access rules [2]

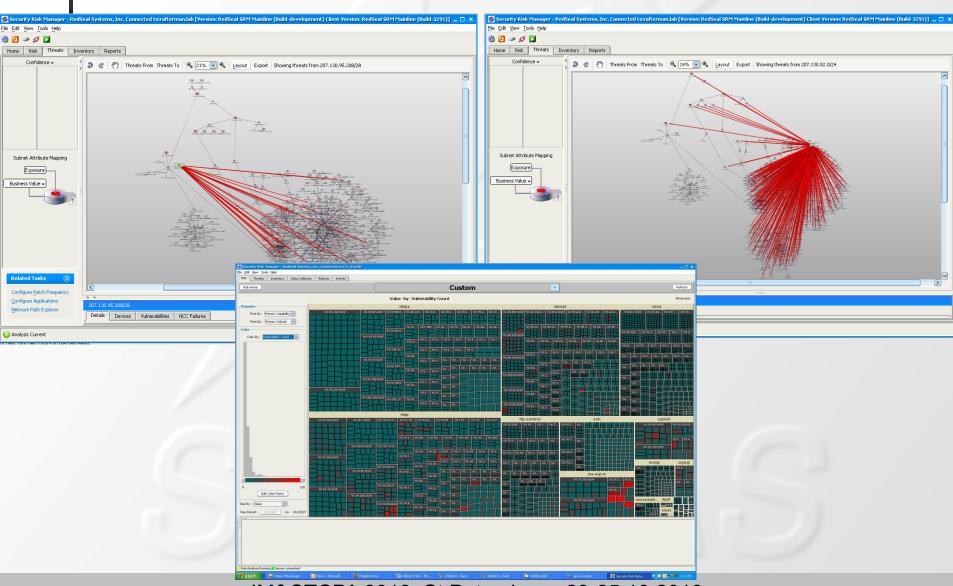


Matrix-based visualization of access rules [3]



- [1] Noel, S., Jacobs, M., Kalapa, P., Jajodia, S.: Multiple Coordinated Views for Network Attack Graphs. Proc. of the IEEE Workshops on Visualisation for Computer Security, IEEE Computer Society, pp.12 (2005)
- [2] Williams, L., Lippmann, R., Ingols, K.: GARNET: A Graphical Attack Graph and Reachability Network Evaluation Tool. Proc. of the 5th International Workshop on Visualisation for Computer Security (VizSec'08), LNCS, Vol.5210, Springer-Verlag, Berlin, Heidelberg, pp.44-59 (2008)
- [3] Noel, S., Jajodia, S.: Understanding Complex Network Attack Graphs through Clustered Adjacency Matrices. Proc. of the 21st Annual Computer Security Applications Conference (ACSAC'05). IEEE Computer Society, pp.160-169 (2005)

Graph-based and treemap-based visualization of attack steps (in Security Risk Manager, RedSeal)



IM&CTCPA 2018, St.Petersburg, 23-25.10.2018

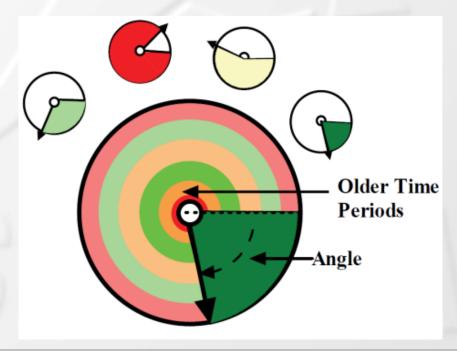




Risk Map displays information on the Risk (R), Vulnerability (V) and Availability (A) status of each network object located on the map, this information is presented in the form of traffic lights



- Each metric is represented by the dial, and its value is reinforced with color to make perception of the value more quickly. The outer ring provides the most current value
- The set of the metrics is represented by the cyber command gauge cluster purposed to support decision making and other specialized security tasks





- It displays three types of trust into one indicator, color is used to outline trust value.
- These parameters are represented by a section of the outer ring of the circle. The small circle in the center represents the overall trust, computed as a weighted sum of the other types of trust

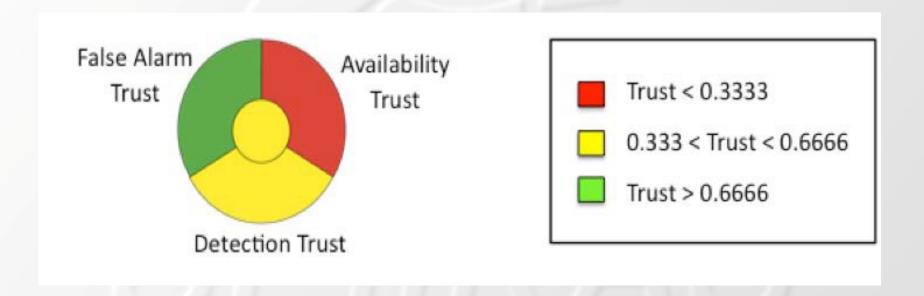


Table of content

- Introduction. Cyber situational awareness, security monitoring and SIEM systems
- Security event correlation
- Technologies of advanced security analytics
- SIEM systems and big data
- State-of-the-art
- Security evaluation framework
- Conclusion



Table of content

- Introduction
- Technologies of advanced security analytics
- State-of-the-art
- Security evaluation framework
- Metrics calculation techniques
- Visual analytics models
- Implementation
- Case study and evaluation
- Conclusion

Main processes and models

Data storage module

External sources

System
specification
Hosts, security
policies,
network topology

Operator
Security
requirements
Malefactor model

Attack Model Generation

Zero-day vulnerabilities, known vulnerabilities

Topological Vulnerability Analysis

Service dependencies

Attack analysis

Stochastic simulation

Security metrics, possible consequences of attacks, the effectiveness of countermeasures

Network weaknesses

Decision support

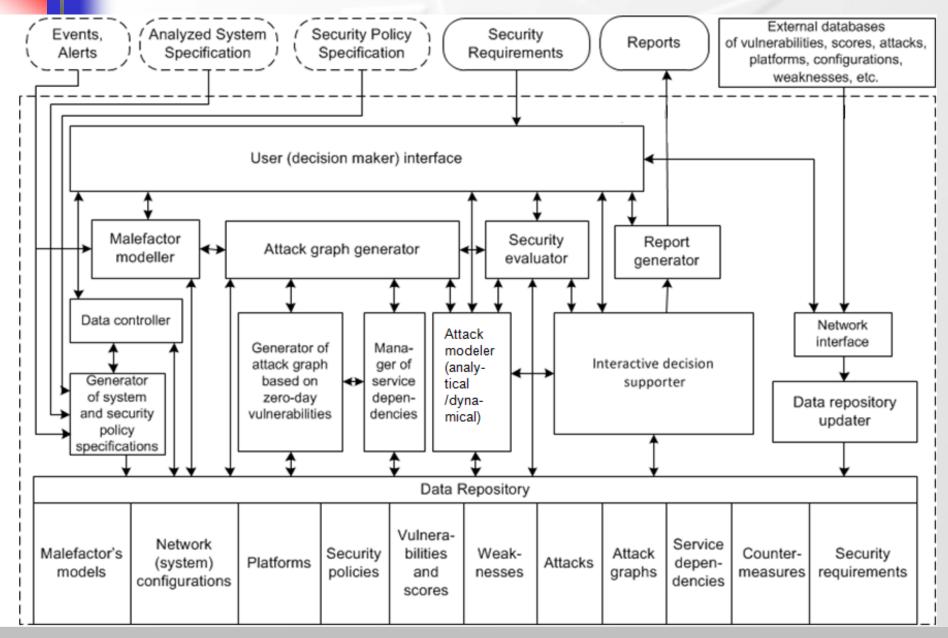
Generating possible countermeasures

Choice of countermeasures

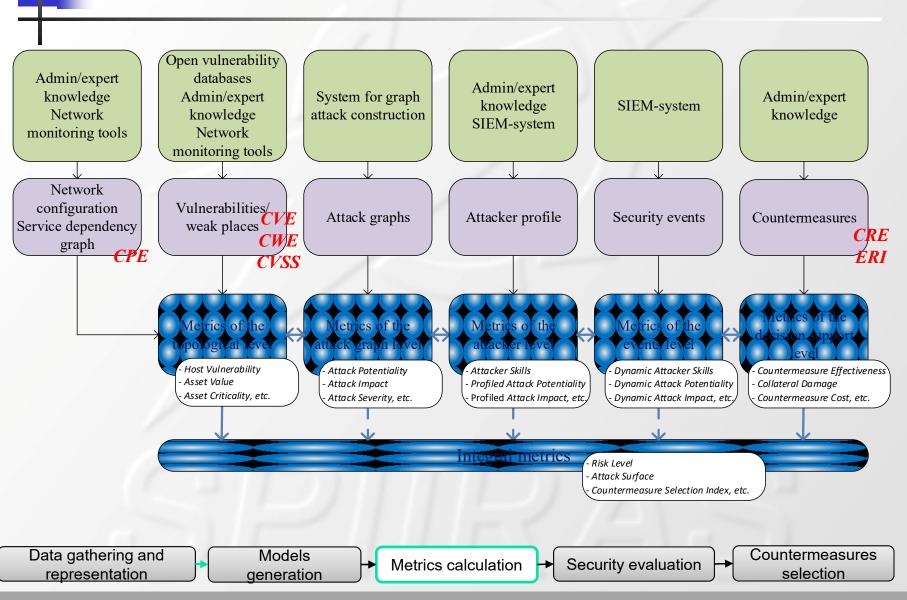
Report generation

Evaluation of countermeasures

Common architecture



Metrics calculations



CRiSIS 2015, Mytilene, Lesvos, Greece, July 20-22, 2015

Main view of the VizSecAnalyzer

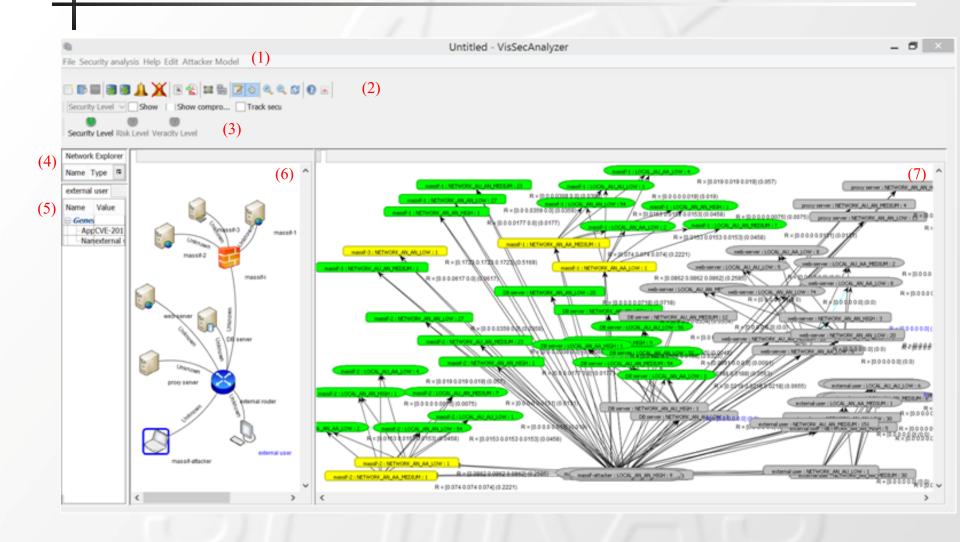


Table of content

- Introduction. Cyber situational awareness, security monitoring and SIEM systems
- Security event correlation
- Technologies of advanced security analytics
- SIEM systems and big data
- State-of-the-art
- Security evaluation framework
- Conclusion





- The characteristic of modern technologies and scientific research in the field of analytical processing of security data in SIEM-systems is given.
- The proposed general approach, architecture, and implemented prototypes of systems for collecting, storing and analytical data processing and security events are presented.
- Future research and development will be aimed at further improving the system architecture, studying the interaction of components with each other for event processing and security information, implementation of analytical processing components, as well as analysis and experimental evaluation of system performance parameters for various event streams and security information.

Questions

Thank you for your attention Questions?

Contact information:

Igor Kotenko (ivkote@comsec.spb.ru)