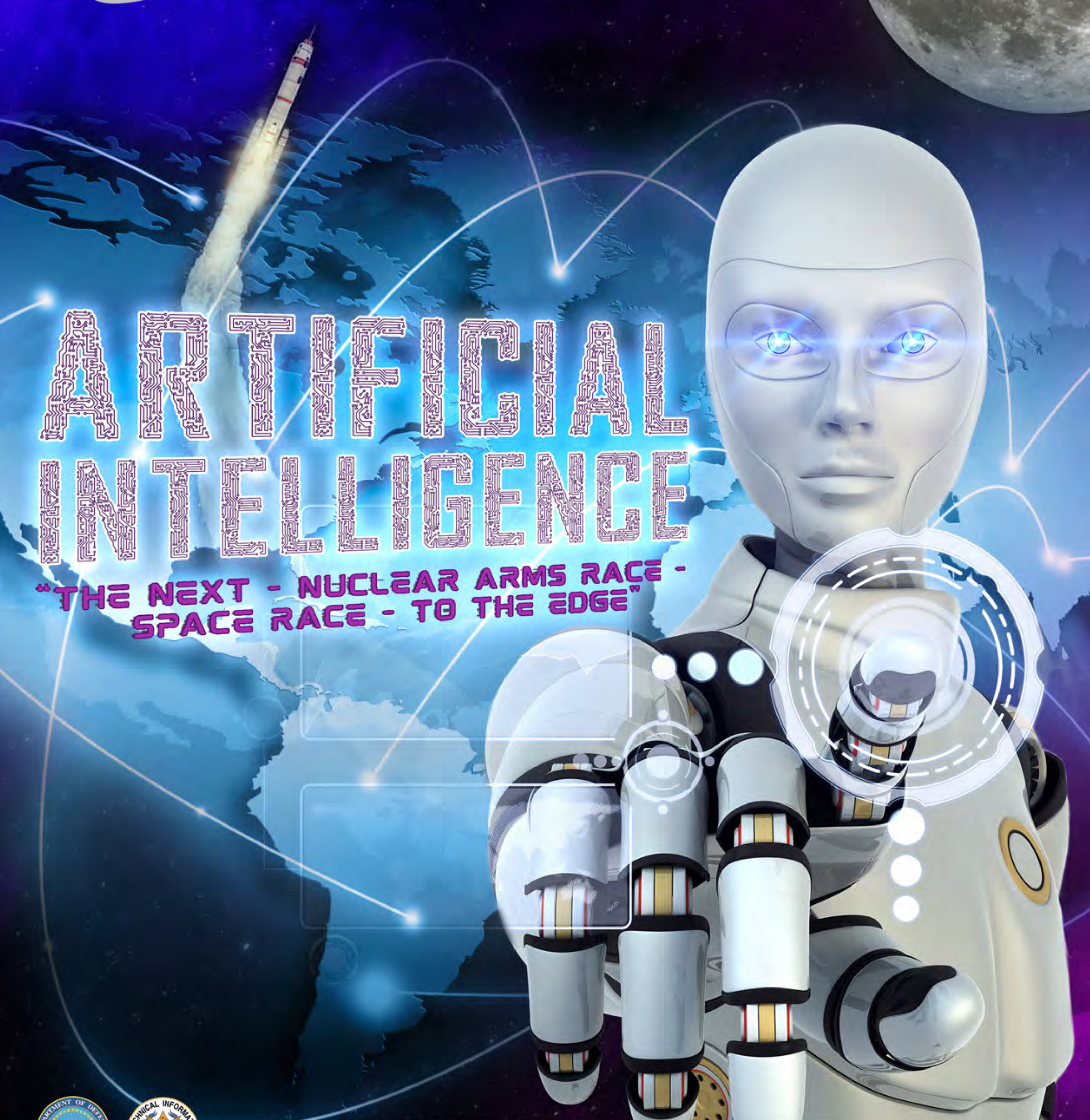
VOLUME 7 NUMBER 1
APRIL 2019

CSIAC JOURNAL

A Quarterly Publication of the Cyber Security & Information Systems Information Analysis Center



ARTIFICIAL INTELLIGENCE

"THE NEXT - NUCLEAR ARMS RACE - SPACE RACE - TO THE EDGE"

DISTRIBUTION STATEMENT A: APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED.

CSIAC

Cyber Security & Information Systems
Information Analysis Center

ABOUT THE CSIAC

As one of three DoD Information Analysis Centers (IACs), sponsored by the Defense Technical Information Center (DTIC), CSIAC is the Center of Excellence in Cyber Security and Information Systems. CSIAC fulfills the Scientific and Technical Information (STI) needs of the Research and Development (R&D) and acquisition communities. This is accomplished by providing access to the vast knowledge repositories of existing STI as well as conducting novel core analysis tasks (CATs) to address current, customer focused technological shortfalls.

OUR MISSION

CSIAC is chartered to leverage the best practices and expertise from government, industry, and academia in order to promote technology domain awareness and solve the most critically challenging scientific and technical (S&T) problems in the following areas:

- Cybersecurity and Information Assurance
- Software Engineering
- Modeling and Simulation
- Knowledge Management/Information Sharing

The primary activities focus on the collection, analysis, synthesis, processing, production and dissemination of Scientific and Technical Information (STI).

OUR VISION

The goal of CSIAC is to facilitate the advancement of technological innovations and developments. This is achieved by conducting gap analyses and proactively performing research efforts to fill the voids in the knowledge bases that are vital to our nation. CSIAC provides access to a wealth of STI along with expert guidance in order to improve our strategic capabilities.

WHAT WE OFFER

We provide expert technical advice and assistance to our user community. CSIAC is a competitively procured, single award contract. The CSIAC contract vehicle has Indefinite Delivery/Indefinite Quantity (ID/IQ) provisions that allow us to rapidly respond to our users' most important needs and requirements.

Custom solutions are delivered by executing user defined and funded CAT projects.

CORE SERVICES

- Technical Inquiries: up to 4 hours free
- Extended Inquiries: 5 - 24 hours
- Search and Summary Inquiries
- STI Searches of DTIC and other repositories
- Workshops and Training Classes
- Subject Matter Expert (SME) Registry and Referrals
- Risk Management Framework (RMF) Assessment & Authorization (A&A) Assistance and Training
- Community of Interest (COI) and Practice Support
- Document Hosting and Blog Spaces
- Agile & Responsive Solutions to emerging trends/threats

PRODUCTS

- State-of-the-Art Reports (SOARs)
- Technical Journals (Quarterly)
- Cybersecurity Digest (Semimonthly)
- RMF A&A Information
- Critical Reviews and Technology Assessments (CR/TAs)
- Analytical Tools and Techniques
- Webinars & Podcasts
- Handbooks and Data Books
- DoD Cybersecurity Policy Chart

CORE ANALYSIS TASKS (CATS)

- Customer tailored R&D efforts performed to solve specific user defined problems
- Funded Studies - $1M ceiling
- Duration - 12 month maximum
- Lead time - on contract within as few as 6-8 weeks

CONTACT INFORMATION

266 Genesee Street
Utica, NY 13502

1 (800) 214-7921

info@csiac.org

/DoD_CSIAC

/CSIAC

/CSIAC

CSIAC is operated by Quanterion Solutions Inc and sponsored by the Defense Technical Information Center (DTIC)
266 Genesee Street Utica, NY 13502 | 1 (800) 214-7921 | info@csiac.org | https://www.csiac.org

ABOUT THE JOURNAL OF CYBER SECURITY AND INFORMATION SYSTEMS

ABOUT THIS PUBLICATION

The Journal of Cyber Security and Information Systems is published quarterly by the Cyber Security and Information Systems Information Analysis Center (CSIAC). The CSIAC is a Department of Defense (DoD) Information Analysis Center (IAC) sponsored by the Defense Technical Information Center (DTIC) and operated by Quanterion Solutions Incorporated in Utica, NY.

Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the CSIAC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the CSIAC, and shall not be used for advertising or product endorsement purposes.

ARTICLE REPRODUCTION

Images and information presented in these articles may be reproduced as long as the following message is noted:

"This article was originally published in the CSIAC Journal of Cyber Security and Information Systems Vol.7, No 1"

In addition to this print message, we ask that you notify CSIAC regarding any document that references any article appearing in the CSIAC Journal.

Requests for copies of the referenced journal may be submitted to the following address:

Cyber Security and Information Systems
266 Genesee Street
Utica, NY 13502

Phone: 800-214-7921
Fax: 315-732-3261
E-mail: info@csiac.org

An archive of past newsletters is available at **https://www.csiac.org/journal/.**

*To unsubscribe from CSIAC Journal Mailings please email us at **info@csiac.org** and request that your address be removed from our distribution mailing database.*

JOURNAL OF CYBER SECURITY AND INFORMATION SYSTEMS

Artificial Intelligence the Next "Nuclear Arms Race" - "Space Race To The Edge"

ARTIFICIAL INTELLIGENCE THE NEXT "NUCLEAR ARMS RACE" - "SPACE RACE TO THE EDGE"

Figuring out where and how Artificial Intelligence (AI) and its various sub-types (Machine Learning, Deep Learning, etc.) fit into our world as we move into the future is difficult.

In some cases it seems straightforward; AI/ML speech recognition is astoundingly good and can be applied across many domains in a meaningful way. Mostly successful demonstrations of autonomous vehicles of all types portend many possible good implementations that are arguably "better" than how we do things now. For the military, using capabilities that are AI-enabled have the potential to keep personnel safe, reduce casualties, and improve mission success rates. Most all of them have to do with the combination of AI and lots of Data, so as to make sure (train, bound, qualify, quantify) they perform as intended. And that combination (AI and Data) is where the difficulty tends to migrate. We are far from the simplistic view of putting tons of data into the hopper, pushing the big red "AI" button, and turning the crank to get the results we want. The absolute truth today is that successful implementation of AI depends primarily on the expertise of people who know how to curate data, tune algorithms, and understand the intent/domain to build goal scenarios. Then, through large numbers of iterations over time, the results of using AI in controlled situations is reviewed, further tweaked and tuned, and pondered as to why "that just doesn't look right" (TJDLR) – a definitively human operation, at least as of now. Deep Learning (DL), one of the most promising and prominent areas of AI research today, is not immune to this combination. For all its promise, DL's heavy dependence on large amounts of pertinent (the truly hard part) data can cause it to react in very unpredictable (from a human perspective) ways.

To get an understanding of some of the thought that goes into "getting to" AI through domain experts, this edition of the Journal highlights three very different views of complex situations where AI might, should, and does intersect with our ability to use AI effectively.

The first article is focused on the impact of quantum computing and cryptography, with a reference to the role that machine learning might play in the future of post-quantum cryptography. This is another possible future intersection between AI and Data that will need domain expertise (human-centric, certainly at first) to determine what kinds of algorithms need to be applied, and what kind of data needs to be provided to move ahead.

The second article represents a view into the domain expertise necessary to include autonomy into a scenario in an effective way. This operationally-focused article highlights the importance of understanding the domain (essentially the frame of reference) which the AI/autonomy must be able to reason within.

Even a straightforward scenario like the one provided shows the immense investment in understanding before it can be augmented with an AI capability.

The levels of expertise necessary to get to a successful full implementation of AI to reach a goal are many and varied. The third entry is a more illustrative step-through article showing a methodology of implementing an AI algorithm on a set of data to reach a goal. While more tutorial, it reveals the many steps involved in getting to an actual result. As frameworks evolve, the steps may be refined and made more streamlined, but they are still steps that must be understood before they can be automated.

Ultimately, that is one of the questions we have to ask of AI. How much of AI can be used to assist human activities, and how much can be used to replace human activities. With each level of automation/intelligence that we levy onto the AI "plate", what are we gaining and what are we losing - and, can we understand the difference?

Features and Operation of an AUTONOMOUS AGENT FOR CYBER DEFENSE

By: Michael J. De Lucia, Allison Newcomb, and Alexander Kott, U.S. Army Research Laboratory

AN EVER INCREASING NUMBER OF BATTLEFIELD DEVICES THAT ARE CAPABLE OF COLLECTING, PROCESSING, STORING, AND COMMUNICATING INFORMATION ARE RAPIDLY BECOMING INTERCONNECTED.

The staggering number of connected devices on the battlefield greatly increases the possibility that an adversary could find ways to exploit hardware or software vulnerabilities, degrading or denying Warfighters the assured and secure use of those devices. Autonomous software agents will become necessities to manage, defend, and react to cyber threats in the future battlespace.

The number of connected devices increases disproportionately to the number of cyber experts that could be available within an operational environment. In this paper, an autonomous agent capability and a scenario of how it could operate are proposed. The goal of developing such capability is to increase the security posture of the Internet of Battlefield Things and meet the challenges of an increasingly complex battlefield.

This article describes an illustrative scenario in a notional use case and discusses the challenges associated with such autonomous agents. We conclude by offering ideas for potential research into developing autonomous agents suitable for cyber defense in a battlefield environment.

MOTIVATION AND CONTEXT

The Internet of Battlefield Things (IoBT) is "a set of interdependent and interconnected entities (e.g., sensors, small actuators, control components, networks, information sources)" that are composed and connected dynamically to support the goals of a military mission (US Army Research Laboratory, 2017; Kott et al., 2016). These "things" will function with varying levels of autonomy in order to adapt to a broad range of mission goals and environments. The sheer number of these things will far exceed the number of humans available to oversee their operation. As a result, the things within the IoBT will require the support of autonomous agents, particularly for the purposes of their cyber defense.

The battlefield is a highly dynamic and uncertain environment, often dominated by adverse conditions degrading the effectiveness of communications and information networks that are the Warfighters' critical tools. Battlefield networks are necessarily mobile and are composed of many heterogeneous devices. Mobility and adversary actions result in topologies that change quickly and frequently. This requires reestablishing connections as configurations change. The lack of infrastructure in battlefield environments severely constrains the amount of bandwidth and computational capabilities available to Warfighters. Reliable, timely communication of accurate information is critical to the successful execution of every mission, but the resource constraints inherent to tactical networks threaten the delivery and assurance of vital information.

As an example, consider a robotic vehicle, an entity within the IoBT that gathers information for dismounted Warfighters. This unmanned vehicle collects and transmits images of buildings and roads as well as meteorological and geographic data. If a specific threat is detected by the robotic vehicle, the transmission of such information to the Warfighters takes priority over all other communication. The adversary may attempt to deploy a malware on the robotic vehicle in order to deny or degrade the vehicle's high-priority communications. An autonomous cyber agent—a software agent—resides on the robotic vehicle, senses the status of its environment, and chooses the best course of action to protect it. In this hypothetical example, the agent recognizes a software module's attempt to connect to a suspicious web site, one that is associated with downloaded malware. Blocking traffic between the robotic vehicle and that site is a course of action the agent chooses as a means to defend the vehicle.

> *"The adversary may attempt to deploy a malware on the robotic vehicle in order to deny or degrade the vehicle's high-priority communications."*

Here, a few words about the terminology are in order. Autonomous agents can be a physical agent or a software agent. Autonomous agents can be physical entities, such as a robot or drone functioning independently and without direct human guidance; and they can be software entities, such as a cyber-defense agent. The focus of this article is software agents, specifically cyber defense agents. In the preceding example, the robotic vehicle is a physical autonomous agent. In addition, one or multiple software agents, such as a cyber defense agent, reside on a physical agent. An autonomous software agent can be defined as software which acts on its own without human intervention (Shiffman, 2012) or as a "self-activating, self-sufficient, and persistent computation" (Shattuck, 2015). For the purposes of this paper, the term "agent" is used to refer to software agents, not physical agents. Furthermore, we use the term "agent" to refer to a specific type of a software agent—an agent that specializes in cyber defense of things within the IoBT.

The remainder of the paper is organized as follows. In the next section, we elaborate on the argument that the enormous size of the future IoBT, along with the sophisticated cyber threats of the future battlefield, will necessitate the wide use of autonomous intelligent cyber defense agents. We then present a notional, illustrative scenario—a few moments in the life of a cyber defense agent. This is followed by a discussion of challenges and research topics that arise from considering the scenario.

AUTONOMOUS CYBER DEFENSE AGENTS ARE A NECESSITY

Technologies such as "machine intelligence and networked communications" have spurred the growth and acceptance of connecting cell phones and other personal devices with everything from household appliances to automobiles. Similarly, the military is applying computational intelligence to its interconnected battlefield devices to make them smarter and thereby, more useful to Soldiers. The IoBT encompasses various sensors, vehicles, communication devices, computers, and information sources. It is certain that the future battlefield will be densely populated with a variety of interconnected devices (Kott et al., 2016).

To gain an appreciation for the size and complexity of the IoBT, consider the following:

It has been estimated (Anonymous, 2017) that the commercial or consumer

Internet of Things (IoT) grew from 2 billion devices in 2006 to 15 billion devices in 2015. The same source estimates that by the year 2020, 200 billion devices will populate the IoT. If the proliferation of IoBT (military devices) follows the growth of IoT (consumer devices), it is clear that human Warfighters, will require an augmentation by autonomous cyber defense agents to monitor and defend battlefield devices.

Additional complexity becomes apparent when one considers that the IoBT will have to operate effectively within environments that it neither owns nor controls. For example, a military force may be operating within a city where the majority of computing devices—the consumer devices—belong to the neutral civilians but are also potentially controlled by the adversary. Furthermore, in the case of IoBT the adversary will actively pursue compromise, capture, or destruction of the battlefield devices.

Given that cyber-attacks will occur frequently and at a high pace that will surpass human ability to respond in a timely fashion, decisions on the most appropriate course of cyber defense actions will have to occur in near real-time.

AN ILLUSTRATIVE OPERATING SCENARIO

In order to illustrate how an autonomous cyber defense agent might operate, we offer a notional operating scenario. In this scenario, Blue refers to friendly forces and Red refers to the adversary. Blue-17, Blue-19, and Blue-23 are peer cyber defense agents. Each agent is installed by a human operator on its respective device within the Blue IoBT (e.g., an Android phone) and is tasked with cyber defense of that device. Blue-C2 is the Command and Control (C2) node that commands, coordinates, and supports all other Blue agents, at least when communications between an agent and the Blue-C2 node are available. There is only one Red agent—Red-35—in our simple scenario.

The protagonist of our scenario is Blue-17, a cyber defense agent that has been installed on a friendly device; it continuously monitors Blue space network and scans event logs looking for suspicious activity. The antagonist is Red-35, a malware agent successfully deployed by the Red forces on the device defended by Blue-17. The events unfold, briefly, as follows.

Blue-17 detects a hostile activity associated with Red-35 and attempts to contact the Blue-C2 for additional remediation instructions. Unfortunately, the communications are heavily contested by the adversary, and response from Blue-C2 is not coming. Therefore, Blue-17 decides to contact peer agents (Blue-19 and Blue-23) in search for relevant information. Although Blue-19 and Blue-23 receive this message from Blue-17, their responses are not arriving to Blue-17. Having heard nothing within a reasonable waiting time, Blue-17 independently formulates and executes a set of actions to defeat Red-35. However, having completed these actions, Blue-17 receives a belated reply from Blue-23. Blue-17 determines that Blue-23 is compromised because the response is suspicious. Given the extreme seriousness of this situation, Blue-17 neutralizes Blue-23 and places a copy of itself on the device that was being protected by Blue-23.

Table 1 provides a hypothetical timeline of these events and the agents' actions. Durations are intended to merely illustrate the flow of time in the scenario and are in no way representative of execution speeds of any actual hardware or software. Following the table, we discuss each step of the scenario in more detail.

Table 1: Hypothetical Timeline of Agents' Actions.

Step	Elapsed Time	Condition/Event	Active Software Agent	Action
1	H = 0 sec	Start up	Blue-17	Monitor network traffic and scan logs
2	H = H + 0.100 sec	Hostile software agent compromises device and network	Red-35	Red-35 infiltrates Blue device and network. Blue-17 does not notice the infiltration.
3	H = H + 0.200 sec	Red-35 begins operations. Suspicious activity detected	Red-35 and Blue-17	Red-35 conducts malicious activities. Blue-17 detects an activity and predicts probable compromise.
4	H = H + 0.22 sec	Compromise suspected	Blue-17	Contacts C2 node
5	H = H + 3.00 sec	No response from C2 node	Blue-17	Contact Blue-19 and Blue-23 agents
6	H = H + 5.00 sec	Message among Blue peer agents	Blue-19 and Blue-23	Receive message from Blue-17
7	H = H + 10.00 sec	Message acknowledgement time out	Blue-17	Choose alternate course of action
8	H = H + 12.00 sec	No communication with peer defensive agents	Blue-17 and Red-35	Block or redirect Red-35 communication. Red-35 is unable to defend itself.
9	H = H + 23.00 sec	Response received from Blue-23	Blue-17	Blue-17 determines that the response is invalid
10	H = H + 28.00 sec	Neutralize compromised Blue agent	Blue-17	Quarantine or destroy Blue-23 software code
11	H = H + 28.3 sec	Replicate and overwrite	Blue-17	Copy to device

Scenario Steps 1–2

In the scenario, Blue-17 passively monitors the inbound and outbound network communications using a lightweight intrusion detection system (IDS) such as FAST-D (Yu and Leslie, 2018). FAST-D is a software that performs intrusion detection using far less computational resources than alternative solutions. Its algorithm uses hash kernels and byte patterns as signatures to examine the packet payload content of all network communications. Additionally, Blue-17 scans the device logs looking for indicators of compromise (privilege escalation, abnormal crashes, failed logins, etc.).

> *"Blue-17 scans the device logs looking for indicators of compromise (privilege escalation, abnormal crashes, failed logins, etc.)."*

Scenario Steps 3–4

Blue-17 sends a message to its C2 node for further remediation instructions and verifications. A C2 node is one that is central (root) and is responsible for the management and tracking of all Blue agents. A C2 node resides in a central location that may be the tactical operations center. The message sent to the C2 node is encrypted to protect the confidentiality and integrity and is in a predefined format for agent messages. This message is split up into many small segments, is blended into normal traffic to masquerade as other legitimate traffic, and sent through different routes within the network in order to avoid an attacker from intercepting or detecting the agent message sent to the C2 node. Lastly, the address of the C2 node changes over time based on a deterministic algorithm, known to all agents to make it more difficult for Red-35 to discover its location.

Scenario Step 5

After some reasonable waiting time passes, and Blue-17 does not receive a reply back from the C2 node, it decides as an alternative action to send out a request to its peer agents (Blue-19 and Blue-23) for their remediation recommendations. Again this message is sent out using an encrypted predefined format for agent messages as previously described in sending a message to the C2 node. The message is sent directly to the peers and is blended into other network traffic. The peer agents are neighbors to Blue-17 and are also be under the management of the C2 node.

Scenario Step 6

Both Blue-19 and Blue-23 have received the message from Blue-17. After some delay, Blue-23 sends a response and recommendation back to Blue-17 using the same method for sending a message to a peer agent.

Scenario Steps 7–8

Within a specified time interval, Blue-17 has not received a response from either its C2 node or its peers (Blue-19 and Blue-23). Blue-17 requested further verification of the threat before taking a destructive action against Red-35. However, since a response was not received, Blue-17 decides to take action on the perceived Red-35 malware agent threat. The Blue-17 agent first isolates the Red-35 malware agent and its communication in a honeypot to observe the actions taken by the attacker. Blue-17 has taken this action since it is not confident in its assessment of the detection of the perceived Red-35 agent.

Scenario Step 9

After some time has passed, and Blue-17 has already taken action, a response from Blue-23 is received. Blue-23's response contains a signature and timestamp that allows Blue-17 to determine the authenticity of the message received. However, as Blue-17 verifies the response message from Blue-23, it determines that the message signature is not valid and rejects the message. Blue-17 concludes that Blue-23 may be compromised.

Scenario Steps 10–11

Blue-17 has discovered that Blue-23 has been compromised. Blue-17 takes action to quarantine Blue-23. Blue-17 clones itself to create a pristine copy of the defensive agent. Blue-17 initiates the overwriting of the Blue-23 agent image with a fresh copy of a defensive agent with the initial state of Blue-17. The agent package is sent via an encrypted message from Blue-17 to the container management of Blue-23. The container management package of the agent uses cryptographic authentication, allowing the overwriting to occur. Blue-23 is restored back to a fresh agent image and is no longer infected.

DISCUSSION OF CHALLENGES AND REQUIREMENTS

Having offered a scenario—simple yet sufficiently illustrative of potential difficulties—we now have a basis for discussing the technical challenges and requirements. One of the requirements illustrated in part by the scenario is that a defensive agent must reside outside of the operating system of the device it is protecting. This arrangement avoids the possibility of the malware providing false information or changing the view of the defensive agent (i.e., Blue-17). Malware can disable processes or deceive (e.g., by providing false information) software such as the anti-virus (AV) software or firewall on a device (Baliga et al., 2007). A logical separation at the hardware level between the operating system being protected and the defensive agent will protect the Blue-17 agent from being compromised by malware infection. The defensive agent will require access in a secure manner to all of the files and state from its outside view, while being protected from any threats affecting the Blue-17 operation or integrity.

Additionally, because the Blue-17 agent resides outside of the protected operating system, Red-35 will not be able to detect Blue-17's presence or any of its actions. A traditional placement alternative for an agent that resides outside of the protected operating system, would be a distributed or network-based sensor. That configuration comes with a tradeoff: an agent (Blue-17) at the network level would not be able to monitor the file system of the protected operating system. Therefore, the Blue-17 agent must reside on the same physical device outside of the operating system being protected.

Also, in order for the agent to move around freely among the devices within the protected network, the agent must be unconstrained by any particular operating system. It is also presumed that the container in which the agent runs has been pre-installed on the device to which agents can migrate freely to, such as in the case with Blue-17 overwriting Blue-23.

Clearly required, as illustrated in our scenario, is a fast, highly reliable and low-resource means of detecting potentially malicious activity. For example, using a low-resource intrusion detection software, Blue-17 was able to detect rapidly and with a significant degree of assurance a suspicious activity performed by a sophisticated agent Red-35. Additional solutions could be employed that use supervised machine-learning approaches, coupled with features such as network packet inter-arrival times, packet sizes, Transport Control Protocol (TCP) flags, and such, to perform detection of malware infiltration. However, in either case it is important to understand the limitations (i.e., inability to detect malware within encrypted communications) of the intrusion detection algorithm chosen to perform detection of malicious communications. It is also important to know the possible ways an attacker could evade (fragmentation attack, encrypted attack, etc.) the IDS. Successful evasion by an attacker will result in a missed attack, also called a false negative. It is also critical for an autonomous agent employing an IDS algorithm to have a low false-positive rate (misclassified legitimate traffic as an attack) and low false-negative rate (missed attack). In a military context a false-positive in an autonomous cyber defense agent will impact the mission by denying legitimate and essential communication.

Another challenging requirement is the need to manage the degree of the agent's autonomy. Blue-17 could be fully autonomous or semiautonomous. In our scenario, Blue-17 is fully autonomous, as defined by the lack of human intervention at any point. Consequently, Blue-17 must be highly confident in the detection event and its resultant course of action. The agent's actions must avoid any adverse reaction, such as degrading network performance or dropping nodes on the network as a mitigation, resulting in access denials. Alternatively, Blue-17 could act as a semiautonomous agent, with varying levels of interaction between the agent and human controllers, which present many challenges of their own (Kott and Alberts, 2017). For example, Blue-17 could detect a potential compromise and then defer to a human analyst (e.g., by contacting the C2 node and waiting for instruction) in a case where there is low to moderate confidence in the detection event.

> "It is also critical for an autonomous agent employing an IDS algorithm to have a low false-positive rate (misclassified legitimate traffic as an attack) and low false-negative rate (missed attack)."

The agent will require the ability to share threat data directly with its peers (e.g., Blue-17 had to share data with Blue-23 and Blue-19) and orchestrate coordinated defensive actions when necessary. Additionally, the agent must also be able to work in an isolated environment and make appropriate decisions independently, as Blue-17 had to do when it failed to receive response from either Blue-C2 or Peer's agents. These agents will need to store pertinent information on detected attacks and outcomes (successful vs. unsuccessful) of the selected mitigation strategies. This information will need to be stored in a compressed format due to the limited resources characteristic of the various devices of IoBT. On the other hand, when the agents return to a less-contested environment where power and bandwidth are less constrained and more reliable, the data would be uploaded to a central repository. Lessons learned (quantitative measures of outcomes) and specifics on detected attacks would be compiled to improve the process of informing other autonomous agents. This arrangement would expand and enrich the agents' knowledge and ability to learn from historical decision-making strategies.

The agent (i.e., Blue-17, Blue-19, or Blue-23) hosted within the IoBT environment must process and synthesize the information it produces or receives from other agents to a subset relevant to the Warfighters' cognitive needs (Kott et al., 2016). For example, of all the alerts produced by the agents, the Warfighter will only need to be aware of a small subset to form a situational awareness of on-going cyber-attacks. This filtered information must be relevant and trustworthy to both the IoBT device and the Warfighter's cognitive needs. Providing incorrect or irrelevant information could cause significant and negative impact to the mission (Kott et al., 2016). Further, information stored by agents on IoBT devices must be distributed and obscured from the adversary. An approach to secure the distributed agent information within an IoBT environment is to split the data into fragments and disperse them among the many devices in a way to thwart the adversary's ability to reconstruct the information based on a number of captured segments (Kott et al., 2016a).

Ideally, the agents' performance would be evaluated in order to refine and share successful strategies with other agents. Performance in this context includes the agents' decision-making value, timing, and the resulting impacts of the courses of action executed (e.g., Blue-17 was successful—what factors contributed to these successes?). This supports the need for agents to be able to learn from their actions as well as the actions of other agents via machine-learning techniques.

> *"personnel who interact with the IoBT devices will not be cyber security experts and will be focused on the execution of the mission without the ability to continuously monitor the health of their devices"*

The agent could employ a combination of supervised and unsupervised machine learning. The lessons learned and outcomes of the course of action taken by an agent could be used with a reinforcement-based machine-learning algorithm. For example, the successful course of action executed by Blue-17 with respect to defeating Red-35 would receive a positive reward. This approach could be used to expand the knowledge of the autonomous agents, thereby improving the agents' performance and effectiveness.

Another requirement of these agents will be trust management between devices. Each device on the network will require software-based logic to participate in the network with a full degree of trust and access. This logic can be preinstalled or can be acquired from a peer node by a device that seeks to join the network in a comply-to-connect mode of operation. Once compliance conditions are met, the agent can be transferred to other network member nodes. For example, in our scenario, Blue-17 needed a way to determine that Blue-23 is no longer trustworthy. At the same time, Blue-17 had to elicit a sufficient degree of trust from the node where Blue-23 resided in order to overwrite the Blue-23 image.

Device-to-device transfer of the agents—such as the move of a copy of Blue-17 to the node originally defended by Blue-23—necessarily raises concern for unintended propagation and behaviors beyond the intended network, as witnessed with the Morris worm (Qing and Wen 2005; Spafford 1989) and the more recent Stuxnet attack (Farwell and Rohozinski, 2011). Findings from studies on limiting the spread of malware in mobile networks (Zyba et al., 2009; Li et al., 2014) could be adapted to manage the propagation of defensive agents. Another potential solution to controlling propagation is to require consensus approval of a certain number of nodes prior to enabling transfer of the agent to a new device. A suggested approach is to define boundary rules to determine whether the agent has been transferred outside its intended network. When the boundary rules evaluate to a true condition (out of bounds), mandatory removal of the agent or a self-destruct sequence would be triggered. The effects of these combined approaches to controlling propagation require additional research.

While autonomous agents should be free to learn, act, and propagate, careful thought should be given to methods that would constrain behaviors within the bounds of legal and ethical policies, as well as the chain of command. For example, it would be undesirable if Blue-17 were to learn that requests to Blue-C2 are generally fruitless and should not be attempted. An agent that is fully autonomous must be able to operate within an appropriate military C2 construct (Kott and Alberts, 2017). It is imperative that a software agent be bounded in its propagation, yet capable to move around freely between authorized devices.

CONCLUSIONS

With a large number of devices within the future IoBT it will be imperative for these devices to be able to defend themselves. Further, personnel who interact with the IoBT devices will not be cyber security experts and will be focused on the execution of the mission without the ability to continuously monitor the health of their devices. Autonomous cyber defense agents will be required to augment and multiply military forces. Such agents will need to possess awareness of, and ability to learn about, threats in near real-time. The agents will need to have the capability to reliably and predictably self-propagate, sense malicious activity, and disseminate information among trusted network members.

REFERENCES

[1] Anonymous. A Guide to the Internet of Things. https://www.intel.com/content/www/us/en/internet-of-things/infographics/guide-to-iot.html accessed 09-25-2017.

[2] Baliga, A., Kamat, P., & Iftode, L. (2007, May). Lurking in the shadows: Identifying systemic threats to kernel data. In *Security and Privacy, 2007. SP'07. IEEE Symposium on* (pp. 246-251). IEEE.

[3] Farwell, J. P., & Rohozinski, R. (2011). Stuxnet and the future of cyber war. *Survival*, *53*(1), 23-40.

[4] Kott, A., Swami, A., & West, B. (2016). The Internet of Battle Things. Computer. 49. 70-75. 10.1109/MC.2016.355.

[5] Kott, A., Swami, A., & West, B. J. (2016a). The Fog of War in Cyberspace. Computer, 49(11), 84-87.

[6] Kott, A., & Alberts, D. S. (2017). How Do You Command an Army of Intelligent Things?. Computer, 50(12), 96-100.

[7] Li, Y., Hui, P., Jin, D., Su, L., & Zeng, L. (2014). Optimal distributed malware defense in mobile networks with heterogeneous devices. *IEEE Transactions on mobile computing*, *13*(2), 377-391.

[8] Qing, S., & Wen, W. (2005). A survey and trends on Internet worms. *Computers & Security*, *24*(4), 334-346.

[9] Shattuck, L. G. (2015). Transitioning to Autonomy: A Human Systems Integration Perspective. *Transitioning to Autonomy: Changes in the Role of Humans in Air Transportation*.

[10] Shiffman D. (2012). The Nature of Code: Simulating Natural Systems with Processing, December 2012.

[11] Spafford, E. H. (1989). The Internet worm program: An analysis. *ACM SIGCOMM Computer Communication Review*, *19*(1), 17-57.

[12] US Army Research Laboratory. Internet of Battlefield Things Collaborative Research Alliance program announcement (2017). https://www.arl.army.mil/www/default.cfm?page=3050 accessed 09-25-2017.

[13] Yu, K. & Leslie, N. O. (2018). FAST-D: Malware and intrusion detection for mobile ad hoc networks (MANETs), Journal of Defence and Security Technologies - Special issue on Predictive Analytics and Analysis in the Cyber Domain.

[14] Zyba, G., Voelker, G. M., Liljenstam, M., Méhes, A., & Johansson, P. (2009). Defending mobile phones from proximity malware. In *INFOCOM 2009, IEEE* (pp. 1503-1511). IEEE.

ABOUT THE AUTHORS

MR. MICHAEL J. DE LUCIA is a computer scientist at the US Army Research Laboratory's (ARL) Network Security Branch. He currently is a Ph.D. candidate in the Electrical and Computer Engineering Department at the University of Delaware and has earned an M.S. in Computer Science from New Jersey Institute of Technology (NJIT), (2006), & B.S. in Information Technology with specialization in Network Security from NJIT, (2005). Mr. De Lucia has over 10 years of experience in Cyber Security and worked at the US Army Communications Electronics Research, Development, and Engineering Center (CERDEC) (2007-2016) and the US Army Communications Electronics Command (CECOM) Software Engineering Center (2006-2007) before coming to ARL.

DR. ALLISON NEWCOMB is a computer scientist at the Army Research Laboratory's Network Security Branch. She has earned a D.Sc. in Information Technology from Towson University (2016), MS in Computer Science from Johns Hopkins University, (2006), & BS in Computer Science from Mississippi College (1992). Dr. Newcomb is a member of ACM and IEEE, and holds the Certified Information Systems Security Professional (CISSP) and Certification and Accreditation Professional (CAP) certifications.

DR. ALEXANDER KOTT earned his PhD in Mechanical Engineering from the University of Pittsburgh, Pittsburgh, PA, in 1989, where he researched AI approaches to invention of complex systems. He serves as the US Army Research Laboratory's Chief Scientist in Adelphi, MD. In this role he provides leadership in development of ARL technical strategy, maintaining technical quality of ARL research, and representing ARL to external technical community. Between 2009 and 2016, he was the Chief, Network Science Division, Computational and Information Sciences Directorate, ARL, responsible for fundamental research and applied development in network science and science for cyber defense. In 2003-2008, he served as a Defense Advanced Research Programs Agency (DARPA) Program Manager. His earlier positions included Director of R&D at Carnegie Group, Pittsburgh, PA. There, his work focused on novel information technology approaches, such as Artificial Intelligence, to complex problems in engineering design, and planning and control in manufacturing, telecommunications and aviation industries. Dr. Kott received the Secretary of Defense Exceptional Public Service Award, in October 2008. He published over 80 technical papers and served as the co-author and primary editor of over ten books.

Call for Papers for Publication

CSIAC is chartered to ***leverage the best practices and expertise from government, industry and academia*** in order ***to promote technology domain awareness and solve the most critically challenging scientific and technical (S&T) problems.***

As part of that mission, CSIAC publishes the *Journal of Cyber Security and Information Systems,* focusing on scientific and technical research & development, methods and processes, policies and standards, security, reliability, quality, and lessons learned case histories. Contributing authors realize the benefits of being published in a highly-respected resource within the technical community, with an enormous reach across the Department of Defense and the broader scientific community (23k+ subscribers). Demonstrate your expertise and accomplishments or pose the challenging questions for further thought in the resource that reaches fellow Subject Matter Experts (SMEs) developing the solutions to support the warfighter.

To Submit an Article

Visit: **https://www.csiac.org/csiac-journal-article-submission-policy/**

CSIAC is currently accepting articles submitted by the professional community for consideration in the following topic areas:

- Advances in AI and machine learning, deep learning, cognitive computing, intelligent agent, chatbot
- AI applications in industry, business, healthcare, and education and training
- Biometric Identity Management
- Trust, resilience, privacy and security issues in AI applications
- Testing and validation of AI and ML applications
- Security automation techniques/real-time incident response
- IoT, Smart Cities, Connected & Autonomous Vehicles (CAV)
- Data-Centric Security
- Data Loss Prevention (DLP)
- Endpoint Security Risk
- DevOps/DevSecOps

(A non-exhaustive set of topics)

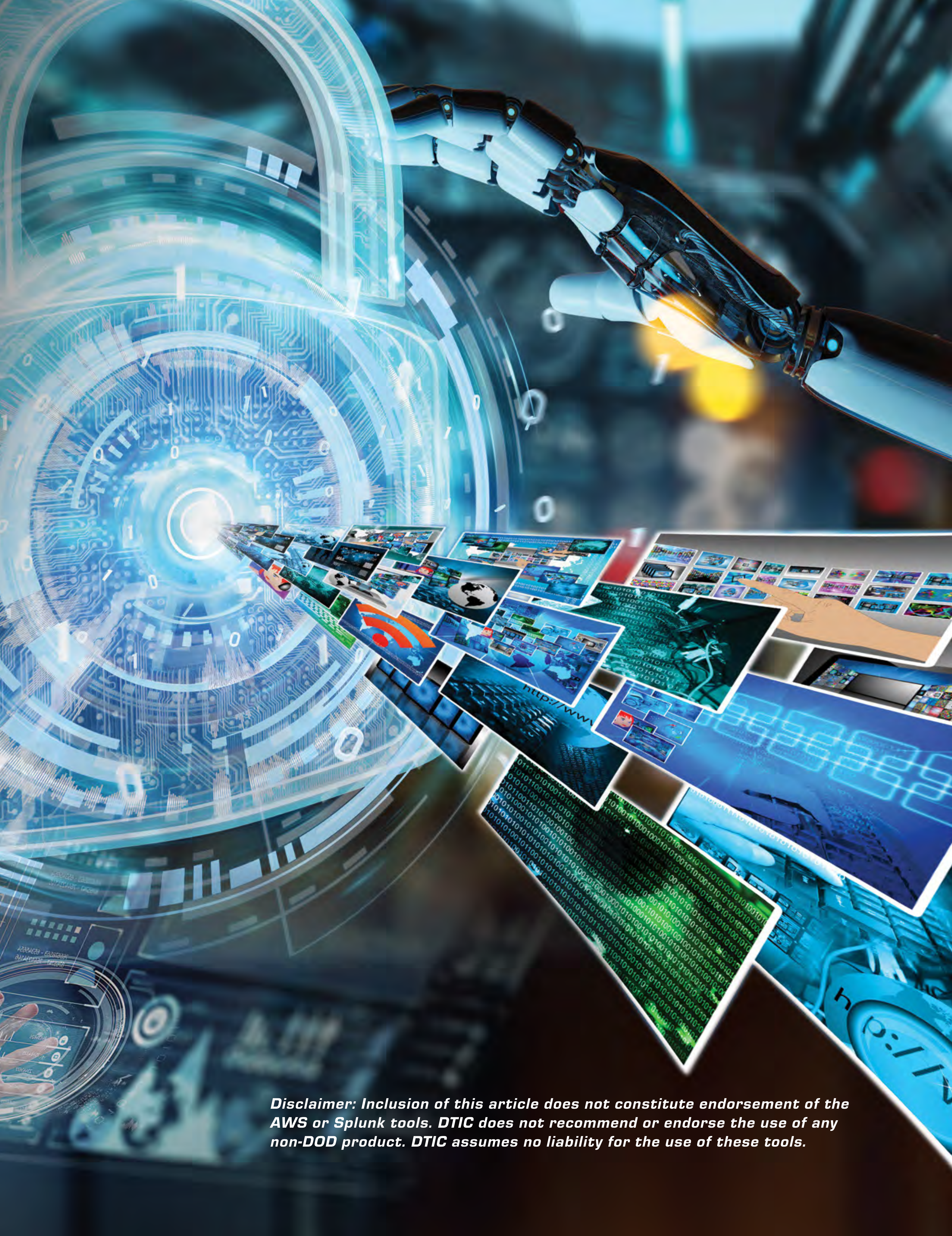
CLOUD SECURITY MONITORING

with AI-ML Infused Technologies

By: Balaji Balakrishnan, Graduate, SANS Technology Institute, MS in Information Security Management

THIS ARTICLE DISCUSSES HOW TO APPLY SECURITY LOG MONITORING CAPABILITIES FOR AMAZON WEB SERVICES (AWS) INFRASTRUCTURE AS A SERVICE (IAAS) CLOUD ENVIRONMENTS.

It will provide an overview of AWS CloudTrail and CloudWatch Logs, which can be stored and mined for suspicious events. Security teams implementing AWS solutions will benefit from applying security monitoring techniques to prevent unauthorized access and data loss. Splunk will be used to ingest all AWS CloudTrail and CloudWatch Logs. Machine learning models are used to identify the suspicious activities in the AWS cloud infrastructure. The audience for this article is the security teams trying to implement AWS security monitoring.



Disclaimer: Inclusion of this article does not constitute endorsement of the AWS or Splunk tools. DTIC does not recommend or endorse the use of any non-DOD product. DTIC assumes no liability for the use of these tools.

1 INTRODUCTION

Organizations are starting to use cloud computing to take advantage of the many benefits it provides such as cost savings, quick time-to-market and on-demand scaling of the environment. As organizations start to use cloud computing, security professionals must update their operations to align with cloud computing models. The References section in this article provides many recommendations on cloud security controls from NIST, cloud deployment models, cloud security references from Cloud Security Alliance, ENISA, and NIST.

In the most recent edition of the Cloud Computing Top Threats in 2016, the report (CIS, 2016) identified 12 critical issues to cloud security. Effective security monitoring mitigates some of the following risks:

- Weak Identity, Credential, and Access Management
- Insecure APIs
- Account Hijacking
- Malicious Insiders
- Advanced Persistent Threats (APTs)
- Data Loss
- Abuse and Nefarious Use of Cloud Services

> "cloud solutions use the DevOps methodology for continuous deployment."

Securing Cloud Services involves conducting a detailed risk assessment and architecting a secure solution to meet the business requirements. Security Monitoring plays a vital role in securing Cloud Services. This article highlights how to implement security monitoring solution for Amazon Web Services (AWS) environments.

1.1 Cloud Security Monitoring Challenges

The primary types of cloud computing solutions are Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS). Amazon Web Services (AWS) has established itself as a leading cloud services provider, with Microsoft Azure and Google Cloud in the distant second and third positions.

AWS innovates at a rapid pace introducing many new services every day. Last year alone, on average AWS customers got access to three new services every day. The different AWS services available and the best practices for securing AWS environments are in the reference section. Some of the best practices include encryption, privileged access management, segregation of resources and monitoring.

This article focuses on implementing security monitoring for AWS workloads. The next subsections highlight the key areas of security monitoring when deploying AWS workloads in addition to traditional data center monitoring.

AWS Management Console Monitoring

Management of AWS instances and resources are performed using the AWS Management Console. Some of the main activities that can be conducted using the AWS Management Console are creating new virtual machines and removing any existing virtual machines and other AWS services. Monitoring the unauthorized access to AWS Management Console is critical since gaining this convenient access to the cloud management plane is like having keys to the cloud kingdom.

Application Programming Interface (API) Access Monitoring

As organizations move towards cloud solutions, they have to adapt to the new DevOps architecture. Realizing the benefits of the cloud platform would be difficult if teams shift and move the current applications as-is to the cloud. The existing application infrastructure has to be rearchitected to suit the cloud deployment models. Ideally, cloud solutions use the DevOps methodology for continuous deployment. This method enables the business to reduce development time and turn around quick solutions. As an example, some AWS environments use AWS CodePipeline for continuous application deployment using DevOps strategies in AWS environments.

DevOps introduces new challenges for security monitoring. The number of API calls is increasing due to automation related to AWS CodePipeline, infrastructure as code and serverless computing. It is critical to monitor these API calls to ensure there is no unauthorized access. It's hard to follow these events using traditional rule and threshold-based monitoring due to the high volume of activities. Machine learning techniques are well suited for monitoring this vast amount of activity by learning different features/characteristics from the data.

AWS Serverless Computing Monitoring

Recently AWS introduced "serverless" computing; serverless computing depends on AWS Lambda to run the application code. In serverless computing, there is no server infrastructure; the focus is on monitoring the AWS Lambda function executions, invocations and other parameters related to the AWS Lambda functions.

AWS Identity and Access Management (IAM) Monitoring

AWS IAM enables organizations to control access to AWS services and specific resources. AWS IAM provides options to configure granular permissions in AWS environments. It is recommended to give the least amount of permissions to manage AWS resources required for performing the job function. As an effective information security control, security teams should use many tools provided by AWS like

Access Advisor. Providing appropriate access prevents any unauthorized access and enables effective monitoring of AWS resources administrative access. Monitoring the different administrative credentials used in the AWS environments is a requirement enforced by various compliance regulations. Machine learning is ideal for controlling the various AWS credentials since it learns from the previous events and understands what is normal to identify anomalies. Financial regulatory requirements like Sarbanes-Oxley mandate organizations to review all privileged access and changes to the AWS environments hosting financial data as part of security compliance monitoring.

1.2 Overall architecture of the proposed solution

The proposed solution for cloud security monitoring is to use a big data analytics solutions such as Splunk, Apache Spark or Amazon Elasticsearch to load all the AWS cloud infrastructure logs. Machine learning models should be used to develop risk scores to identify the most suspicious events. Then, based on the incident, the security team should take further action using automation(lambda functions) (or) email to alert the security team for manual analysis.

It is a challenge to manually baseline and configure AWS infrastructure security monitoring rules due to the changes in AWS environments. Machine learning techniques like Supervised Learning algorithms explained later in this article can handle the security monitoring challenges of cloud security monitoring by automatically learning from data to understanding anomalies and high-risk events. Machine learning models can be used to build baselines and develop risk scores to identify suspicious events using identity authentication information, location information and activity type.

In this article, Splunk will be used to ingest all AWS CloudTrail and CloudWatch Logs for implementing the AWS security monitoring use cases. Machine learning models are applied to identify the suspicious activities in the AWS cloud infrastructure. The latest version of Splunk 6.5 has a built-in machine learning toolkit which supports various machine learning algorithms. Machine learning models will be applied using the Splunk Machine Learning toolkit. These steps involved in using machine learning algorithms are as follows:

a. Visualize and combine data cleansing with smart feature engineering,
b. Choose right metric/method for estimating model performance
c. Tune the parameters.

Summary of the key concepts proposed are:

1. Collect all of the AWS log data from Cloudtrail and CloudWatch to Splunk
2. Apply machine learning models to build baselines and develop the risk scores instead of manual rules/thresholds.

Some factors which make this implementation feasible are:

a. Advancement in big data technologies which enables information security teams to store all types of data at scale.
b. Many machine learning solutions are becoming available like Microsoft Azure ML Studio, Amazon Machine Learning, Databricks Spark, Splunk Machine Learning toolkit.

By having a centralized open source big data analytics solution, security teams can apply machine learning and other statistical techniques to any data set. The major advantage of this solution is that once a successful method is identified using machine learning, similar challenges can be solved using the same approach. For example, if a technique is helpful in identifying suspicious access attempts from AWS cloud-based infrastructure identity and access authentication data, the same method can be applied to identify suspicious access attempts for other applications and cloud infrastructures like Microsoft Azure and Google Cloud. The next section describes machine learning techniques and the two use cases are implemented using Splunk.

> *"Machine learning techniques like Supervised Learning algorithms can handle the security monitoring challenges of cloud security monitoring by automatically learning from data to understanding anomalies and high-risk events"*

1.3 Risk Scoring Methodology

Risk scoring is not a new concept; it has always been in use in the information security community to prioritize the most critical vulnerabilities and issues to resolve. In traditional data center monitoring, the risk scoring methodology relies on understanding the corporate environment to identify suspicious events. An example of this type of process is creating an unauthorized access alert to critical server asset events based on an understanding of authorized administrators who have access. Detecting malicious events based on the known bad patterns and assigning risk scores to known bad patterns is useful for threats which are already seen and known to the information security community. The References section has some examples for developing risk scores manually using static rules and thresholds in AWS environments.

The challenge with these types of standard risk scoring based monitoring is keeping up with the rapid pace of new API calls and permissions that are being rolled out by AWS. Some of the criteria that are relevant to cloud security monitoring are the identity, data access, the action performed, and geo-location. By leveraging these criteria (features) in combination with historical data, machine learning techniques can learn the environment and identify anomalies for further investigation. Machine learning models can provide risk scores based on the learning from previous data. In this article, a Linear regression algorithm is used as an example to develop a machine learning model which predicts risk scores. Linear regression algorithms will predict numeric values. The Linear regression algorithm models the relationship between continuous output variable with the features (input, explanatory variables) using the linear function. The following section on machine learning explains the algorithm in detail. The model learns from the data; this is efficient for this AWS use case compared to manually updating the rules/thresholds for the risk scores.

1.4 Machine learning

Machine learning has two major types: Supervised and Unsupervised Learning. In Supervised Learning, the machine learning algorithm will learn from the data and labels (classification) provided. The resultant model will try to predict the label (classification) given a set of features (Astroml, 2015). Some Classification Algorithms commonly used are Neural Networks, Random Forests, Support Vector Machines (SVM), Decision Trees, Logistic Regression, and Naive Bayes. An example of Supervised Learning is providing a set of dog and cat pictures to machine learning algorithm with labels indicating if the picture is cat or dog. The Supervised Learning algorithm will learn from the dog and cat pictures and create a predictive model. Applying new pictures to the predictive model will predict if the provided picture is a dog or a cat as seen in Figure 1:

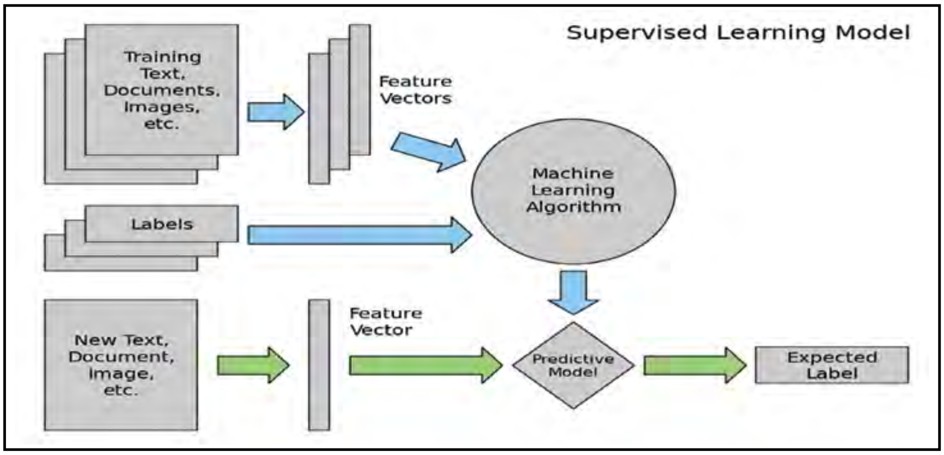

Supervised Learning Model
Training Text, Documents, Images, etc.
Feature Vectors
Machine Learning Algorithm
Labels
New Text, Document, Image, etc.
Feature Vector
Predictive Model
Expected Label

Figure 1- Supervised Learning Model (AstroML, 2015)

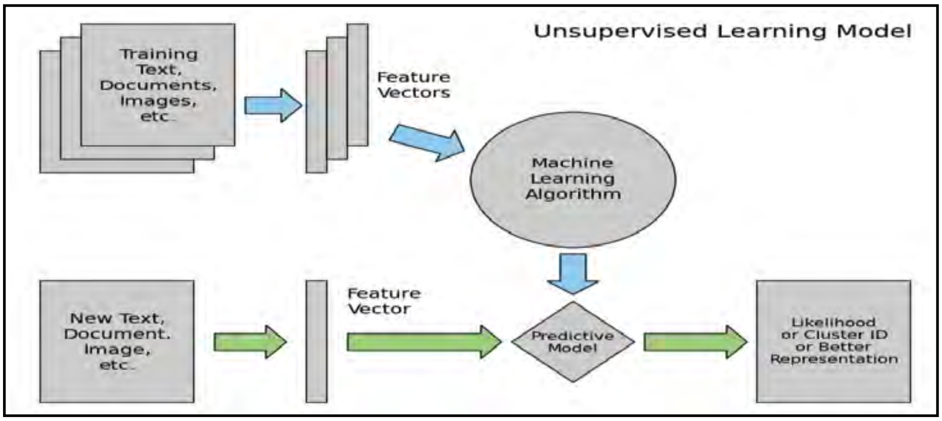

Unsupervised Learning Model
Training Text, Documents, Images, etc.
Feature Vectors
Machine Learning Algorithm
New Text, Document, Image, etc.
Feature Vector
Predictive Model
Likelihood or Cluster ID or Better Representation

Figure 2 - Unsupervised Learning Model

In this article, a Supervised Learning technique using a linear regression algorithm will be used to predict risk scores for AWS cloud infrastructure events.

In Unsupervised Learning, the model tries to understand the data based on the features without any labels and the tasks are to identify patterns and anomalies from data. Unsupervised Learning comprises tasks such as dimensionality reduction, clustering, and density estimation (AstroML, 2015). An example of Unsupervised Learning is providing a set of dog and cat pictures to the machine learning algorithm; it will cluster the cat and dog pictures as separate groups as depicted in Figure 2.

Unsupervised Learning algorithms will be useful to identify the principal features in the dataset. It is also very helpful to provide different vantage points based on various features. In the example of dog and cat pictures, using Unsupervised Learning techniques will be useful to understand how the several facial features will be the most helpful to classify by segregating data based on those facial features. In our use case of AWS cloud infrastructure events, separating the data based on the location of logins can provide insight on whether it is an important feature. Some Common Unsupervised algorithms are K-Means Clustering, Hierarchical clustering, and Hidden Markov models. Figure 3 highlights the different algorithms in the Splunk Machine Learning toolkit (Splunk, 2016):

Machine learning should only be applied to use cases that are applicable and produce results. Machine learning is very data hungry and ingesting a lot of data for creating machine learning models will produce definitive results. Since the machine learning algorithms require a lot of data to provide useful models, significant patience is needed to obtain results. There are a lot of logs generated

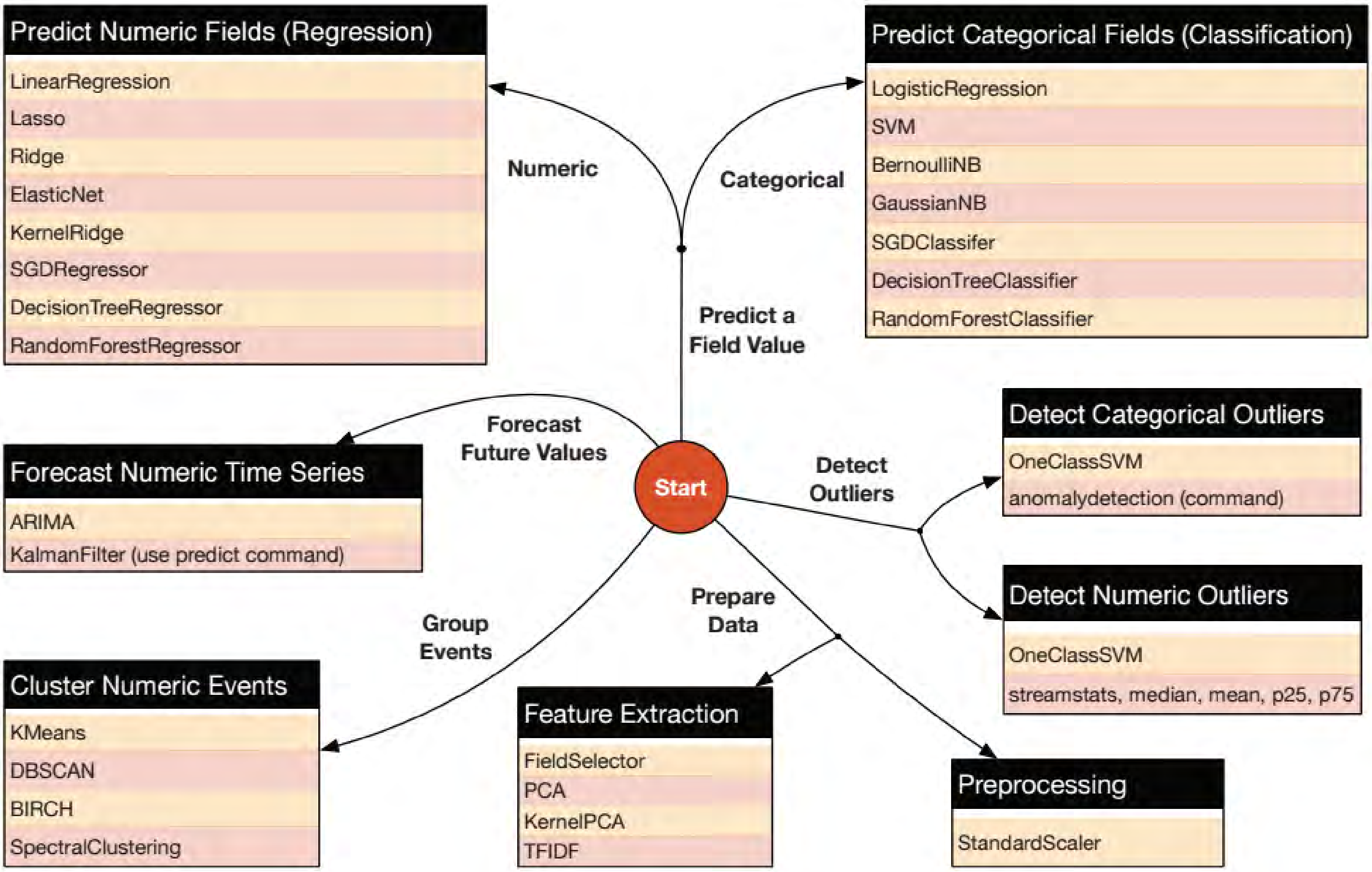


Figure 3 Splunk Machine Learning Algorithms

in AWS environments; creating the models with proper algorithms, and with a large amount and variety of data will ensure there are no overfitting problems. Another aspect to keep in mind is every AWS environment is different. As an example, most of the environments use different AWS Virtual Private Cloud (VPC) configurations to segregate AWS resources according to specific business needs. Creating machine learning models with data from the same AWS environments will produce the best results.

Limited use cases have been using Machine Learning for a long time, but recent developments in technologies (like Splunk and Apache Spark) make it feasible to deploy machine learning algorithms quickly over different data types and use features from many different data sets.

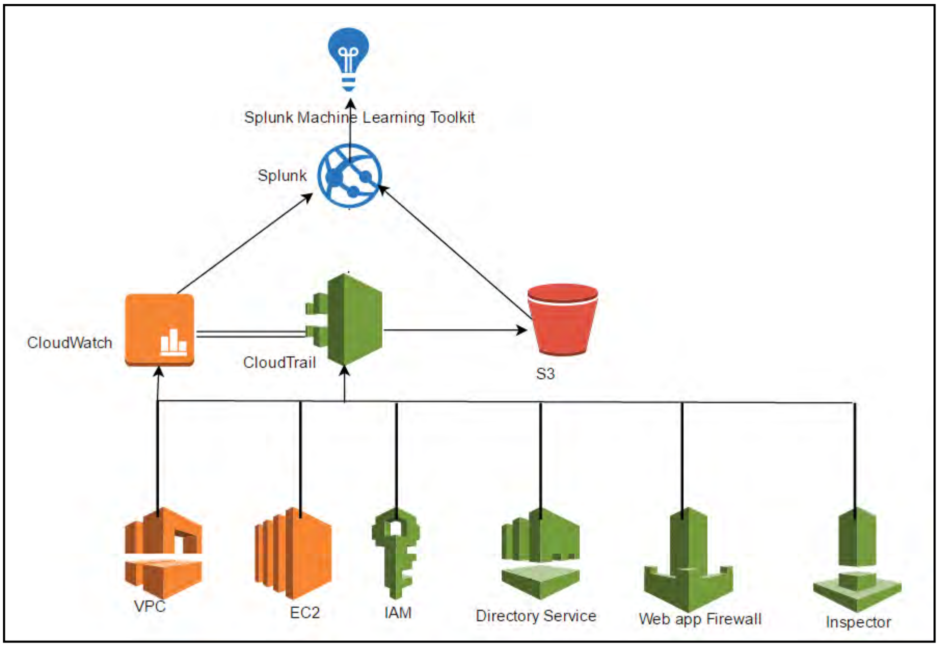


Figure 4 Lab Setup

2 LAB SETUP

Figure 4 highlights the logical structure of the lab created for this article.

Appendix A explains all the steps followed for the initial lab setup. In the lab configuration, Splunk is configured to receive logs from AWS Cloudtrail. The Splunk Machine Learning Toolkit is installed and set.

For further testing, generated additional logs by adding users, adding instances, launching new environments from AWS QuickStart. With the on-going collection of logs, machine learning examples will be applied in next section to calculate risk scores and detect suspicious events.

3 MACHINE LEARNING - PROCESS

The steps below highlight the methodology for applying machine learning techniques using Splunk Machine Learning Toolkit. The same process can be used with any machine learning solution and can apply to any security monitoring use case. This solution can also be implemented using Apache Spark MLLib libraries. One challenge is parsing and normalizing the AWS Cloudtrail JSON data files in Apache Spark. AWS has released the open source code to convert AWS CloudTrail logs to a Spark Data Frame (Github, 2016). After loading the data to Apache Spark data frames, the data can be used by Apache Spark MLLib libraries.

One aspect to remember during machine learning is data cleansing. Data cleansing ensures that the data is consistent and uniform. In many cases, the data should be extracted and formatted before being fed to machine learning algorithms. Splunk inherently addresses the data cleansing by indexing data at ingestion time and extracts relevant fields and provides a natural mapping from JSON format into standard columns. Machine learning algorithms can use the data directly from these columns. Splunk saves a lot of time in data cleansing and formatting when compared to many open source solutions like Apache Spark. Figure 5 highlights the steps involved in the machine learning process.

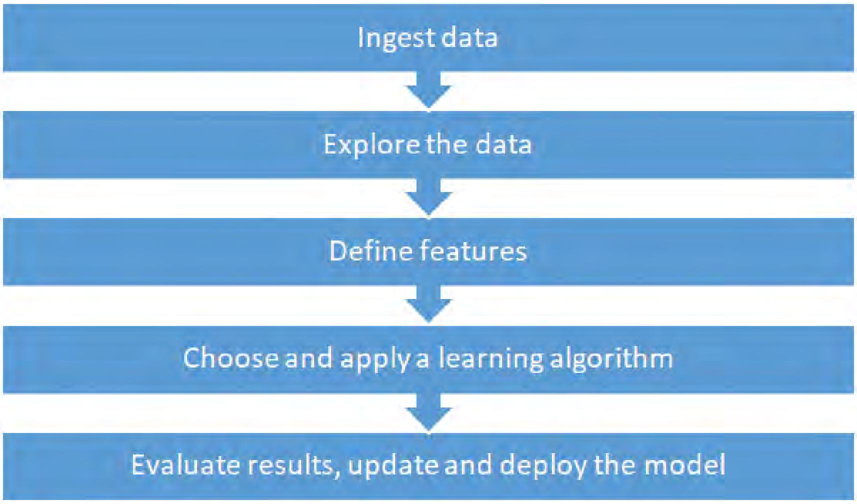

Ingest data
Explore the data
Define features
Choose and apply a learning algorithm
Evaluate results, update and deploy the model

Figure 5 Machine Learning Process

3.1 Ingest data to Splunk and understand the data

Security teams must collect all the AWS logs in a central place. Even if the organization could not implement any active monitoring, the logs will be useful for forensic analysis at a later point when an incident occurs.

In the initial lab setup, Splunk was configured to ingest data from Cloudtrail and Cloudwatch logs. The Splunk AWS app can be used to explore and understand the log events to identify features for machine learning.

3.1.1 AWS CloudTrail

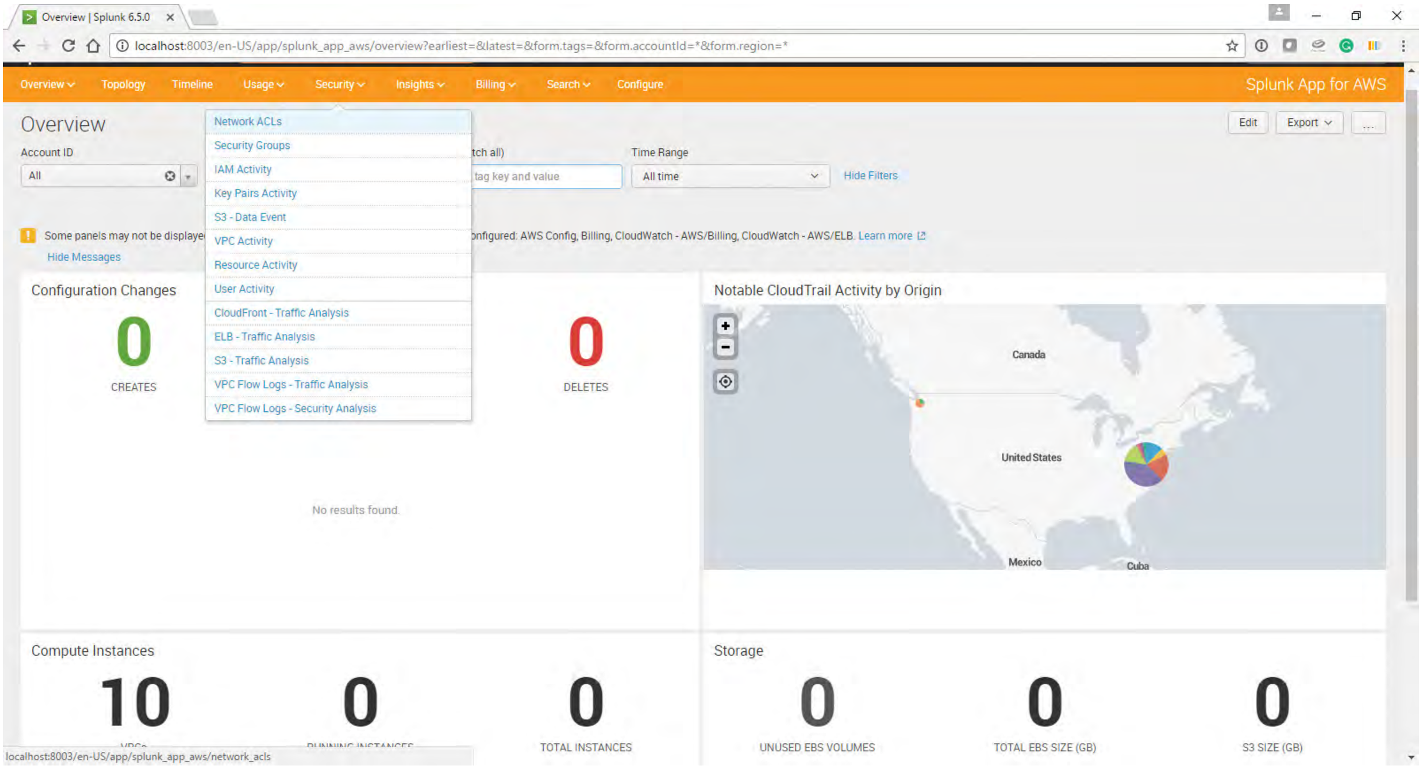
AWS Cloudtrail creates logs of all the API access requests, AWS resources access and AWS console login access information. It is important to understand the AWS Cloudtrail log data to efficiently design features for machine learning algorithms. AWS Cloudtrail User Guide (AWS, 2014) provides excellent reference and examples of different types of log events. The Cloudtrail API Call log contains two parts, Record Body Contents, and userIdentity Element. The types of events analyzed are:

- aws_cloudtrail_notable_network_events
- aws_cloudtrail_iam_change
- aws_cloudtrail_errors
- aws_cloudtrail_change
- aws_cloudtrail_delete_events
- aws_cloudwatch_sns_events
- aws_cloudtrail_auth
- aws_cloudtrail_iam_events
- aws_cloudtrail_ec2_events

3.2 Explore the data

In this particular use case, the Splunk App for AWS can be used to study the AWS log data. The Splunk App for AWS gives critical operational and security insight into the Amazon Web Services account. Figure 6 shows different dashboard options available on the Splunk App for AWS. These panels can be used to understand the relevant AWS logs which would help in determining any suspicious activity. The scroll down highlights the different security dashboards available.

The dashboard in Figure 7 highlights various user activities in the AWS environment. Looking at the data using different fields(features) will help the security team to understand the relevant fields in the log data.



localhost:8003/en-US/app/splunk_app_aws/overview?earliest=&latest=&form.tags=&form.accountId=*&form.region=*

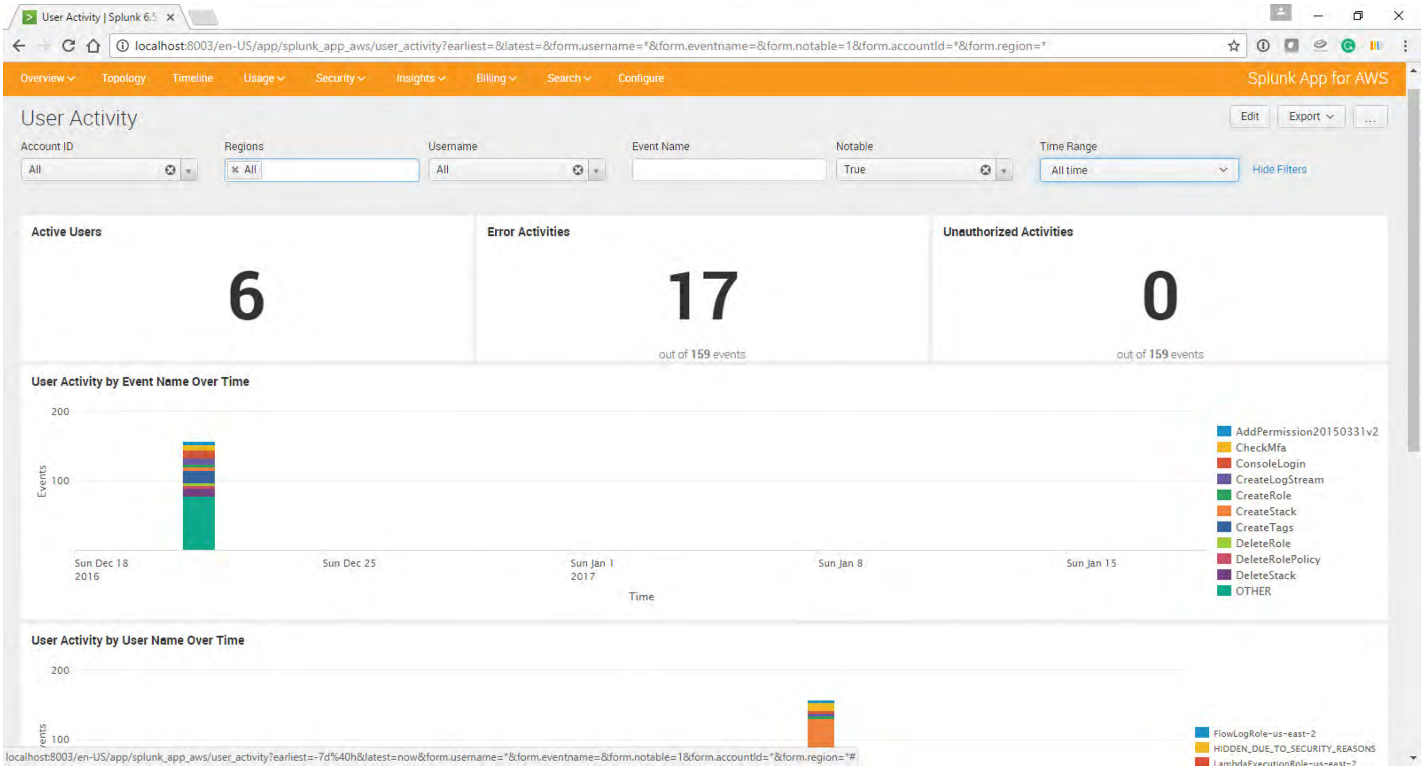
Account ID: All

Time Range: All time

Some panels may not be displaye… …onfigured: AWS Config, Billing, CloudWatch - AWS/Billing, CloudWatch - AWS/ELB.

localhost:8003/en-US/app/splunk_app_aws/network_acls

Figure 6 Splunk App for AWS Dashboards



localhost:8003/en-US/app/splunk_app_aws/user_activity?earliest=&latest=&form.username=*&form.eventname=&form.notable=1&form.accountId=*&form.region=*

User Activity

Account ID: All
Regions: × All
Username: All
Event Name:
Notable: True
Time Range: All time

Active Users: 6

Error Activities: 17 out of 159 events

Unauthorized Activities: 0 out of 159 events

localhost:8003/en-US/app/splunk_app_aws/user_activity?earliest=-7d%40h&latest=now&form.username=*&form.eventname=&form.notable=1&form.accountId=*&form.region=*#

Figure 7 Splunk App for AWS Security Dashboard - User Activities

Cyber Security & Information Systems Information Analysis Center

CSIAC
266 Genese Street
Utica, NY 13502
Phone: 1-800-214-7921

Build and Opera

ORGANIZE

Lead and Govern

EO 13800: Strengthening Cybersecurity of Fed Nets and CI	EO 13636: Improving Critical Infrastructure Cybersecurity	PPD 41: United States Cyber Incident Coordination	PPD 21: Critical Infrastructure Security and Resilience	National Cyber Strategy	U.S. I
National Defense Strategy (NDS)	2019 National Intelligence Strategy	DoD Cloud Strategy	National Military Strategy (NMS)	DoDD 8000.01 Management of the DOD Information Enterprise	

ORGANIZE

Design for the Fight

NIST SP 800-119 Guidelines for the Secure Deployment of IPv6	Common Criteria Evaluation and Validation Scheme (CCEVS)
CNSSP-11 Nat'l Policy Governing the Acquisition of IA and IA-Enable IT	DFARS Subpart 208.74, Enterprise Software Agreements
DoDD 5000.01 The Defense Acquisition System	DoDD 7045.20 Capability Portfolio Management
DoDD 8115.01 IT Portfolio Management	DoDI 5000.02 Operation of the Defense Acquisition System
DoDI 5200.44 Protection of Mission Critical Functions to Achieve TSN	DoDI 7000.14 Financial Management Policy and Procedures (PPBE)
DoDI 8115.02 IT Portfolio Management Implementation	DoDI 8310.01 Information Technology Standards in the DoD
DoDI 8330.01 Interoperability of IT and National Security Systems (NSS)	DoDI 8510.01 Risk Management Framework for DoD IT
DoDI 8580.1 Information Assurance (IA) in the Defense Acquisition System	RMF Knowledge Service
MOA between DoD CIO and ODNI CIO Establishing Net-Centric Software Licensing Agreements	DODAF (Version 2.02) DoD Architecture Framework
CJCSI 5123.01H Charter of the JROC and Implementation of the JCID	Joint Publication 6-0 Joint Communications System
CNSS National Secret Fabric Architecture Recommendations	MOA Between DoD and DHS (Jan. 19, 2017, requires CAC)

Develop the Workforce

CNSSD-500 Information Assurance (IA) Education, Training, and Awareness	NSTISSD-501 National Training Program for INFOSEC Professionals
CNSSI-4000 Maintenance of Communications Security (COMSEC) Equipment	NSTISSI-4011 National Training Standard for INFOSEC Professionals
CNSSI-4012 National IA Training Standard for Senior Systems Managers	CNSSI-4013 National IA Training Standard For System Administrators (SA)
CNSSI-4014 National IA Training Standard For Information Systems Security Officers	NSTISSI-4015 National Training Standard for System Certifiers
CNSSI-4016 National IA Training Standard For Risk Analysts	DoDD 8140.01 Cyberspace Workforce Management
DoD 8570.01-M Information Assurance Workforce Improvement Program	DoDI 8550.01 DoD Internet Services and Internet-Based Capabilities

Partner for Strength

NIST SP 800-144 Guidelines on Security and Privacy in Public Cloud Computing	NIST SP 800-171, R1 Protecting CUI in Nonfederal Systems and Organizations
CNSSP-14 National Policy Governing the Release of IA Products/Services…	CNSSI-1253 Security Categorization and Control Selection for Nat'l Security Systems
CNSSI-1253F, Atchs 1-5 Security Overlays	CNSSI-4007 Communications Security (COMSEC) Utility Program
CNSSI-4008 Program for the Mgt and Use of Nat'l Reserve IA Security Equipment	DoDD 5205.13 Defense Industrial Base Cyber Security / IA Activities
DoD 5220.22-M, Ch. 2 National Industrial Security Program Operating Manual (NISPOM)	ICD 503 IT Systems Security Risk Management and C&A

ENABLE

Secure Data in Transit

FIPS 140-2 Security Requirements for Cryptographic Modules	NIST SP 800-153 Guidelines for Securing Wireless Local Area Networks
CNSSP-1 National Policy for Safeguarding and Control of COMSEC Material	CNSSP-15 Use of Pub Standards for Secure Sharing of Info Among NSS
CNSSP-17 Policy on Wireless Communications: Protecting Nat'l Security Info	CNSSP-19 National Policy Governing the Use of HAIPE Products
CNSSP-25 National Policy for PKI in National Security Systems	NSTISSP-101 National Policy on Securing Voice Communications
NACSI-2005 Communications Security (COMSEC) End Item Modification	CNSSI-5000 Voice Over Internet Protocol (VoIP) Computer Telephony (Annex I, VoSIP)
CNSSI-5001 Type-Acceptance Program for VoIP Telephones	NACSI-6002 Nat'l COMSEC Instruction Protection of Gov't Contractor Telecomm's
CNSSI-7003 Protected Distribution Systems (PDS)	DoDD 8100.02 Use of Commercial Wireless Devices, Services, and Tech in the DoD GIG
DoDD 8521.01E Department of Defense Biometrics	DoDI 4650.01 Policy and Procedures for Mgt and Use of the Electromagnetic Spectrum
DoDI 8100.04 DoD Unified Capabilities (UC)	DoDI 8420.01 Commercial WLAN Devices, Systems, and Technologies
DoDI 8523.01 Communications Security (COMSEC)	DoDI S-5200.16 Objectives and Min Stds for COMSEC Measures used in NC2 Comms
CJCSI 6510.02E Cryptographic Modernization Plan	*CJCSI 6510.06C Communications Security Releases to Foreign Nations*

Manage Access

HSPD-12 Policy for a Common ID Standard for Federal Employees and Contractors	FIPS 201-2 Personal Identity Verification (PIV) of Federal Employees and Contractors
CNSSP-3 National Policy for Granting Access to Classified Cryptographic Information	CNSSP-16 National Policy for the Destruction of COMSEC Paper Material
CNSSI-1300 Instructions for NSS PKI X.509	NSTISSI-3028 Operational Security Doctrine for the FORTEZZA User PCMCIA Card
CNSSI-4001 Controlled Cryptographic Items	CNSSI-4003 Reporting and Evaluating COMSEC Incidents
CNSSI-4005 Safeguarding COMSEC Facilities and Materials, amended by CNSS-008-14	CNSSI-4006 Controlling Authorities for COMSEC Material
DoDI 1000.25 DoD Personnel Identity Protection (PIP) Program	DoDI 5200.01 DoD Information Security Program and Protection of SCI
DoDI 5200.08 Security of DoD Installations and Resources and the DoD PSRB	DoDI 8520.02 Public Key Infrastructure (PKI) and Public Key (PK) Enabling
DoDI 8520.03 Identity Authentication for Information Systems	DoDM 1000.13, Vol. 1 DoD ID Cards: ID Card Life-cycle

Assure Information Sharing

DoDI 8320.02 Sharing Data, Info, and IT Services in the DoD	DoDI 8582.01 Security of Unclassified DoD Information on Non-DoD Info Systems
DoD Information Sharing Strategy	United States Intelligence Community Information Sharing Strategy
CJCSI 6211.02D Defense Information System Network: (DISN) Responsibilities	*CJCSM 3213.02D, Joint Staff Focal Point*

ANTIC

Understand th

FIPS 199 Standards for Security Categorization of Federal Info. and Info. Systems
NIST SP 800-60, Vol 1, R1 Guide for Mapping Types of Info and Info Systems to Security Categories
NIST SP 800-101, R1 Guidelines on Mobile Device Forensics
CNSSP-28 Cybersecurity of Unmanned National Security Systems

Prevent and D and Prevent Attac

FIPS 200 Minimum Security Requirements for Federal Information Systems
NIST SP 800-53 R4 Security & Privacy Controls for Federal Information Systems
NIST SP 800-61, R2 Computer Security Incident Handling Guide
NIST SP 800-128 Guide for Security-Focused Configuration Mgt of Info Systems
DoDI 5200.39 CPI Identification and Protection within RDT&E
DoDI 8530.01, Cybersecurity Activities Support to DoD Information Network Operations
DoDM 5105.21V1, SCI Admin Security Manual: Info and Info Sys Security
CJCSM 6510.01B Cyber Incident Handling Program
DTM 17-007, Defense Support to Cyber Incident Response

ABOUT TH

- This chart organizes cybersec Strategic Goal and Office of Pr Key). Double-clicking on the b authoritative publicly accessib
- *Policies in italics indicate the distribution or no authoritative currently available.*
- The linked sites are not contro chart. We check the integrity you may occasionally experier problems at the source site or document. Please let us know longer valid.
- CNSS policies link only to the implemented by its website de
- Boxes with red borders reflect
- Note: Users of the iPad, iPhon can view this Chart but that its because of Apple's decision n products. For those who desire there are apps in the iTunes s
- For the latest version of this ch ia_policychart.html. You can s any updates to this document.

ate a Trusted DoDIN

Cybersecurity-Related Policies and Issuances

Developed by the DoD Deputy CIO for Cybersecurity

Last Updated: February 28, 2019
Send questions/suggestions to
info@csiac.org

nt'l Strategy for Cyberspace

NIST Framework for Improving Critical Infrastructure Cybersecurity

2017 National Security Strategy

CNSSP-24 Policy on Assured Info Sharing (AIS) for National Security Systems(NSS)

DoDI 8500.01 Cybersecurity

2018 DoD Cyber Strategy

DoD Defending Networks, Systems and Data Strategy

DoD Information Technology Environment Strategic Plan

IPATE

e Battlespace

NIST SP 800-59 Guideline for Identifying an Information System as a NSS

NIST SP 800-92 Guide to Computer Security Log Management

NISTIR 7693 Specification for Asset Identification 1.1

DoDI S-5240.23 Counterintelligence (CI) Activities in Cyberspace

elay Attackers
kers from Staying

NIST SP 800-37 R1 Guide for Applying the Risk Mgt Framework to Fed. Info. Systems

NIST SP 800-53A R4 Assessing Security & Privacy Controls in Fed. Info. Systems & Orgs.

NIST SP 800-124, R1 Guidelines for Managing the Security of Mobile Devices in the Enterprise

CNSSAM IA 1-10, Reducing Risk of Removable Media in NSS

DoDI 8551.01 Ports, Protocols, and Services Management (PPSM)

DoD O-8530.1-M CND Service Provider Certification and Accreditation Program

CJCSI 6510.01F Information Assurance (IA) and Computer Network Defense (CND)

CJCSM 6510.02 IA Vulnerability Mgt Program

PREPARE

Develop and Maintain Trust

CNSSP-12 National IA Policy for Space Systems Used to Support NSS

CNSSP-21 National IA Policy on Enterprise Architectures for NSS

NSTISSD-600 Communications Security (COMSEC) Monitoring

CNSSI-5002, National Information Assurance (IA) Instruction for Computerized Telephone Systems

DoDD 3020.40 Mission Assurance

DoDD 3100.10 Space Policy

DoDI 8581.01 IA Policy for Space Systems Used by the DoD

DoDD 5144.02 DoD Chief Information Officer

Strengthen Cyber Readiness

NIST SP 800-18, R1 Guide for Developing Security Plans for Federal Information Systems

NIST SP 800-30, R1 Guide for Conducting Risk Assessments

NIST SP 800-126, R3 SCAP Ver. 1.3

NIST SP 800-137 Continuous Monitoring

NIST SP 800-39 Managing Information Security Risk

DoDD 3700.01 DoD Command and Control (C2) Enabling Capabilities

DoDD S-3710.01 National Leadership Command Capability

DoDI 8560.01 COMSEC Monitoring

Joint Special Access Program (SAP) Implementation Guide (JSIG)

Sustain Missions

CNSSP-18 National Policy on Classified Information Spillage

CNSSP-22, IA Risk Management Policy for National Security Systems

CNSSP-300 National Policy on Control of Compromising Emanations

CNSSI-1001 National Instruction on Classified Information Spillage

CNSSI-4004.1, Destruction and Emergency Protection Procedures for COMSEC and Class. Material

CNSSI-7000 TEMPEST Countermeasures for Facilities

NSTISSI-7001 NONSTOP Countermeasures

DoDD 3020.26 DoD Continuity Policy

DoDD 3020.44 Defense Crisis Management

DoDI 8410.02 NetOps for the Global Information Grid (GIG)

Defense Acquisition Guidebook RMF for DoD IT

NSA IA Directorate (IAD) Management Directive MD-110 Cryptographic Key Protection

AUTHORITIES

Title 10 Armed Forces (§§2224, 3013(b), 5013(b), 8013(b))

Title 14 Cooperation With Other Agencies (Ch. 7:§§ 141,144,145,148,149,150)

Title 32 National Guard (§102)

Title 40 Public Buildings, Property, and Works (Ch. 113: §§11302, 11315, 11331)

Title 44 Federal Information Security Mod. Act, (Chapter 35)

Title 50 War and National Defense (§§3002, 1801)

Clinger-Cohen Act, Pub. L. 104-106

UCP Unified Command Plan (US Constitution Art II, Title 10 & 50)

NATIONAL / FEDERAL

Computer Fraud and Abuse Act Title 18 (§1030)

Federal Wiretap Act Title 18 (§2510 et seq.)

Stored Communications Act Title 18 (§2701 et seq.)

Pen Registers and Trap and Trace Devices Title 18 (§3121 et seq.)

Foreign Intelligence Surveillance Act Title 50 (§1801 et seq)

Executive Order 13231 as Amended by EO 13286 - Critical Infrastructure Protection in the Info Age

Executive Order 13526 Classified National Security Information

Executive Order 13587 Structural Reforms To Improve Classified Nets

Executive Order 13691 Promoting Private Sector Cybersecurity Information Sharing

NSD 42, National Policy for the Security of Nat'l Security Telecom and Information Systems

PPD 28, Signals Intelligence Activities

NSPD 54 / HSPD 23 Computer Security and Monitoring

A-130, Management of Fed Info Resources

FAR Federal Acquisition Regulation

Ethics Regulations

National Strategy to Secure Cyberspace

Summary of the 2018 DoD Artificial Intelligence Strategy

NIST Special Publication 800 Series

NIST SP 800-125A, R1, Security Recommendations for Hypervisor Platforms

NISTIR 7298, R2, Glossary of Key Information Security Terms

CNSSD-502 National Directive On Security of National Security Systems

CNSSD-900, Governing Procedures of the Committee on National Security Systems

CNSSD-901 Nat'l Security Telecomm's and Info Sys Security (CNSS) Issuance System

CNSSI-4009 Cmte on National Security Systems Glossary

OPERATIONAL

CYBERCOM Orders

JFHQ-DODIN Orders

SUBORDINATE POLICY

Security Configuration Guides (SCGs)

Component-level Policy (Directives, Instructions, Publications, Memoranda)

Security Readiness Review Scripts (SRRs)

Security Technical Implementation Guides (STIGs)

IS CHART
urity policies and guidance by
rimary Responsibility (see Color
ox directs users to the most
e source.
document is marked for limited
public-facing hyperlink is

lled by the developers of this
of the links on a regular basis, but
nce an error message due to
the site's decision to move the
if you believe the link is no

CNSS site, per restrictions
sign.
recent updates.
e or iPod Touch may find they
hyperlinks are inoperable,
ot to fully support certain Adobe
e a workaround for this issue,
tore for less than $1.00.
hart go to http://iac.dtic.mil/csiac/
sign up to be alerted by e-mail to

Color Key - OPRs

■	ASD(NII)/ASD(C3I) /DOD CIO	■	NIST	■	USD(I)
■	CNSS/NSTISS	■	NSA	■	USD(P)
■	DISA	■	OSD	□	USD(P&R)
■	DNI	■	CYBERCOM	■	Other Agencies
■	JCS	■	USD(AT&L)	□	Recently updated policy and/or link
■	NIAP	■	USD(C)	□	Expired, Update pending

Distribution Statement A: Approved for Public Release. Distribution is unlimited.

The Splunk App for AWS allows security practitioners to understand and explore the data to determine the fields that will be helpful in identifying the suspicious AWS activities. The Splunk Machine Learning Toolkit can use these fields as features while developing the model.

3.3 Use Case 1 – Detecting Suspicious AWS Console Logins

3.3.1 Define features

In this case study, the “AwsConsoleSignIn” events were explored using domain expertise on AWS Cloud security with the goal of understanding which fields will be beneficial to determine any suspicious login to AWS Console.

- Some of the relevant fields identified are:
- sourceIPAddress
- userAgent
- userIdentity.arn
- eventTime
- responseElements.ConsoleLogin

In the above example, understanding and exploring the various logs from a security perspective enabled to identify the features. The Splunk Machine Learning Toolkit has algorithms like Principal Component Analysis (PCA) which can be used to explore and define features mathematically using the data. It will be useful to understand the data from a different vantage point which will assist with determining unusual activity.

3.3.2 Choose and apply a learning algorithm

This section highlights the Splunk Machine Learning Toolkit commands required to create, score and test the model.

AwsConsoleSignIn.csv is generated using the AWS logs from the lab environment.

Splunk command can be used to export the events with defined features as CSV:

```
* sourcetype=”aws:cloudtrail”
eventType=AwsConsoleSignIn |
table sourceIPAddress, userAgent,
userIdentity.arn, eventTime,
responseElements.ConsoleLogin
```

Security professionals should add risk scores to these events. Risk scores should be assigned based on the security domain knowledge and the environment. Ideally, security professionals should review and assign risk scores to the events for a month or more depending on the environment. The Machine learning model requires a significant amount of labeled risk score data to achieve useful results.

A sample record from the AwsConsoleSignIn.csv is highlighted in Figure 8.

Create New Model – AwsConsoleSignIn

The goal of this machine learning model is to predict a risk score to identify the highest set of suspicious events. The inputs for creating this Supervised Learning Model is the data with assigned risk scores and the algorithm LinearRegression. The output will be a model which will mathematically predict the risk scores by learning from the data.

In the Machine Learning Toolkit app configuration highlighted in Figure 9, choose “Assistants -> Predict Numeric fields” and in the “Enter a search” provide the input file AwsConsoleSignIn.csv using the command:

```
| inputlookup AwsConsoleSignIn.csv
```

AwsConsoleSignIn.csv is provided as input file. The search loads all the records in the AwsConsoleSignIn.csv file for analysis. The following options are chosen to create the model:

Algorithm: LinearRegression.

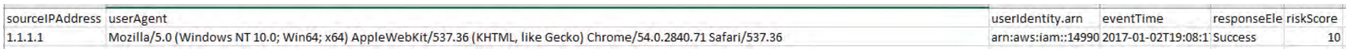
Field to predict: riskScore

Fields to use for predicting: “sourceIPAddress”, “userAgent”, “userIdentity.arn”, “eventTime”, “responseElements.ConsoleLogin”

The resultant set of Splunk commands are below:

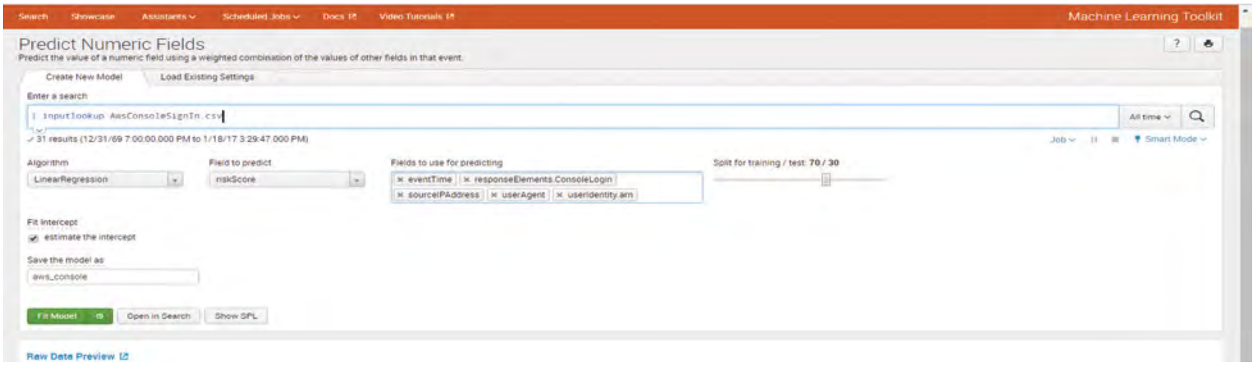
```
| inputlookup AwsConsoleSignIn.csv
| fit LinearRegression fit_
intercept=true “risk_score” from
“sourceIPAddress”, “userAgent”,
“userIdentity.arn”, “eventTime”,
“responseElements.ConsoleLogin”
into “aws_console”
```

Supervised Learning is used in this particular example and the machine learning algorithm LinearRegression is provided with AWS log data and



sourceIPAddress	userAgent	userIdentity.arn	eventTime	responseEle	riskScore
1.1.1.1	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/54.0.2840.71 Safari/537.36	arn:aws:iam::14990	2017-01-02T19:08:1	Success	10

Figure 8 Sample Record AWSConsoleSignIn



Predict Numeric Fields

Predict the value of a numeric field using a weighted combination of the values of other fields in that event.

Enter a search: | inputlookup AwsConsoleSignIn.csv

✓ 31 results (12/31/69 7:00:00.000 PM to 1/18/17 3:29:47.000 PM)

Algorithm: LinearRegression

Field to predict: riskScore

Fields to use for predicting: eventTime, responseElements.ConsoleLogin, sourceIPAddress, userAgent, userIdentity.arn

Split for training / test: 70 / 30

Fit Intercept: ☑ estimate the intercept

Save the model as: aws_console

Figure 9 Splunk Machine Learning Toolkit

actual risk scores. The resultant model aws_console will try to predict the risk score given the set of features "sourceIPAddress", "userAgent", "userIdentity:arn" , "eventTime", "responseElements:ConsoleLogin".

Evaluate results and update the model

In the configuration highlighted in Figure 10, data is split 70 % for training and 30 % for testing and evaluating the model. Allocating 30 % of data for testing and evaluation of the model helps understand the accuracy of the model.

After fitting the model, Splunk Machine Learning Toolkit performs the necessary computations used for measuring the performance of the model.

Plot actual vs. predicted values on a line chart as depicted in Figure 11 helps security teams to understand the efficiency of the model.

These commands will apply the model to the data set in AwsConsoleSignIn.csv and plot the actual vs. predicted values to understand the accuracy of the model.

```
| inputlookup AwsConsoleSignIn.csv
| apply " aws_console."
| table _time, " risk_score ",
"predicted(risk_score)"
```

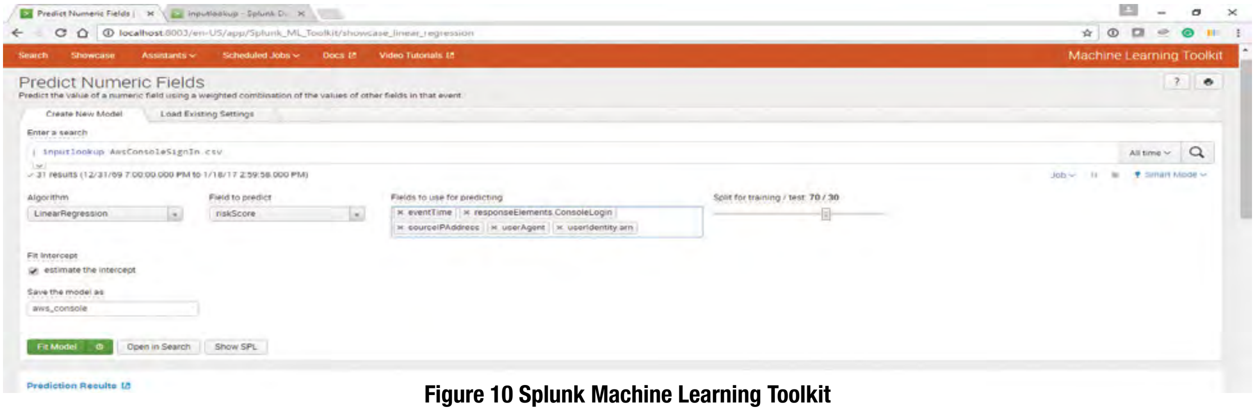
These commands will apply the model to the data set in AwsConsoleSignIn.csv to calculate the R^2 and root mean squared error (RMSE).

These values assist with measuring the accuracy of the model.

```
| inputlookup AwsConsoleSignIn.
csv
| apply " aws_console "
| `regressionstatistics("risk_score
", "predicted(risk_score)")`
```

Root Mean Square Error and R^2 values provide an idea of the magnitude of the error. R^2 presents an indication of the effectiveness of a set of predictions to the actual values. The value is between 0 and 1. The value near 0 indicates the model is not a good fit, as seen in Figure 12.

After analyzing the performance, if the performance is not satisfactory, additional features can be extracted. As an example, including Geolocation data from Maxmind Geodatabase for the new context of Geolocation for the IP addresses involved will help improve the effectiveness of the model.



Predict Numeric Fields

Predict the value of a numeric field using a weighted combination of the values of other fields in that event

Enter a search: | inputlookup AwsConsoleSignIn.csv

✓ 31 results (12/31/69 7:00:00.000 PM to 1/18/17 2:59:58.000 PM)

Algorithm: LinearRegression

Field to predict: riskScore

Fields to use for predicting: eventTime, responseElements.ConsoleLogin, sourceIPAddress, userAgent, userIdentity.arn

Split for training / test: 70 / 30

Fit Intercept: ☑ estimate the intercept

Save the model as: aws_console

Figure 10 Splunk Machine Learning Toolkit

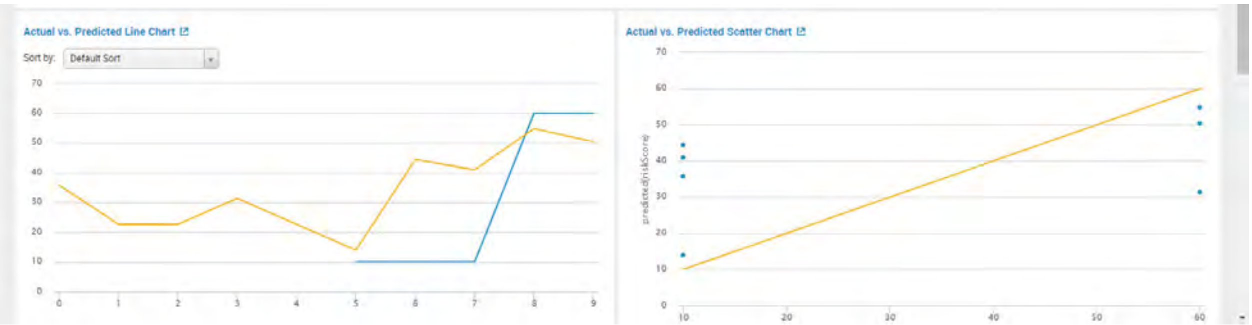
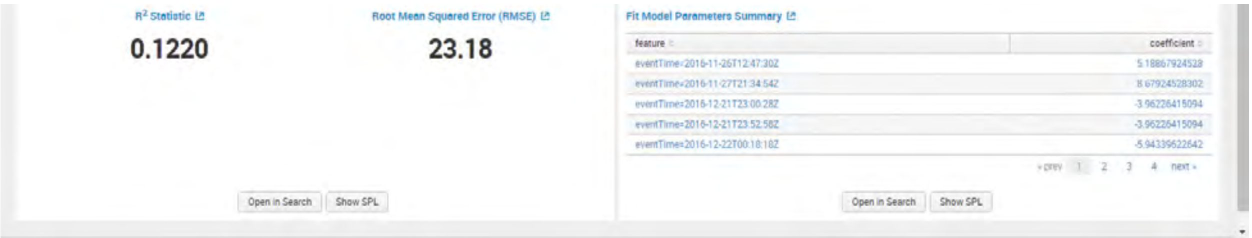


Figure 11 Splunk Machine Learning Toolkit - Actual vs. Predicted Chart



R² Statistic: 0.1220

Root Mean Squared Error (RMSE): 23.18

Fit Model Parameters Summary

feature	coefficient
eventTime=2016-11-26T12:47:30Z	5.18867924528
eventTime=2016-11-27T21:34:54Z	8.67924528302
eventTime=2016-12-21T23:00:28Z	-3.96226415094
eventTime=2016-12-21T23:52:58Z	-3.96226415094
eventTime=2016-12-22T00:18:18Z	-5.94339622642

Figure 12 Splunk Machine Learning Toolkit - RMSE

Once the performance is satisfactory, security team should deploy the model using the apply <model> command. After implementing the model, security analysts should continually tune the model based on the feedback from the security analysts who are analyzing the results.

3.4 Use Case 2 – Detecting Suspicious API calls

3.4.1 Define features

The “AwsApiCall” events were explored using domain expertise on AWS Cloud security with the goal of understanding which fields will be beneficial to determine any suspicious AWS API calls.

Some of the relevant fields identified are:

- sourceIPAddress
- eventSource
- eventName
- userIdentity:arn
- eventTime
- userAgent
- userIdentity:type

3.4.2 Choose and apply a learning algorithm

AwsAPICall.csv is generated using the AWS logs from the lab environment.

Splunk command can be used to export the events with defined features as CSV:

```
* sourcetype=”aws:cloudtrail”
eventType=AwsAPICall | table
sourceIPAddress, eventSource ,
eventName , userIdentity.arn,
eventTime, userAgent, userIdentity.
type
```

Security professionals should add risk scores to these events. The risk scores should be assigned based on the security domain knowledge and the environment. A sample record from the AwsAPICall.csv is shown in Figure 13.

Create New Model – AwsAPICall

The goal of this machine learning model is to predict a risk score to identify the highest set of suspicious API calls. In the Machine Learning Toolkit app, choose Assistants -> Predict Numeric fields, and in the search box provide the input file AwsConsoleSignIn.csv using the command:

```
| inputlookup AwsAPICall.csv
```

AwsAPICall.csv is provided as input file. The search loads all the records in the AwsAPICall.csv file for analysis. The following options are chosen to create the model:

Algorithm: LinearRegression.

Field to predict: riskScore

In the configuration highlighted in Figure 14, the fields used for predicting are: “sourceIPAddress”, “eventSource” , “eventName” , “userIdentity.arn”, “eventTime”, “userAgent”, “userIdentity.type”

The resultant set of Splunk commands are below:

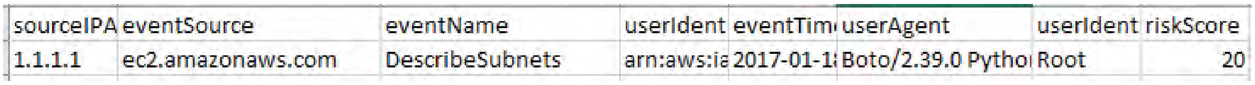
```
| inputlookup AwsAPICall.csv
| fit LinearRegression fit_
intercept=true “riskScore” from
“sourceIPAddress”, “eventSource”
, “eventName” , “userIdentity.
arn”, “eventTime”, “userAgent”,
“userIdentity.type”  into “aws_
apicall”
```

Supervised learning is used in our example, the machine learning algorithm LinearRegression is provided with AWS log data and actual risk scores. The resultant model aws_console will try to predict the risk score given the set of features “sourceIPAddress”, “eventSource” , “eventName” , “userIdentity.arn”, “eventTime”, “userAgent”, “userIdentity.type”

Evaluate results and update the model

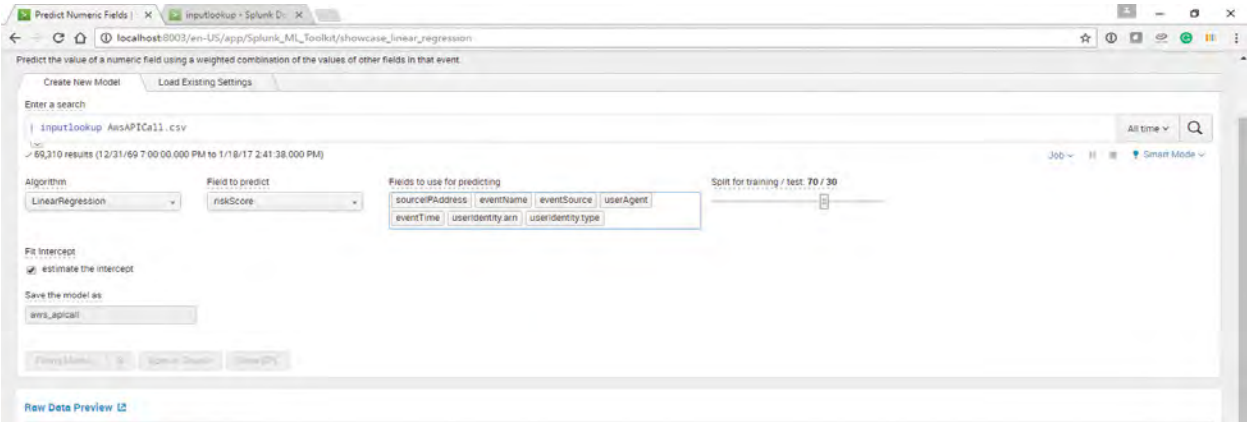
The data is split 70 % for training and 30 % for testing and evaluating the model. This helps understand the accuracy of the model.

After fitting the model, Splunk Machine Learning Toolkit performs the necessary computations used for measuring



sourceIPA	eventSource	eventName	userIdent	eventTim	userAgent	userIdent	riskScore
1.1.1.1	ec2.amazonaws.com	DescribeSubnets	arn:aws:ia	2017-01-1	Boto/2.39.0 Pytho	Root	20

Figure 13 Sample Record AWSAPICAll



Predict Numeric Fields | ×
inputlookup - Splunk D ×
localhost:8003/en-US/app/Splunk_ML_Toolkit/showcase_linear_regression

Predict the value of a numeric field using a weighted combination of the values of other fields in that event

Create New Model
Load Existing Settings

Enter a search

```
| inputlookup AwsAPICall.csv
```

✓ 69,310 results (12/31/69 7:00:00.000 PM to 1/18/17 2:41:38.000 PM)

Algorithm: LinearRegression

Field to predict: riskScore

Fields to use for predicting: sourceIPAddress, eventName, eventSource, userAgent, eventTime, userIdentity.arn, userIdentity.type

Split for training / test: **70 / 30**

Fit Intercept

☑ estimate the intercept

Save the model as: aws_apicall

Figure 14 Splunk Machine Learning Toolkit

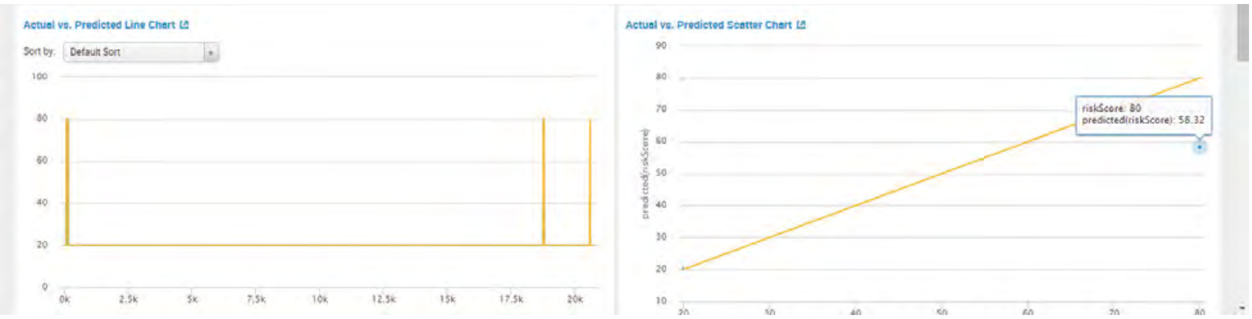
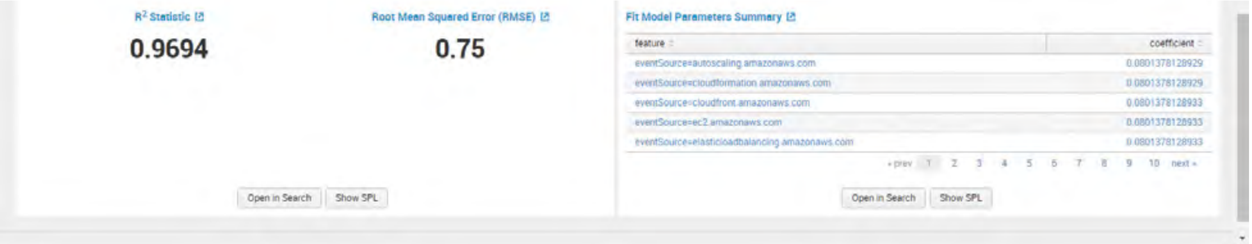


Figure 15 Splunk Machine Learning Toolkit - Plot actual vs. predicted value



R² Statistic

0.9694

Root Mean Squared Error (RMSE)

0.75

Fit Model Parameters Summary

feature	coefficient
eventSource=autoscaling.amazonaws.com	0.0801378128929
eventSource=cloudformation.amazonaws.com	0.0801378128929
eventSource=cloudfront.amazonaws.com	0.0801378128933
eventSource=ec2.amazonaws.com	0.0801378128933
eventSource=elasticloadbalancing.amazonaws.com	0.0801378128933

Figure 16 Splunk Machine Learning Toolkit - RMSE

the performance of the model. Plot actual vs. predicted values on a line chart as depicted in Figure 15 helps understand the efficiency of the model.

Root Mean Square Error and R^2 values provide an idea of the magnitude of the error. R^2 presents an indication of the effectiveness of a set of predictions to the actual values. The value is between 0 and 1. The value near 1 indicates the model is a good fit.

After analyzing, if the performance is not satisfactory, additional features can be extracted. As an example, including AWS VPC information for the context of VPCs involved in an event will help improve the effectiveness of the model.

Once the performance is satisfactory, deploy the model using the apply <model> command. After implementing the model, security analysts should continually tune the model based on the feedback from the security analysts who are analyzing the results.

This section highlights how Splunk Machine Learning toolkit could be used to create, evaluate and deploy Machine Learning models. The two models generated are examples with very few data records created in the lab to prototype the Splunk Machine Learning toolkit functionality. Security teams should test, tune and deploy the Machine Learning models according to the AWS environment.

There are only two use cases discussed in this article. Another practical use case that might be useful is determining anomalous network traffic sessions between various AWS VPCs. Threat modeling with inputs from adversary Tools, Techniques, and Procedures (TTPs) can be used to identify additional security monitoring use cases in AWS environments. After determining the use case, the methodology discussed in this article can be used to evaluate features and apply Machine Learning model to new use cases.

Machine learning is very data hungry and ingesting a lot of data for creating machine learning models will produce useful results. Also, if multiple data sources are used to extract features, greater fidelity can be achieved. As an example, including Geolocation data for the additional context of the IP addresses involved will help improve the effectiveness of the model.

4 CONCLUSION

This article highlights how to implement machine learning techniques for AWS logs. Machine learning techniques were applied in this article to identify suspicious events in IaaS environments. Identity is the new perimeter and using machine learning techniques to identity data in combination with other telemetry data will help security professionals identify suspicious events. As a first step, the security team members should understand the monitoring requirements, understand the data and evaluate the suitable methods. The security team should consider if machine learning is appropriate for the nature of the logs, explore, visualize and select the features as inputs to creating the model. After building and testing the model, the security team should apply the model to real-time traffic(data). After using the model, the security team should periodically evaluate the results and tune the model.

Many machine learning solutions are becoming available like Microsoft Azure ML Studio, Amazon Machine Learning, Databricks Spark, Splunk Machine Learning toolkit. All of these machine learning tools make the implementation of machine learning models very intuitive and easy to implement with simple user interfaces. These user interfaces encapsulate the mathematics and coding involved in traditional machine learning application languages like R.

Using Amazon Machine Learning for security monitoring was explained with demos in the AWS re Invent 2016 Conference (Videos from re Invent 2016 security and compliance sessions, 2016). As the cloud implementations evolve, the security teams should also learn the advantages and new ways of implementing security operations and security monitoring activities. Automation and machine learning are two key areas in the cloud that give an edge to defenders.

A defender has to detect only one of the attacker's activity before successful completion of attacker's objectives. As defenders, the goal is to deploy defense in depth strategy by placing preventive and detective controls at every layer to introduce high cost for an attacker to achieve his objectives. Machine Learning can be one useful tool in the defense in depth strategy to detect suspicious activity. After identification of the suspicious activity, using forensics, the security teams could be able to track and trace any activity performed by the attacker and take remediation actions.

Some other use cases that might benefit from this solution are Risk Management, Security Automation/ Orchestration, User/Network Behavior Analytics, Fraud Detection, Threat Hunting, Threat Intelligence aggregation from various sources, and Incident Response/Forensic Analysis.

REFERENCES

[1] The NIST Definition of Cloud Computing. (n.d.). Retrieved January 01, 2017, from http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf

[2] US Government Cloud Computing Technology Roadmap. (n.d.). Retrieved January 1, 2017, from http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.500-293.pdf

[3] Cloud Computing Synopsis and Recommendations. (2012). Retrieved January 1, 2017, from http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-146.pdf

[4] Cloud Computing Risk Assessment. (n.d.). Retrieved January 01, 2017, from http://www.enisa.europa.eu/act/rm/files/deliverables/cloud-computing-risk-assessment

[5] NIST Cloud Computing. (n.d.). Retrieved January 01, 2017, from http://csrc.nist.gov/groups/SNS/cloud-computing/

[6] Cloud Security Alliance. (n.d.). Retrieved January 01, 2017, from http://www.cloudsecurityalliance.org/

[7] Cloud Computing Top Threats in 2016. (2016). Retrieved January 1, 2017, from https://downloads.cloudsecurityalliance.org/assets/research/top-threats/Treacherous-12_Cloud-Computing_Top-Threats.pdf

[8] (2016,). AWS Well-Architected Framework. Retrieved January 4, 2017, from http://d0.awsstatic.com/whitepapers/architecture/AWS_Well-Architected_Framework.pdf

[9] (2016,). Hardening AWS Environments and Automating Incident Response for AWS Compromises. Retrieved January 4, 2017, from https://s3-us-west-2.amazonaws.com/threatresponse-static/us-16-Krug-Hardening-AWS-Environments-and-Automating-Incident-Response-for-AWS-Compromises-wp.pdf

[10] (2016,). Cloud security monitoring: Challenges and guidance. Retrieved January 4, 2017, from http://searchcloudsecurity.techtarget.com/tip/Cloud-security-monitoring-Challenges-and-guidance

[11] (2016,). About the Splunk App for AWS. Retrieved January 4, 2017, from http://docs.splunk.com/Documentation/AWS/latest/Installation/Abouttheapp

[12] (2016, August). Amazon Web Services: Overview of Security Processes. Retrieved January 4, 2017, from https://d0.awsstatic.com/whitepapers/aws-security-whitepaper.pdf

[13] (2015, October 6). Splunk App for AWS: Making the invisible, visible. Retrieved January 4, 2017, from http://blogs.splunk.com/2015/10/06/splunk-app-for-aws/

[14] (2016, July). Best Practices for Managing Security Operations in AWS. Retrieved January 4, 2017, from http://www.slideshare.net/AmazonWebServices/best-practices-for-managing-security-operations-in-aws-aws-july-2016-webinar-series

[15] (2014,). Use Your AWS CloudTrail Data and Splunk Software to Improve Security and Compliance in AWS. Retrieved January 4, 2017, from http://www.slideshare.net/AmazonWebServices/aws-partner-webcast-use-your-aws-cloudtrail-data-and-splunk-software-to-improve-security-and-compliance-in-aws

[16] (2016,). Add a CloudTrail input for the Splunk Add-on for AWS. Retrieved January 4, 2017, from http://docs.splunk.com/Documentation/AddOns/released/AWS/CloudTrail

[17] (2015, September). AWS Cloudtrail Splunk. Retrieved January 4, 2017, from https://github.com/xueshanf/aws-cloudtrail-with-splunk

[18] (2016, June). AWS Cloud Adoption Framework. Retrieved January 4, 2017, from https://d0.awsstatic.com/whitepapers/AWS_CAF_Security_Perspective.pdf

[19] (2016,). Cloud Security Resources. Retrieved January 4, 2017, from https://aws.amazon.com/security/security-resources/

[20] (2015, May 15). How to Receive Alerts When Specific APIs Are Called by Using AWS CloudTrail, Amazon SNS, and AWS Lambda. Retrieved January 4, 2017, from https://aws.amazon.com/blogs/security/how-to-receive-alerts-when-specific-apis-are-called-by-using-aws-cloudtrail-amazon-sns-and-aws-lambda/

[21] (2015, February 5). How to Receive Alerts When Your IAM Configuration Changes. Retrieved January 4, 2017, from https://aws.amazon.com/blogs/security/how-to-receive-alerts-when-your-iam-configuration-changes/

[22] (2016,). AWS Security Monitoring & Compliance Validation from Adobe. Retrieved January 4, 2017, from https://

conf.splunk.com/files/2016/slides/you-cant-protect-what-you-cant-see-aws-security-monitoring-and-compliance-validation-from-adobe.pdf

[23] (2016,). Splunk - Cloud Is a Journey. Make Splunk Your Partner. Retrieved January 4, 2017, from http://www.slideshare.net/AmazonWebServices/partner-solutions-splunk-cloud-is-a-journey-make-splunk-your-partner

[24] (2015,). AWS July Webinar Series - Troubleshooting Operational and Security Issues in Your AWS Account using CloudTrail. Retrieved January 4, 2017, from http://www.slideshare.net/AmazonWebServices/july-webinar-series-troubleshooting-operational-and-security-issues-in-your-aws-account-using-cloud-trail-20150729

[25] (2015). (SEC318) AWS CloudTrail Deep Dive. Retrieved January 4, 2017, from http://www.slideshare.net/AmazonWebServices/sec318-aws-cloudtrail-deep-dive

[26] (2015). (SEC308) Wrangling Security Events in The Cloud. Retrieved January 4, 2017, from http://www.slideshare.net/AmazonWebServices/sec308-wrangling-security-events-in-the-cloud

[27] A. (2015). Awslabs/timely-security-analytics. Retrieved January 01, 2017, from https://github.com/awslabs/timely-security-analytics

[28] Marko, K. (2015) AWS security management: In need of automation. Available at: http://markoinsights.com/2015/12/27/aws-security-mgmt/ (Accessed: 7 January 2017).

[29] Cassidy, S. (2016) Solutions. Available at: https://www.defensestorm.com/cybermind/security-logging-on-aws/ (Accessed: 7 January 2017).

[30] Chan, J. (2010) Announcing security monkey - AWS security configuration monitoring and analysis. Available at: http://techblog.netflix.com/2014/06/announcing-security-monkey-aws-security.html (Accessed: 7 January 2017).

[31] Creating CloudWatch alarms for CloudTrail events: Examples (2017) Available at http://docs.aws.amazon.com/awscloudtrail/latest/userguide/cloudwatch-alarms-for-cloudtrail.html (Accessed: 7 January 2017).

[32] Creating and updating your cloudtrail. (2016). Retrieved January 7, 2017, from http://docs.aws.amazon.com/awscloudtrail/latest/userguide/cloudtrail-create-and-update-a-trail.html

[33] (2015, August). CloudWatch Logs Subscription Consumer + Elasticsearch + Kibana Dashboards. Retrieved January 4, 2017, from https://aws.amazon.com/blogs/aws/cloudwatch-logs-subscription-consumer-elasticsearch-kibana-dashboards/

[34] (2016,). Tutorial: Using Amazon ML to Predict Responses to a Marketing Offer. Retrieved January 4, 2017, from http://docs.aws.amazon.com/machine-learning/latest/dg/tutorial.html

[35] (2016, October). Building Event-Driven Batch Analytics on AWS. Retrieved January 4, 2017, from https://aws.amazon.com/blogs/big-data/building-event-driven-batch-analytics-on-aws/

[36] Now available: Videos from re Invent 2016 security and compliance sessions (2016) Available at https://aws.amazon.com/blogs/security/now-available-videos-and-slide-decks-from-reinvent-2016-security-and-compliance-sessions/ (Accessed: 8 January 2017).

[37] (2016) Available at: http://conf.splunk.com/sessions/2016-sessions.html (Accessed: 8 January 2017).

[38] ML-SPL quick reference guide Preprocessing StandardScaler (no date) Available at https://docs.splunk.com/images/e/ee/MLTKCheatSheet.pdf (Accessed: 8 January 2017).

[39] (2014) AWS CloudTrail user guide. Available at: http://awsdocs.s3.amazonaws.com/awscloudtrail/latest/awscloudtrail-ug.pdf (Accessed: 8 January 2017).

[40] How to Receive Alerts When Specific APIs Are Called by Using AWS CloudTrail, Amazon SNS, and AWS Lambda. (2015). Retrieved January 01, 2017, from https://aws.amazon.com/blogs/security/how-to-receive-alerts-when-specific-apis-are-called-by-using-aws-cloudtrail-amazon-sns-and-aws-lambda/

[41] How to Receive Alerts When Your IAM Configuration Changes. (2015). Retrieved January 01, 2017, from https://aws.amazon.com/blogs/security/how-to-receive-alerts-when-your-iam-configuration-changes/

[42] Machine learning algorithm cheat sheet for Microsoft Azure Machine Learning Studio. (n.d.). Retrieved July 28, 2016, from https://azure.microsoft.com/en-us/documentation/articles/machine-learning-algorithm-cheat-sheet/

[43] Binary Classification: Network intrusion detection. (n.d.). Retrieved July 28, 2016, from https://gallery.cortanaintelligence.com/Experiment/Binary-Classification-Network-intrusion-detection-2?share=1

[44] MLlib: Scalable Machine Learning on Spark. (n.d.). Retrieved July 28, 2016, from http://stanford.edu/~rezab/spark-workshop/slides/xiangrui.pdf

[45] Balakrishnan, B. (2016) Applying machine learning techniques to measure critical security controls. Available at: https://www.sans.org/reading-room/whitepapers/critical/applying-machine-learning-techniques-measure-critical-security-controls-37247 (Accessed: 19 January 2017).

[46] Balakrishnan, B. (2015) Insider threat mitigation guidance. Available at: https://www.sans.org/reading-room/whitepapers/monitoring/insider-threat-mitigation-guidance-36307 (Accessed: 19 January 2017).

[47] Blum, D. (2016) Discovering agile cloud security | security architects partners. Available at: http://security-architect.com/discovering-agile-cloud-security/ (Accessed: 19 January 2017).

[48] Introduction to AWS CodePipeline (2017) Available at: https://aws.amazon.com/devops/continuous-delivery/ (Accessed: 22 January 2017).

[49] What is DevOps? - Amazon web services (AWS) (2017) Available at https://aws.amazon.com/devops/what-is-devops/ (Accessed: 22 January 2017).

[50] AWS Identity and Access Management (2016) Available at https://aws.amazon.com/iam/ (Accessed: 22 January 2017).

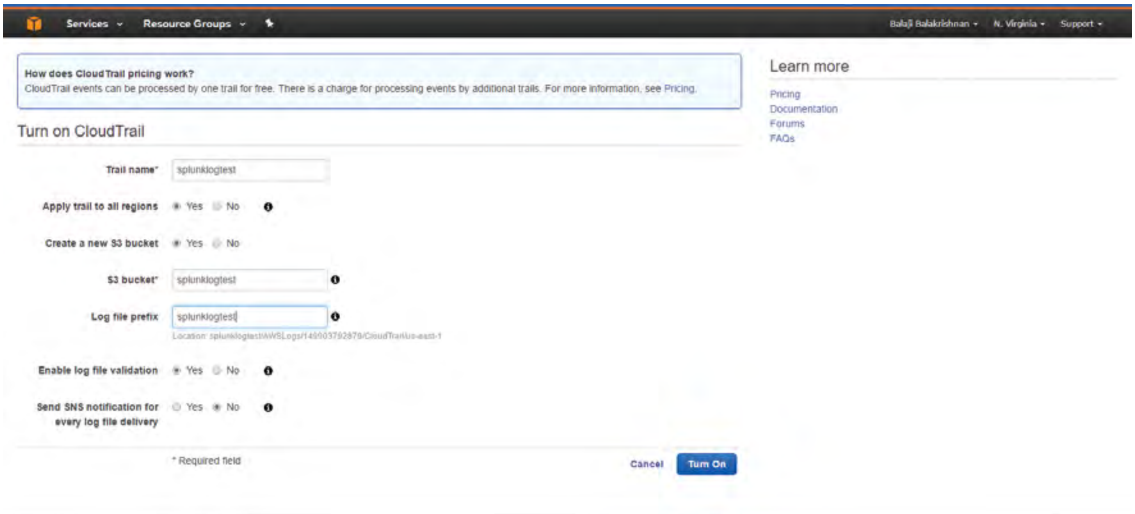
[51] Brownlee, J. (2016) Metrics to evaluate machine learning Algorithms in python. Available at: http://machinelearningmastery.com/metrics-evaluate-machine-learning-algorithms-python/ (Accessed: 12 February 2017).

ABOUT THE AUTHOR

BALAJI BALAKRISHNAN has more than 18 years' experience in IT and Information security domain specializing in security operations and incident response. He has worked in major financial services organizations and has led 24/7 SOCs/incident response teams.

APPENDIX A - LAB SETUP

- Build a test AWS environment
- Create Free Tier AWS account
- Enable Cloudtrail



How does CloudTrail pricing work?
CloudTrail events can be processed by one trail for free. There is a charge for processing events by additional trails. For more information, see Pricing.

Turn on CloudTrail

Trail name* splunklogtest

Apply trail to all regions ● Yes ○ No

Create a new S3 bucket ● Yes ○ No

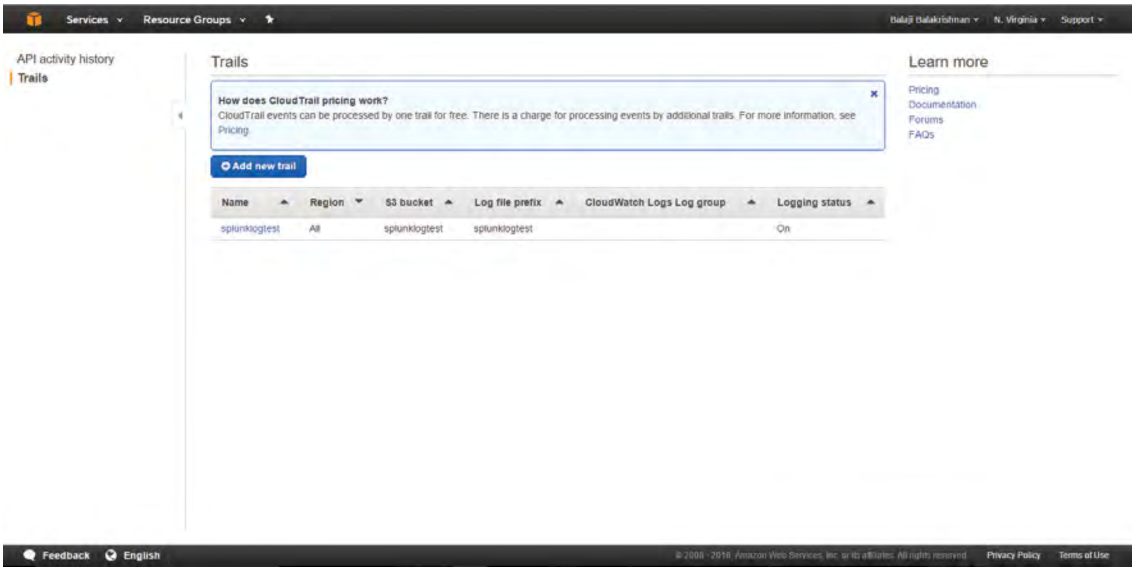
S3 bucket* splunklogtest

Log file prefix splunklogtest
Location: splunklogtest/AWSLogs/149903792879/CloudTrail/us-east-1

Enable log file validation ● Yes ○ No

Send SNS notification for every log file delivery ○ Yes ● No

* Required field

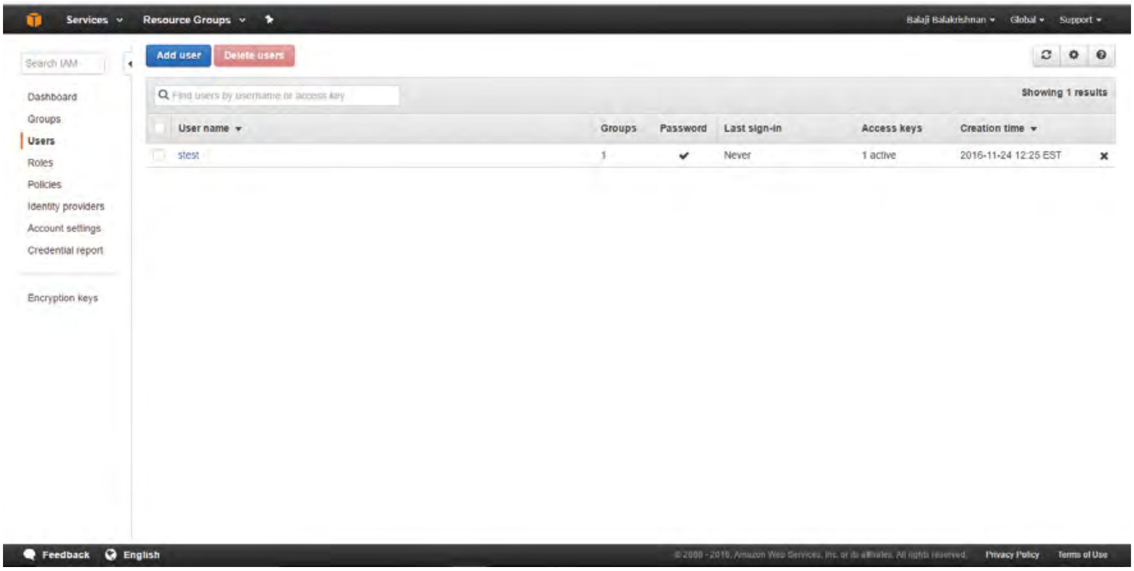


Trails

How does CloudTrail pricing work?
CloudTrail events can be processed by one trail for free. There is a charge for processing events by additional trails. For more information, see Pricing.

Name	Region	S3 bucket	Log file prefix	CloudWatch Logs Log group	Logging status
splunklogtest	All	splunklogtest	splunklogtest		On

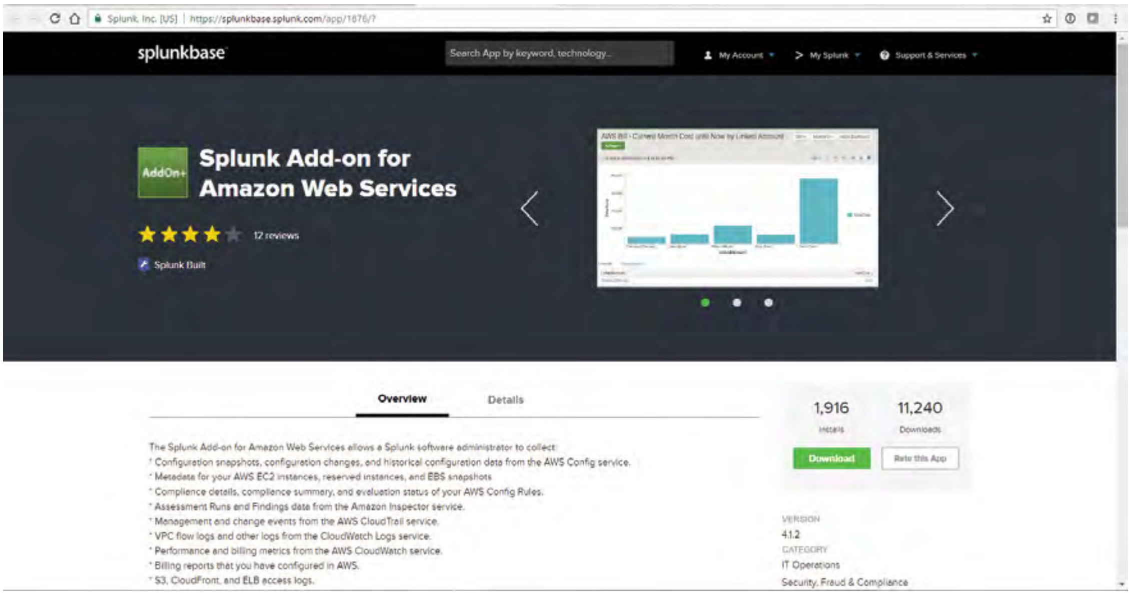
- Create accounts using IAM



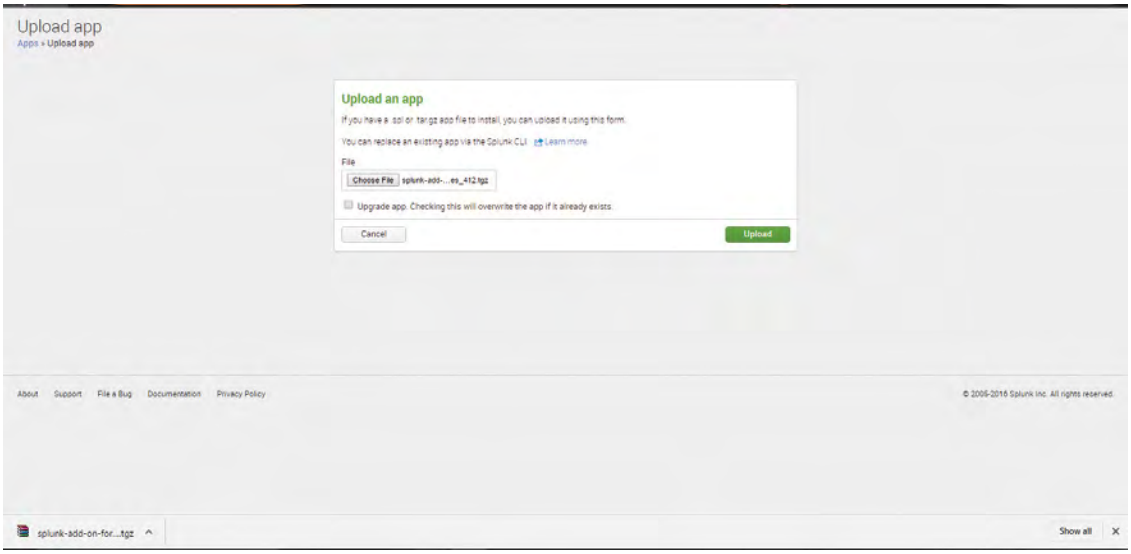
Showing 1 results

User name	Groups	Password	Last sign-in	Access keys	Creation time
stest	1	✓	Never	1 active	2016-11-24 12:25 EST

- Build a Splunk test environment –
  - Getting AWS Cloudtrail logs to Splunk
  - Install Splunk Enterprise 6.5
  - Download and Install Splunk Add-on for Amazon Web Services



The Splunk Add-on for Amazon Web Services allows a Splunk software administrator to collect:
* Configuration snapshots, configuration changes, and historical configuration data from the AWS Config service.
* Metadata for your AWS EC2 instances, reserved instances, and EBS snapshots
* Compliance details, compliance summary, and evaluation status of your AWS Config Rules.
* Assessment Runs and Findings data from the Amazon Inspector service.
* Management and change events from the AWS CloudTrail service.
* VPC flow logs and other logs from the CloudWatch Logs service.
* Performance and billing metrics from the AWS CloudWatch service.
* Billing reports that you have configured in AWS.
* S3, CloudFront, and ELB access logs.



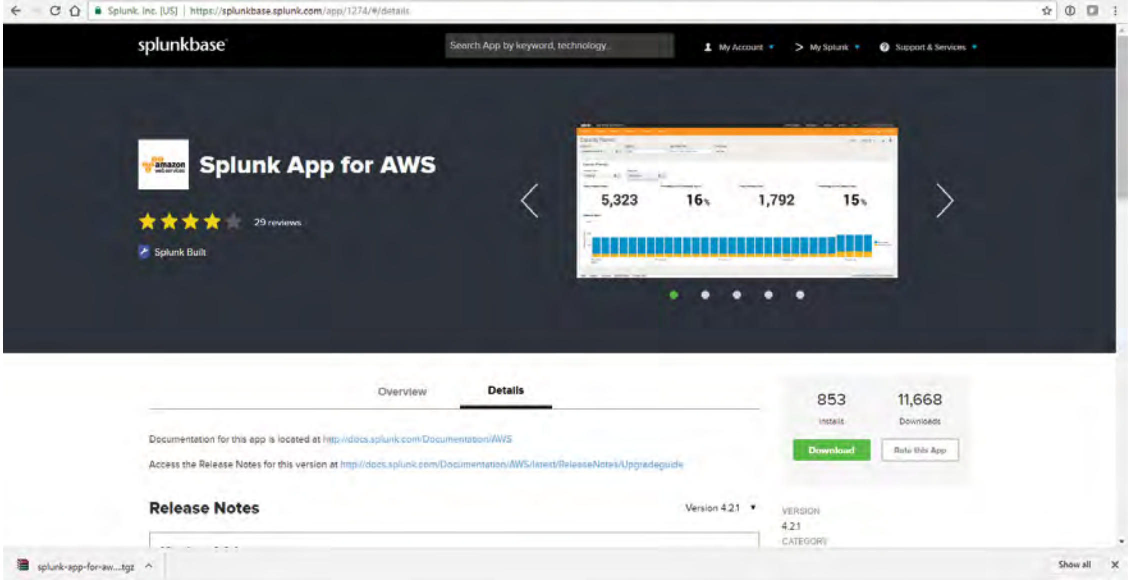
If you have a .spl or .tar.gz app file to install, you can upload it using this form.

You can replace an existing app via the Splunk CLI.

File: splunk-add-...es_412.tgz

□ Upgrade app. Checking this will overwrite the app if it already exists.

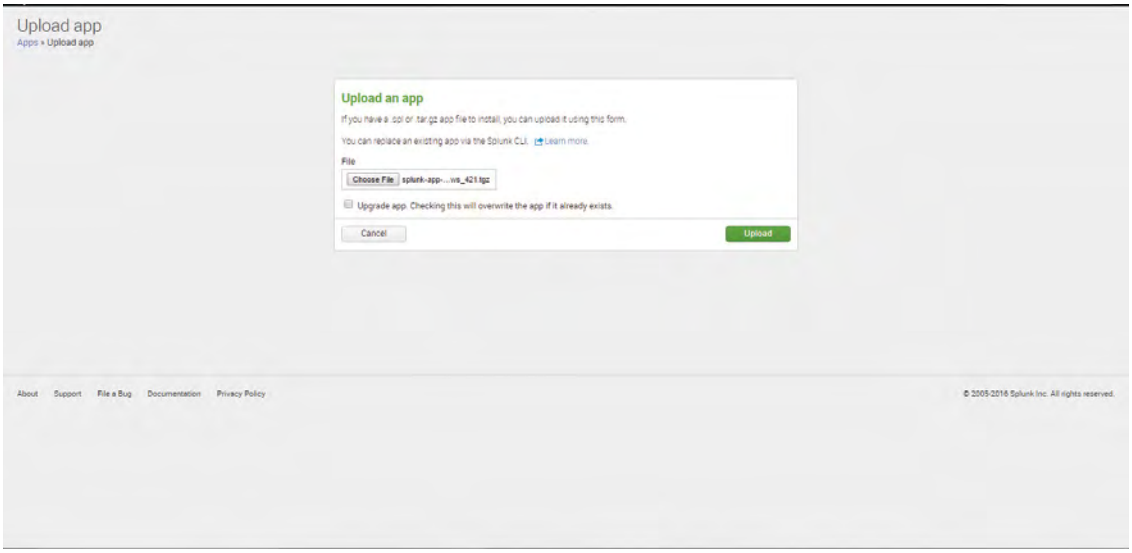
- Download and Install Splunk App for AWS



Documentation for this app is located at http://docs.splunk.com/Documentation/AWS

Access the Release Notes for this version at http://docs.splunk.com/Documentation/AWS/latest/ReleaseNotes/Upgradeguide

- Install the app and restart Splunk



Upload an app

If you have a .spl or .tar.gz app file to install, you can upload it using this form.

You can replace an existing app via the Splunk CLI. Learn more.

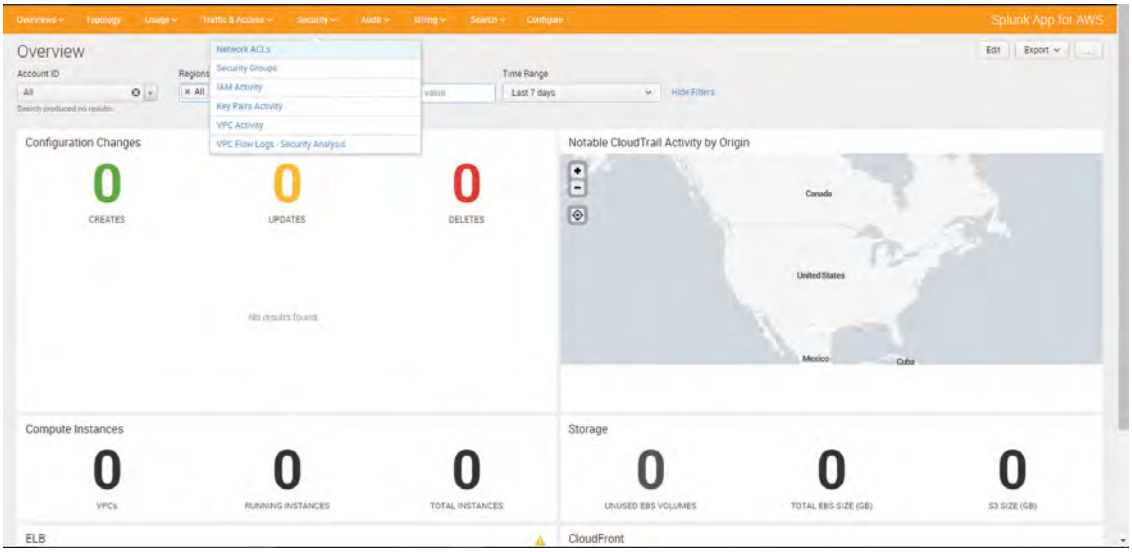
File: Choose File splunk-app-...ws_421.tgz

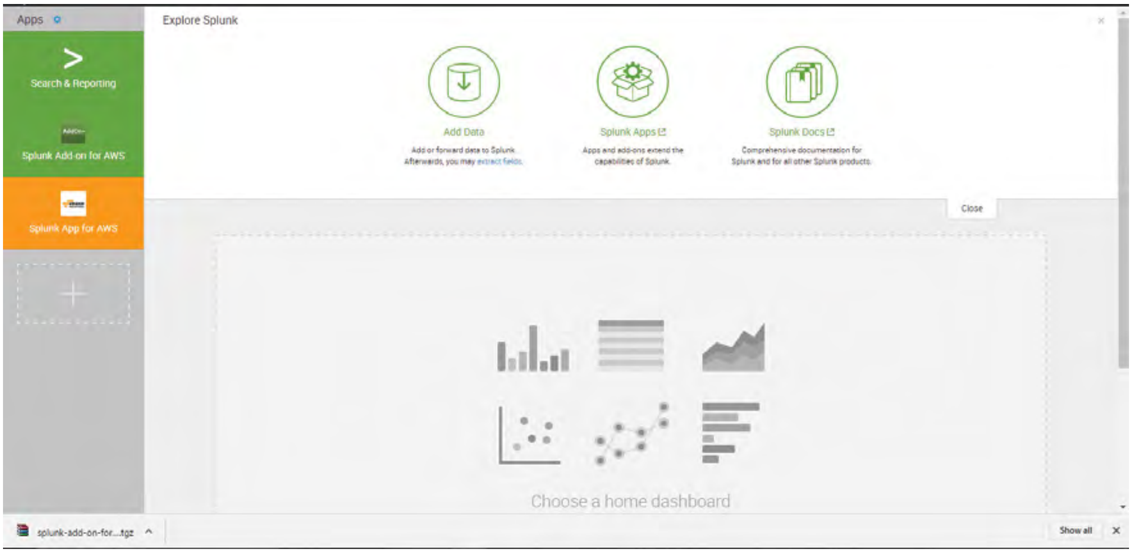
☐ Upgrade app. Checking this will overwrite the app if it already exists.

Cancel | Upload

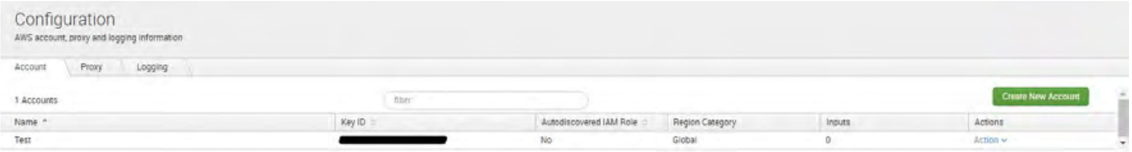
- Splunk App for AWS

The Splunk App for AWS gives you significant operational and security insight into your Amazon Web Services account and infrastructure.




Apps
Search & Reporting
Splunk Add-on for AWS
Splunk App for AWS
Explore Splunk
Add Data
Add or forward data to Splunk. Afterwards, you may extract fields.
Splunk Apps
Apps and add-ons extend the capabilities of Splunk.
Splunk Docs
Comprehensive documentation for Splunk and for all other Splunk products.
Close
Choose a home dashboard
splunk-add-on-for...tgz
Show all

- Configured Splunk Add-on with AWS credentials

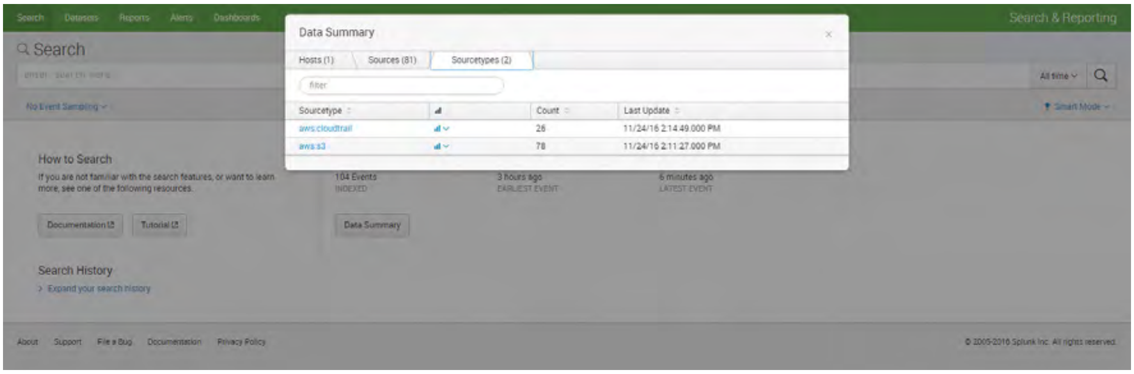


Configuration

AWS account, proxy and logging information

1 Accounts

Name	Key ID	Autodiscovered IAM Role	Region Category	Inputs	Actions
Test	[redacted]	No	Global	0	Action



Data Summary

Sourcetype		Count	Last Update
aws:cloudtrail		26	11/24/16 2:14:49.000 PM
aws:s3		78	11/24/16 2:11:27.000 PM

How to Search

If you are not familiar with the search features, or want to learn more, see one of the following resources.

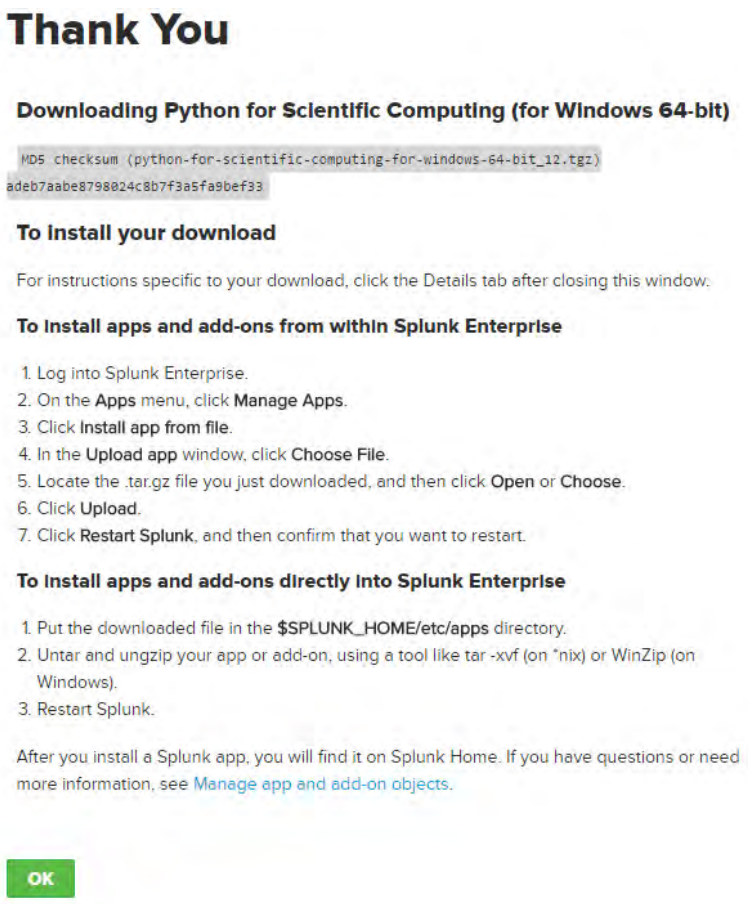
104 Events INDEXED

3 hours ago EARLIEST EVENT

6 minutes ago LATEST EVENT

© 2005-2016 Splunk Inc. All rights reserved.

- Installing Machine Learning Toolkit
  - Download and Install Python for Scientific Computing



Thank You

Downloading Python for Scientific Computing (for Windows 64-bit)

```
MD5 checksum (python-for-scientific-computing-for-windows-64-bit_12.tgz)
adeb7aabe8798024c8b7f3a5fa9bef33
```

To install your download

For instructions specific to your download, click the Details tab after closing this window.

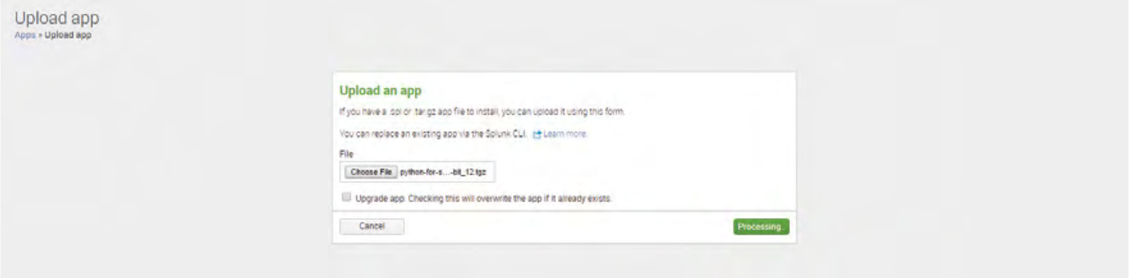
To install apps and add-ons from within Splunk Enterprise

1. Log into Splunk Enterprise.
2. On the **Apps** menu, click **Manage Apps**.
3. Click **Install app from file**.
4. In the **Upload app** window, click **Choose File**.
5. Locate the .tar.gz file you just downloaded, and then click **Open** or **Choose**.
6. Click **Upload**.
7. Click **Restart Splunk**, and then confirm that you want to restart.

To install apps and add-ons directly into Splunk Enterprise

1. Put the downloaded file in the **$SPLUNK_HOME/etc/apps** directory.
2. Untar and ungzip your app or add-on, using a tool like tar -xvf (on *nix) or WinZip (on Windows).
3. Restart Splunk.

After you install a Splunk app, you will find it on Splunk Home. If you have questions or need more information, see Manage app and add-on objects.



Upload app

Upload an app

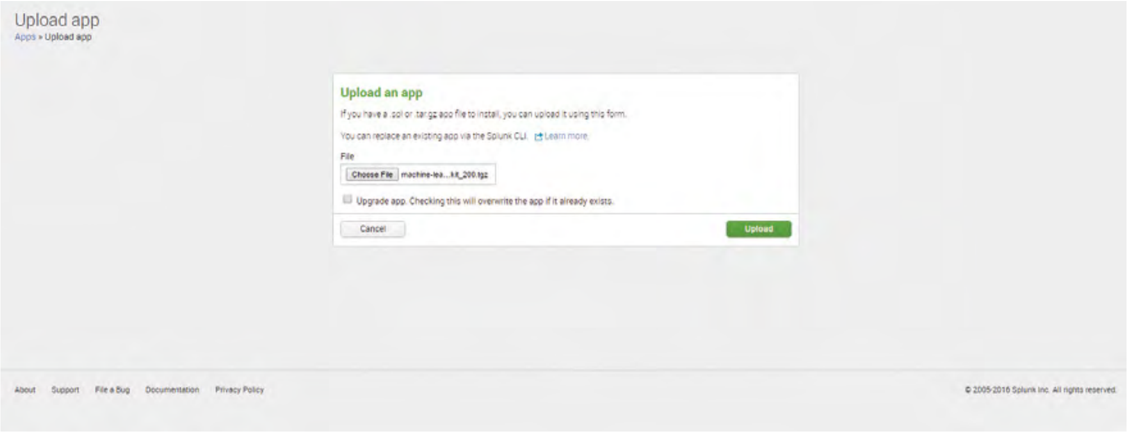
If you have a .spl or .tar.gz app file to install, you can upload it using this form.

You can replace an existing app via the Splunk CLI. Learn more.

File: python-for-s...-bit_12.tgz

□ Upgrade app. Checking this will overwrite the app if it already exists.

- Download and Install Machine Learning Toolkit



Upload an app

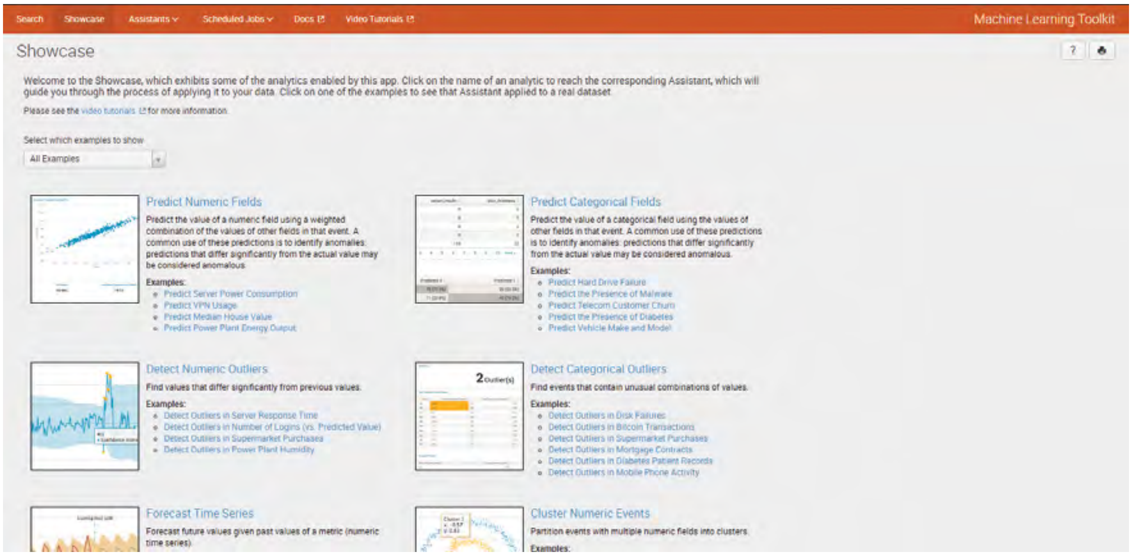
If you have a .spl or .tar.gz app file to install, you can upload it using this form.

You can replace an existing app via the Splunk CLI. Learn more.

File: Choose File machine-lea....kit_200.tgz

☐ Upgrade app. Checking this will overwrite the app if it already exists.

Cancel | Upload



Showcase

Welcome to the Showcase, which exhibits some of the analytics enabled by this app. Click on the name of an analytic to reach the corresponding Assistant, which will guide you through the process of applying it to your data. Click on one of the examples to see that Assistant applied to a real dataset.

Please see the video tutorials for more information.

Select which examples to show: All Examples

Predict Numeric Fields

Predict the value of a numeric field using a weighted combination of the values of other fields in that event. A common use of these predictions is to identify anomalies: predictions that differ significantly from the actual value may be considered anomalous.

Examples:
- Predict Server Power Consumption
- Predict VPN Usage
- Predict Median House Value
- Predict Power Plant Energy Output

Predict Categorical Fields

Predict the value of a categorical field using the values of other fields in that event. A common use of these predictions is to identify anomalies: predictions that differ significantly from the actual value may be considered anomalous.

Examples:
- Predict Hard Drive Failure
- Predict the Presence of Malware
- Predict Telecom Customer Churn
- Predict the Presence of Diabetes
- Predict Vehicle Make and Model

Detect Numeric Outliers

Find values that differ significantly from previous values.

Examples:
- Detect Outliers in Server Response Time
- Detect Outliers in Number of Logins (vs. Predicted Value)
- Detect Outliers in Supermarket Purchases
- Detect Outliers in Power Plant Humidity

Detect Categorical Outliers

Find events that contain unusual combinations of values.

Examples:
- Detect Outliers in Disk Failures
- Detect Outliers in Bitcoin Transactions
- Detect Outliers in Supermarket Purchases
- Detect Outliers in Mortgage Contracts
- Detect Outliers in Diabetes Patient Records
- Detect Outliers in Mobile Phone Activity

Forecast Time Series

Forecast future values given past values of a metric (numeric time series).

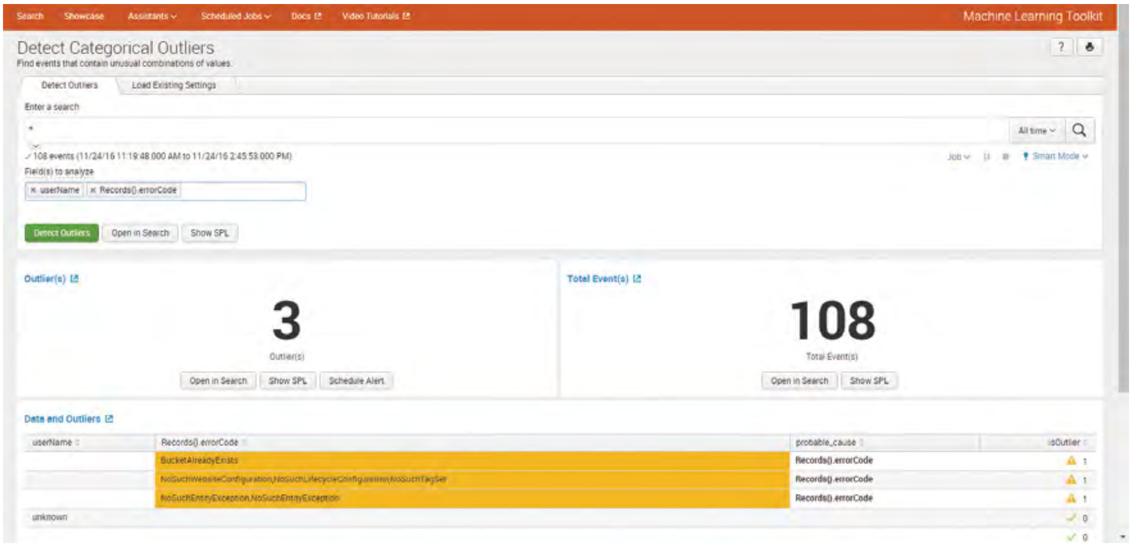
Cluster Numeric Events

Partition events with multiple numeric fields into clusters.

Examples:

- This step completes the lab setup.

Unsupervised Learning Example – Detecting Categorical Outliers



Detect Categorical Outliers

Find events that contain unusual combinations of values.

Enter a search: *

✓ 108 events (11/24/16 11:19:48.000 AM to 11/24/16 2:45:53.000 PM)

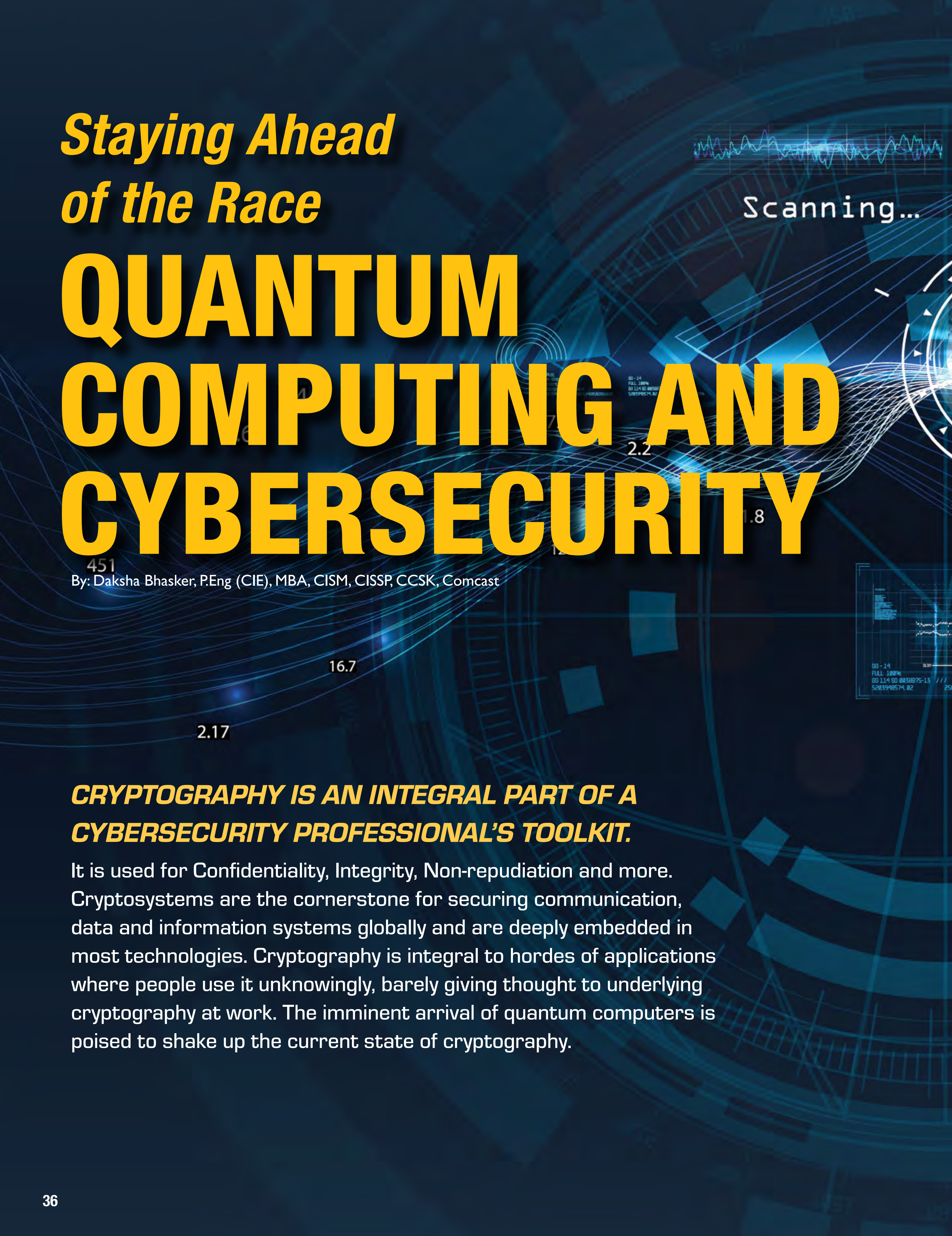
Field(s) to analyze: × userName × Records{}.errorCode

Outlier(s): 3 Outlier(s)

Total Event(s): 108 Total Event(s)

Data and Outliers

userName	Records{}.errorCode	probable_cause	isOutlier
	BucketAlreadyExists	Records{}.errorCode	⚠ 1
	NoSuchWebsiteConfiguration,NoSuchLifecycleConfiguration,NoSuchTagSet	Records{}.errorCode	⚠ 1
	NoSuchEntityException,NoSuchEntityException	Records{}.errorCode	⚠ 1
unknown			✓ 0
			✓ 0



Staying Ahead of the Race

QUANTUM COMPUTING AND CYBERSECURITY

By: Daksha Bhasker, P.Eng (CIE), MBA, CISM, CISSP, CCSK, Comcast

CRYPTOGRAPHY IS AN INTEGRAL PART OF A CYBERSECURITY PROFESSIONAL'S TOOLKIT.

It is used for Confidentiality, Integrity, Non-repudiation and more. Cryptosystems are the cornerstone for securing communication, data and information systems globally and are deeply embedded in most technologies. Cryptography is integral to hordes of applications where people use it unknowingly, barely giving thought to underlying cryptography at work. The imminent arrival of quantum computers is poised to shake up the current state of cryptography.

Quantum computing
28
Lorem ipsum dolor sit amet consectetur adipiscing elit Sed sagittis lorem ac enim viverra in fermentum urna pellentesque. Donec gravida luctus sapien quis fringilla tortor porta ac. Morbi tincidunt nunc sit amet sollicitudin consequat turpis risus scelerisque lectus ut consectetur eros nulla tincidunt felis.
78.1
17.1
16.1
18.2
6.6
23.9
17840
0
0.00XX
GK-100%
000000000000000.02341
7593-3923-2305-1103-4051
GLK-OFF ///
9.5
13.5
13
57%
90%
ARTIFICIAL INTELLIGENCE
75

Every emerging technology has two faces, one that revs up the potential for technical evolution propelling humanity into the future, the other a harbinger of novel security vulnerabilities and attacks inconceivable before. Quantum computing is no different in this regard and has breathed life into theoretical mathematics from the late nineties, making them potentially the most powerful codebreaking tools ever developed. In Quantum computing, a quantum bit (qubit) can hold multiple values simultaneously whereas a classic bit holds a single binary value of 0 or 1, theoretically enabling a single qubit to take part in multiple computations simultaneously. This makes quantum computers extremely adept at solving certain types of problems, that classical computers cannot. Quantum computers with the power to solve mathematics underlying classical cryptosystems are on their way.

A race for quantum supremacy between various players, industry, government and academia, to build the most capable quantum computer is underway. Physicists, computer scientists, engineers, mathematicians and geeks have a head start, working on quantum computing for well over two decades. In recent years it has become increasingly clear that cybersecurity professionals have an important role to play in addressing potential vulnerabilities in the development and implementation of this incredible technology.

The principle underpinning the effectiveness of cryptography is that the work effort, resources and time needed for cryptoanalysis is either infeasible with the technology available at the present time, or the time taken to defeat the applied cryptography is significantly larger than the meaningful useful lifespan of the encrypted information. This principle has been defeated before. The Turing Bombe in World War II defeated the German Enigma machines' encryption scheme. 40-bit encryption was common in software released before 1999, especially those based on RC2 and RC4 algorithms. The 40-bit key cipher system approved for use in the 1990s was defeated by the end of the 20th century, when a single PC could search a 2^{40} key space in a matter of hours, ushering in 128-bit encryption [1]. And thus, through the numerous ciphers and algorithms that have been deprecated and retired, over the years, new ones instated with larger key sizes and progressively harder mathematical problems underlying them. Today, we stand yet again, at a similar threshold of cryptographic evolution for popularly used crypto-schemes with the anticipated prowess of quantum computing capable of rendering them no longer secure.

Quantum computers are astonishing contraptions that stabilize sub-atomic particles called qubits that are fragile and can lose their data if disturbed. In quantum computing a qubit is the basic unit of quantum information and operations are conducted by manipulating its quantum mechanical properties. Quantum computers are maintained at zero degrees kelvin and isolated from disturbances through noise, temperature change, electrical fluctuations or vibrations. Therefore, they are neither small nor portable at this time.

In 1996 Grover's algorithm, proved that quantum computers could implement search functions in $O\sqrt{N}$ time, where N is the size of the function's domain [2]. This essentially means that a symmetric key can be brute forced in square-root of the time it previously could. It is estimated that a quantum computer with 2953 qubits is able to brute force AES128 [3]. The proposed quick fix to build resistance against quantum attacks targeting symmetric cryptography is to double the key length. This mitigates speed up by square-root time from quantum attacks. This means that to have the same level of protection as AES256, a standard commonly used today, AES512 would need to be deployed. Keep in mind that where secret keys are not pre-shared securely, key exchange of symmetric key cryptography often leverages asymmetric public key cryptography.

In 1997 Peter Shor published a quantum algorithm that performs prime factorization of integers and solves discrete logarithm problems in polynomial time. Cryptographic algorithms such as RSA, ECC and Diffie-Hellman that depend on the inability of classical computers to complete such calculations are now broken by quantum computers. In fact, in 2001, a group in IBM demonstrated Shor's algorithm on a 7-qubit factorizing the prime number 15 using nuclear magnetic resonance techniques to manipulate the qubits [4]. It is estimated that 2048-bit RSA requires 4096 qubits and 224 ECC requires 1300 to 1600 qubit quantum computers to break respectively [5]. Essentially both the technology and mathematics to break cryptography as we know it has been proven. The only missing element is the number of qubits in universal quantum computers available for operations today.

This brings us back to the racetracks where the community is watching with bated breath, the exciting race for the proclaimed arrival of the "Supreme" universal quantum computer. There are numerous companies such as IBM, Intel and Google [6], building and stabilizing the quirky qubits at zero degrees kelvin and error correction algorithms for consistent performance quantum computer. Rigetti has publicly announced a 128 Qubit Universal quantum computer for deployment in 2019 [7]. In a mere two years, the technology has developed from 3 qubit computers to 72 qubit computers (figure -1). D-wave and Fujitsu have 2048-qubit and 1024-qubit annealing quantum computers respectively [8], however they are special purpose computers intended to solve specialized problems [9]. Notably, quantum computers are not to be mistaken for faster computers, in fact, they are good at solving a completely different set of problems that classical computers are not good at, such as, optimization (e.g the travelling salesman problem), machine learning, biomedical simulations and financial services among others [10].

In response, the NIST 2016 report on post-quantum cryptography published the following (Table-1), impact of quantum computing on common cryptographic algorithms [11].

The cryptosystems listed in Table-1 are pervasive in our environment today: and the post quantum crypto-stack includes: web browsers, certificates, Tor, Signal, imessaging, software updates, mobile phones connecting to cell towers, credit cards transactions, secure boot, code signing, secure password hashing, checksums, encrypted disks, file

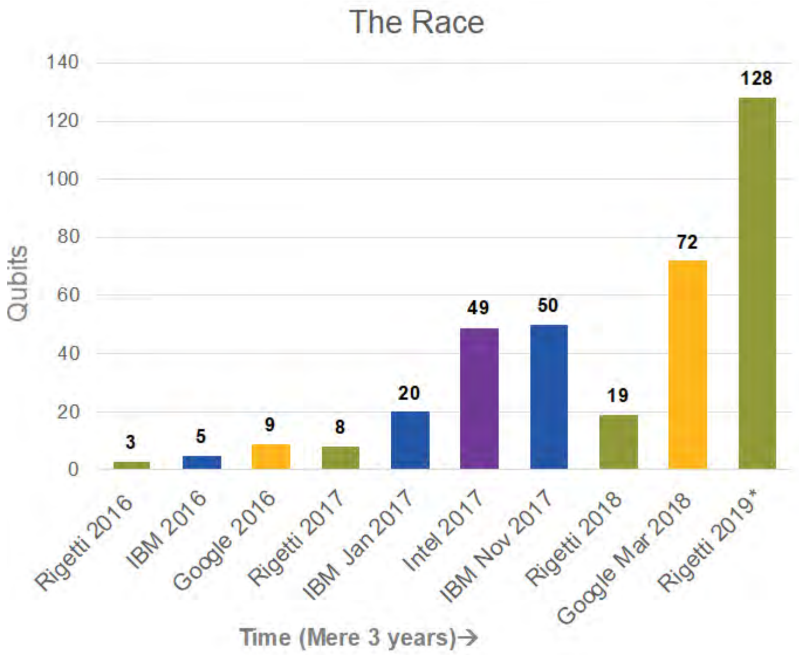


Figure 1

systems, databases, digital signatures, key exchange, protocols as TLS, IPSec, SSH S/MIME, DNSSec, and lower level modules as gnu multi-precision libraries (GMP), Number theory libraries (NTL), AES block ciphers, hash functions, random number generators, HSMS, TPMs and much more [12] [13].

Table 1 [11]

Cryptographic Algorithm	Type	Purpose	Impact from large scale quantum computer
AES	Symmetric Key	Encryption	Larger key sizes needed
SHA-2, SHA-3	-------------	Hash Functions	Larger output needed
RSA	Public Key	Signatures, Key establishment	No longer secure
ECDSA, ECDH (Elliptical Curve Cryptography)	Public Key	Signatures, Key Exchange	No longer secure
DSA (Finite Field Cryptography)	Public Key	Signatures, Key Exchange	No longer secure

NIST is currently working on analyzing, rigorously testing various candidates for post-quantum algorithms and expects to release draft standards between 2022 and 2024. NIST is concerned with the long-term viability and robustness of an algorithm, therefore its diligence is very thorough, and will among other parameters, evaluate an algorithms' efficiency, confidence against cryptoanalysis, usability and interoperability before releasing standards [14].

As of 2016 the National Security Agency (NSA), in anticipation of the transition to quantum resistant algorithm has retired for use by organizations that run classified and unclassified national security systems (NSS) and vendors that build products used in the NSS the following: ECDH and ECDSA with NIST P-256, SHA-256, AES-128, RSA with 2048-bit keys, Diffie-Hellman with 2048-bit keys [15]. The NSA announcement of the adoption of the commercial national security algorithm suite (CNSA) has been issued with the intent to enable NSS developers and operators to plan, budget, design and build new cryptography into their systems [15]. The NSS has further advised that the following set of public cryptographic standards (Table-2) be used to protect the NSS until acceptable standards for quantum resistant cryptography become available and are approved for use:

This is an interim measure and in the longer term the NSA expects all its systems and suppliers to use standardized quantum resistant algorithms [15].

In the meantime, as of 2016, NIST has proposed a "hybrid mode" [12] and stated that "a focus on maintaining crypto-agility is imperative" [16]. A hybrid mode is a transition or migration step to post-quantum cryptography where such a mode combines a classical algorithm with a post-quantum one. Cryptographic agility or crypto-agility is the ability to easily make changes to cryptographic algorithms and protocols used in a system without having to rebuild the system [17].

Table 2 [15]

Algorithm	Usage
RSA 3072-bit or larger	Key Establishment, Digital Signature
Diffie-Hellman (DH) 3072-bit or larger	Key Establishment
ECDH with NIST P-384	Key Establishment
ECDSA with NIST P-384	Digital Signature
SHA-384	Integrity
AES-256	Confidentiality

Based on this information, there are several steps the community should take to further secure itself. In the past 10 years, our industry has been through several significant cryptographic updates: SHA-1, MD-5, RSA-1024 [18]. Some of these updates took as many as 10 years, and several retired protocols continue to lurk around in disparate software, systems and infrastructure. If history is any indicator of how effectively (or ineffectively) we handle change of cryptographic protocols, we can anticipate a challenge in updating fast enough to meet the quantum era.

Some experts speculate that there is close race between NIST's post-quantum cryptographic standards expected in 2022 – 2024 and the development of a universal quantum computer that can break classical cryptography forecasted to arrive between 2023 – 2033. In fact, considering change management windows for cryptography and deeply embedded cryptosystems in our technical ecosystems, some analysts suggest that the race to secure our data and infrastructure is a challenging one. It has been reported, that hackers and intelligence agencies are actively harvesting encrypted data today, with the intent to decrypt in the future once a capable quantum computer arrives [19]. Information assets (such as customer contracts, intellectual property etc.), whose lifespan exceeds 5 to 10 years as of today, are at potential risk of exposure by advances in quantum computing, if not adequately protected.

Regardless of the exact time of arrival of the crypto-defeating quantum computer, our entire cryptographic algorithm ecosystem is inevitably coming up for one or several cryptographic updates imminently, only with shorter time windows in which to execute these changes than ever before in the past. In the time preceding standards announcement by the NIST on quantum-resistant cryptography, we have opportunities to proactively prepare ourselves in numerous ways.

Risk Assessment: Security is essentially about managing risks aligned with the business objectives of an organization through various security controls, where cryptography is one control, albeit a critical one. Risk assessment is the first step to understanding risks to the business. There are numerous risk assessment frameworks available. Select one that is suitable for your organization and industry for quantum risk assessment. The Global Risk Institute describes a six-phase quantum risk framework to assess, evaluate, implement and integrate with the organizations' cyber risk assessment and management framework. [20]

- Phase 1 – Identify and document information assets, and their current cryptographic protection
- Phase 2 – Research the state of emerging quantum computers and quantum-safe cryptography.
- Phase 3 – Identify threat actors, and estimate their time to access technology.
- Phase 4 – Identify the lifetime of your assets, and the time required to transform the organization's technical infrastructure to a quantum-safe state
- Phase 5 – Determine quantum risk by calculating whether business assets will become vulnerable before the organization can move to protect them
- Phase 6 – Identify and prioritize the activities required to maintain awareness and to migrate the organization's technology to a quantum-safe state.

> *"Risk assessment is the first step to understanding risks to the business."*

Hybrid schemes/Crypto-experimentation: There are numerous quantum safe algorithms available today that have never been broken, such as NTRU [21], NewHope, McEliece [22] among many others. Google has an experimental web browser Canary that uses the NewHope algorithm [22]. Cryptographers caution that post-quantum algorithms are under study. Cryptographers and cryptanalysts are learning about them, and they should not be positioned as post-quantum secure as yet [23]. However, experimenting with PQC, alongside classical cryptography, offers invaluable insight into the effort, implementation, interoperability and operational aspects of the new mathematics underlying the impending post-quantum cryptographic (PQC) suites.

Crypto-Agility: Crypto-agility or cryptographic agility, is the capacity for an information security system to adopt an alternative to the original encryption method or cryptographic primitive without significant change to system infrastructure [24]. Besides the looming quantum threat there are myriad of drivers for crypto-agility such as emergence of new cryptographic standards, new legislation such as GDPR, discovery of a new vulnerability in a cryptographic function that requires an update fix. Business needs where old legacy devices with weak cryptography needs to be managed through a prolonged retirement lifecycle. Exposure is significantly reduced if crypto-agility design principles are in place. Integrate crypto-agility as a core premise of the cryptography lifecycle management.

Crypto-agile abstraction essentially decouples the implementation of cryptographic suites from the application, solution or infrastructure itself. The algorithms are not hardcoded and systems are able to change cryptography dynamically or with a simple update. Usage of cryptography suites must be manageable separately from the solution and the cryptography must be portable across multiple devices without having to rebuild new systems [25]. Cryptographic change is constant and the slower an entity is to respond to adopt a cryptography update, the greater the vulnerability to a potential cyberattack.

The new attribute, crypto-agility has been added to the suite of elements under the umbrella of **cryptographic life cycle management**. All components of a cryptographic system such as keys, certificates, algorithms, ciphers are implemented, optimized and collectively managed to the desired security objective of the asset it is protecting. For all systems, solutions, software, firmware and infrastructure there needs to be an approach that assumes that the cryptographic algorithms will change during the course of its useful lifetime and have an inbuilt capability to seamlessly incorporate this cryptographic change or update. It makes sense to plan for crypto-agility as part of cryptographic life-cycle management and build products and solution with crypto-agile abstraction and architectures in mind to prepare for our collective quantum crypto-futures.

For new products, software, infrastructure and solutions being developed, ensure specifications that respect crypto-agile architectures and design principles. Use the latest cryptographic schemes approved and the most secure postures as publicized by the standards and regulatory bodies based on the industry in question. Consider scenarios that may require the solutions to be capable of doubling the key length, hash strings or similar crypto-applications via a simple configuration change. Architect products and infrastructure such that they are capable of running both classical and quantum secure algorithms in parallel, to support the previously mentioned "hybrid" approach to PQC.

Supply Chain: Integrate quantum safe and hybrid cryptographic criteria into your procurement pipeline. Ensure that the products, equipment and software that is purchased today is capable of supporting cryptographic changes in the near future. Include crypto-agility, crypto-agile abstraction, design principles and architectures as part of your requirement specifications to third party suppliers. Require your suppliers to architect solutions such that when NIST announces PQC standards their solutions can transition and support them without a rebuild or a repurchase. Include crypto-agility and PQC readiness criteria into your contracting language. Discuss the vendors' consideration of quantum-safe algorithms and evaluate maturity of the vendors' product roadmaps in this regard.

Data Life Cycle Management: For information assets ensure data classification policies are utilized effectively throughout the organization. If there are data assets that are known to be of critical value such as trade secrets, customer and employee PII, financial information and records, intellectual property and the like, that essentially has a life span into years where a quantum computer may exist, consider tagging them and applying quantum resistant cryptographic (or hybrid) discipline around its storage, transit and management throughout its lifecycle. Know where your crown jewels are, the current state of cryptographic protection applied and pre-emptively plan a strategy to manage the risks of the quantum cryptography era.

Inventory Applications that use Cryptography: Armed with the inventory of information assets that must be protected, knowing what applications, protocols, infrastructure interact with the data along with the cryptography in place will enable prepare for necessary algorithm swaps or changes in parameters that are integral to crypto-agility [26]. There are tools available in the market that enable this, such as Infosec Global's Agilescan [27].

Industry Standards, Audits and Compliance: Most industries grapple with an alphabet soup of standards such as GDPR, PCI Compliance, HIPAA. Many of these standards require an array of specific security controls requiring auditability. Several standards require assurance that implemented cryptography is current and secure. One can anticipate that as PQC is standardized, industry standards will cascade and fold updated PQC requirements into their own standards. It is advisable to assess the impact of managing cryptographic updates for an ecosystem of technologies, in context to compliance and auditability.

"It is advisable to assess the impact of managing cryptographic updates for an ecosystem of technologies."

Business Continuity Planning: While PQC readiness is certainly a better antidote in the face of a probabilistic zero-day attack from quantum computing, it is impossible to be a 100% prepared for a new technology with newly emerging quantum attack vectors. As with any novel threat, ensure business continuity plans concern themselves with the impacts from PQC. Business drivers and risk appetites of certain businesses may drive PQC strategy to be skewed more towards mitigation and remediation rather than prevention. Conducting table top exercises on PQC impacts as relates to business continuity planning and recovery would be valuable.

Incident Response Plan: Throughout history we have seen cryptography succumb to attacks in various ways, brute force, poor implementation, side channel attacks, mathematical developments among others. In the current climate of accelerated crypto-updates, it is recommended that the SOC be up to date on crypto-agile concepts and the vulnerabilities of classic cryptography to quantum attacks. Gartner suggests including cryptographic alternatives and algorithm swap-out procedures in incident response plans [26].

Quantum Products Market Awareness: The quantum computing related products and services marketplace is extremely active. Morgan-Stanley has forecasted a market potential of $5 to $10 billion annually over the next ten years [28], with a CAGR of 24.6% over 2018-2024 projected by Homeland Security Research [29]. There are new and innovative products and services becoming available, such as, cryptographic cipher scanning tools, quantum-proof digital certificates poised to remake the PKI industry, PQC ready hardware security modules to name a few. There may be solutions available that enable the PQC transition for your organization on certain fronts.

As discussed, cybersecurity professionals have a lot of work ahead in preparation for the entrance of quantum computing into the world of cryptography. While not typically quantum scientists, cybersecurity professionals are specialists in their own right and post quantum cybersecurity needs the attention of every cybersecurity professional, not in quantum time, but NOW. At this time preparation is our best defense against the future quantum attack on our information assets and infrastructure.

REFERENCES

[1] R. Curley, Cryptography Cracking Codes, New York: Britannia Educational Publishing, 2013.

[2] L. K. Grover, "A fast quantum mechanical algorithm for database search," Symposium on the Theory of Computing (STOC), Murray Hill, 1996.

[3] B. L. ,. M. R. ,. a. R. S. Markus Grassl, "Applying Grover's algorithm to AES: quantum resource estimates," arXiv:1512.04965v1 , 15 December 2015. [Online]. Available: https://arxiv.org/pdf/1512.04965v1.pdf. [Accessed 5 Nov 2018].

[4] "IBM's Test-Tube Quantum Computer Makes History - First Demonstration of Shor's Historic Factoring Algorithm," IBM, 19 Dec 2001. [Online]. Available: https://www-03.ibm.com/press/us/en/pressrelease/965.wss. [Accessed 5 Nov 2018].

[5] C. Z. John Proos, "Shor's discrete logarithm quantum algorithm for elliptic curves," University of Waterloo, 22 Jan

2004. [Online]. Available: https://arxiv.org/pdf/quant-ph/0301141v2.pdf. [Accessed 5 Nov 2018].

[6] Quantum Computing Report, "Quantum Computing Report," 13 Oct 2018. [Online]. Available: https://quantumcomputingreport.com/scorecards/qubit-count/. [Accessed 5 Nov 2018].

[7] C. Rigetti, "The Rigetti 128-qubit chip and what it means for quantum," Rigetti, 8 August 2018. [Online]. Available: https://medium.com/rigetti/the-rigetti-128-qubit-chip-and-what-it-means-for-quantum-df757d1b71ea. [Accessed 5 Nov 2018].

[8] P. Teich, "Google Joins the Quantum Race," EE Times - Tirias Research, 3 Sep 2018. [Online]. Available: https://www.eetimes.com/author.asp?section_id=36&doc_id=1333058. [Accessed 5 Nov 2018].

[9] D-wave, "Quantum Computing Applications," D-wave, [Online]. Available: https://www.dwavesys.com/quantum-computing/applications. [Accessed 5 Nov 2018].

[10] Gartner, "The CIO's Guide to Quantum Computing," Gartner, 29 Nov 2017. [Online]. Available: https://www.gartner.com/smarterwithgartner/the-cios-guide-to-quantum-computing/. [Accessed 5 Nov 2018].

[11] NIST, "Report on Post-Quantum Cryptography," April 2016. [Online]. Available: https://csrc.nist.gov/csrc/media/publications/nistir/8105/final/documents/nistir_8105_draft.pdf. [Accessed 5 Nov 2018].

[12] D. Moody, "The Ship has Sailed, The NIST Post-Quantum Cryto "Competition"," 2017. [Online]. Available: https://csrc.nist.gov/CSRC/media/Projects/Post-Quantum-Cryptography/documents/asiacrypt-2017-moody-pqc.pdf. [Accessed 5 Nov 2018].

[13] M. W. e. al, "Post-Quantum Crypto for Dummies," Wiley, Weinheim, 2018.

[14] Post Quantum Cryptograhy Team - NIST, "A Quantum World and how NIST is preparing for future crypto," March 2014. [Online]. Available: https://csrc.nist.gov/CSRC/media/Events/ISPAB-MARCH-2014-MEETING/documents/a_quantum_world_v1_ispab_march_2014.pdf. [Accessed 5 Nov 2018].

[15] Information Assurance Directorate, "Cryptome.org," Jan 2016. [Online]. Available: https://cryptome.org/2016/01/CNSA-Suite-and-Quantum-Computing-FAQ.pdf. [Accessed 5 Nov 2018].

[16] NIST, "Report on Post-Quantum Cryptography - NISTIR 8105," Computer Security Division, Applied and Computational Mathematics Division Information Technology Library, 2016.

[17] IETF, "Guidelines for Cryptographic Algorithm Agility and Selecting Mandatory-to-Implement Algorithms," Nov 2015. [Online]. Available: https://tools.ietf.org/html/rfc7696. [Accessed 5 Nov 2018].

[18] K. Martin, "Waiting for quantum computing: Why encryption has nothing to worry about," TechBeacon, 15 Aug 2018. [Online]. Available: https://techbeacon.com/waiting-quantum-computing-why-encryption-has-nothing-worry-about. [Accessed 5 Nov 2018].

[19] M. Schwartz, "Post-Quantum Crypto: Don't Do Anything," Bank Info Security, 22 Feb 2017. [Online]. Available: https://www.bankinfosecurity.com/quantum-crypto-dont-do-anything-a-9737. [Accessed 5 Nov 2018].

[20] J. M. Dr. Michele Mosca, "A Methodology for Quantum Risk Assessment," Global Risk Institute, 5 Jan 2017. [Online]. Available: https://globalriskinstitute.org/publications/3423-2/. [Accessed 5 Nov 2018].

[21] Onboard Security, "NTRU Post Quantum Cryptography," Onboard Security, 2018. [Online]. Available: https://www.onboardsecurity.com/products/ntru-crypto. [Accessed 5 Nov 2018].

[22] A. W. William Buchanan, "Will quantum computers be the end of public key encryption?," Journal of Cyber Security Technology, vol. 1, no. 1, pp. 1-22, 2017.

[23] B. Schneier, "Google's Post-Quantum Cryptography," 12 July 2016. [Online]. Available: https://www.schneier.com/blog/archives/2016/07/googles_post-qu.html. [Accessed 5 Nov 2018].

[24] J. Henry, "What is CryptoAgility," Cryptomathematic, Aug 2018. [Online]. Available: https://www.cryptomathic.com/news-events/blog/what-is-crypto-agility. [Accessed 5 Nov 2018].

[25] I. G. Tomislav Nad, "Cryptography Lifecycle," in PrimeKey Tech Days 2018, 2018.

[26] D. M. Mark Horvath, "Better Safe Than Sorry: Preparing for Crypto-Agility," Gartner, 12 April 2018. [Online]. Available: https://www.gartner.com/doc/3645384/better-safe-sorry-preparing-cryptoagility. [Accessed 10 Nov 2018].

[27] Infosec Global, "Agilescan," Infosec Global, 2018. [Online]. Available: https://www.infosecglobal.com/solutions/threat-detection/agilesca. [Accessed 31 Jan 2019].

[28] P. Bajpai, "Quantum Computing: What It Is, And Who The Major Players Are," www.Nasdaq.com, 26 March 2018. [Online]. Available: https://www.nasdaq.com/article/quantum-computing-what-it-is-and-who-the-major-players-are-cm939998. [Accessed 5 Nov 2018].

[29] [29] Homeland Security Research, "Quantum Computing Market & Technologies – 2018-2024," Jan 2018. [Online]. Available: http://old.homelandsecurityresearch.com/Quantum+Computing+Market+and+Technologies. [Accessed 5 Nov 2018].

ABOUT THE AUTHOR

DAKSHA BHASKER, P.ENG (CIE), MBA, CISM, CISSP, CCSK, is a Senior Cybersecurity Architect at Comcast. Daksha has over fifteen years of experience in the telecommunications service provider industry with roles in both business management and technology development, accountable for complex solutions architectures and security systems development. Her security work spans carrier scale voice, video, data and security solutions. Prior to joining Comcast she worked at Bell Canada developing their cyber threat intelligence platform and securing cloud deployments. She has worked on security controls for Sarbanes Oxley compliance and security risk management in complex deals with large enterprise customers. Daksha holds an M.S in computer systems engineering from Irkutsk State Technical University, Russia, and an MBA in electronic commerce from the University of New Brunswick, Canada. She contributes to security standards development and maintains an interest in security research, analysis and authorship.

Cyber Security & Information Systems
Information Analysis Center

Need Specialized Technical Support with Easy Contract Terms?

Core Analysis Task (CAT) Program

A Pre-Awarded, Pre-Competed Contract Vehicle.

CSIAC provides Subject Matter Expert (SME) support on an as-needed basis to quickly address technical requirements with minimal contracting effort. CSIAC provides such solutions via the utilization of our Core Analysis Task (CAT) service/capability. CSIAC is a competitively awarded contract with Indefinite Delivery/Indefinite Quantity (ID/IQ) provisions that allow us to rapidly respond to our users' most important needs and requirements. Custom solutions are delivered by executing user-defined and funded CAT projects without the need for further competition.

Through the CAT program, CSIAC is a pre-competed contracting vehicle, enabling the DoD and other agencies to obtain technical support for specific projects/programs that fall within one of the CSIAC technology areas. As with any inquiry, the first four hours are free. If the scope requires a CAT, CSIAC will assist with the development of a Performance of Work Statement (PWS) to be approved by the Contracting Officer's Representative (COR).

Key Advantages of working with CSIAC:

Expansive Technical Domain
The CSIAC's broad technical scope provides numerous pre-qualified resources for potential projects, and is especially valuable for today's information system challenges that frequently cross multiple domains.

Comprehensive STI Repositories
As a consolidation of three predecessor Information Analysis Centers (IACs), CSIAC has a wealth of expertise, data and information to support the successful completion of CATs.

Expansive Subject Matter Expert Network
CSIAC is able to leverage reach-back support from its expansive SME Network, including technical experts from the CSIAC staff, team members, or the greater community, to complete CATs.

Minimal Start-Work Delay
Not only does CSIAC provide DoD and other government agencies with a contract vehicle, but as a pre-competed single award CPFF IDIQ, work can begin in just a matter of weeks.

Apply the Latest Research Findings
CSIAC draws from the most recent studies performed by agencies across the DoD, leveraging the STI holdings of the Defense Technical Information Center (DTIC). The results of all CSIAC CATs and other DoD-funded efforts are collected and stored in DTIC's STI repository to support future efforts by the CSIAC and others.

How To Get Started

If you have a need for CSIAC technical support, the first step is to contact us. All Technical Inquiries are free to the customer for up to four hours of service. If the scope of the support is more extensive and requires a CAT, CSIAC will assist with the development and submission of the task description and related contract documents. CATs may be awarded as either Cost Plus Fixed Fee (CPFF) or Firm Fixed Price (FFP) delivery orders.

Inquiries may be submitted by email to **info@csiac.org**,
or by phone at **1-800-214-7921**.

Please visit our website for more information:
https://www.csiac.org/services/core-analysis-task-cat-program/

Who We Are

The Cyber Security Information Systems Information Analysis Center (CSIAC) is the DoD's Center of Excellence in Cyber Security and Information Systems, covering the following technical domains:

- Cybersecurity
- Software Engineering
- Modeling and Simulation
- Knowledge Management/ Information Sharing

CSIAC is chartered to leverage best practices and expertise from government, industry, and academia to solve the most challenging scientific and technical problems. The Center specializes in the collection, analysis, synthesis, and dissemination of Scientific and Technical Information (STI) to produce solutions in support of the defense community.

Our Team

Quanterion Solutions Incorporated is the prime contractor responsible for operating the CSIAC. In addition to Quanterion, customers also have access to the other members of the CSIAC team which include leading technology corporations as well as prestigious academic institutions that perform cutting edge research activities to expand our knowledge base.

Cyber Security & Information Systems
Information Analysis Center

266 Genesee Street
Utica, NY 13502

1-800-214-7921
https://www.csiac.org

Cyber Security and Information Systems
Information Analysis Center
266 Genesee Street
Utica, NY 13502

THE CENTER OF EXCELLENCE IN CYBER SECURITY AND INFORMATION SYSTEMS

Leveraging the best practices and expertise from government, industry, and academia in order to solve your scientific and technical problems

CSIAC
Design and Development Process for Assured Software - Volume 1
Journal of Cyber Security & Information Systems
Read the Journal
Upcoming Events
Share and Invite
Read the Latest Journal

https://www.csiac.org/journal/

To unsubscribe from CSIAC Journal Mailings please email us at ***info@csiac.org***
and request that your address be removed from our distribution mailing database.