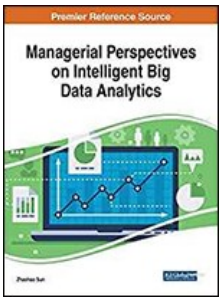

Managerial Perspectives on Intelligent Big Data Analytics

by Zhaohao Sun

Chapter 5: Using Intelligent Agents Paradigm in Big Data Security Risks Mitigation

Mihai Horia Zaharia,
"Gheorghe Asachi" Technical University,
Romania

ABSTRACT

Big data has a great potential in improving the efficiency of most of the specific information society instruments. Yet, because it uses the newly introduced cloud technology support, it may need continuous improvements especially in the security assurance area. In this chapter, a possible solution based on the intelligent agent paradigm in securing the big data infrastructure is presented. This approach will also require some changes at the general strategy level. The main accent is on using big data techniques and tools to ensure data security. Unfortunately, due to some security-related issues at the global level, the business environment must increase the amount of resources driven to this area.

INTRODUCTION

The potential of the big data approach is far from being fully exploited nowadays. This happens due to the slow adoption of cloud based architectures that are big data main support technologies. The interest in cloud adoption is high at the level of economical and research based environments (HARDY, 2016). Yet, there are a lot of problems that make this transition slow. One consists in the higher costs involved by rewriting the commonly used applications as native ones. To solve this transition the most common method is to virtualize the real machine together with its applications. This solves the cloud transition and decreases long time maintenance costs. Unfortunately this means an inefficient use of the cloud resources, so the porting of the used application or movement to newly appeared native similar cloud applications is required (Badola, 2015).

The cloud solution has not yet reached its maturity as technology. As a result there are a lot of problems concerning its security (Khana & Al-Yasirib, 2016). The same security problem appears in the majority of applications because the process of designing and implementing a secure application is at least twice expensive than making an application with a reasonable security level. So, many common applications are not adapted to the asymmetric informational war that is ongoing nowadays at global level (Lasconjarias & Larsen, 2015). This must involve fundamental changes at any levels of the information society beginning with the application design level (Mumtaz, Alshayeb, Mahmood, & Niazi, 2018) and ending with paradigm changes in global security approach.

The public cloud based systems or private ones will be the computing nodes in future global cyberspace (Sharma, 2016). If we take this into account, then there is a need of fundamental changes in the used methods for security risks mitigation. A good idea will be to use some of the military grade protocols. If this approach is chosen it may involve a significant deployment cost increase. It is true that military approaches are better prepared from a security point of view (Eggen, et al., 2013) and in many cases there is a physical isolation between public communication networks and their internal ones. Yet, as normal, there are still two situations when this isolation is partially broken. One situation concerns the needed intersection points among the public domains and the military ones. Inhere there are a lot of problems generated due to disrespecting the internal security procedures by the people or targeted attacks of the third parties (private or from governmental level). The other one refers to the cases when, sometimes, the existing public communication network is used to handle high sensitive information. In these cases even if supplementary security measures are taken, no one can guaranty a full security.

To make the situation even worse, a new era in designing cryptographic algorithm begins due to the fragility of current approaches when quantum computing is used (Mood, 2016). Besides that, the information flows grow exponentially and the dedicated existing structures specialized in information security assurance begin to be overcome. As a result, major changes at the level of used cyber defense paradigms are required no matter if the military public or private domains are analyzed.

To design a new cyber defense paradigm for big data specific infrastructure some aspects must be taken into account:

- The redesign of the main big data architecture by increasing the security involvement.
- The use of the Artificial Intelligence (AI) based assisted decision support systems.
- The use of the AI dedicated entities that will emulate at low and middle level a security expert administrator.
- To make a leap from a static based structure of the cyber defense networks to a dynamic one that will have a better reaction time and collaboration in handling the current attacks.

Background

The secure design of cloud based application methodology is still at the beginning. One reason is the extreme fluidity of the related software frameworks and technologies that are in a continuous transformation. The SOA specific security design patterns (Erl, 2009) are redesigned due to the necessary refactoring, specific to transition from service based architectures to hybrid or full micro services based architectures. As a result, the experts must improve, adapt or propose new security related design patterns (Amato, Mazzocca, & Moscato, 2018). Efforts in securing code are done at any level, even at the object communication related one (Mourad, Laverdiere, & Debbabi, 2008).

One possibility is to use decision support systems in handling the cloud, both hardware and software security. In the future cyberspace the decision support systems will play a major role in driving the society. They are not new in the governmental or private sector (Arnott, Lizama, & Song, 2017; Sun, Sun, & Meredith, 2012). This is due to the complex problems that must be simultaneously managed at higher decisional levels which are almost impossible to be handled by human mind. Unfortunately, these solutions have their own risks. Since the beginning of human society, the information validity verification was crucial in making correct (or at least as good as possible) decisions. The dimension of nowadays informational flows is beyond human comprehension. Hence, some automated processing of the primary data is required. Until now statistical method were used to detect the injection of false information in the system. These methods are now deprecated because the existing malicious software already has all the needed ability to fully intoxicate with false data (e.g. data injection) any system (Hu, Wang, Han, & Liu, 2018). Moreover, the process of validation using humans from various governmental security systems (classical approach) is so slow that, until an injection of false data in the system is detected and corrected, many real-time critical decisions may be done based on false information. This is a common aspect of the existing asymmetric informational war. This is another reason for the need in changing the main used cyber defense paradigms. The problem can be partially solved if some measures are taken on both levels: human and informational systems.

At the human level this means that the human experts involved in various security related structures must be reeducated to properly use the existing informational tools in order to increase the communication and validation speed but without renouncing to use the existing standard techniques.

At the informational system a quality of information service method must be used (Shrivastavaa, Sharmab, & Shrivastavac, 2015). The approach is not new and the security system classifies, since the beginning of time, the used human informational sources. This must be translated into a rating for various automated gathered data. Unfortunately, in the human approach any informational source cannot be seen as fully trusted even if their existing trust rates are very high. This will imply that AI techniques must be used in the analysis of all gathered data according to a set of rules that is continuously improved. Due to the already mentioned magnitude of informational flows, it is clear that most of the existing informational systems are not efficient enough. Yet, there is a solution to exponentially increase the ability of the system in properly handling these complex operations. The big data specific methods that are nowadays used in the economic related studies can be retrofitted with some effort in solving the previously mentioned problems. So, the evolution of the assisted decision systems must be directed in using these newly emerged tools, but without disbanding existing solutions. This double approach is highly costing but will provide a supplementary redundancy that is critical at these levels of decision.

The big data security problem is reduced in fact to its main support associated risks - the cloud technology. The security risks are as big as are the complexity of hardware and software of the cloud. At any layer or even tier various types of problems may occur (Singh & Chatterjee, 2017). This diversity can be handled only by the use of a variable grain architecture intelligent autonomous entity, such as an intelligent agent. In order to implement it, the big data native techniques, as information retrieval or knowledge extraction, must be used. As a result, the agent will be implemented as service composition (Zaharia, 2014).

The agent based approach can be an integrator also for insuring the cyber physical systems security (Ashibani & Mahmoud, 2017) because it has the needed granularity at the software level. Moreover, the autonomous, intelligent and, if it needed, partially supervised behavior can hide the inherent heterogeneity at the hardware level under a common view (Tao, Zuo, Liu, Castiglione, & Palmieri, 2018). This is important especially when the already deployed systems must be emerged into the big data stream.

Handling Worldwide Informational Attacks

Due to the inherent distributed architecture of the Internet, most types of attack are almost instantly replicated all around the world. This will always be a problem. The human administrators are organized in well structured networks that handle the distributed attack in the same way an elastic mesh will handle a high speed ball (NIST, 2017). In fact there are more overlapped meshes that are interlinked in order to properly handle the attack. The system works acceptable under the current conditions. Unfortunately, the future cyberspace will provide too much computing power and access to high speed communication channels so this static approach may be insufficient. In this case the security structures must be capable of increasing their reaction speed and also all available resources should be temporarily accessible as much as possible nearby the attack origin zone. One possibility is to use a scheme similar to the immune system of the mammals (Hamon & Quintin, 2016). This will involve that the dedicated agents migrate using the communication network around the initial zone. This may

be difficult because, in most cases, the communication networks themselves, or access gates may be highly compromised. Another problem is that these agents will require a lot of supplementary computing power in order to do their job and this may not be possible. Finally, it is clear that all the agents can mostly do their main aims, but cannot adapt quickly enough, if the attacker manages the attack in real time. Thereupon, supplementary human resources (experts) must also be joined if they are in the area (US-CERT, 2017). It is possible that if they exist, they may have enough knowledge to offer a real help, but they are either yet unaware of the attack or do not have the required credentials to act at this level. In this situation their level of access can be highly increased in order to help, if this it is required. Inhere a problem of grant and revocation of higher access level emerges. Also, consider the problem of communication with affected areas. This can be solved using a virtual security assistant agent that may receive the needed credentials. This is possible because these agents are not fully under the control of their owners. Moreover, the agents can use different, more secured communication channels to interoperate.

This solution may also involve supplementary security access risks but there is no perfect choice in solving this problem. Overall, it is possible that the benefits given by this approach could overcome the inherent risks. This solution represents in fact the change from a static method of handling the problems to a dynamic one.

The specialists involved in this supplementary reaction force can be selected, with their consent, from the private security companies, from the white hat community experts and also, in high-risk situations, from the army and any governmental security related bodies that may be involved. The problem of paying these services can be easily handled due to the recording of all activity at the level of the virtual agent associated to each one.

This approach may be in a partial contradiction with the political main directions, yet any government must accept the fact that in a fully information based society that interacts using a common place – the cyberspace – the higher goal is for this place to remain stable. The old approach in securing most of their geographical influence zones may highly decrease the economic efficiency of the global market concept.

MAIN FOCUS OF THE CHAPTER

Cyber Defense Paradigm Changes

Most of the cyber defense related strategies are territorial oriented, as it is natural. The current solutions are based on concentric circles of protection (US Department of Defence, 2015). These circles concern various strategic infrastructures, each of them with its proper clearance levels and proper access protocols. This was enough until now. Unfortunately, in the future, some aspects concerning the current approach will appear as follows.

Credential Checking

In most cases this is the most important problem in software systems, as well as in human interaction. Most of the attacks are based on social engineering (Lord, 2017) or on altering the credential associated to an entity at any granularity level of software architecture.

The social engineering attack can be easily handled using two approaches. First, it is based on laws and internal constraints specific to each organization. As already known, the human nature is volatile enough to avoid, in various degrees, these rules (Howarth, 2014). The motivations are complex and are driven by the lack of attention due to fatigue or lack of fully understanding of the internal protocols and laws. To properly solve the problem, a simplified version of a virtual security officer can be used as a supplementary control to mitigate the possible risks, by direct interaction with its user. This cannot solve everything because the ultimate decision must remain at the user level, but at least its decision is recorded and submitted to the upper decisional levels by the virtual agent.

In case of software impersonation (object ACL modification, rights elevation, ticket or cookie tampering and login account breach) another version of virtual security agent may be used (Harris & Maymi, 2016). In this case the structure and complexity of the agent may be used depending on the level of complexity of the supervised problem. As previously stated, this supplementary cross check may involve a significant use of the system resources, but in the context of cloud computing this will mostly drive only to costs increase. This is very good because it fits perfectly with the way of computing the level of security associated with a piece of information. The golden rule in information security is that information is well protected when the cost associated in gaining it overcomes its intrinsic value. Thus, the decision of spending supplementary computing in order to protect the information will become more easy to take and with a smaller margin of error.

Required Support Mechanisms

At the level of each cloud provider a dedicated set of services that will be the basis of the virtual agents must be installed (Zaharia, 2014). As already mentioned above, in case of emergency, they may have access to the alternate, more secured

communication network in order to properly help in handling any cybernetic attack. Given the fact that their main role is to help in cyber infrastructure, the costs may be split between the government, and also the private body that maintains the cloud. Still, due to their importance, the main control must remain at the governmental level.

The security provider agents will be designed using services from each level of cloud and big data architecture. The primary layer of services will be on top of the hypervisor level inside cloud architecture. These services are required for basic monitoring at the lower levels. Because cloud also offers virtualization, a possibility may be to create some minimal agents inside virtual machines. Unfortunately, this is not a feasible approach because the variety of virtualized software is too high. Instead, the agents will externally monitor all communication of each virtual machine. This will involve supplementary security risks at the virtual machine level. Because the virtualization is used to help the transition from existing applications to new ones, cloud native, this risk can be neglected. The virtualization will be used in new cloud based applications but, in this case, the basic services offered at the hypervisor level will give enough information to the agents in order to ensure a good security. Because most of the cloud solutions are open, inserting an agent layer can be easily done. Until know the agent has similar abilities with the Internet security solutions. These ones also use a swarm based approach in order to improve the rate of detection. The proposed virtual agent will use a similar approach, but a more general one. It will have a basic function as an Internet security (possibly, specific services from experienced providers can be used in agent construction) but it will also use the inference rules extraction mechanisms provided by services on upper architectural layers. This will give the agent enough computing power to be more reactive to the complex attacks, such as impersonation based ones.

All virtual agents will communicate indirectly by using collective intelligence, continuously improved by big data specific information retrieval instruments.

Big Data Architecture

A big data specific architecture is usually based on the standard multi-layer multitier approach but it also complies with the specifics of big data support (cloud and SOA). It involves the use of dedicated tiers needed for inter-tier or inter-layer translation, security or even some parts from business logic layer (Erl, 2009).

In Figure 1 a simplified architecture of big data is presented. The big data consists in fact of a worldwide big analytics system. The data can be provided from the existing databases or by interacting with live data streams (e.g. economic ones) and gives the user a global view analysis of his queries. In order to maintain functional a system of this magnitude, supplementary layers are required: the Integration, the QoS the System Management and the big data Governance. From user's point of view these layers are transparent because there are specific to the internal control of the system.

At the business logic level the use of analysis frameworks begin to increase (Inoubli, Aridhi, Mezni, Maddouri, & Nguifo, 2018). It may not be a fine tuned solution, because the granularity of application may thus significantly decrease its immediate flexibility or life time one. Yet it is a cheap and quick solution and these arguments make this approach so used. Of course this raises two major classes of problems. One concerns the inefficient use of the resource (a general approach versus a dedicated one), the other refers to the security aspects because the frameworks designers are, in most of the cases, focused mainly on giving the needed results and on leaving the security aspect at the minimum because they expect those problems to be handled only outside their framework.

A concurrency based market will provide many and sometimes different solutions to solve the same problems. So it is possible that even in the future a cyberspace similar approach will be used. This will further increase the problems associated with the framework use in rapid application development (Martin, Raponi, Combe, & Pietro, 2018; Pekka Pääkkönen, 2015). Therefore, new rules related to increased granularity, flexibility, interoperability must be added when designing a framework. Also, there should be more focus on security aspects such as code execution monitoring, data path analysis for virtual execution and the list may continue.

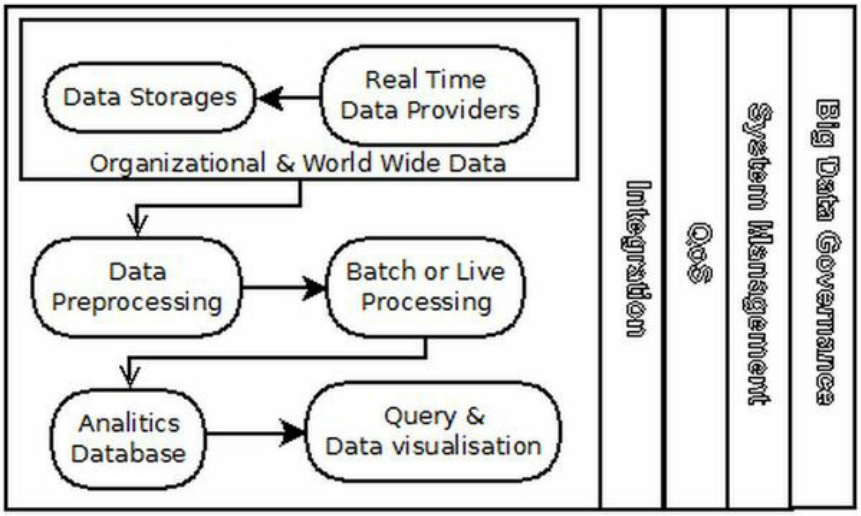

Figure 1: Typical big data architecture

At the upper levels of the software architecture (commonly known as "front end") the security problems are increased because, in most of the situations, these applications are executed in less secured environments, no matter if the system consists either of a mobile one or an Internet enabled user workstation.

The main problem of a worldwide big data system concerns the fact that it is composed by various hardware and software pieces that are literally forced to work together by some supplementary software layers. There was no uniform view in designing this system because this is not possible. As a result, at least from security related analysis point of view, the general model must be taken into account. Because the modern technologies used in constructing various applications that stand in each layer of the model are also designed using the multi-layer multitier approach, a natural conclusion is to continue the design and analysis of this subject in the same manner. The main aim will be to find a way in increasing the security of the system as a whole. Unfortunately, due to its heterogeneous nature, this means that the security analyst must consider each component or subcomponent as a black hole and also not make false assumptions concerning the security of any component. It is enough to look at the continuous emergence of the vulnerabilities reports to justify it.

So, from big data security point of view, two different zones (that will enclose various layers from classic architectural model) with different security constraints will emerge. This may not comply with the standard approach but when the multilayer architectures were designed the security aspects were at the same level of importance as the others. Nowadays, the security analysis must be made from the attacker's point of view. As a result, the statistics regarding the most used entry points or attacked zones must be taken into account (Broeders, et al., 2017). This happens because the software architect must maintain a common view of all layers and equilibrium among them with respect to the development costs. The attacker usually uses all resources to find an entry point.

The first zone will enclose the lower layers beginning with the hardware support, communication network and persistence area and in many cases even the business layer (or at least a part of it). The risks in attacking these areas are high but so is the reward. This zone usually handles the most valuable organization data.

The other zone concerns the user levels (hardware and software) and it is easier to be penetrated, but in most situations, only user personal data is exposed. Of course, when the internal organizational rules concerning information security are now well known or respected by the user, the attacker can gain supplementary information (that is sometimes vital) needed to attack the other zones by the use of social engineering based attacks or the other standard methods.

This is possible due to the local view of security analysis because each software provider tries to mitigate only the reported security problems and there is no one to make a continuous analysis of the system as a hole. It is also clear that doing this type of analysis is impossible for a human mind. It is the same limit that appears in microprocessor design where the people only make the design and the test at the general high level and all details are handled by specific software. So the only natural way in increasing the big data security is to design a global approach that will significantly increase the control of the system for the administrators.

Secure Big Data Architecture

If we analyze the Big Data Architecture proposed by NIST (NIST, 2015), it seems that the classical multilayer - multitier architecture was redesigned according to the new existing support hardware and software architecture. As a result, the user interface zone is redefined as Data Usage, a more generic term, because there is the idea of continuously increasing the

homogeneity of used interfaces, thus simplifying the design and implementation at the front end level.

The Data Transformation Layer encloses both the business and local persistence layer. This will take the advantage of centralized management and security provided by the cloud.

The third layer is similar to adaptation layer used in federative systems architecture, because it provides the interoperability with any type of existing system external data providers. This will provide backward compatibility with older infrastructures that must be also connected to the big data main stream.

This approach is tuned in with new and future big data architecture but it may confuse sometimes the new generations of system architects, due to transition from older solutions, that are more diverse, to the more homogeneous big data approach and because the life cycle for a information system may be (especially in database solutions) over thirty years.

There is also a deeply covered problem concerning the model security because the uncoupling given by service oriented architectures and the new stratification of the architecture may induce the idea that if anyone will handle its security related aspects at the needed level, the system will be secure. This is exactly the expected behavior from an attacker point of view because this thinking will significantly increase both his rate of success and the time until breach discover and risk mitigation.

Due to the high complexity of the big data systems, the classic solution of human administrators handling any security aspects is not feasible especially if we take into account the virtualization. The virtualization has a lot of advantages in auto-isolation of something that cannot be clearly classified as malicious. Unfortunately, if a lot of processes or even full servers with their applications are virtualized without having automatic data path checking, it can hide existing security many problems. Moreover, many virtualization technologies that seem to be attack-proof due to virtualization (e.g. Dockker & Kubernetes) are not (Martin, Raponi, Combe, & Pietro, 2018; Kozhirbayev & Sinnott, 2017).

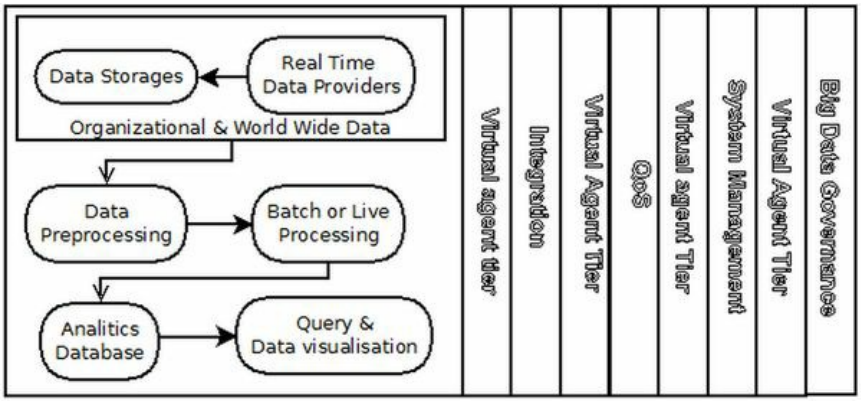

Figure 2: Secure big data architecture

In this context, one possible solution is to create virtual assistants for information security officers. In order to be efficient this solution must be a global one going across the architectural layers of the big data infrastructure. Hence, a new architecture for big data may be the one depicted in Figure 2, and it is based on using intelligent agents to insure security on top of already existing methods.

Inhere the bridge tiers between each layer (presentation, business logic and persistence) are not depicted. Virtually at each big data specific processing stage some security related virtual agents must reside in order to gain full control of the system. This is required because all the pieces are interrelated and the encapsulation paradigm is not enough if the system complexity is high. Securing the interaction between Big Data systems and old data repositories is also important because there are a lot of problems concerning the transition and interoperability. It is true that, in time, most of the existing infrastructure will be gradually replaced; yet, until this process is completed, the security related aspects are very important. These old infrastructures also have the problem of database migration into cloud, which is expensive and difficult.

As previously mentioned, the Security Intelligent Agents layer will provide the needed global security mechanism. This is possible because the big data architecture provides almost anything as a service, thus the dedicated agent orchestration can be automated (Elshawi, Sakr, Talia, & Trunfio, 2018).

Artificial Immune Systems in Big Data

The Artificial Immune Systems (AIS) is a research direction within the field of computational intelligence that emerged in the early 1990s and is based on the studies over the mammalian immunology. Most of the research is based on various solutions implemented using genetic algorithms. Inhere dominant classes are clone based selection algorithms, immune network

algorithms and negative selection algorithms (Brownlee, 2011). Due to the inherent complexity of genetic algorithms and because of parallel architecture generalization at the processor or video processor level, new classes of parallel algorithm are being developed (Cao, et al., 2018).

Due to the inherent diversity existing nowadays in the cloud, that represents the main big data support technology, one of the most interesting solutions in assuring security may be the use of the biology inspired systems, such as the immune one. It has some unique features that make it best fitted for the job.

The latest solutions begin to reevaluate de immune system approach in order to use it as a system architecture design patterns. As a result, the original model is reanalyzed. In the case of biological immune system of a new born, he will inherit the mother's immune system. Due to its high adaptability, after a period the immune system will continue to adapt to the new host body and also to gain, on long term, specific particularities depending on the illnesses specific to his host. Consequently, an immune system can be considered unique for each human. In fact, only a part of it will be different, otherwise no vaccine would ever succeed. An immune system also has abilities such as (Diogo A.B. Fernandes, 2017):

1. **Distributed Detection:** Due to the large amount of T cells that are distributed all over the body and quickly gather around the entry point of the infection.
2. **Self-Regulation:** Because no central coordination and control are present. In fact, this is a simplistic view only at the mobile units because the human immune system is too complex to be modeled only at the mathematical level. The role of the brain is not yet clarified in the direct or indirect control of this system, but from the biological inspired systems point of view the following model is enough at this stage.
3. **Approximate Detection and Pattern Matching**: The granularity at the specific receptors is lower than the one specific to most of the attackers; as a result these receptors will only bind to portions of antigen peptides. That will avoid the need for absolute detection of every pattern. There is a price paid for this solution. There are sometimes risks of autoimmune diseases or some illnesses cannot be recognized at all but in most cases it is the best approach. From nature's point of view the death of some part of the population is a simple illness negotiation problem if the other part will adapt and survive.
4. **Diversification:** That is assured by constant clone-based selection and hyper mutation gives its unique adaptability at the mobile agents' level.
5. **Anomaly Detection:** Refers to the ability of identifying the host cells as friendly, in comparison with anything else that will be classified as unknown patterns, thus being considered enemies.
6. **Learning and Memorizing:** Is gained by its continuous adaptability to various attacks and the new constructed antigens remain forever available after creation, resulting in future immunity in front of the same type of attack.
7. **Self-Protection:** Is indirectly obtained because the protection is offered to the whole organism, therefore to itself.

When new rules are created in centralized approach, some of the old rules may be considered as obsolete and deleted. Also, the centralized solution decreases the reaction speed if the incident is located far away or if the attack already isolates the center. The immune based approach must be a hybrid one. The newly gathered rules must be submitted to the center, but each agent that fights and cooperates with others by making continuous adaptation to the attacker's strategy and exchanges rules with other peers must retain its own knowledge. This particular knowledge must be also indexed at central points in order to be used, if a local rule seems inappropriate.

In most computer systems the security systems are not shuffled at intimate level with the security system components (Stewart, Tittel, & Chapple, 2005) because, ultimately, it is a question of economic efficiency. This type of protection exists at the top level secure operating systems and applications but the costs are prohibitive in many cases.

The proposed solution is based, at a higher level, on this approach. Due to the high complexity of a big data system there will be a cooperation between various agents or even agent societies using the pattern of an immune system, but at middle or lower level there are various degrees of autonomy needed in order to handle the local problems and also to process the user or the expert input related to risk mitigation.

The distributed detection will be automatically provided by dedicated agents' custom made from scratch for some particular cases or assembled from libraries provided by the internet security solution providers.

The self regulation is implicit because, at each level the agents are destroyed if they have no jobs, in order to maintain at reasonable levels the system load, given by the security framework. This is possible because the knowledge of each agent is stored before its destruction.

The approximate detection and pattern matching is a simple question of using rule sets for inference engines.

The diversification is immediate because, with every solved problem new rules are added to knowledge database, thus the new created agents (when a new similar attack appears) will inherit all needed rules so they will be better adapted in handling the specific problem.

The anomaly detection is also a simple question of verifying the agent credential (certification authority, hash based token, access control lists or any required combination).

Learning and memorizing are also implicit abilities given by the use of artificial intelligence at the level of any required agent.

The self protection is partially based on the encapsulation and access rights maintained at the level of operating system, but also on constant and sometimes reciprocal supervising among agents.

Hence, the use of this approach in designing a global level security system may be a complex but not an unfeasible task for the big data system architects.

Big Data Persistence Layer

Due to inherent different types of data sources and various technologies that begin to be emerged in the Big Data mainstream, a possible solution for databases' integration is based on intelligent agents grouped in architectural tiers similar to the one specific to federative systems. In this tier, the agent will be used only as design pattern concept applied over service based architectures that are cloud native. This high level architectural design level approach is required due to the various service providers existing nowadays on the market (Erl, 2009). At this level the intelligent agents must be also used to supervise inter-tier or inter-layer communications in order to provide load balancing and security mechanisms. The decision to create a distinct tier for service security may or may not depend on the complexity of the required architecture. From the chosen deployment architecture the persistence level can be located on the same cloud or virtual machine, or outside the current cloud. In both situations, to ensure a good protection level, basic services from Internet security solution providers can be encapsulated at the agent level in order to ensure homogenous software architecture.

A better resource handling, in terms of system load, can be acquired if a dedicated tier will be implemented. Typically, at the persistence level access layer, this may not be required. In real life, there are many applications (especially mobile related ones) that may have a minimal business logic layer that is executed locally and then remotely access the persistence level. It is true that, in theory, the majority of database systems may provide some degree of scalability on some internal load balancing in order to provide a homogeny quality of service for each client. But some order in accessing this layer may be chosen by taking into account supplementary parameters, such as client priority, or its connection quality. Also, inhere the agent based tier may provide increased flexibility in changing the communication specific protocols.

Resource Agent Based Monitoring

Monitoring all kinds of resources could be more important than it may initially seem. For instance, monitoring energy consumption for a resource can provide early advertisements regarding an unknown activity that appears. The complexity of the actual malicious eco system makes the dedicated internet security solutions overloaded sometimes, until it totally neglects some class of attacks. One reason is that the use of the swarm intelligence (cloud knowledge based) provides a highly reactive structure, but in case of an unknown class of attack (in early stages) this approach is not efficient due to the inherent lack of information regarding the subject. Given the existing system complexity, the human experts cannot handle all new emerged attacks, consequently supplementary means of suspicious activity detection must be provided. Monitoring the energy consumption at any physical device level can be an asset, especially in the context of newly emerged IoT/IoE and fog computing directions (Zhang, Zhou, & Fortino, 2018). For instance, making profiles for normal activity and monitoring each resource spike can help the administrator to monitor the Wireless Sensor Networks – WSN (Ziwen, Yuhui, & Li, 2018). The same observation is also valid for monitoring the memory, disk and CPU. Inhere there is a special problem which is difficult to handle, yet it concerns monitoring, from the security point of view, of the hypervisor of a virtual machine, because it is an important security hole that is already used by the hacker. Hence, a data path automatic validation mechanism must be deployed; probably using an AI based system. This checking may be done differently, depending on the architectural chosen solution.

There is a problem regarding the performance counters provided by the operating system for processor load in terms of fine tuning over the process (Shao, Li, Gu, Zhang, & Luo, 2018). For instance, the processor can be reported at higher load due to overload of one core but, at the same time, the others may have lower load. Hence, new ways of measuring it must be developed. A possible solution is to use a centralized architecture. Another, more efficient, possibility would be to monitor the performance counters for each application. In the context of cloud, where there is no bottleneck problem, the last one seems

more appropriate. All modern operating systems provide performance counters for any application or its processes, thus there will be no other problem in monitoring virtual machines' use of the resources. All gathered data can be stored by one agent and can be accessed and processed by another, if so. The application/process induced load is usually monitored only by load balancing reasons, but it can be used for tampering or for unauthorized application detection. This profile can also be used to have internal cloud related execution cost estimation. This may be an asset in the years to come, because it will give the owner the possibility of fine tuning its cloud related costs by updating or changing any application that proves to be economically inefficient.

In this context, dedicated hardware monitoring agents must be used. When an activity spike is detected, more specialized agents may begin to monitor the suspect device in order to see if the alert was real or not. This process can be also used in application behavior monitoring. If the problem proves to be real, the help of a human expert can be asked by an interface agent. The same way, we can use timestamp patterns dedicated agents (Ho, Kao, & Wu, 2018).

To make the automatic rule extraction, a local agent must reside on each virtual or real machine in order to collect all required information. The agent must have the ability to use the available parallelism in order to ensure a good performance of the process. The gathered information must be stored only locally. The other involved agent must extract the specific behavior pattern using data mining techniques. This agent can be implemented from scratch or will simply call an available data mining service in order to obtain the rules. In the second case, native big data instruments can be used (Elshawi, Sakr, Talia, & Trunfio, 2018). It will locally store the gathered rules and it will continuously enhance them. These rules will be also be uploaded to a central database where they will be further processed using full power of the big data. As a result, the local agent will receive updates on its rules. The method is not new and it is already used in Internet security related products that use a cloud to process all information gathered by all the worldwide installed clients. The third agent will make a continuous assessment using the rules database in order to observe if a file or a set of files begin to breach its normal timestamp pattern and then the suspicious file can be sent for further verification to an antivirus antimalware dedicated service.

An agent for traffic monitoring is also required. The content of the traffic is usually assessed by dedicated applications or services from most of the security problems. Its role in the system is to monitor the traffic overload and to detect if this is legitimate or not. Most of the Internet security solutions do not take into account this aspect in their detection schemes. This will also involve the need for communication pattern extraction for each process or application.

Due to the high load involved in the monitoring of all involved applications, a system based by training and then offset/error based monitoring triggering must be used in all monitoring schemes.

Of course, this can raise a problem regarding the costs, but the decision in using this approach may depend strictly on the importance of protected data. Also, it is possible that some of the rules created by actors with economic power to be freely distributed among the community.

Intelligent Agent as a Virtual Information Officer

There are two major sub-domains where this approach is used. One concerns the social network monitoring and control and has, as main purpose, the information control. The other is targeted on data monitoring and gathering at various levels. Both approaches are sensible of attack using the same methods. Nowadays, solutions follow the existing structures at the governmental security related departments. This stratification, that has strong reasons, must be overlapped at the information systems levels because this will further decrease the risks of false information injection or data manipulation. The other reason in doing that concerns the long time needed to adapt the informational systems to the new cryptographic solutions that will resist to quantum based attacks, and also to change the cyber defense paradigms (Mood, 2016).

Implementing a new cyber defense paradigm will have two distinct problems. One concerns the paradigm modification, then the redesign and implementation of new informational system, accordingly. This may be the easiest part, if enough resources (measured in time, money and experts) are allocated. The other major problem refers to the involved organizational changes at various departmental levels and also to the gathering of new abilities and procedures by the human security officers. This last problem is more time consuming and, in fact, will strongly decrease the speed of paradigm shift implementation.

In this context, the need for a more complex piece of software that will first help the human to adapt and then will be a common electronic assistant, is obvious (Zaharia, 2016). Therefore, an AI information officer must have the following basic abilities:

- To use direct audio visual interaction with its human partner.
- Each human will have one or more dedicated virtual agents who, in time, will learn to better help the human expert in his work.
- To have minimal auto-programming abilities in terms of modifying the used inference rules at the explicit request of the

human.

- To have the proper means to certify that the user is the designated one, not a third undesired party.
- Its credentials must be provided by a third trusted party that will be the only point that can modify those access levels.
- To use highly secured informational channels, according to its existing credentials.
- To have a backup zone that resides in a highly secured zone.
- To send a copy of any data provided to its human partner, to the higher levels of decision.

The higher authority control of the agent should be limited only to suspending its functions or destroying the agent, but without destroying or altering its primary database.

Until the emergence of cloud and big data, this type of software complexity was at least economically unfeasible in order to be implemented, but in the current context, a dedicated layer in the big data mainstream may be designed.

The big data systems already provide all needed tools to assembly this type of agent. In Figure 3 a composition flow for the agent is presented.

For instance, the user can say to the interface agent: "Please initiate a full monitoring over the network communication between the user x and the server y and provide a daily report". As a result the virtual officer will receive the request after the translation receiving from Human Computer Interface Agent. Using an inference engine the rules generated from the user request will be solved. In this particular case a clone of traffic monitoring agent will be created in order to separately handle the traffic between two specified machines. Than the feature extraction agent will extract the rules that define the supervised information flow. Full monitoring will be translated by the use of all available sets of rules for vulnerability assessment in the process. The daily report will present a summary of traffic over monitored communication. In case of some anomalies the instant report will be submitted to the security officer.

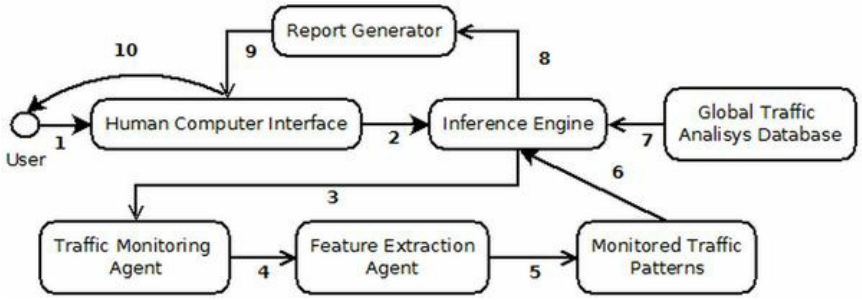

Figure 3: Interaction flow among agents

Customer Satisfaction Evaluation Agent

The problem of automatic evaluation for customer satisfaction is more complex (Wang, Du, Chiu, & Li, 2018). One possibility is based on indirect measurements by the use of information retrieval techniques over public sources, such as social networks. Inhere there are problems of handling the natural language but the most common solution is based on simple associations between brand name and some attributes clustered in bad or good perception. These solutions are also applied in measuring the 'sentiment' or 'emotion' using intelligent agents or not. The approaches provide only simple answers about the general perception, but in many cases they are over evaluated from the correctitude or efficiency point of view (Ragini, Anand, & Bhaskar, 2018).

In the big data context this problem is mostly focused on quality of service or micro-service because the future cyberspace is based on it (Elshawi, Sakr, Talia, & Trunfio, 2018; Sun J., Sun, Li, & Zhao, 2012). As a result, a more detailed measurement process regarding service QoS must be used by measuring various specific factors at different architectural levels of a big data application (Jatoth, Gangadharan, Fiore, & Buyya, 2017).

Measuring service security perception will require the use of direct or indirect measurements. In the case of direct measurements they can be done manually (using experts) or automatically, using AI based solutions. In this case, specific agents can be used in penetration testing. The indirect measurements are based on gathering user community perception. There also are two other possibilities. One is to use dedicated providers. In this case there is an increased economic efficiency, but the measurement area is limited to their offer. In the other case, specific instruments are used in order to gather user satisfaction, especially when some supplementary parameters must be analyzed and there are no external providers. This is,

as expected, a solution with higher costs; but choosing it will depend only on the general economic SWOT analysis.

Another important parameter concerns service availability. The fact that cloud guarantees this for its basic services does not automatically apply for custom based services provided by any analyzed application; this is why this investigation must also be done.

The cloud will provide some advantages, as automatic scalability and performance for any hosted service, but at a price. As a result, the parameter concerning the average price – performance must be computed and analyzed, too.

The user perception is also important because not all the users take decisions based only on economic efficiency. This happens only in case of economic actors but there, common clients captive or not may choose based on their internal set of values and perception to use a service or not.

The common user does not take into account in his analysis various components of an application (such as a service or a set of services) but mostly the application as a whole; and sometimes the decision is even simpler and based only on some specific features from the nice to have category.

Gathering user perception can be also done by dedicated market analysis providers or by using specific tools for an application or product evaluation. Unfortunately, when a more complex model for user perception is required, it is dependent on a variable degree on the application specific. In this case, a test must be designed with the help of a team that must enclose at least one psychology and one sociology expert. This may further increase the cost of the solution.

One conclusion is that a more flexible instrument will be an asset. The AI can also help in designing such solution, especially if an inference engine is used because, due to its nature, it is highly customizable by simply adding or changing the used set of rules. If we combine this observation with the use of an agent as design pattern, it can be concluded that an intelligent agent based on inference engine can provide all means to solve the problem. In order to apply this in the big data concept, the agent must be constructed using services, thus naturally integrating into the big data support technology.

Issues, Controversies, Problems

The big data security is still an open problem (Amato, Mazzocca, & Moscato, 2018). The complex infrastructure of services which resides at the cloud base makes possible breaches at any level. Using the proposed system may further increase the system complexity, making a centralized view of the system more difficult. Yet, there are some inherent advantages. Using autonomous service based agents with cognitive abilities to maintain control and system security partially solves the problem of reciprocal check among security parties. For an attacker the system will present itself not as a very well stratified architecture where it is enough to compromise iterative existing security layer, but as a conglomerate of dynamic entities that continuously check one each other or temporary cooperate if it is necessary, making it very hard to attack it.

In case of massive cyber nodes that usually consist of datacenter tier three or four, new dedicated interface security agents must be used in order to continuously interact with their human companions that work in local or external security providers. Moreover, all involved personnel activity must be supervised with dedicated agents but they will also interact with dedicated interface agents that will help on various levels in their work.

Fault Tolerance of the Security System

One of the security problems is cost related (no matter how it is computed). As a result, in most of cases, assuring a fault tolerant security infrastructure is not feasible. So any hacker that manages to disable a specific service may further proceed into the system. The natural solution is to maintain multiple levels of security located on each layer or even tier of the software architecture. The existing software and hardware infrastructure already have, from the original design, implemented various security levels that succeed, with the help of human supervisor, to manage with a reasonable reaction time most of the security breaches. Unfortunately, the cloud and its virtualization support begin to create autonomous systems with higher complexity that cannot be handled manually by human experts. The answer is to further increase the AI involvement on each layer or inside those systems to better model a virtual security officer. It is clear that even that approach cannot harness the wave of problems, but at least with some human supervision (to monitor and change used rules) a significant problem alleviation can be gained with the expected increase in costs.

SOLUTIONS AND RECOMMENDATIONS

The previously proposed solutions conclude that the whole interaction over Internet must be changed beginning with the used security paradigm and ending with the use of a solution consisting of a set of security agents crossing all levels, homogeneously deployed on each part of the big data hardware and software system. In here, the agents are differentiated depending on their role in the global cyberspace. These service based agents can have various sources depending on their

producer or owner, so there is a chance that in some areas there may too many different agents doing the same job. This may appear unfeasible form an economic perspective, but as it can be already seen, in the future cyberspace the computer power or communication network may become cheaper when more and more countries will join the global cyberspace. In this context, the security and system tight control will became more important than the associated costs.

AI Associated Security Risks

Manny artificial intelligence (AI) based solutions are used inside Big Data architectural levels (Villaronga, Kieseberg, & Li, 2018; Passalis & Tefas, 2018). The supervised learning can be used in handling the transaction fraud detection problems (Carcillo, et al., 2018).

The simple statistical methods, such as regression, are still efficient in analysis regarding customer lifetime value (Cox, Kartsonaki, & Keogh, 2018).

Of course that using AI without proper analysis (due to the high pressure of needed changes) may drive the global system to a so-called "singularity point" but this problem is over estimated (Bossmann, 2016). The main risk is at the public domain levels where the control of the AI complexity is still weak. A better solution in avoiding this may be found if governments begin to adopt a set of rules that would clearly state the limits that a public AI application must have. This process in not even started so there is serious risk that, without quick and decisive measures at the governmental levels, the free market will produce uncontrollable AI based entities. Unfortunately, such risks already exist at the governmental and military level because many governments already used AI dedicated entities to control the informational market and in some cases the autonomy level of this entities is too high (Fang, 2016). Therefore, the need for a global accepted standard of AI based entity limits is strongly required.

FUTURE RESEARCH DIRECTIONS

The future research regarding the presented solution will involve its testing with the help of a dedicated intelligent framework, based on the cloud specific technologies. As for the domain level, there are many possibilities. One can be the use of quantum based communication that was already tested at satellite level (Liao, et al., 2018). This would provide complete protection in front of the Man in the Middle specific attack (Qin, et al., 2017) due to its properties. In this case the use of cryptographic algorithm may not be necessary or at least the need for new enforced solutions is significantly decreased. Given the fact that this type of communication would be at least one decade prohibitive for most of the economic actors from the global market, some alternative solutions must be proposed, too. Inhere two different main directions emerge. One consists of the continuous improvement at the application architectural level by the introduction of more security related tiers based on technology related design patterns (Dong, Peng, & Zhao, 2010; Abramov, Sturm, & Shoval, 2012). The other is based on the continuous improvement and replacement of the commonly used cryptographic algorithms. This is required due to the quantum computing that nowadays enters a new era (Chen, et al., 2016). The third less evident problem concerns the effort put in enforcing the specific standards at governmental level. The USA already started this process (Ross, Viscuso, Guissanie, Dempsey, & Riddle, 2016). But, without a strong international cooperation, many security breaches may appear due to the inherent interoperability at the global level that is specific to the information society.

CONCLUSION

In this chapter some solutions regarding new ways of assuring big data cyber defense using intelligent agents as architectural design paradigm were presented. In fact, new general software architecture for big data systems was proposed according to the newly suggested security paradigm change. From architectural point of view the main idea is that the agents can be created, used or destroyed anywhere necessary in the system (the use of AIS solution).

In order to be efficient the system will require agents that will supervise execution at all possible levels (including resource monitoring), agents that will act as information security officers, agents that will be used in interfacing the main system with any human security expert.

This approach may seem very expensive, but in the future cyberspace, computing resources will not be an economical problem. Instead, the problem of maintaining the system under control for security related reasons will became more important. In this context, the proposed solutions may help in solving most of the expected problems.

REFERENCES

Abramov, J., Sturm, A., & Shoval, P. (2012). *Evaluation of the Pattern-based method for Secure Development. Information and Software Technology*, 54(9), 1029–1043. doi:10.1016/j.infsof.2012.04.001

Amato, F., Mazzocca, N., & Moscato, F. (2018). *Model driven design and evaluation of security level in orchestrated cloud*

services. *Journal of Network and Computer Applications*, 106, 78–89. doi:10.1016/j.jnca.2017.12.006

Arnott, D., Lizama, F., & Song, Y. (2017). *Patterns of business intelligence systems use in organizations*. *Decision Support Systems*, 97, 58–68. doi:10.1016/j.dss.2017.03.005

Ashibani, Y., & Mahmoud, Q. H. (2017). *Cyber physical systems security: Analysis, challegnes and solutions*. *Computers & Security*, 68, 81–97. doi:10.1016/j.cose.2017.04.005

Badola, V. (2015, October 1). *Cloud migration: benefits and risks of migrating to the Cloud*. Retrieved from *http://cloudacademy.com/blog/cloud-migration-benefits- risks/*

Bossmann, J. (2016, October 21). *Top 9 ethical issues in artificial intelligence*. Retrieved from *https://www.weforum.org/agenda/2016/10/top-10-ethical- issues-in-artificial-intelligence/*

Broeders, D., Schrijvers, E., Sloot, B., Brakel, R., Hoog, J., & Ballin, E. H. (2017). *Big Data and security policies: Towards a framework for regulating the phases of analytics and use of Big Data*. *Computer Law & Security Review*, 33(3), 309–323. doi:10.1016/j.clsr.2017.03.002

Brownlee, J. (2011). *Clever Algorithms: Nature-Inspired Programming Recipes*. Morrisville: LuLu.

Cao, B., Zhao, J., Po Yang, Z. L., Liu, X., Kang, X., Yang, S., … Anvari-Moghaddam, A. (2018). *Distributed parallel cooperative coevolutionary multi-objective*. *Future Generation Computer Systems*, 82, 256–267. doi:10.1016/j.future.2017.10.015

Carcillo, F., Pozzolo, A. D., Borgne, Y.-A. L., Caelen, O., Mazzer, Y., & Bontempi, G. (2018). *SCARFF: A scalable framework for streaming credit card fraud detection with spark*. *Information Fusion*, 41, 182–194. doi:10.1016/j.inffus.2017.09.005

Chen, L., Jordan, S., Liu, Y.-K., Moody, D., Peralta, R., Perlner, R., & Smith, D. (2016). *Report on Post-Quantum Cryptography*. Gaithersburg, MD: National Institute of Standards and Technology. doi:10.6028/NIST.IR.8105

Cox, D., Kartsonaki, C., & Keogh, R. H. (2018). *Big data: Some statistical issues*. *Statistics & Probability Letters*, 136, 111–115. doi:10.1016/j.spl.2018.02.015 PMID:29899584

Diogo, A. B., & Fernandes, M. M. (2017). *Applications of artificial immune systems to computer security: A survey*. *Journal of Information Security and Applications*, 35, 138–159. doi:10.1016/j.jisa.2017.06.007

Dong, J., Peng, T., & Zhao, Y. (2010). *Automated verification of security pattern compositions*. *Information and Software Technology*, 52(3), 274–295. doi:10.1016/j.infsof.2009.10.001

Eggen, A., Hauge, M., Hedenstad, O. E., Lund, K., Legasp, A., Seifert, H., & Simon, P. (2013). *Coalition Networks for Secure Information Sharing (CoNSIS)*. In *MILCOM 2013 - 2013 IEEE Military Communications Conference* (pp. 354–359). San Diego, CA: IEEE. doi:10.1109/MILCOM.2013.68

Elshawi, R., Sakr, S., Talia, D., & Trunfio, P. (2018). *Big Data Systems Meet Machine Learning Challenges: Towards Big Data Science as a Service*. *Big Data Research*, 1-11.

Erl, T. (2009). *SOA Design Patterns*. New York: Prentice Hall PTR.

Fang, L. (2016, April 14). *The CIA Is Investing in Firms That Mine Your Tweets and Instagram Photos*. Retrieved from The intercept: *https://theintercept.com/2016/04/14/in-undisclosed-cia- investments-social-media-mining-looms-large/*

Hamon, M. A., & Quintin, J. (2016). *Innate immune memory in mammals*. *Seminars in Immunology*, 28(4), 351–358. doi:10.1016/j.smim.2016.05.003 PMID:27264334

Hardy, Q. (2016, December 25). *Why the Computing Cloud Will Keep Growing and Growing*. Retrieved from *https://www.nytimes.com/2016/12/25/technology/why-the- computing-cloud-will-keep-growing-and-growing.html?_ r=0*

Harris, S., & Maymi, F. (2016). *CISSP® All-in-One Exam Guide* (7th ed.). New York: McGraw-Hill Education.

Ho, S. M., Kao, D., & Wu, W.-Y. (2018). *Following the breadcrumbs: Timestamp pattern identification for cloud forensics*. *Digital Investigation*, 24, 79–94. doi:10.1016/j.diin.2017.12.001

Howarth, F. (2014, September 2). *The Role of Human Error in Successful Security Attacks*. Retrieved from

SecurityIntelligence: *https://securityintelligence.com/the-role-of-human-error- in-successful-security-attacks/*

Hu, L., Wang, Z., Han, Q.-L., & Liu, X. (2018). *State estimation under false data injection attacks: Security analysis and system protection*. *Automatica*, 87, 176–183. doi:10.1016/j.automatica.2017.09.028

Inoubli, W., Aridhi, S., Mezni, H., Maddouri, M., & Nguifo, E. M. (2018). *An experimental survey on big data frameworks*. *Future Generation Computer Systems*, 1–19.

Jatoth, C., Gangadharan, G., Fiore, U., & Buyya, R. (2017). *QoS-aware Big service composition using MapReduce based evolutionary algorithm with guided mutation*. *Future Generation Computer Systems*, 1–11.

Khana, N., & Al-Yasirib, A. (2016). *Identifying Cloud Security Threats to Strengthen Cloud Computing*. *Procedia Computer Science*, 94, 485–490. doi:10.1016/j.procs.2016.08.075

Kozhirbayev, Z., & Sinnott, R. O. (2017). *A performance comparison of container-based technologies for the Cloud*. *Future Generation Computer Systems*, 68, 175–182. doi:10.1016/j.future.2016.08.025

Lasconjarias, G., & Larsen, J. A. (2015, December 17). *New Research Division Publication - NATO's Response to Hybrid Threats*. Retrieved from NATO Defence college: *http://www.ndc.nato.int/download/downloads.php?icode=471*

Liao, S.-K., Cai, W.-Q., Handsteiner, J., Liu, B., Yin, J., Zhang, L., … Pan, J.-W. (2018). *Satellite-Relayed Intercontinental Quantum Network*. *Physical Review Letters*, 120(3), 030501–030505. doi:10.1103/PhysRevLett.120.030501 PMID:29400544

Lord, N. (2017, February 28). *Social Engineering Attacks: Common Techniques & How to Prevent an Attack*. Retrieved from Digital Guardian: *https://digitalguardian.com/blog/social-engineering- attacks-common-techniques-how-prevent-attack*

Martin, A., Raponi, S., Combe, T., & Pietro, R. D. (2018). *Docker ecosystem – Vulnerability Analysis*. *Computer Communications*, 122, 30–43. doi:10.1016/j.comcom.2018.03.011

Mood, D. (2016, February 24). *Post-Quantum Cryptography: NIST plan for the future*. Retrieved from pqcrypto2016.jp: *https://pqcrypto2016.jp/data/pqc2016_nist_announcement. pdf*

Mourad, A., Laverdiere, M.-A., & Debbabi, M. (2008). *An aspect-oriented approach for the systematic security hardening of code*. *Computers & Security*, 27(3-4), 101–114. doi:10.1016/j.cose.2008.04.003

Mumtaz, H., Alshayeb, M., Mahmood, S., & Niazi, M. (2018). *An empirical study to improve software security through the application of code refactoring*. *Information and Software Technology*, 96, 112–125. doi:10.1016/j.infsof.2017.11.010

NIST. (2015). *Reports on Computer Systems Technology 1500-5. Big Data Public Working Group*. Gaithersburg, MD: National Institute of Standards and Technology.

NIST. (2017, March 1). *Cybersecurity Framework Overview*. Retrieved from NIST: *https://www.nist.gov/file/354081*

Passalis, N., & Tefas, A. (2018). *Learning bag-of-embedded-words representations for textual information retrieval*. *Pattern Recognition*, 81, 254–267. doi:10.1016/j.patcog.2018.04.008

Pekka Pääkkönen, D. P. (2015). *Reference Architecture and Classification of Technologies, Products and Services for Big Data Systems*. *Big Data Research*, 2(4), 166–186. doi:10.1016/j.bdr.2015.01.001

Qin, B., Huang, J., Wang, Q., Luo, X., Liang, B., & Shi, W. (2017). *Cecoin: A decentralized PKI mitigating MitM attacks*. *Future Generation Computer Systems*, 1–11.

Ragini, J. R., Anand, P. R., & Bhaskar, V. (2018). *Big data analytics for disaster response and recovery through sentiment analysi*. *International Journal of Information Management*, 42, 13–24. doi:10.1016/j.ijinfomgt.2018.05.004

Ross, R., Viscuso, P., Guissanie, G., Dempsey, K., & Riddle, M. (2016). *Protecting Controlled Unclassified Information in Nonfederal Systems*. Gaithersburg, MD: National Institute of Standards and Technology. doi:10.6028/NIST.SP.800-171r1

Shao, Y., Li, C., Gu, J., Zhang, J., & Luo, Y. (2018). *Efficient jobs scheduling approach for big data applications*. *Computers & Industrial Engineering*, 117, 249–261. doi:10.1016/j.cie.2018.02.006

Sharma, C. S. (2016). *Securing Cyberspace: International and Asian Perspectives*. New Delhi: Pentagon Press.

Shrivastavaa, S., Sharmab, A., & Shrivastavac, D. (2015). *An Approach for QoS Based Fault Reconfiguration in Service*. *Procedia Computer Science*, 46, 766–773. doi:10.1016/j.procs.2015.02.145

Singh, A., & Chatterjee, K. (2017). *Cloud security issues and challenges: A survey*. *Journal of Network and Computer Applications*, 79, 88–115. doi:10.1016/j.jnca.2016.11.027

Stewart, J. M., Tittel, E., & Chapple, M. (2005). *CISSP: Certified Information Systems Security Professional Study Guide* (3rd ed.). San Francisco: Sybex.

Sun, J., Sun, Z., Li, Y., & Zhao, S. (2012). *A Strategic Model of Trust Management in Web Services*. *Physics Procedia*, 24(B), 1560-1566.

Sun, Z., Sun, J., & Meredith, G. (2012). *Customer Decision Making in Web Services with an Integrated P6 Model*. *Physics Procedia*, 24(B), 1553-1559.

Tao, M., Zuo, J., Liu, Z., Castiglione, A., & Palmieri, F. (2018). *Multi-layer cloud architectural model and ontology-based security service framework for IoT-based smart homes*. *Future Generation Computer Systems*, 78, 1040–1051. doi:10.1016/j.future.2016.11.011

US-CERT. (2017, January 24). *Critical Infrastructure Cyber Community Voluntary Program*. Retrieved from US-CERT: *https://www.us-cert.gov/ccubedvp*

US Department of Defence. (2015, June). *2015 - The National Military Strategy of the United States of America*. Retrieved from Aquisition Community Connection: *http://www.jcs.mil/Portals/36/Documents/Publications/2015_National_Military_Strategy.pdf*

Villaronga, E. F., Kieseberg, P., & Li, T. (2018). *Humans forget, machines remember: Artificial intelligence and the Right to Be Forgotten*. *Computer Law & Security Review*, 34(2), 304–313. doi:10.1016/j.clsr.2017.08.007

Wang, J.-N., Du, J., Chiu, Y.-L., & Li, J. (2018). *Dynamic effects of customer experience levels on durable product satisfaction: Priceand popularity moderation*. *Electronic Commerce Research and Applications*, 28, 16–29. doi:10.1016/j.elerap.2018.01.002

Zaharia, M. H. (2014). *Generalized Demand-Driven Web Services*. In Z. Sun, & J. Yearwood (Eds.), *Handbook of Research on Demand-Driven Web Services: Theory, Technologies, and Applications* (pp. 102-134). IGI Global. doi:10.4018/978-1-4666-5884-4.ch005

Zaharia, M. H. (2016). *A Paradigm Shift in Cyberspace Security*. In B. A. Hamid & R. Arabnia (Eds.), *Emerging Trends in ICT Security* (pp. 443-451). Morgan Kaufmann.

Zhang, P., Zhou, M., & Fortino, G. (2018). *Security and trust issues in Fog computing: A survey*. *Future Generation Computer Systems*, 88, 16–27. doi:10.1016/j.future.2018.05.008

Ziwen, S., Yuhui, L., & Li, T. (2018). *Attack localization task allocation in wireless sensor networks based on multi-objective binary particle swarm optimization*. *Journal of Network and Computer Applications*, 112, 29–40. doi:10.1016/j.jnca.2018.03.023