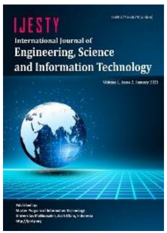
International Journal of Engineering, Science & InformationTechnology (IJESTY)
Volume 1, No. 2 (2021) pp. 62-71
ISSN 2775-2674 (online)
Website: http://ijesty.org/index.php/ijesty
DOI : https://doi.org/10.52088/ijesty.v1i2.101
Research Paper, Short Communication, Review, Technical Paper



A Noble Security Analysis of Various Distributed Systems

Sworna Akter, Md. Alamgir Hossain*, Md. Mojibur Rahman Redoy Akanda

Department of Computer Science and Engineering, Prime University, Dhaka, Bangladesh
**Corresponding author E-mail: alamgir.cse14.just@gmail.com*

Manuscript received 23 Feb 2021; revised 2 March. 2021; accepted 15 March 2021. Date of publication 2 April 2021

Abstract

Distributed systems increased the performance of a system by allowing applications to be processed in parallel. It helps us to share resources like printers, computers, storage facilities, data, files, web pages, networks, and thus it reduces the cost by shaing a single resource among several users instead of multiple resources. Since multiple machines communicate with each other through the network so they can be easily affected by the attacker and the full system should be corrupted. Users can access remote and local resources but they also may not aware of which machines their processes are running on. So, secure communiction of a distributed system is the most important issue. We need to analyze various vulnerabilities and can take proper protection of the system. In this paper, we discuss different architectural styles of distributed systems. Also, hold up different threads and protection mechanisms to get rid of these threads. The most focusing part of this paper is a summary of various distributed systems security. Here we hold up different distributed system security techniques like SSSE algorithm, BLCS architecture, ODIS algorithm, two eavesdropper model, dynamic cuckoo filter, etc. Finally, we summarize the full discussion like used methodology or architecture, advantages, disadvantages, accuracy, and future work in a table by which an author can easily gather knowledge about the summary of this paper. To better understanding, we show the result of various papers with a visual representation. I think it helps authors to know about security techniques and grow interested to work in this area.

Keywords: *Distributed System, Security, Distributed Algorithms, Raft Consensus Mechanism, Hybrid Signature Algorithm.*

1. Introduction

After the invention of a high-speed computer network, we can connect hundreds of machines in Local Area Networks (LANs) and can transfer information from one machine to another one in a few microseconds. We can transmit a large amount of data between machines at rates of 100 million to 10 million bits/sec. In Wide Area Networks (WANs) we can connect millions of machines all over the world and the speed varying from 16kbps (kilo-bits per second) to gigabits per second. So, it is easy to connect many computers or computing systems with high speed in the network. Instead of a centralized system, we can use many independent computers that can communicate through networks to transmit messages. Though there is a number of computers in the network users seem that they use a single system. The goal of a distributed system is the following [1]:

a. Making resource accessible
b. Distribution transparency
c. Openness
d. Scalability

Assets can be things like printers, PCs, storerooms, information, documents, site pages, and organizations. The principle objective of a disseminated framework is to make it simple for clients and applications to get to these distant assets. In a distributed system, processes and resources are physically distributed among multiple machines. The goal of a distributed system is to make these distributed processes and resources to users as if they use a single system. This is called the transparency of a distributed system. There are various types of transparency: Access transparency, location transparency, migration transparency, relocation transparency, concurrent transparency, failure transparency, replication transparency. An open conveyed framework offers administrations as indicated by standard principles. The scalability of a distributed system can be measured in three different dimensions. a. size scalability b. geographical scalability and c.administrative scalability.
In this paper, we discuss security issues of distributed systems and focus on the security of various distributed systems and summarize these techniques. I also discuss different architectural styles, possible attacks of distributed systems, and hold up some mechanisms to protect from these attacks.

2. Literature Review

2.1 Structural Styles

There are four significant styles of engineering for distributed systems [2]They are:
a. Layered architectures
b. Object-based architectures
c. Data-centered architectures
d. Event-based architectures
In layered architectures, components are organized in a layered fashion. A component can call the underlying layer but not the other way. In object-based architectures, each object corresponds to a component, and these components are connected through a procedure call mechanism. In data-centered architectures, processes communicate through a common repository. Cycles convey through the spread of occasions which likewise convey information in occasion based structures.

2.2 Security Issues of Distributed System

In a distributed system, multiple nodes communicate with each other through the network. This communication should be secure otherwise full system may be affected by attackers. So, it is most important to ensure the security of distributed systems. The security of disseminated frameworks can be isolated into two sections. One section concerns the correspondence between clients or cycles that are dwelling on various machines. For ensuring secure communication we should ensure a secure channel that means authentication, message integrity, and confidentiality. Another part concerns authorization.

Our objective is to protect data and services from an attacker. So we need to familiar with threads. There are many types of security threads includes [2]:
a. Interception,
b. Interruption,
c. Modification,
d. Fabrication etc.
When an unauthorized party can access to data and services it is referred to as interception. When the services or data become unavailable, unusable, destroyed and so on, this situation referred to as interruption. Modification involve unauthorized changing of data. Creation alludes to the circumstance where extra information are produced that ordinarily not exist.

So, to protect data and services from attacker we need some mechanisms. There are various security mechanisms and important security mechanisms are [2]:
a. Encryption
b. Authentication
c. Authorization
d. Auditing etc.
Encryption transforms data into something another so that an attacker can not understand the original messages. Authentication is used to verify the claimed identity of a user, client, server, host, or other entity. After a user has been authenticated, it is necessary to check whether the client is authorized to perform the requested action. Auditing tools are used to track which clients accessed what and which way.

2.3 An Overview

In 2020, Guoqi Xie, Gang Zeng, and Renfa Li. [3] develop a safety mechanism to handle the error of safety-critical automotive applications within its deadline. In the creation of automobiles, the main thing is protected driving. There is various safety component like safety belt, airbag, brake-by-wire, and so on are invented that can enhance the safety of the automobile system. Before starting authors studied road vehicles standard that was reveled for automobile software developers. In this paper, the authors mentioned that safety should be enhanced by reducing risk by reducing the probability or the severity of harm or both. They also mentioned in their paper that the probability of risk will be 100% if safety-critical automobile applications don't complete properly within its deadline. In this paper, the authors proposed an existing Backward Safety Enhancement (BSE) algorithm and a new Forward Safety Enhancement (FSE) algorithm. BSE is a backward safety enhancement algorithm that tries to migrate each task to another Electronic Control Unit (ECU) which generates maximum reliability value from exit to entry task. On the other hand, FSE handles the recovery process from entry to exit task. FSE can migrate each task to another ECU that can generate maximum reliability value. Then they proposed to combine the BSE and FSE algorithm for more safety enhancement. The authors also proposed Repeated Backward Safety Enhancement (RBSE) and Repeated Forward Safety Enhancement (RFSE) algorithms in their paper. RBSE and RFSE handle the recovery process by primary backup repetition. These algorithms try to add a new replica in backward and forward for each task to an available ECU that can generate maximum reliability value among all available ECUs. For more safety enhancement they proposed to combine RBSE and RFSE algorithm.

Finally, they proposed a new approach called the Stable Stopping-based Safety Enhancement (SSSE) algorithm. This algorithm is a combination of the above four algorithms that are the existing BSE algorithm, proposed FSE algorithm, proposed RBSE algorithm, and proposed RFSE algorithm. SSSE is an intermingling calculation that implies when the unwavering quality worth arrives at a consistent express the calculation can stop.

Authors showed their paper that after applying the BSE algorithm reliability value of automobile application is enhanced from 0.947716 to 0.95294318. Applying their proposed FSE calculation unwavering quality worth improved from 0.95294318 to 0.95475549. Using the FSE algorithm the end time of the exit task is not ended at 100. Therefore they apply RBSE and RFSE algorithms. Then they got enhanced reliability from 0.95475549 to 0.97584149 by using RBSE and got from 0.97584149 to 0.97586917 by using RFSE. Finally, they apply the SSSE approach to reaches a stable and fixed value. Applying the SSSE algorithm the reliability value is enhanced and got the same value (0.97586917) at the first round of RFSE, the second round of RBSE, and the second round of RFSE. So, the value 0.97586917 is a stable and fixed value. In this paper, the authors discussed some disadvantages of the SSSE algorithm. There is too many unnecessary Message Receiving Interrupts (MRIs) bring ECU load that is considerable and interrupt triggered switch overhead between tasks. In the automotive system, ECUs running at relatively low clock speed with small memory space. Therefore, although the SSSE algorithm enhances the reliability of the automotive application, the load of ECU affects the reliability of task scheduling.

Authors mentioned in their paper that they adopted the real-life parallel automotive application with 31 tasks and using SSSE algorithm time required to calculate the result is 1s. To confirm the advantages of proposed algorithms they use parallel automotive applications with 100 tasks and analyze the result. They showed that it also takes 1s to calculate the result using the SSSE approach. As future work, authors said in their paper that further study the safety enhancement of distributed automotive embedded systems with considering multiple automotive applications with different deadlines.

In 2020, Hui Yang, Kaixuan Zhan, Michel Kadoch, Yongshen Liang, and Mohamed Cheriet, propose an architecture named brain-like distributed control security (BLCS) to protect the security of the cyber-physical system (CPS) against cyber-attack for fog radio and optical networks (F-RON) [4]. In their paper authors illustrated the functional modules of BLCS architecture including various controllers and a brain-like knowledge base and also described the internetworking procedures in distributed control security modes based on brain-like security (BLS). In terms of average mistrust rate, path provisioning latency, packet loss probability, and blocking probability they evaluated the efficiency of their proposed architecture on a software-defined network tested. Authors illustrated their proposed architecture with graphical representation and it can help anyone to understand their proposed work easily. In their architecture, they showed different sections of their architectures and labeling them as a physical plane, cyber plane, control plane, SDN controller, and trust management. The cyber plane is a distributed multi-domain wireless and optical network that is deployed with radio, optical, and fog computing resources. In the physical plane, authors showed physical entities like the robotic arm and camera that are interconnected with the cyber plane. In the SDN controller when one domain gathers its cyber information in a physical module other domains' contain partly virtual topology and routing behavior in a virtual module to teach for route verification. SDN controller contains a control model and it is the core engine of it. The path computes element includes intra-domain and inter-domain computation that executes cross-domain path calculation. Parser allocates continuous radio and optical spectrum and finds out the routing path and it contains the messages of the packet in, states reply, flow modification, and features reply. After calculating the path or network the parser controls the spectrum bandwidth and center frequency for CPS services by setting all the related antennas and optical switches.

In this paper authors mentioned that trust management is an important module of their proposed architecture. This trust management model has three modules named knowledge base (KB), brain-like learning, and evaluation module. KB is updated periodically based on the statistics about the state that are collected from the physical cyber. Brain-like learning module builds the identification states behavior relational network (ISB-RN) of KBs for trust calculation based on partial information. The assessment module computes the regulator's trustworthiness, dependability, and joint effort by the ISB-RN which can decide a regulator is trusted to execute a digital activity or not.

The authors also described the brain-like learning scheme with a suitable diagram. There are two types of controller, one is the leading secure controller and another is the receiver. The secure controller is the leader of multiple controllers. It periodically synchronizes the network information by which it can realize the consistency among multiple controllers. Insecurity authentication process SC2 receives the authentication request and partial information of SC1. Then SC2 can confirm whether SC1 is trusted or not. In the control plane, the author provides different relations in different space called knowledge base input for the relation network (RN) model. Then RN learns to infer unknown relations and they are able to provide more specific input directly into RN and separate these inputs by using tags. They use long short term memory (LSTM) to process the information. These settings use some prior knowledge. After processing the input they can understand how information is connected. Thus, each layer provides architecture and the model can learn to partition information and can calculate the interaction between partitioned information. The motor rotates to control the physical arm by calculating the motor rotating angle. Thus the relationship among the calculating unit, motor, and automobile manufacturing is build up. After these processes verify which SC is trustable. When one object sends a request then avoid untrusted SCs and use a greedy algorithm to route the safe path. The authors also described the whole process with necessary example in their paper.

In the performance analysis and result discussion part, authors evaluated their BLCS architecture on their SDN testbed with heavy traffic load and compare it with other schemes in terms of normal question rate, way provisioning dormancy, bundle misfortune likelihood, and obstructing likelihood. They showed that BLCS can maintain a low average mistrust rate. The scheme without BLCS shows a high average mistrust rate when the ratio of malicious controllers increases the average mistrust rate raises faster. The reason behind this high rate is without performing a trusted authentication connection established directly. In terms of path provisioning latency, they showed their proposed scheme can decrease the path provisioning latency. The reason is that in their scheme the trusted authentication and path provisioning can perform quickly. In terms of packet loss probability, their proposed scheme shows less packet loss probability because their scheme can choose a route with low packet loss. In blocking probability, their proposed scheme also reduces the blocking probability by selecting trusted controllers and perform faster path provisioning. All of these experimental results have been showing in their paper using a graph by which a reader can differentiate easily the result using their proposed scheme and without using this scheme. As future work, authors said that to enhance the security, reliability, and accuracy of rapidly growing CPs architectures need research more sophisticated, optimization techniques with unsupervised learning and reinforcement learning BLCSs.

In 2020, Fangyu Li, Rui Xie, Zengyan Wang, Lulu Guo, Jin Ye, Ping Ma, and Wenzhan Song [5] authors proposed an online distributed Internet of things (IoT) security monitoring algorithm to deal with the "big data" issues in IoT security. Their proposed algorithm expertly handles the complex streaming multinational time series. To extract intrinsic data structure data science techniques such as IPS and streaming data modeling has been proposed by the authors. Authors showed graphical views of the workflow of their proposed OIDS algorithm with streaming big data that includes different stages named streaming big data, influential point selection (IPS), data structure modeling (DSM), online distributed monitoring (ODM), and consensus hypothesis testing (CHT).

In the methodology and algorithm section, the author's details described the principles of their proposed ODIS algorithm by introducing every key step and also summarized the whole algorithm. In the symbol and notation, step authors introduced symbols that are used for matrices, operations, collection of the set, scalar, vector, transpose matrix, different norms, identity matrix, etc. In multidimensional time series modeling steps, the authors said that multidimensional time series modeling effectively extracts the temporal dependent information from the streaming multidimensional data which helps to understand and monitor the status of the IoT system. They introduced the VAR family in this section that is used to reveal the complex dependence structure in the

streaming time-series data. It can quantify the complex temporal and cross-sectional relationships among the multidimensional time series.
Their next step was big data influential point selection. This point only accounts for a small amount of the whole dataset but they represent the data structure. By influential points selection authors reduce the processing time and energy consumption. They find out the influential point by the following:

a. They extend the linear VAR model to the form of a general streaming non-linear model for observation up to time t.
b. For a given function they denote a column vector with a given length and then estimate coefficient matrix.
c. They value the importance of a data point through its predicted value and it is called Mahalanobis distance.
d. To reduce the data authors select a subset data and estimate the model coefficient matrix by using the subset data. The authors mentioned that if the subset is much less than the data size, the ISP will greatly save the computational time and cost.
e. Then the authors defined ISP according to the selection rule with the selection threshold.

Authors said that if the root of Mahalanobis distance is greater than the threshold then the data point selected as an influential point. The authors also show the influential points with the necessary diagram. Authors propose an online streaming IPS adapting a single-pass streaming algorithm. When new data arrives, collect the first batch of data points to calculate the sample covariance matrix. Then they replaced the streaming IPS and the corresponding selection rule with another one and it makes the streaming IPS scalable in the big data setting. In short, when the streaming data arrives sequentially they update the estimate of the model coefficient matrix.

In distributed online monitoring, the streaming IPS and VAR modeling are integrated into the asynchronous distributed computing environment. Then they describe how they monitor and summarize their work by an algorithm. Finally, based on the distributed monitoring result authors develop the statistical hypothesis testing strategy to distinguish the attack status from the normal status in a quantitative way. The authors said that after this distributed hypothesis testing their system can obtain a unified decision.

In their experimental setup, they use beaglebone blackboards (BBB) to implement a real IoT system. For attack detection and monitoring their experiment applies to volumetric DoS cyber-attack. Volumetric DoS attacks consume available network bandwidth between the target and the internet. In experiment 1, they showed a DoS attack pattern with a suitable diagram and it happens in the 20s then there is a 5s interval. The authors compare the modeling performance with different sampling methods such as ISP, Vanilla, and Bernoulli. Vanilla and ISP sampling method generates small modeling errors than Bernoulli sampling but due to the noise of real testbed the Vanilla results are affected by noise and ISP gives better robustness. In experiment 2, authors use more strength DoS attack and compare the modeling method and also got ISP to give better performance than others.

Authors also experiment with the computation efficiency of ISP, Vanilla, and Bernoulli and showed that ISP is not time-saving for big data but it gives better modeling accuracy than others. In this paper, the authors discussed some disadvantages. One of the disadvantages is ISP has additional computations, so it is a little slower. Another disadvantage is when the attack is more strength ISP doesn't give a good performance.

In 2019, Ruyan Wang, Hanyong Liu, Honggang Wang, Qing Yang, and Dapeng Wu, proposed a distributed blockchain-based security [6]. To automatically achieve system confidentiality, integrity, and authenticity they use smart contracts that act as trusted third party. The authors studied research challenges related to security and privacy issues and then found potential solutions. The authors also evaluate the effectiveness of their proposed architecture. In this paper, the authors showed that their proposed architecture has four layers. These are the physical layer, distributed database layer, communication layer, and interface layer. The physical layer consists of two parts. These are sensors and smart devices and as the sensor collects the patient's physiological data, only legitimate nodes entered into the system. This layer passes data to the upper layer. Distributed database layer stores data that are generated by smart devices.

In a distributed layer, each sensor and smart devices generate data and in each sensor, a signature is generated by a smart contract to protect data integrity. Distributed database layer stores these data. Different wireless networking technologies like Bluetooth, 4G, and Wi-Fi are adopted in the communication layer for data communication among smartphones. The interface layer allows different applications/ devices to communicate with each other that make collaborative decisions. Authors studied different security challenges and address them proposed potential solutions that consist of three parts including authentication based on a Merkle tree, a hybrid signature algorithm, and Raft consensus mechanism. Merkle tree is used to verify data and identity accuracy by generating a hash value. The hybrid signature algorithm is the combination of the Digital Signature Algorithm (DSA) and Rivest-Shamir-Adleman (RSA) that can speed up computation and ensure data integrity. A raft consensus mechanism is used to achieve a consistent confirmation of data records. The authors compare the running time of DSA, RSA, and hybrid signature algorithms. The authors showed that the hybrid signature algorithm has a signature creation time of 0.226s, a signature verification time of 0.081s, and a total signature time of 0.307s which was 0.4s faster than RSA. As advantages authors showed that there is increased memory utilization, fewer data errors, and improved node authentication efficiency because node only maintains a Merkle root hash without saving an entire dataset in the authentication process. They also showed proposed hybrid signature modes use a double key for signature and verification and it provides better performance in data privacy.

In 2019, Omar Hussein, proposed a paper named "Identification of Threats and Vulnerabilities in Public Cloud-based Apache Hadoop Distributed File System" where the author identifies, exposes, and discusses security threats and vulnerabilities in public cloud-based Hadoop Distributed File System (HDFS) [7]. Apache Hadoop (AH) technology can store and process large scale datasets. Science there is large data sets it can easily corrupt and showed many vulnerabilities. In this paper author first, discuss current HDFS related security issues. Then he discusses security threats and vulnerabilities and also shows how to identify these threats and vulnerabilities.

After studying current HDFS related security author conclude that HDFS security focuses on four countermeasures against unauthorized data access, data leakage, modification, or corruption: (1) Development of trusted architecture (2) authentication and authorization mechanism (3) cryptography (4) virtualization-based security. In this paper author identified and explained two categories of threats to HDFS that reduce HDFS's security-based behavior. The identification was: (1) vulnerable execution environment (2) over-reliance on perimeter defenses. The author also identified, explained, and illustrated four vulnerabilities in HDFS: (1) unreliable and attackable architecture (2) unprotected meta-data file (3) load restriction on access control lists information and (4) optional adaption of Kerberos authentication protocol. These weaknesses are pertinent to shortcoming in the HDFS organization figuring climate, engineering plan, and embraced security methodology. In this paper author finally, conclude that HDFS security is still immature.

In 2020, S. Kruglik [8], proposed a new type of eavesdropper and the possible framework to become secure against the attack for distributed storage systems (DSS) in his paper named "Security Issues in Distributed Storage Networks". To recover temporarily or permanently unavailable nodes author focused on information theoretic secrecy by introducing random symbols into stored data. Author consider a DSS that store information symbols and form a code matrix C. For locally recoverable codes author assume that the value of alpha symbol will be 1. Recovery set contains r symbols and the content of failed node can be recovered by measuring no more than r nodes. There are many generalizations of locally recoverable codes among them author discuss one generalization [9]. Code matrix C is called regenerating if it permits recovery of information symbols and repair of failed node by downloading symbols. The bandwidth that is needed to repair that code is called repair bandwidth and there exist a cut-set bound on the message size that forms a trade-off between the storage space and the repair bandwidth. Author mentioned two ways to ensure secrecy in distributed storage. First one is precoding information and random symbols by maximum rank distance codes and another one is direct mixing of information and random symbols using the storage code. In the paper, author consider two different types of eavesdropper (l1, l2) and (0, 0, l3) that coincide for locally recoverable codes and they are unable to change the node content. In (l1, l2), eavesdropper observe two nodes l1, l2 where data is stored on l1 node and data is downloaded during the repair of l2 nodes. Author assumes l2=0 without case of generality in case of regenerating codes and locally recoverable codes. When eavesdropper control some subset of nodes that model corresponds to the case. In (0, 0, l3) eavesdropper can observe l3 symbols stored at any node. This model corresponds to the case when eavesdropper control communication infrastructure of DSS and can observe only small portion of data from all servers. Author then illustrate the difference between two eavesdroppers with toy example. Author then illustrate the locality security of locally recoverable codes and regenerating codes. Author also proposed a MBR array codes as resistant to eavesdropper and prove that it securely stores the maximum possible amount of information in case of minimum repair bandwidth.

In 2019, V. Pleskach and M. Pleskach and Olena O. Zelikovska [10], told that the important areas for improving information's security in distributed information system (DIS) is the development and implementation of an information security management system (ISMS) and it is the piece of general administration framework which depends on the business hazard approach in making, executing, working, checking, breaking down, keeping up and improving data security. Considering this aspect authors described the significance, features, other important aspects of implementing the ISMS in a DIS, also analyze the current state of application of ISMS in DIS in Ukraine and finally formulating the proposals regarding the need to implement ISMS in their paper. Authors hold on various data leakage informations in their paper. Among them some information are pointed below:

1. In 2019, Facebook's latest leak includes data on 420 million phone numbers.
2. In October 2019, 20 million of instagram users contact data had stolen and exposed.
3. There is greatest loss of data prevail in medical industry. In Singapore, hackers broke into the patients' database at SingHelth's largest network of clinics and stole 1.5 million peoples' data.

In 2019, Ukraine priorities the development of electronic information resources, e-services and digitalization and authors emphasize on careful security system implementation.

Distributed system covers large number of computers which is part of a network and the channel is most vulnerable place. Therefore, authors told to take measures like signature analysis technologies, advanced antivirus protection systems, correctly organized replication in different time zones, daily backup and modern firewalls to protect information, hardware and software. Authors also told to increase the skills of users as in 2018 International Telecommunication Union pointed out that there is lacks of professional standards in Ukraine. Finally in their paper author calls for urgent actions depend on these view.

In 2019, Zoltan Andras Lux, Felix Beierle, Sebastian Zickau and Sebastian Gondor [11], proposed a full-text search framework based on the publicly available metadata on the hyper ledger indy ledger for retrieving matching credential types. By using full-text search engine and maintaining a local copy of the ledger their proposed solution able to find credential types. Their solution takes textual input from the user. Authors also proved the feasibility of their concept by implementing and evaluating a prototype of the full-text credential metadata search service. In Implementation step, authors hold up three steps for clear discussion. These are full-text search engine, DIMS API, frontend. As full-text search engine authors use Apache Solr as it is an open source and totally free to use software. Apache Solr also supports sharding and replication and gives good solution in case of very large ledger. Authors use DIMS API to implement and demonstrate proof of concept SSI issuer and web application verifier. Authors also use it to create local copy of the Indy ledger. Authors also thinks about the search functionality on the user interface (UI) side and they told that for querying need a single HTTP Get request and relevant records of ledger are returned as response. Authors conducted performance benchmarks of their full-text search service on a commodity machine and tested the throughput using 400 open connections and a single thread. They evaluate the performance of five query: (1) schemas or credential definitions by schema name, (2) schema by schema name, (3) transactions from a specific person with a typo in the name, (4) credential by schema name, (5) credential definition by attribute name. Their Solr instance contained the first 47312 domain transactions. Their standard showed that a single node Solr setup can serve up to 16000 requests per second based on the type of query. Authors conclude that their full-text search service is more performant than necessary. Authors also depicted their result graphically. Authors mentioned some points as future works of their work. They told about integrity verification of chain. They told to use Markle tree specific verification process. This verification process should integrate into the regular update ledger mechanism looking for new transactions. Authors also told the idea of overlay using. It may provide additional information about schemas for SSI capable applications and services.

In 2020, Allen Starke and J. McNair [12], proposed a state of the art WAM-SDN system for large scale network management. In their paper authors examined about the necessities for enormous scope remote appropriated WAM-SDN and furthermore give early benchmarking and execution examination dependent on half breed circulated and decentralized design. Authors first discuss about the necessity of Software Defined Network (SDN) for Wireless and Mobile (WAM) systems. Authors told that SDNs can be used to quickly assemble new services and infrastructure to meet the objectives of dynamically changing environment. In centralized architecture, controller is one entity and responsible for path routing, policy implementation, partitioning the network and other administrative functionality and it leads to single point of failure. For large scale dynamic wireless networks this architecture creates some challenges. So, WAM-SDN and heterogeneous SDN solutions must be fault tolerant. Authors told that to produce efficient, reliable and trustworthy SDN architecture, the coordination among distributed controllers must be considered. Authors showed their proposed hybrid distributed and decentralized SDN controller architecture graphically. This architecture will monitor network performance metrics in data plane and performance metrics in the control plane. Author told many controllers needed to keep the switch to controller delay lower. They conducted experiments using mininet and extensions to emulate the data plane environment in a linear and ring topology. They conducted their test on ADM A6-6310 APU with 4 cores. Authors illustrate the comparison of the internal packet processor service times and round trip times in terms of wireless centralized and distributed controller architectures. Utilizing the distributed design can reduce LLDP and reactive packet processing times by 55% and 52%.

In 2018, Baozhou Luo and Wenjun Zhu and P. Li and Zhijie Han [13], proposed a distributed Dynamic Cuckoo Filter system based on Redis cluster which has capability to automatically scale the storage network data. Authors illustrate their work in the paper named "Distributed Dynamic Cuckoo Filter System Based on Redis Cluster". Their method can store only finger print information of data. Creators utilized the predictable hashing calculation to build Redis bunch and furthermore utilize intensive correspondence instrument to Redis group to accomplish the information sharing and productive use of multi-machine channels. Authors also told that their proposed scheme can take into account the time and space efficiency, may greatly improve the massive data retrieval performance as well as improve the reliability and availability of Redis cluster. They showed the overall framework of Distributed Dynamic Cuckoo Filter (DDCF) system based on Redis cluster with figure. The system consists of a system center module and a plurality of DDCF system sub-nodes and are distributed on a circular Ring with 232 uniform distribution points. The system center module is used to complete the system initialization, node mapping between data and storage nodes and the global management of the sub-nodes. Framework sub-hubs is utilized to compute the information unique mark data and pail up-and-comer address data to finish the comparing information tasks like supplement, inquiry and erase. System center module consist of some parts and Redis cluster sub-node contain several parts and authors also describe their functionality and flow of their work. Authors hold up two algorithm, one is insertion and another one is for query and delete operation algorithms and describe in detail. Authors also discuss the time complexity of deletion and query operation. The time complexity of query operation is O(b) and delete operation is O(nb).

3. Methodology

The reason behind this work is day by day security analysis of distributed systems has become a more important research sector. We need to increase the security of this system. When I started to work I was gone through some steps. At first, I was gone to the "IEEE Xplore" site by using Google. Then I have search papers using the keywords "Security analysis of distributed systems". As a result, I was found 6375 papers. I wanted to work on very recent work. That's why I set the years 2019 and 2020 and got 1360 papers among 6375 papers. Among them, I have select 100 papers and focused on some different work like Distributed Automotive Embedded Systems, Cyber Physical Systems, Distributed IoT system, Distributed Architecture based on Blockchain, Apache Hadoop Distributed File System, Distributed Storage Network, Wireless Distributed SDNs, and Distributed Dynamic Cuckoo Filter System. Finally, I take 10 papers to explore. I studied them and analyze their used methodology by which they protect their system's privacy, also analyze the accuracy of their work, the advantages, and disadvantages of their system and also future work. Finally, I summarize them in this paper and from it, one can gather knowledge about the security of the distributed systems and able to analyze new techniques.

4. Results and Discussion

In this part, the detailed analysis and results is discussed by using table and figure. Table 1 includes the details of methodology they used, advantages and disadvantages. Table 2 showed the summary of their accuracy and their future plans.

Table 1. Summary of Used Methodology, Advantages and Disadvantages

Paper Reference	Used Methodology or Architecture	Advantages	Disadvantages
[3]	BSE, FSE, RBSE, RFSE, SSSE algorithms.	SSSE algorithm enhances the reliability of the automotive application and takes 1s to calculate the result.	Too many unnecessary MRIs bring ECU load and this load affects the reliability of task scheduling.
[4]	F-RON, BLCS architecture, SDN controller, parser, KB, brain like learning, RN.	BLS can find vindictive regulators and security directing, while at the same time decreasing the normal doubt rate, way provisioning inertness, bundle misfortune likelihood, and hindering likelihood.	No
[5]	ODIS algorithm, VAR model, single pass streaming algorithm. ISP, Vanilla and Bernoulli sampling methods.	ODIS is suitable for the real-time large scale multidimensional streaming IoT security monitoring. It shows promising performance in terms of cyber-attack detection and monitoring.	ISP is a little slower and when the attack is more strength ISP doesn't give good performance
[6]	Merkle tree, a hybrid signature algorithm and Raft consensus mechanism, DSA, RSA.	Provide increased memory utilization, fewer data errors and improved node authentication efficiency. It also provides better performance in data privacy.	No
[7]	Security issues, Apache Hadoop	This paper tackles the security state of Apache Hadoop Distributed File System (HDFS).	No
[8]	Proposed two eavesdropper model and showed secrecy against it.	Securely stores maximum possible amount of information in case of minimum repair bandwidth.	No
[10]	Said to take measures like signature analysis technologies, advanced antivirus protection systems, correctly organized replication in different time zones, daily backup and modern firewalls to protect information, hardware and software	The real implementation of ISO/IEC 27001, is able to significantly reduce operating costs and losses from breaches of information security. It also improves the reputation of Ukrainian entities that process personal data, and thus improve Ukraine's situation in the worldwide digital records and data security.	No
[11]	Apache Solr software, JSON-LD.	Proposed solution able to find matching credential types.	No
[12]	Hybrid distributed and decentralized SDN controller architecture, linear and ring topology,	To provide system fault tolerance in the event of controller failures and attacks, offload network administrative functions to the cloud or other controllers for energy preservation, offload controller functions to multiple nodes for load balancing, allow for	No

		distributed resolution of node failures and attacks.	
[13]	Hashing algorithm, query and delete operation algorithms.	Automatically scale the storage network data, improve the massive data retrieval performance as well as improve the reliability and availability of Redis cluster.	No

Table 2. Summary of Accuracy and Future Work

Paper Reference	Accuracy	Future Work
[3]	Applying BSE, FSE, RBSE, RFSE and SSSE algorithms authors got 0.95294318, 0.95475549, 0.97584149, 0.97586917 and 0.97586917 value accordingly. The value 0.97586917 is a stable and fixed value.	Further study the safety enhancement of distributed automotive embedded systems with considering multiple automotive applications with different deadlines.
[4]	Can maintain a low average mistrust rate, decreases the path provisioning latency and can choose a route with low packet loss.	To enhance the security, reliability, and accuracy of rapidly growing CPs architectures need research more sophisticated, optimization techniques with unsupervised learning and reinforcement learning BLCSs.
[5]	ISP gives better performance than Vanilla and Bernoulli and also gives better modeling accuracy than others.	Need to research more sophisticated techniques to enhance the performance with increasing attack strength.
[6]	The hybrid signature algorithm has a signature creation time of 0.226s, a signature verification time of 0.081s, and a total signature time of 0.307s which was 0.4s faster than RSA.	No
[7]	HDFS security is still immature.	No
[8]	Securely stores maximum possible amount of information in case of minimum repair bandwidth.	Will combine two mentioned eavesdropper model.
[10]	No	Implementation of their proposed methods.
[11]	A single node Solr setup can serve up to 16000 requests per second based on the type of query.	Will use Markle tree specific verification process as future work which helps to integrate into the regular update ledger mechanism looking for new transactions and also use the idea of overlay.
[12]	The distributed design can reduce LLDP and reactive packet processing times by 55% and 52%.	No
[13]	The time complexity of query operation is O(b) and delete operation is O(nb).	No

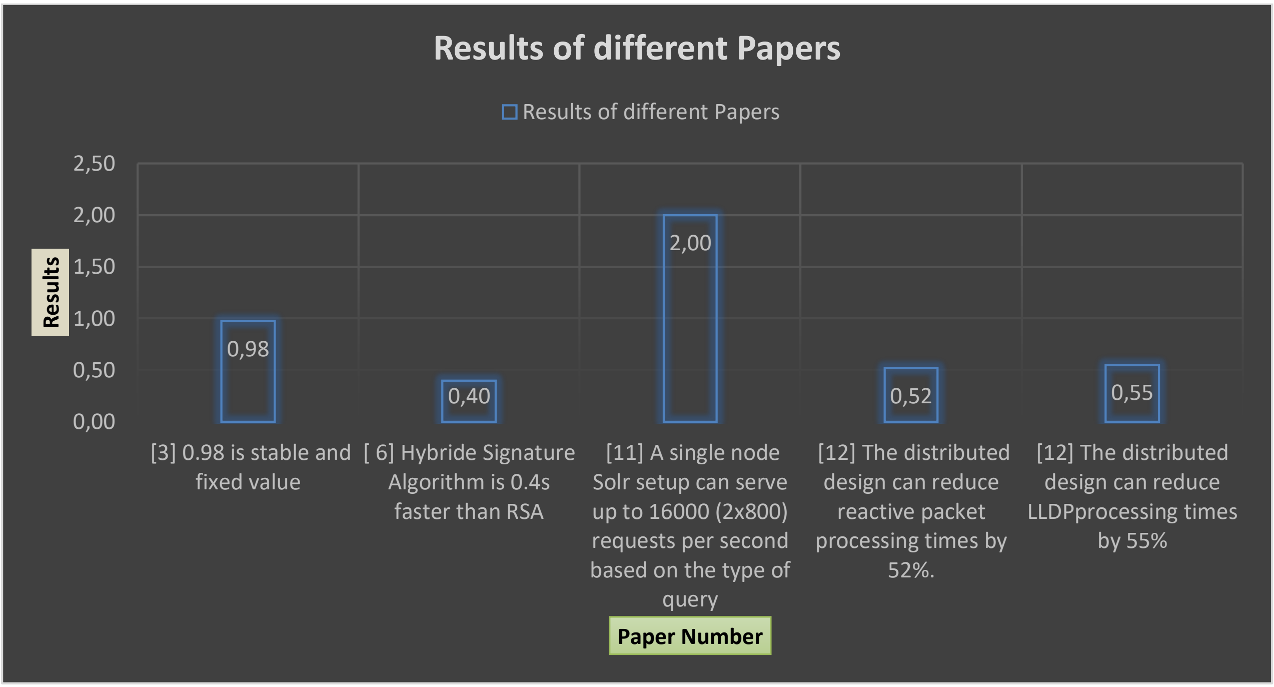


Fig 1. Results of Various papers

Fig.1 showing that third paper got stable and fixed value of 0.98 for SSSE algorithm after applying different algorithm. The total signature time for the hybrid signature algorithm is 0.307s, which is 0.4s faster than RSA. A single node Solr setup can serve up to 16000 requests per second based on the type of query. The Hybrid distributed and decentralized SDN controller architecture can reduce LLDP and reactive packet processing times by 55% and 52%.

5. Conclusion

In this paper, we summarized various distributed system's security-based works and also uphold their used technology, the accuracy of their work, advantages, disadvantages, and their future work. Finally, I briefed all of these in Table 1 and Table 2 as if one can easily take knowledge by seeing it at a glance.

In the future, the authors will be benefited from this and they are eager to develop a new ways for the evolvement of strengthening distributed systems.

References

[1]. Firdhous, M., *Implementation of Security in Distributed Systems – A Comparative Study*. International Journal of Computer Information Systems, 2011. **2**(2).

[2]. Andrew S. Tanenbaum, M.V.S., *Distributed Systems: Principles and Paradigms*. 2006: Prentice Hall, Netherlands.

[3]. Guoqi Xie, G.Z., *Safety Enhancement for Real-Time Parallel Applications in Distributed Automotive Embedded Systems: A Stable Stopping Approach.* IEEE Transactions on Parallel and Distributed Systems, 2020. **31**(9): p. 2067 - 2080.

[4]. Hui Yang, K.Z., Michel Kadoch, Yongshen Liang, Mohamed Cheriet, *BLCS: Brain-Like Distributed Control Security in Cyber Physical Systems.* IEEE Network, 2020. **34**(3): p. 8-15.

[5]. Fangyu Li, R.X., Zengyan Wang, Lulu Guo, Jin Ye, Ping Ma, Wenzhan Song, *Online Distributed IoT Security Monitoring With Multidimensional Streaming Big Data.* IEEE Internet of Things Journal, 2020. **7**(5): p. 4387 - 4394.

[6]. Ruyan Wang, H.L., Honggang Wang, Qing Yang, Dapeng Wu, *Distributed Security Architecture Based on Blockchain for Connected Health: Architecture, Challenges, and Approaches.* IEEE Wireless Communications, 2019. **26**(6): p. 30-36.

[7]. Hussein, O., *Identification of Threats and Vulnerabilities in Public Cloud-Based Apache Hadoop Distributed File System*, in *15th International Computer Engineering Conference (ICENCO), 2019*. 2019.

[8]. Kruglik, S., *Security Issues in Distributed Storage Networks*, in *IEEE 21st International Symposium on "A World of Wireless, Mobile and Multimedia Networks*. 2020.

[9]. S. B. Balaji, M.N.K., M. Vajha, V. Ramkumar, B. Sasidharan, P. V. Kumar, *Erasure coding for distributed storage: an overview.* Science China Information Sciences, 2018. **61**(10).

[10]. Zelikovska, V.P.a.M.P.a.O.O., *Information Security Management System in Distributed Information Systems*, in *IEEE International Conference on Advanced Trends in Information Theory (ATIT)*. 2019.

[11]. Zoltan Andras Lux, F.B., Sebastian Zickau and Sebastian Göndör, *Full-text Search for Verifiable Credential Metadata on Distributed Ledgers*, in *Sixth International Conference on Internet of Things: Systems, Management and Security (IOTSMS)*. 2019.

[12]. Allen Starke, J.M., *Load Balanced Controller Association in Wireless Distributed SDNs*, in *IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. 2020, University of Florida.

[13]. Baozhou Luo and Wenjun Zhu, P.L.a.Z.H., *Distributed Dynamic Cuckoo Filter System Based on Redis Cluster*, in *IEEE 4th International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing, (HPSC) and IEEE International Conference on Intelligent Data and Security (IDS)*. 2018.