A Novel Approach for Compliance Assurance using IoT

Mr. Ritz Sebastian
Department of Computer Science and Engineering,
Rajagiri School of Engineering and Technology,
Ernakulam, India
Ritz.Sebastian@gmail.com

Ms. Shimmi Asokan
Department of Computer Science and Engineering,
Rajagiri School of Engineering and Technology,
Ernakulam, India
shimmi_a@rajagiritech.edu.in

***Abstract*— Compliance is known as one of the most expensive areas, irrespective of domain and geography. Technology paradigms like Cloud computing, Big Data and IoT are evolving with more opportunities. The significant advancements brought lots of challenges in parallel. The need for user authorization is the key requirement of compliance. Though, there are several methods applied to ensure compliance, many security breaches are reported these days. Data breach is one among the top of those concerned areas. This elaborates the need of compliance and all organizations to focus on their customer data evidently with extra vigilance. Witnessing capability of Internet of Things is leveraged to enhance the compliance assurance servicing area. The novel approach, proposed in this paper, is for user authorization in real time for the key transactions. This would be a key step in the area of proper authorization of any user even after an initial authentication. The model is discussed in detail as a proof of concept with couple of trade-off as assumptions and is implemented and the results are discussed.**

Keywords— Compliance; Assurance; Internet of Things, Internet of Things witnessing; Cloud Computing; AWS;

I. INTRODUCTION

Information technology with various paradigms like cloud computing, IoT, Big Data and Automation are in its peak. The momentary progresses on invention of these areas are giving lots of space for future advancements in all the domains. The compliance requirement stands as an unavoidable need to all those organizations who handle customer's data and their own confidential information. People are identified as the most risky element in any domain compared to other elements like process and technology.

Current information systems use multi-factor authentication as the key for authorization and authentication. But it is proved from recent report that, data breach issues still exists and it is a big challenge to the public domain. In financial institutions, the root cause of the breaches happens online where the current authentication is insufficient. Again when the study further expanded, it is found that no system does a real time authentication throughout a session. That is, once initially authorized till the session expires. This issue opens an opportunity to improve the compliance assurance needs around data access, against the key element such as people by using a witnessing environment. Data access requirements are well defined by Sarbanes-Oxley Act for all US public company boards, management and public accounting firms [1] and HIPAA for healthcare domain [2]. Even though, meeting those data access requirements accountable at all the levels in an organization, in reality, the accountability lies with an organization's responsibilities and an inevitable part to sustain in the business vertical. The other motivations for this work are :

- Normal human-beings use sense organs to perceive an environment (Sight, Hearing, Taste, Smell and Touch)
- ' In the future, intelligence services might use the internet of things for identification, surveillance, monitoring, location tracking, and targeting for recruitment ', says James Clapper, US director of national intelligence
- IoT is an emerging and promising area which is used for sensing data/information

Compliance is still considered to be open area of research as well in the industries like financial and medical where customer's credentials are managed. People, process and technologies are source of data compliance issues. Among them people are still risky elements. New technologies can be leveraged for these kinds of issue resolutions. This can be improved with proper surveillance using IoT witnessing environments. The model proposed here, can be implemented as a solution to the compliance assurance where any key transactions are performed. The key transactions referred here are important financial transactions or any type of confidential data access.

II. RELATED WORK

The work reported in the literature, focuses on the areas such as root causes of the data breaches, how the compliance assurance needs are ensured, the role of multi-factor authentication and advancements in security, and how the new paradigms are leveraged for these key problems.

The reports ISAE 3402/SSAE 16, discusses on compliance issues in special cases of outsourcing relationships are focus on the auditing part [3]. As per the PCI compliance report on 2014, two-thirds of organizations did not adequately test the security of all in-scope systems [4]. Verizon Trend Micro published various reports and the data breach report shows

that eighty nine percent of breaches had a financial or espionage motive [7].

Security considerations for IoT from the perspectives of cloud tenants, end-users, and cloud providers, in the context of wide-scale IoT proliferation, working across the range of IoT technologies (be things or entire IoT subsystems) were analyzed. The current state of the cloud supported IoT study shows that it is not completely secure [9]. The 2016 year summit organized by AWS considered the increasing complexity of mobility and system connectivity. This increases difficulty in managing risk and security and in demonstrating compliance. They provide cloud enabled services and are continuously putting efforts to address the real world issues in the area of compliance's [5]. The Matthias et.al developed a reference model and it support companies in managing and reducing risk and compliance efforts on the solid basis of a systematic literature review and practical requirements by analyzing Cloud Computing Service offers [8]. In another work, the process compliance checking using model checker, a framework is modeled and created in simulator for the compliance needs and assurance, considering data as the key element [13].

Every work concentrated on either the authentication side or authorization part. The needs for ccompliances around data authentication and authorization, by considering the data classifications, are not tried to address with innovative or new paradigms like IoT. This opportunity is leveraged and this paper proposes an innovative solution.

III. Proposed Model

The new model proposed, ensures the user authorization requirements as per any compliance controls. The design solution improves the environment to a better compliant environment due to the following reasons:

1) The compliance assurance is using IoT witnessing environment

2) Re-authorization of key transactions and real time monitoring is done using pre-configured and pre-approved IoT devices in passive mode

3) It is proactive preventive approach to the problem

4) Both the computing and external environments are protected to avoid the data breach issues

For the development of proof of concept of the model, four types of users and three environments are considered. The users considered are : Normal User, the user who works from office, home or pre-approved remote place. Delegated User is the user who is delegated by original user, works from office, home or remote place. Snooper is the user who might be snooping the details while the normal user or delegated user works from different location like office. Hacker is the user who is a threat to the system itself and breach the data mostly from remote place. Three Environments are office, home and remote/unknown location.

IV. Modeling and Simulation of Design using UPPAAL

As a first step, a formal model of the system is created and simulated using a model checking tool. In modeling, two basic data types are defined for analysis purposes. They are public and private. A channel is defined for authentication and authorization. A parallel channel is used for capturing the IoT witnessing device status. These two statuses are verified before granting the access to the restricted data. Here restricted data is private. This rule can extended to any level of confidentiality. Apart from this, data integrity and availability are modeled and verified. UPPAAL SMC tool [13] is used for timed automata modeling, simulation and model checking. Computational Tree Logic (CTL) formula is used to write input specification for system verification. This work used 64-bit version of OS with 8GB memory capacity. The simulation experiment executed on Windows 7 - 32 bit platform with UPPAAL- 4.0.14. (Academic Version). The system had the Java version 6 (e.g. J2SE Java Runtime Environment) or newer installed and properly configured on the system.

In model checking, model and specification is required. The system is represented as a model M, and expected system properties are described as a specification S respectively. A choice to model hybrid system is timed automata. Temporal logic, such as Computation Tree Logic, which extends proposition logic with temporal operators, is a good choice of description of system properties. The design and build of simulation include the development of network of automata and specifications.

A. Formal modeling of the system

The design and build of simulation includes the development of network of automata and specifications. The network of automata includes process automata, login operations automata and common module automata.

The whole model M= Process, loginOperations, mediaDataAndLog (Common module), can be checked against any specification S. The Process automata, which is designed in process compliance work [14]. is modified according to the design requirement and mediaDataAndLog automata is reused. loginOperations automata is newly added to simulation design. The Login Operations model includes nine key states which is shown in Fig.1. They are Free, GetCredentials, VerfiyCredentials, logCredentials-AttemptStatus, decideDataType, publicData, privateData, shareDataforRead and returnTo. Free State means that process is idle and it does not need any type of data access. This model sets the following flags which are declared as global :

1) Access flag-By default it is set to false. It means that none of the processes got access to any type of data available

2) dataType flag-By default it is set to zero. It means that no datatype is identified. The dataType =1 stands for Public and 2 for Private data

3) gotCredentials flag-By default it is set to false. It means that no credentials is available

4) gotCredentialsLogged flag-By default it is set to false. It means that credentials verification details are not logged

5) thruVerifyCredentials flag-By default it is set to false. It means that process is not through with credentials verification checks

6) dataAvailable flag-By default it is set to false. It means that no data is available It is used for availability check

7) clusteringData flag-By default it is set to false. It means that no data is available in cluster mechanism

8) loadbalancingData flag-By default it is set to false. It means that no data is available in load balancing mechanism

9) faulttoleranceData flag-By default it is set to false. It means that no data is available in fault tolerance mechanism

Also this process model uses release, takeModify, takeRead and takeDelete channels to synchronize with other sub models.

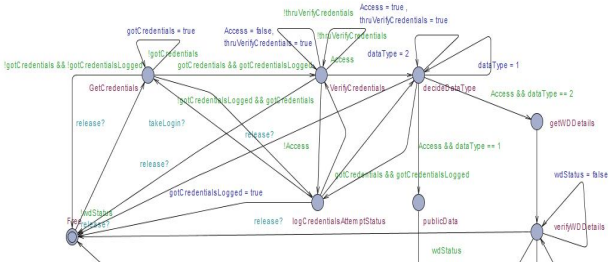

Fig. 1. Login Operations Automata

All other states have well defined functions and are transitioned to next state based on the input and actions. Those functions are briefly discussed below:

1) GetCredentials is the state to get the credentials. If it is success, gotCredentials is set to true and otherwise it is set to false

2) VerfiyCredentials is the state at which the Credentials are tried to be verified. If it is success, thruVerifyCredentials set to true and otherwise it is set to false. If all the prior checks are passed and cleared, another key flag Access is set

3) logCredentialsAttemptStatus is the state at which all the logs are made for all the GetCredentials, VerfiyCredentials and decideDataType functions

4) decideDataType is the state at which the data type is identified. In this state, it will decide whether datatype is Public or Private

5) getWDDetails In this state, it will identify the witnessing device data for an user

6) verifyWDDetails is the state at which the witnessing data verified

7) publicData is the state at which the data type is 1 and Access is true.

8) privateData is the state at which the data type is 2 and Access is true.

a) Both the states publicData and privateData can be further enhanced according to the GRC needs. But here it is not elaborated

9) shareDataforRead is the state at which the data availability check is triggered using the synchronization flag mediaLog

10) returnTo is a state which is reached after a successful or unsuccessful operations it will come back to this state and route it to free state.

Also this model uses guarding flags gotCredentials, gotCredentialsLogged, Access, DataType. Synchronization flags are release, takeRead (takeModify and takeDelete) and mediaLog to invoke other sub models and moving the transition to earlier state called model's state instance.

B. Specifications

After modeling system, formal specifications are developed for which the model has to be verified. The properties that are identified for this solution to be verified are:

1) Standard Properties

a) A[] not deadlock

- This property is used to ensure that system has no dead-lock situation at any point of time

b) E<> modifyOperations.logCredentialsAttemptStatus

- The reachability property is used for any state, for example, logCredentialsAttemptStatus that can be reached

2) Confidentiality Check properties

a) E<> modifyOperations.privateData and Access == true and dataType == 2

- Private Data is accessible only when Access = true , dataType = 2 (code used for identifying the private data) and wdStatus = true

b) E<> modifyOperations.privateData and Access == true and dataType == 2

- Private Data is not accessible when Access = true and dataType = 2 and wdStatus = false

3) Integrity Check properties

a) modifyOperations.Free -> modifyOperations.Free | modifyOperations.GetCredentials

- From state Free to Free or GetCredentails states shall be reached and no data change is expected unless until a function or a state modified the data

b) E<> modifyOperations.publicData | modifyOperations.privateData

- Data ooperations like add, modify or delete, can be done on both Public and Private

4) Availability Check properties

a) E<> modifyOperations.privateData and Access == true and dataType == 2 and dataAvailable == true

- The property verify the availability check for any type of data based on the credential status

After modeling the system, the specifications developed for confidentiality, integrity and availability against data compliance requirements are verified against models. All the properties are satisfied for all the models.

V. REAL-TIME IMPLEMENTATION

The system implementation includes both hardware and software. Raspberry Pi 3, GPS Sensor and C270 Logitech Webcam are the additional key hardware components other than AWS and hosting environment with LAMP configuration in GoDaddy. The software languages and tools include AWS IoT, AWS S3 Services, AWS CLI, Open CV2, Python 2.7.9, Linux Shell script, PHP, MySQL. The system works and gives a close to real time experience and ambient monitoring details and logs. Here the architecture is limited only to two witnessing devices and this can be modified. The system uses two separate channels for passive and active dataset prepared for authentication. This enhances the security of system. Bit Bucket is used for configuration management.

A. Architecture

The capabilities of modern sensing devices, features of cloud and IoT are used in the solution architecture. The main components of architecture, Fig. 2. are: It includes pre-configured and pre-approved IoT devices. These things are registered in cloud (Here it is in Amazon Web Service). Web pages are built (Login and only limited to few data transaction to imitate the real world scenario) and maintained in private hosting area. Authentication queue and service are running in hosted area. HTTPS and MQTT are the protocols used and internet is used as the channel. Whenever a key transaction is performed by the user, the things are challenged to witness again and again in real time to grant authorization

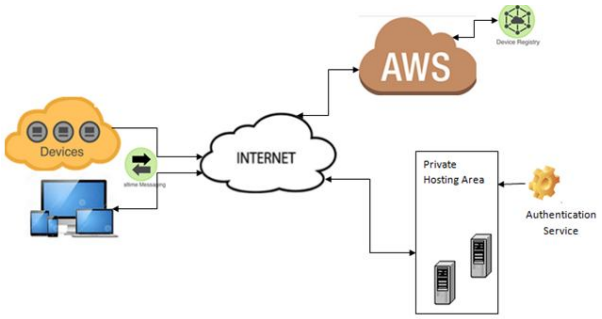

Fig. 2. Architecture

In the architecture, the number of witnessing devices restricted to two. But this can be modified as per the convenience of cost or accuracy. This model recommends keeping the confidential data in the private hosting area with all the security features. The main computation of authentication and authorization are distributed at the restricted environment. This can be enhanced according to the business needs and load. This architecture supports the minimum or only required data transmissions which ensure the efficient utilization of channels in two ways. First the real-time re-authorization is limited to only the critical transactions and then data stream is not fully shared at the end to end level.

B. Design of active mode of authentication and passive mode of authorization

There are two types of authentication and authorization. They are active and passive. The active authentication requires a user response, but the passive response is not prepared by user. Here passive responses are prepared by the IoT things. The username and password are the elements transferred and provided by the user. Location details captured by GPS sensor and face recognition using webcam are done initially as part of passive authorization. Then these data are verified at controller and send as the passive authentication details via cloud to the hosting environment. The following diagram Fig 3. shows the application flow from the end user request till final response.

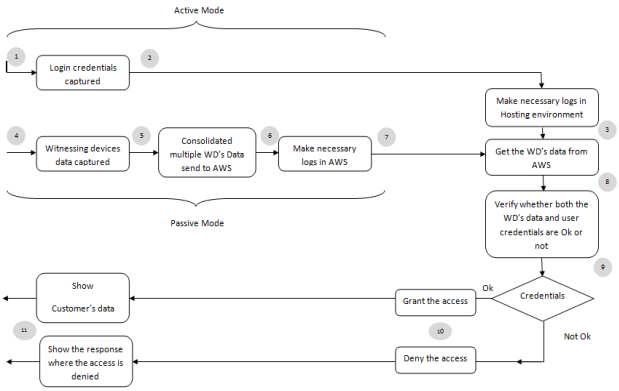

Fig. 3. Application Flow Diagram

The steps in the application flow diagram are :

- Step 1-User browse the URL link and it brings default page for login
- Step 2-Login form captured the user credentials and it will be send to private hosting location
- Step 3-In hosting location, it creates necessary login details
- Step 4-Witnessing devices capture respective data
- Step 5-At controller, the data from witnessing devices are consolidated and send to AWS
- Step 6-At AWS end, necessary logs are created
- Step 7-WD's status are pulled at AWS end from the hosting environment

  Note: Steps 1 to 3 and 4 to 7 are run in parallel
- Step 8-At hosting environment, both the credentials and WD's data are consolidated
- Step 9-Now at hosting environment, the liveners of connection, the credentials and WD's data are verified before granting access
- Step 10-Based on the consolidated status the responses are prepared. If the final status is Ok, customer data

page is created else the restricted status pages is prepared

- Step 11-The final response is shared to the end user based on the final status.
    - If all the user credentials are verified, the confidential data is displayed and it is available for operations. If the credentials are not Ok, the confidential data is not availed to the end user.

VI. RESULTS

Proposed proof of concept (POC) model of the system is built. The model is tested to verify the expected results against actual results. The model built is verified for both simulation and experimental designs. A system with minimum of three scenarios shall be checked for compliance using pre-configured and pre-approved IoT witnessing techniques against normal user, delegated user, snooper and hacker. The expected result is that the system needs to show only confidential information to the legitimate user. If user is not authorized, confidential information shall not be displayed to the end user. The Table I., details the test cases defined and used for simulation and experimental design testing for user authorization.

TABLE I. SYSTEM TEST CASES

Test Case No.	Test Cases	Expected Results	Actual Results
TC001	Right person login from correct location (Normal user)	Should be able to login and view the CCD	User was able to login and view the result
TC002	Right person login from in-correct location (Normal user)	Should be able to login and not able to view the CCD, but able to view other pages	User was able to login and view only the public pages. He was restricted to CCD
TC003	A person logins with others credentials (Hacker)	Should be able to login and not able to view the CCD, but able to view other pages	User was able to login and view the only the public pages. He was restricted to CCD
TC004	Delegated person logged in the credentials or enters premise (Delegated user)	Only authorized person should be able to view the CCD	The person was already authorized, able to view the CCD
TC005	While the authorized person logged in other person (Non-Authorized) enters (Snooper)	When a snooper identified, the confidential pages should not be displayed &	Snooper was identified and was not able to view the CCD
TC006	While the authorized person logged in, internet connection lost due to some reasons	When the connection is available, detect it and show error screen rather than currently displayed sensitive data	Offline situation is identified and error screen displayed at client side

Test cases are designed in such a way that, it is limited to only the scope of problem statement and what is necessary. But it does not mean that it can limited to these test cases only. The testing and evaluation of system is divided into three sections for the ease of understanding.

A. *System test results and evaulation*

The test cases defined against the requirement of user authentication and authorizations for ensuring confidentiality are all verified with expected results. The direct dependencies are internet connectivity, availability of AWS services, availability of hosting environment and knowledge of tools or languages. The failure is due to the limitation of hardware elements and this can be eliminated by using appropriate devices. Except the performance issues, model and architecture holds good for small to large scale organizations. The responses to witnessing data give close to real-time experience. Fig. 4. shows the test result summary. Four series of tests were executed manually and its outputs are used for plotting the graph.

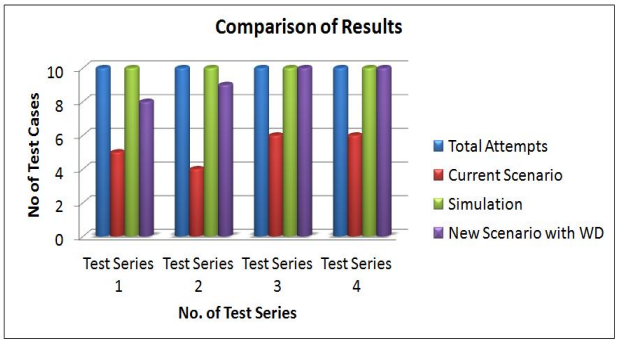

Fig. 4. Comparison of Test Results

Each series of tests contain system test cases repeated for multiple users. In each series there were totally ten test cases included randomly. Here current scenario means that those test cases are executed using only user credentials.

Simulation test results are obtained from UPPAAL. New scenario means those test cases, which considered both user credentials and witnessing devices data for authentication. Current scenario results could be easily breached technically. In initial test series run with new scenario, very simple face recognition and Google location causes bit low accuracy in user identification. The outcomes are considerably improved by using better algorithms and hardware units.

B. *Simulation test results and evaulation*

Proposed system is modeled using UPPAAL. Actual results obtained are as per the expectation. The test cases defined against the requirement of user authentication and authorizations for the standard specifications and Confidentiality, Integrity and Availability (CIA) specifications mentioned in design section were all verified. The Table II., summaries expected and actual simulation test results.

TABLE II. SIMULATION TEST RESULTS

Specifications Verified	Expected Results	Actual Results
Standard	The system shall not get into deadlock state	Verified for deadlock and no such situation is found
	All the states shall be reached	All the states are verified for reachability
Confidentality	Private Data shall be accessible only when regular authentication is successful, data type correctly identified and witnessing devices status are Ok	Private data could be access only when the system satisfied regular authentication, data type identification and witnessing devices
	Private Data shall not be accessible when any of the above conditions are not Ok	Private data could not be access when any one of conditions failed
Integrity	The data in the model shall be accurate and consistent at any state unless until it is modified	Integrity specifications are tested and results were verified for modified and un-modified data
Availability	If the credentials are verified, the data shall be available at any point of time	The availability of both public and private data were verified
	If the credentials are not verified, the data shall not be available	The results are verified for non-availability for unauthenticated users

C. *Experimnetal test results and evaulation*

The quality of proposed model can be evaluated using various types of measurement which can be accuracy or performance. Since this is a POC, quality measurement is limited to accuracy only. Decision support accuracy metrics that are popularly used are Reversal rate, Weighted errors, Receiver Operating Characteristics (ROC) and Precision Recall Curve (PRC), Precision, Recall and F-measure (PRF). After evaluating the feasibility of applying the evaluation techniques, PRF is selected for the evaluation of system accuracy [15].

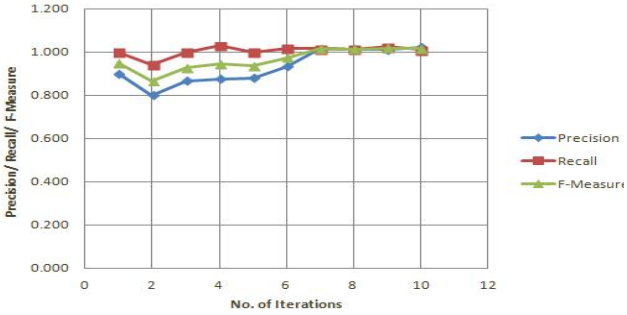

Fig. 5. Experimental Test Results

The test's accuracy of the system is measured using precision, recall F-measure method. In the experiments performed on various algorithms and devices, the proposed model achieves high precision, recall and F-measure. Further the results shown in Fig.5. That, the quality of accuracy increased after seventh iteration because of the usage of better algorithms and devices for image recognition and geo-location identification.

VII. CONCLUSION AND FUTURE SCOPE

Compliance is still a challenging area. The improvisations with the help of IoT help in improving compliance needs. The POC results give a confidence level and could be a solution for the recent reported data breach issues. The architecture can support the needs of compliance for any type of organization with very minimal time and cost impact with definite results. The organization does not need any additional manpower or training towards the implementation. The future scope of the work are (i) machine learning and pattern recognition shall be included to improve the efficiency of compliance checking and assurance and (ii) optimization of the system using better performing hardware or software components.

REFERENCES

[1] Addison-Hewitt Associates, 2002. B2B Consultancy, "The Sarbanes-Oxley Act 2002" [Online]. Available :http://www.soxlaw.com/

[2] Steve Alder, Andy Kelleher, 2017. "HIPAA Journal - latest HIPAA compliance news" [Online]. Available : http://www.hipaajournal.com/

[3] Gerhard Knolmayer and Petra Asprion, "Assuring compliance in IT Subcontracting and Cloud Computing. ", Global Sourcing 2011, LNBIP 91, (pp. 21-45). Springer. 2011 .

[4] Verizon, 2017. "2016 Data Breach Investigations Report." [Online]. Available : http://www.verizonenterprise.com/verizon-insights-lab/dbir/.

[5] Amazon, 2016 "Amazon Web Services Risk and Compliance.". [Online]. Available : http://aws.amazon.com/compliance/awswhitepapers/

[6] David Mitchell, 2016. ICTN 6870 601. "Regulations to Reduce Data Breaches." [Online]. Available :http://www.infosecwriters.com/Papers/DMitchellRegulations.pdf

[7] Martens, Benedikt and Teuteberg, Frank, "Risk and Compliance Management for Cloud Computing Services: Designing a Reference Model." In AMCIS 2011 Proceedings - All Submissions. Paper 228, 2011.

[8] Matthias Flittner, Silvia Balaban, Roland Bless, "CloudInspector: A Transparency-as-a-Service Solution for Legal Issues in Cloud Computing", International Conference on Cloud Engineering Workshops, 2016 @ IEEE.

[9] Jatinder Singh, Thomas Pasquier, Jean Bacon, Hajoon Ko, David Eyers., "Twenty Security Considerations for Cloud-Supported Internet of Things".IEEE INTERNET OF THINGS JOURNAL, VOL. 3, NO. 3, 2016@IEEE.

[10] Kirsty Searle, "Compliance in the spotlight Challenges and opportunities for corporate compliance functions". Deloitte LLP, 2013.

[11] Dinah Barrett, , Chris Astley, "Compliance in the Cloud using Security by Design." KPMG UK LLP, 2016.

[12] Stephane Alberth, Bernhard Babel, Daniel Becker, Georg Kaltonbrunner,, Thomas Poppensiekar, Sebastian Schneider, Uwe Stageman, , Torsten Wegner., "Compliance and Control 2.0". McKinsey Working Papers on Risk, Number 33, 2012 .

[13] G. Behrmann, A. David, and K. Larsen. "A tutorial on Uppaal. Lecture Notes in Computer Science", pages 200–236, 2004.

[14] Ritz Sebastian, Shimmi Asokan, "Process Compliance Checking Using Model Checker" in, Second International Conference on Inventive Communication & Computational Technologies (ICICCT), 2017 @IEEE. DOI: 10.1109/ICICCT.2017.7975220.

[15] F.O. Isinkaye, Y.O. Folajimi, B.A. Ojokoh, " Recommendation systems: Principles, methods and evaluation". Egyptian Informatics Journal Volume 16, Issue 3, Pages 261-273. November 2015,