Using Blockchain to Preserve Chain of Custody (CoC): Cloud Forensics Analysis

Georg Grabner[1*], Ali Ahmed[2], Nilufar Baghaei[3]
[1] School of Computing, University of Liverpool, United Kingdom
[*]Corresponding Author, georg@grabner.nl
[2] School of Engineering and Computer Science, Victoria University of Wellington, New Zealand
[3] School of Information Technology and Electrical Engineering, University of Queensland, Australia

Abstract

In the era of cloud computing, the focus of cyber criminals has shifted from traditional IT infrastructures to the cloud. Cyber forensic investigations in the cloud present several challenges due to legal obligations, technical limitations, and the dynamic nature of cloud resources. For instance, access to digital assets is crucial for practical analysis but is often hindered, as investigators may not have complete access to disk images or log files. The mainstream forensic frameworks and solutions for identifying, retrieving, and preserving evidence in the cloud are lacking. Yet, NIST has identified 65 cloud forensics challenges, with 13 related to log files alone. The preservation and integrity of forensic evidence are the backbones of a digital forensic investigation, and the chain of custody proves the authenticity of the evidence. However, cloud environments are not designed to handle digital evidence with integrity, leading to severe issues when presenting evidence in a court of law. This paper delves into the challenges of cloud forensics investigation, focusing on log files, and aims to identify potential solutions and frameworks to support the chain of custody. For that purpose, the authors evaluate the suitability of blockchain technology for maintaining a proper chain of custody of log files in the cloud. A prototype is implemented to serve as a proof of concept.

1. Introduction

Digital forensics aims to identify, preserve, analyse, and present digital evidence in a legally acceptable manner, with the preservation and retrieval of digital evidence being integral components of the process [1]. In traditional forensic investigations, investigators are granted full access to the devices involved in a crime, such as a PC. Maintaining a proper chain of custody is critical to the process [2]. However, the complexity of preserving evidence and maintaining a chain of custody is heightened in emerging computing environments, such as the cloud and the Internet of Things, where evidence may be stored in many different places [2].

Digital forensics professionals must be familiar with cloud computing technologies and the laws and regulations surrounding data privacy, protection, and access to determine the most effective and legally compliant way to access the needed evidence [3]. Additionally, cloud forensics presents significant challenges, particularly concerning log files and maintaining the integrity and chain of custody [4, 5]. Currently, there are no prevalent solutions or frameworks for cloud forensics.

This paper addresses two research questions: RQ1) What are the challenges related to the chain of custody in cloud forensics investigations, particularly concerning log files? And RQ2) How can blockchain technology be implemented to address these challenges?

2. background and Literature Review

Digital evidence integrity and chain of custody are vital in digital forensics [6]. However, cloud computing's multi-tenant structure and distributed processing present challenges for collecting and preserving digital evidence, compromising the admissibility of evidence in court if the chain of custody is altered [6]. Cloud models also impact the access to resources during a forensic investigation, with SaaS models relying solely on Cloud Service Providers for forensic data [6].

Few frameworks address the challenges of cloud

10.18293/SEKE2023-038

forensics analysis, particularly in maintaining the chain of custody. For instance, a Forensic Readiness Approach suggests a reference architecture consisting of a forensic database component for gathering and storing evidence and a core module that encrypts data and manages the analysis process, including the chain of custody [7]. However, no technical solutions are offered. Similarly, a client-server model captures all data and log files and stores them on a forensic server outside the IaaS cloud environment but fails to address the requirements for preserving a proper chain of custody [4]. Another methodology proposes selecting appropriate cloud services to support the forensic process, focusing on IaaS services, without offering technical solutions [8].

Blockchain is a decentralised technology that uses a distributed peer-to-peer network to store data in a secure ledger without the need for a central authority [9]. Transactions are validated and synchronised by all nodes in the network using cryptographic techniques to ensure the immutability of the data [10,11]. Blockchain applications include cryptocurrency, where it forms the backbone of a secure transaction network.

Several studies have used Blockchain technology to address the challenges of maintaining the chain of custody in forensics [12,13]. For instance, the study in [12] proposes using Blockchain to ensure the immutability of the evidence and the authenticity of the chain of custody process. In their model, evidence is admitted to the network, and all participants are uniquely identified and authorised. Similarly, the study in [13] presents a Java Blockchain implementation for preserving digital evidence in cloud environments.

Forensics in cloud environments pose several challenges, including the absence of direct access to the infrastructure and the volatility of the evidence. NIST has identified 65 issues related to digital forensics in the cloud, and log files and chain of custody are considered the main challenges [14]. Log files, one of the most crucial evidence artefacts in an investigation, must be retrieved, preserved, and maintained securely with their hash values to ensure their integrity. Research in cloud forensics has shown that log files retrieval, preservation, and maintaining the chain of custody are the primary concerns in this field [6, 7, 15, 16].

Blockchain technology offers several benefits for maintaining and tracking the chain of custody in forensic investigations. It provides integrity, transparency, authenticity, security, and auditability, making it a reliable source of evidence in court. The consensus mechanism in blockchain ensures that multiple nodes verify every transaction, maintaining the integrity of the data. Blockchain technology is transparent, allowing easy tracking of the chain of custody. The authenticity of the data is guaranteed as every transaction is verified and authenticated by multiple nodes. The technology also provides a high level of auditability, ensuring the accuracy and integrity of the data stored on the blockchain.

In [12], a blockchain-based solution for digital forensics chain of custody is presented, with a prototype created using Hyperledger Composer and acceptable performance results. However, the work lacks a complete and optimised end-to-end solution. Similarly, the work in [17] explores the potential of blockchain technology in supporting digital forensics and investigations, discussing various blockchain implementations and introducing the concept of Digital Witnesses. While the authors conclude that blockchain-based CoC offers a new level of forensic readiness, more work is needed in data governance and standardisation. Finally, [18] presents a blockchain-based system for secure chain-of-custody transfer and record-keeping, which offers tamper-proof transaction records and increased process efficiency. However, the system does not store the actual evidence.

In [19], a digital chain of custody framework is proposed to integrate blockchain technology into digital forensics. The framework uses smart locks to store evidence, private blockchain to store evidence metadata, and a peer-to-peer network for communication. The framework is implemented using Ethereum nodes, and its performance is evaluated, showing acceptable transaction throughput. Similarly, [20] proposed a blockchain-based e-Chain-of-Custody (e-CoC) ledger, managed by a trusted entity, to ensure the integrity of digital evidence in cyber investigations. The e-CoC ledger provides a tamper-proof record of the chain of custody of digital evidence and can be easily implemented for forensic software developers. While some blocks are sent to a secure public blockchain, the work highlights the need for further research into data governance and standardisation of the admissibility of digital evidence.

In [21], an Information Chain of Custody Model based on blockchain technology is proposed to ensure privacy for sensitive health data. Limitations were identified, but the proposed solution provides traceability and control over information by the owner. The work in [22] proposes MF-Ledger, a blockchain-enabled digital multimedia forensics investigation architecture that addresses issues and challenges in digital forensic investigations. The implementation of MF-Ledger shows the potential of blockchain technology to protect and secure the digital forensic chain of custody. In [23], the authors suggest standardising smart contracts for

a secure and reliable chain of custody process, but note that a holistic approach, including robust evidence and participant management, is necessary. The work in [24] evaluates the effectiveness of fuzzy hashing algorithms in preserving the integrity of digital evidence in image forensics compared to conventional cryptographic hash algorithms, showing that fuzzy hash-based blockchain effectively supports the chain of custody process.

3. Proof of concept

A Chain of Custody process must adhere to the generally accepted four principles of digital evidence namely 1) **Principle 1**: No action taken by anyone should change the evidence, 2) **Principle 2**: In circumstances where access to the original data is needed it must be clear what the implications will be, and provide a statement of evidence why these actions have to be done. Also, the individual must be competent to handle the data, 3) **Principle 3**: An audit trail of all events or processes being executed on the evidence must be created and preserved and be able to be examined by an independent third party. And must also be able to achieve the same results with the same process; 4) **Principle 4**: The person in charge of the investigation must ensure the application of these principles.

Given our investigations, the following requirements regarding a chain of custody process are identified 1) Integrity: the evidence has not been altered or corrupted during the transfer; 2) Traceability: the evidence must be traced from the time of its collection until it is destroyed, 3) Authentication: all the entities interacting with a piece of evidence must provide an irrefutable sign as recognisable proof of their identity, 4) Verifiability: the whole process must be verifiable from every entity involved in the process, and 5) Tamper-proof: Changeovers of evidence cannot be altered or corrupted. To demonstrate those requirements and based on the principles outlined before, a prototype was developed with three main components: a private blockchain, an evidence retrieval agent, and an evidence storage and registration agent. The overall design of the prototype is shown in Figure 1.

The prototype has been implemented using Python on Ubuntu 18.04.3 TLS. For the Blockchain part, Ethereum and Hyperledger were used due to the availability of good documentation and community support. It is worth noting that the Hyperledger blockchain implementation is based on Hyperledger fabric 1.2 with Composer 2.0. The prototype uses the OwnCloud [1] platform as a PaaS server from which the forensic evidence in the form of log files is retrieved.

[1] https://owncloud.com

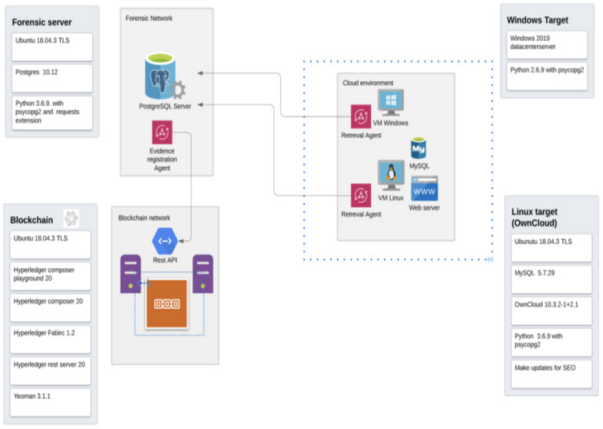


Figure 1. Overall Prototype Logical View

Part of the evaluation of the prototype is to code-review the retrieval and registration agents (i.e. static analysis). Figure 2 demonstrates the results of the code quality review on the registration agent component. The retrieval agent component is also analysed with SonarQube, and the issues found are shown in Figure 3.

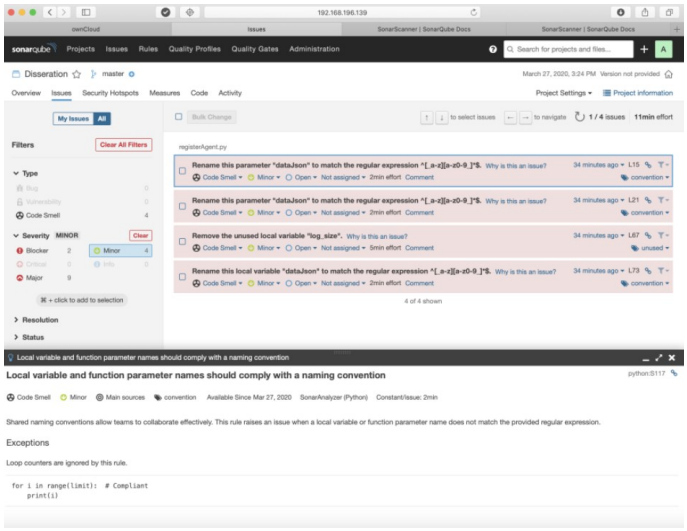


Figure 2. SonarQube Analysis of the Register Agent.py

Security-wise, various serious problems were discovered. The root cause of these issues is that the authentication details were embedded in the scripts, which is highly dangerous and must not be allowed in operational settings. A part of the technical evaluation entails analysing the system's operational efficiency. To

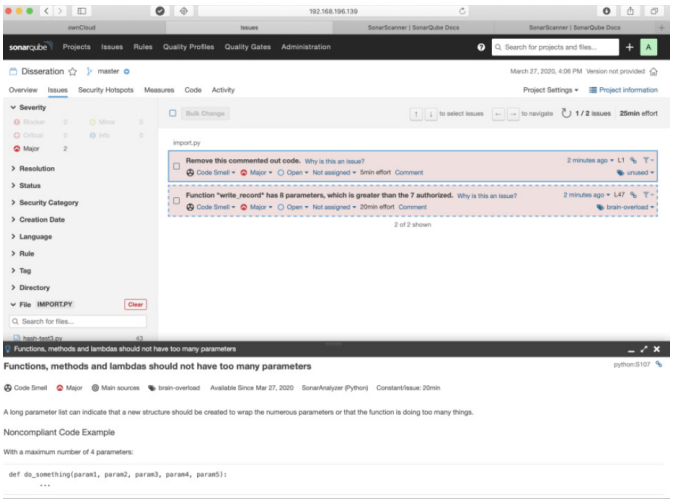
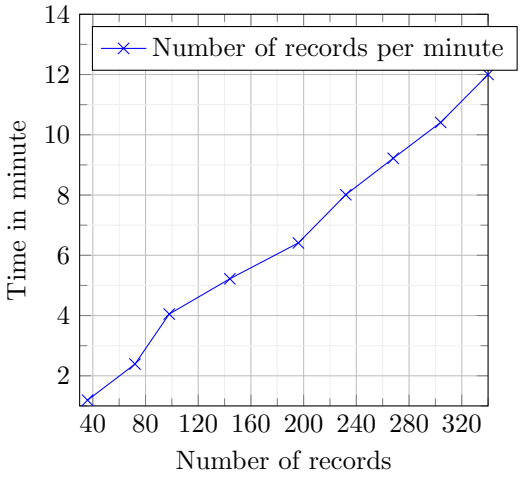


Figure 3. SonarQube Analysis of the Retrieval Agent Import.py

this end, thirty-six evidence records are sent to the Hyperledger REST interface in batches. The time taken to record these entries onto the blockchain is measured, as shown in Figure 4.

Figure 4. Run-time Performance Analysis



An experiment was conducted to assess the impact of record size on the system's performance during runtime, as depicted in Figure 4.

The experiment yielded two expected outcomes. Firstly, the processing time also increased as the number of records increased. This is because blockchain operations are computationally intensive and require more time to compute hash values as the number of log files increases. Secondly, we expected that the system's overhead would remain consistent regardless of the size of the records. The results confirmed our expectations, as the overhead introduced by the system remained consistent regardless of record size. The results in Figure 5 align with the average latency of Ethereum, and Hyperledger reported in [25]. This is a promising finding as it suggests that the system's additional functions for evidence gathering and handling are lightweight and do not significantly impact system performance. Overall, the results of this experiment demonstrate the importance of considering record size when assessing the performance of blockchain-based systems for forensic investigations.

The purpose of the prototype was to demonstrate proof of concept, not to create a fully functional system. Hence, code review and run-time performance analysis were deemed sufficient evaluation criteria. Given the developed system, its evaluation, and the literature survey in this paper, the original research questions are answered as follows. ***RQ1 - What are the current challenges regarding the chain of custody, particularly concerning log files in the cloud forensics investigation?***
The identification and preservation phase and maintaining the chain of custody of log files in cloud environments pose several challenges, per the literature survey and examination of prototype results. These challenges include: 1) Volatility of data: The data in cloud environments is highly volatile due to the elastic nature of the cloud, making it challenging to preserve evidence as virtual machines or docker images are removed or deployed; 2) Overwhelming amount of logging data: In an enterprise IaaS environment with hundreds or thousands of virtual machines, each virtual Windows machine alone can generate 86 different log files, leading to a significant challenge in monitoring and preserving all log files, 3) Multi-tenancy and Privacy concerns: In multi-tenant environments like cloud, isolating and preserving evidence without interfering with other tenants' data or processes is challenging, along with privacy concerns, 4) Inaccessibility and Multi-Jurisdiction issues: In cloud environments, investigators do not have unrestricted or physical access to cloud storage, making it difficult to determine the location of physical data due to the distributed nature of cloud storage. Additionally, suppose the data resides in a different geographical location. In that case, jurisdiction also becomes a challenge, and 5) Integrity of log data and lack of logging standard: Maintaining the integrity of log data during preservation is challenging due to the different logging formats of cloud infrastructure and applications and the lack of standardisation on the minimum requirements for logging and retrieval

of log file metadata. To ensure integrity, SHA-256 is recommended.

The proposed solutions in [12, 26] primarily concentrate on preserving the chain of custody in the evidence handling process during the investigation phase, not during the retrieval and evidence registration phase, where evidence is submitted manually to the blockchain. On the other hand, solutions like those in [4, 27] focus on the automatic retrieval and storage of logs. However, these solutions do not specifically address the challenges in the chain of custody process.

RQ2 - How can the characteristics of blockchain technology provide solutions for these challenges, and in what way can blockchain technology be implemented to address these challenges?

The theoretical ideal solution for maintaining the Chain of Custody is a blockchain due to its distributed nature and cryptographic chaining of each transaction. This provides immutability and provenance for each evidence record and the necessary audit trails requirements such as integrity, transparency, authenticity, security, and auditability of digital evidence [12]. For a blockchain-based forensic network, a private blockchain is the best option, as it can be placed outside the client cloud environment under the control of a third-party investigator or auditor. Regarding blockchain implementations, Hyperledger Fabric and Ethereum are the best and most widely used open-source options. Different parties, such as cloud service providers, customers, and investigators, can be implemented as participating organisations within the blockchain, and strict security measures can be applied to restrict participant actions.

4. Conclusion and Future work

The current design of cloud environments does not prioritise the integrity and handling of digital evidence. However, incorporating Blockchain technology holds promise in addressing the challenges of maintaining a chain of custody for log-file evidence in the cloud. The proposed system addresses the challenge of volatile data by storing the log-file contents in a database. The system acknowledges that the large volume of data still poses a challenge but attempts to address multi-tenancy and privacy by assigning unique IDs for virtual machines and clients and storing evidence records in separate databases or servers for each customer.

The proposed solution partially addresses the challenges of inaccessibility and multi-jurisdiction by retrieving the evidence to a forensic server database accessible to the investigators. The implementation also partially solves the issues of log-data integrity and the lack of a logging standard by using a unique evidence ID and hash generated during log-file retrieval and saving metadata with the evidence record. The forensic server also records information about who, when, and where the evidence was retrieved. Despite this, the large volume of log data remains a challenge, but blockchain implementations such as Bitcoin [10, 11] have demonstrated their ability to handle vast amounts of transactions.

Future work in this research field should focus on finding solutions to the high latency experienced by the system as the number of records increases dramatically, which is crucial in digital forensics investigations where data being acquired could be in terabits.

References

[1] J. Benner, "Establish a transparent chain-of-custody to mitigate risk and ensure quality of specialized samples," *Biopreservation and biobanking*, vol. 7, no. 3, pp. 151–153, 2009.

[2] G. Kumar, R. Saha, C. Lal, and M. Conti, "Internet-of-forensic (iof): A blockchain based digital forensics framework for iot applications," *Future Generation Computer Systems*, vol. 120, pp. 13–25, 2021.

[3] O. Aktera, A. Aktherb, M. A. Uddinc, and M. M. Islamd, "Cloud forensics: Challenges and blockchain based solutions," *International Journal of Wireless and Microwave Technologies*, pp. 1–12, 2020.

[4] M. E. Alex and R. Kishore, "Forensics framework for cloud computing," *Computers & Electrical Engineering*, vol. 60, pp. 193–205, 2017.

[5] P. Purnaye and V. Kulkarni, "A comprehensive study of cloud forensics," *Archives of Computational Methods in Engineering*, pp. 1–14, 2021.

[6] S. Simou, C. Kalloniatis, S. Gritzalis, and H. Mouratidis, "A survey on cloud forensics challenges and solutions," *Security and Communication Networks*, vol. 9, no. 18, pp. 6285–6314, 2016.

[7] L. De Marco, N.-A. Le-Khac, and M.-T. Kechadi, "Digital evidence management, presentation, and court preparation in the cloud," *Security, Privacy, and Digital Forensics in the Cloud*, p. 283, 2019.

[8] S. Simou, C. Kalloniatis, S. Gritzalis, and V. Katos, "A framework for designing cloud forensic-enabled services (cfes)," *Requirements Engineering*, vol. 24, no. 3, pp. 403–430, 2019.

[9] S. Nakamoto and A. Bitcoin, "A peer-to-peer electronic cash system," *Bitcoin*, vol. 4, 2008.

[10] P. Tschannen and A. Ahmed, "On the evaluation of the security usability of bitcoin's apis," in *Proceedings of the Evaluation and Assessment in Software Engineering*, ser. EASE '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 405–412.

[11] P. Tschannen, , and A. Ahmed, "Bitcoin's apis in open-source projects: Security usability evaluation," *Electronics*, vol. 9, no. 7, 2020.

[12] A. H. Lone and R. N. Mir, "Forensic-chain: Blockchain based digital forensics chain of custody with poc in hyperledger composer," *Digital Investigation*, vol. 28, pp. 44–55, 2019.

[13] M. Pourvahab and G. Ekbatanifard, "Digital forensics architecture for evidence collection and provenance preservation in iaas cloud environment using sdn and blockchain technology," *IEEE Access*, vol. 7, pp. 153 349–153 364, 2019.

[14] M. Herman, M. Iorga, A. M. Salim, R. H. Jackson, M. R. Hurst, R. Leo, R. Lee, N. M. Landreville, A. K. Mishra, Y. Wang *et al.*, "Nist cloud computing forensic science challenges," National Institute of Standards and Technology, Tech. Rep., 2020.

[15] H. Alobaidli, Q. Nasir, A. Iqbal, and M. Guimaraes, "Challenges of cloud log forensics," in *Proceedings of the SouthEast Conference*, 2017, pp. 227–230.

[16] R. Neware and A. Khan, "Cloud computing digital forensic challenges," in *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*. IEEE, 2018, pp. 1090–1092.

[17] H. Al-Khateeb, G. Epiphaniou, and H. Daly, *Blockchain for Modern Digital Forensics: The Chain-of-Custody as a Distributed Ledger*. Cham: Springer International Publishing, 2019, pp. 149–168.

[18] M. Chopade, S. Khan, U. Shaikh, and R. Pawar, "Digital forensics: Maintaining chain of custody using blockchain," in *2019 Third International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, 2019, pp. 744–747.

[19] L. Ahmad, S. Khanji, F. Iqbal, and F. Kamoun, "Blockchain-based chain of custody: Towards real-time tamper-proof evidence management," in *Proceedings of the 15th International Conference on Availability, Reliability and Security*, ser. ARES '20. New York, NY, USA: Association for Computing Machinery, 2020.

[20] X. Burri, E. Casey, T. Bollé, and D.-O. Jaquet-Chiffelle, "Chronological independently verifiable electronic chain of custody ledger using blockchain technology," *Forensic Science International: Digital Investigation*, vol. 33, p. 300976, 2020.

[21] W. Silva and A. C. B. Garcia, "Where is our data? a blockchain-based information chain of custody model for privacy improvement," in *2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, 2021, pp. 329–334.

[22] A. A. Khan, M. Uddin, A. A. Shaikh, A. A. Laghari, and A. E. Rajput, "Mf-ledger: Blockchain hyperledger sawtooth-enabled novel and secure multimedia chain of custody forensic investigation architecture," *IEEE Access*, vol. 9, pp. 103 637–103 650, 2021.

[23] P. Santamaría, L. Tobarra, R. Pastor-Vargas, and A. Robles-Gómez, "Designing the chain of custody process for blockchain-based digital evidences," in *Blockchain and Applications, 4th International Congress*, J. Prieto, F. L. Benítez Martínez, S. Ferretti, D. Arroyo Guardeño, and P. Tomás Nevado-Batalla, Eds. Cham: Springer International Publishing, 2023, pp. 225–236.

[24] M. Ali, A. Ismail, H. Elgohary, S. Darwish, and S. Mesbah, "A procedure for tracing chain of custody in digital image forensics: A paradigm based on grey hash and blockchain," *Symmetry*, vol. 14, no. 2, p. 334, Feb 2022.

[25] S. Pongnumkul, C. Siripanpornchana, and S. Thajchayapong, "Performance analysis of private blockchain platforms in varying workloads," in *2017 26th International Conference on Computer Communication and Networks (ICCCN)*, 2017, pp. 1–6.

[26] H. Al-Khateeb, G. Epiphaniou, and H. Daly, "Blockchain for modern digital forensics: The chain-of-custody as a distributed ledger," in *Blockchain and Clinical Trial*. Springer, 2019, pp. 149–168.

[27] A. Păţraşcu and V.-V. Patriciu, "Logging framework for cloud computing forensic environments," in *2014 10th International Conference on Communications (COMM)*. IEEE, 2014, pp. 1–4.