IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 13 November 2022, accepted 12 January 2023, date of publication 25 January 2023, date of current version 30 January 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3239814

RESEARCH ARTICLE

A Secure and Privacy-Preserving E-Government Framework Using Blockchain and Artificial Immunity

NOE ELISA[1,5], (Student Member, IEEE), LONGZHI YANG[1], (Senior Member, IEEE), FEI CHAO[2], (Member, IEEE), NITIN NAIK[3], (Member, IEEE), AND TOSSAPON BOONGOEN[4]

[1]Department of Computer and Information Sciences, Northumbria University, NE1 8ST Newcastle upon Tyne, U.K.
[2]Department of Artificial Intelligence, School of Informatics, Xiamen University, Xiamen 361005, China
[3]School of Informatics and Digital Engineering, Aston University, B4 7ET Birmingham, U.K.
[4]School of Information Technology, Mae Fah Luang University, Chaing Rai 57100, Thailand
[5]Department of Computer Science and Engineering, The University of Dodoma, Dodoma 41218, Tanzania

Corresponding author: Longzhi Yang (longzhi.yang@northumbria.ac.uk)

This work was supported in part by the Commonwealth Scholarship Commission under Grant CSC-TZCS-2017-717, and in part by the Royal Academy of Engineering Industry Academia Partnership Programme under Grant IAPP1\100077.

ABSTRACT Electronic Government (e-Government) systems constantly provide greater services to people, businesses, organisations, and societies by offering more information, opportunities, and platforms with the support of advances in information and communications technologies. This usually results in increased system complexity and sensitivity, necessitating stricter security and privacy-protection measures. The majority of the existing e-Government systems are centralised, making them vulnerable to privacy and security threats, in addition to suffering from a single point of failure. This study proposes a decentralised e-Government framework with integrated threat detection features to address the aforementioned challenges. In particular, the privacy and security of the proposed e-Government system are realised by the encryption, validation, and immutable mechanisms provided by Blockchain. The insider and external threats associated with blockchain transactions are minimised by the employment of an artificial immune system, which effectively protects the integrity of the Blockchain. The proposed e-Government system was validated and evaluated by using the framework of Ethereum Visualisations of Interactive, Blockchain, Extended Simulations (i.e. eVIBES simulator) with two publicly available datasets. The experimental results show the efficacy of the proposed framework in that it can mitigate insider and external threats in e-Government systems whilst simultaneously preserving the privacy of information.

INDEX TERMS E-Government, blockchain, artificial immune system, insider threat, privacy-preserving.

I. INTRODUCTION

E-Government uses digital technologies to deliver public services to individuals, agencies, businesses, and other affiliates in order to improve efficiency, participation, accountability, transparency, and shared responsibilities with various stakeholders [1]. This significantly improves the inclusiveness of government services by ensuring full access to services without the need for physical visits, among other advantages. In general, e-Government is one of the most complex information systems, requiring efficiency, security, and privacy protection [2], [3]. However, various privacy and security breaches are frequently reported around the world as a result of, amongst others, the disclosure of sensitive information, inappropriate sharing and mishandling of private information, and sophisticated attacks on e-Government systems [3], [4].

Most existing commonly used e-Government systems, such as websites and electronic identity management systems (eIDs), are centralised, with all data processed and computed through central servers [2], [4]. Centralised services frequently have a single point of failure, making the systems

The associate editor coordinating the review of this manuscript and approving it for publication was Chi-Yuan Chen.

vulnerable to cyber attacks such as malware, worms, denial of service (DoS), and distributed denial of service attack (DDoS). Furthermore, insider threat is becoming an increasingly critical challenge in many organisations around the world, including e-Government systems; because it originates from a trusted account, it cannot be detected using external security measures such as firewalls [5]. According to a recent insider threat survey published in 2019, 20% of cyber security attacks and 15% of information theft were initiated by insiders within an organisation, with a single insider costing an organisation an average of $11.45 million per year [6].

This paper reports a blockchain-based decentralised secure and privacy-preserving e-Government framework. Blockchain has recently emerged as a key technology for secure data sharing and storage in trust-free and decentralised systems [7], [8], [9], [10], [11], [12]. It enables the development of highly secure and privacy-preserving decentralised applications in which information is not controlled by a centralised host or third parties. Transactions are encrypted and stored in linked blocks (i.e. ledgers), which are distributed across the network in a verifiable and immutable manner using blockchain [13]. This means that once information is added to the chain, it cannot be removed or changed in the future [14]. Because of the immutability nature of blockchain, adding invalid transactions must be avoided [10]. Unwanted traffic, such as spyware, worms, ransomware, and spam, can be extremely expensive and financially disastrous [2], [6]. As a result, such traffic must be identified and prevented from being added to the e-Government blockchain.

As a result, this work proposes an anomaly detection system for identifying and mitigating unwanted traffic in e-Government systems using artificial immune systems (AISs). In a nutshell, an AIS is a computational model created by simulating the behaviour and operation of the biological human immune system. Given that the biological immune system is essentially a decentralised system that functions through agents, the application of AISs in a decentralised e-Government system is therefore appealing to maximally realise the benefits of a fully decentralised system. One particular implementation of AIS is dendritic cell algorithm (DCA), which has been successfully applied for anomaly detection in computer networks with competitive performances demonstrated [15], [16], [17]. DCA works naturally with streaming data, such as network traffic, and exhibits useful properties such as self-organisation, scalability, and adaptability [16]. Consequently, DCA is adopted to the proposed e-Government system herein, but the application of other decentralised intrusion detection approaches and corresponding comparative studies of these approaches, remains as a piece of future work.

The proposed framework was validated and evaluated using the Ethereum Visualisations of Interactive, Blockchain, Extended Simulations (i.e. eVIBES simulator) [18]. The open source eVIBES simulator offers off-chain (sideDB) data storage, which is crucial for e-Government systems since it allows for the storing of items like contacts, photos, and other data/information that are too large to be saved in the blockchain or that must be destroyed or updated in the future. Furthermore, Ethereum is widely used to implement blockchain applications, and comparable systems to the one proposed herein are highly likely to be implemented using the Ethereum protocol, facilitating a fair comparative study with related objectives. The simulated framework was tested using two publicly available datasets, including CERT [5] and UNSW_NB15 [19].

The experimental results confirm the efficacy and competitiveness of the proposed e-Government framework in terms of both efficient decentralised Governmental services and effective secure and privacy-preserving responses to breaches and threats, with the added benefit of potentially increasing trust and accountability of public services due to the transparency provided by blockchain. The contribution of this paper is threefold: 1) proposing a decentralised e-Government framework that innovatively integrates consortium blockchain and DCA in the framework, 2) designing and implementing algorithms for the operations of consortium blockchain to allow effective and efficient e-Government services, and 3) embedding DCA for internal and external intrusion detection as an extra layer for secure and privacy-preserving e-Government services based on consortium blockchain.

The rest of this paper is structured as follows. Sec. II presents the theoretical background and relevant applications of this study. Sec. III details the proposed e-Government framework. Sec. IV reports the performance evaluation and discusses the experimental results. Sec. V concludes this study and points out probable future work.

II. BACKGROUND

A. E-GOVERNMENT SYSTEMS

The advancement of e-commerce systems, which has shifted the focus of economy from goods to services through the use of information and communications technology, is primarily responsible for the development and adoption of e-Government systems [20]. Almost all countries have created websites to convey information to their citizens and other stakeholders, according to a United Nations report on the evolution of e-Government [21]. A citizen-centred, business-focused, and environmentally conscious e-Government system can result in increased transparency and convenience, increased revenue and efficiency, and reduced corruption and operational costs. [21].

Digital identity (eID) is a critical e-Government service that allows individuals to be verified when accessing services from various Government departments [22]. The eID is a simple online method for citizens, businesses, and other organisations to electronically prove their identities. An individual's eID can be used in a variety of sectors, including taxation, national insurance, education, telephony services, banking services, and so on, as well as to fulfil various roles such as civil servant, lawyer, and so on, depending on the context. eIDs can also be employed to authenticate and authorise citizens to use e-services outside of their home countries. The

task of issuing and validating eID is typically assigned to a single organisation, which is also in charge of the dissemination of information to other government departments or nations [22].

Each department or agency must implement efficient access control due to the sensitivity of the information contained in e-Government networks. To protect and keep e-Government systems operating properly, information security technology must be used. E-Government networks must be properly protected to guarantee the security, integrity, and availability of the information or data [23]. It should be noted that in existing e-Government systems, information or data collected from individuals, businesses, and organisations is almost always stored in centralised databases and servers [2], [23].

E-Government systems are frequently classified into four groups based on their interaction and interdependence with their users: Government to Citizen (G2C), Government to Business (G2B), Government to Government (G2G), and Government to Employees (G2E) [1]. The G2C entails the interaction between Government and citizens by using online electronic applications. The G2B involves the interaction between the Government and business firms in an effort to provide more transparency and better business environments. The G2G comprises interactions between Government departments, authorities and agencies locally, regionally or nationally in order to share the information and services available amongst public bureaucracies. The G2E supports the interaction between the Government and its employees by using online applications to make their communications more effective and efficient.

B. BLOCKCHAIN

Blockchain is a peer-to-peer (P2P) distributed database (also known as a ledger) that keeps track of a list of ever-expanding records called blocks that are connected linearly and chronologically and secured by utilising public-key cryptography and cryptographic hashing [13]. Instead of adding to the centrally maintained database in a standard centralised system, such technology adds new information to a block and makes it accessible to all nodes in the distributed network. Although blockchain was primarily developed to share digital currency, with Bitcoin serving as a representative example [13], it has evolved far beyond financial transactions and can now record any type of information or data, such as self-executing digital smart contracts powered by Ethereum [24], as well as general enterprise solutions based on IBM Hyperledger Fabric [25]. Decentralisation, transparency, and immutability are the three fundamental characteristics of blockchain that make it incredibly safe, reliable, and impermeable.

1) BLOCKS

A typical block is composed of a header and a list of transactions performed in that block, as shown in Fig. 1. The block header contains metadata such as the time stamp, nonce, and version. The time stamp indicates when the block was created; the nonce is a random number generated by a consensus algorithm for the computation of the hash value of a block; and the version is the version number of the blockchain. It is worth noting that each block contains references to the previous block hash (or parent) and the next block hash (or child), allowing a chain of blocks to be formed from the first to the current block, as illustrated in Fig. 1. These hash values are generated by hashing nonces typically with the secure hash algorithm (256 bits) (SHA256).

The blockchain application hard-codes the first block, known as the genesis block, by inserting some random data [13]. Whilst there are only one parent and one child for each block, a valid block may momentarily have two or more children if many network peers append blocks at the same time, creating multiple branches from the same parent [14]. This condition is known as a ''fork'' and can be resolved by designating the chain that eventually outpaces the others as the valid blockchain and declaring all other shorter chains invalid (i.e. orphan). If the formed branches are all of the same length, the process of adding new blocks for all the to-be-validated chains continues until one branch becomes longer than the others and thus valid.

A Merkle tree is used to connect all transactions within a block [13], which is an inverted binary tree. To build a Merkle tree, pairs of transactions are hashed recursively until they form only one root node at the top of the tree, known as the Merkle root [13], as shown in the lower part of Fig. 1. More precisely, a Merkle root is the hash of all the transactions that comprise a block in a blockchain network. Any minor change to the transaction data will cause the Merkle root hash to change, resulting in an invalid record. If the number of transactions is odd, the last transaction hash is duplicated to create an even number of transactions, resulting in a balanced tree. Because the hash value of the current block header is linked and stored in the next block, any change to a block will result in a different hash, which will be propagated throughout the network to invalidate that block [13]. Based on this technique, the blockchain is decentralised and distributed and does not require an intermediary or trusted third party to monitor and validate the transactions.

The private keys provided to the blockchain participants are used to digitally sign and verify the transactions they have actually made. Since the blockchain is immutable, as was already mentioned, once data is added into the network it cannot be altered or removed. Therefore, a blockchain is extremely difficult to hack due to the connectivity and share of all transactions across the network. The precise number of nodes that must be compromised in order to successfully hack a blockchain depends on the chosen consensus process, as is briefed in the following subsection.

2) CONSENSUS MECHANISM

To validate transactions, nodes in a blockchain network run a consensus algorithm together. Several consensus algorithms are readily available, including Proof of Work (PoW), Proof of Stake (PoS), Delegated Proof of Stake (DPoS), and Proof

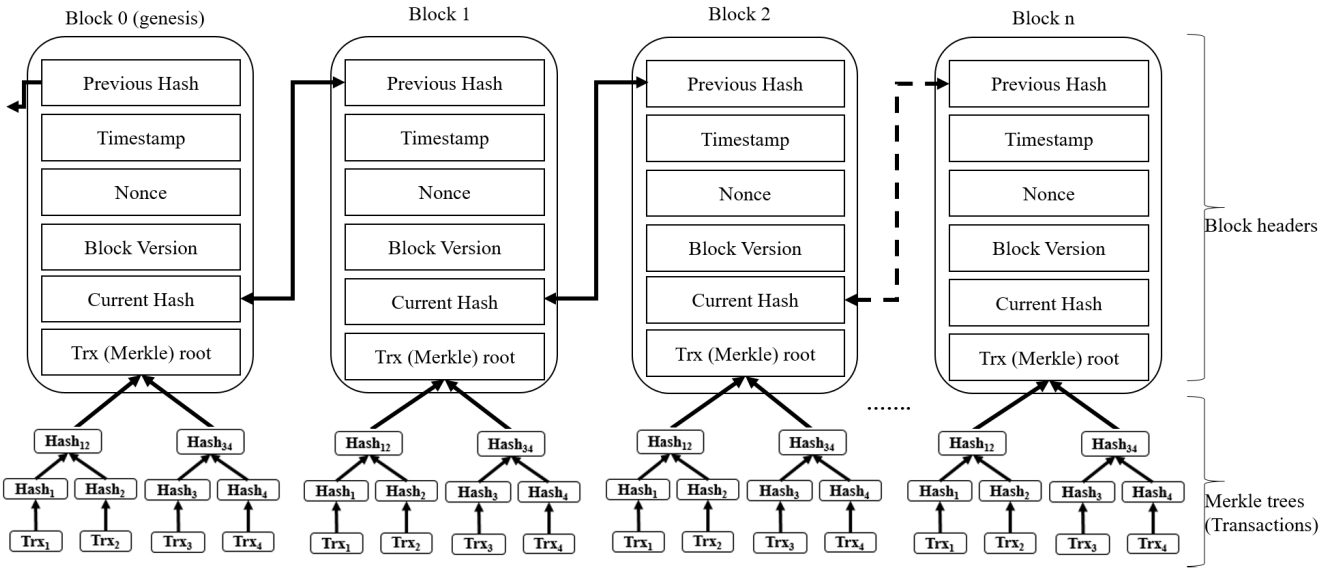


FIGURE 1. An example ledger of incorruptible blocks [14].

of Difficulty (PoD) (PoD) [14], [26], as well as the Byzantine Fault Tolerance (BFT) algorithm. For example, Bitcoin uses PoW, whereas Ethereum and Bitshare employ PoS and DPoS, respectively [26]. When it comes to PoW and Bitcoin, the nodes who want to add (or mine) a new block to the blockchain network are known as miner nodes. To do this, they must first solve a challenging mathematical puzzle that calls for a lot of computing power. The first miner to finish the riddle will add a new block and receive Bitcoin prizes [14].

In contrast to PoW, PoS selects the node that builds a new block deterministically based on its stake (or wealth) [26]. PoS conserves the energy used in PoW to solve the mathematical puzzle, and the only requirement for a node to be a validator of new transactions and blocks is the wealth of the node. The DPoS uses delegates in an effort to address the issue of consensus [26]. A panel of few trusted delegates who will witness and validate the blocks are created by DPoS using a real-time voting and reputation mechanism. Along with forbidding malevolent nodes from adding blocks, the witnesses also have the authority to generate new blocks and add them to the blockchain network. As a general rule in PoS and DPoS, network stakeholders are not supposed to purposefully make poor judgments for the network.

3) TYPES OF BLOCKCHAINS

There are four types of blockchains: public (or permissionless), private (or permissioned), consortium (or semi-public and semi-private) and hybrid.

a: PUBLIC BLOCKCHAIN

The public blockchain is available for any node to see, edit, and audit. A decentralised consensus method, like PoW employed by Bitcoin, is used to reach consensus and make decisions [13]. Participants who can add new transactions to the distributed ledger are determined by the computational power of network nodes. Participants are incentivised every time when new transactions are added to the blockchain network, effectively motivating users to spend more computational effort in order to have a chance to add transactions to the ledger. The more users a public blockchain network has, the more secure the network is because it creates a network of trusted individuals between the participants.

b: PRIVATE BLOCKCHAIN

A private blockchain is typically owned by a single organisation that is in charge of granting new users access to the network. Only a few individuals within the organisation have the authority to validate transactions and blocks before adding them to the blockchain network. Although it is more centralised than a public blockchain, the computational power required in a private blockchain network is significantly lower.

c: CONSORTIUM BLOCKCHAIN

A consortium blockchain is managed by a pre-selected group of nodes that are in charge of resource access control [12]. The goal of a consortium blockchain is to eliminate the private blockchain's individual/single autonomy by having multiple entities or organisations perform consensus for the benefit of the entire network of peers. Incentives are not required for this network because only pre-selected nodes are permitted to validate transactions and perform consensus. Because of the pre-selected set of nodes, it is partially private, partially public, and partially decentralised. As a result, it offers the benefits of public blockchain in terms of efficiency and scalability whilst still allowing for some central safeguarding and monitoring, as used in private blockchain. Consortium blockchains, such as Hyperledger fabric amongst others [25], are intended to meet the needs of businesses in which a group of cooperating agencies use blockchain to

improve the delivery of services. All consortium blockchain consensus participants are well-known and reputable, effectively preventing malicious users from participating.

d: HYBRID BLOCKCHAIN

A hybrid blockchain combines the advantages of a private and public blockchain to create a peer-to-peer network that is secure, transparent, and privacy-preserving. Participants in a hybrid blockchain can freely join the blockchain network and participate in the consensus process. In general, the participants of a hybrid blockchain agree on which data should be kept private and which must be made public when necessary. As a result, network participants determine access privileges and implement information access control. A hybrid blockchain adds transparency to business operations without jeopardizing the security and privacy of information shared among participants; however, it also has a private network element for important information that must be kept private, though any information created in the network is verified by the public network to maintain transparency.

C. BLOCKCHAIN FOR E-GOVERNMENT

IoT, smart homes and cities, educational systems, supply chains, Industry 4.0, and healthcare are just a few domains and applications where blockchain has been extensively applied for security, trust, and privacy preservation [27], [28], [29], despite the fact that it was initially developed for transferring digital currencies. To investigate the potential of blockchain technology in providing effective public services to people and organisations, many countries throughout the world have proposed a variety of blockchain initiatives [30], as summarised in Table 1. These initiatives typically each concentrate on a specific electronic online service, such as e-health, e-land registration, or e-residency, and each of these systems is created independently.

The systematic application of blockchain in e-Government systems is still in its early stages, and there is no common e-Government framework that effectively integrates all e-services, security measures, and so on into a single system [30], despite the promising aspects as listed in Table 1. The blockchain-based e-Government platforms created by many nations may make it harder for people to communicate across international borders for information sharing and collaboration. In fact, the blockchains proposed in these initiatives are either permissioned (private) or permissionless (ublic) [31]. Note that consortium blockchain is invented to meet the needs between collaborative organisations, which is exploited in this study for a decentralised, secure, and privacy-preserving e-Government framework to support e-Government internationally.

D. THREAT DETECTION

1) CYBER THREATS

Since the widespread use of the Internet in the 1990s, external cyber threats have been the focus of cybersecurity research. These are incoming network traffics that deviate from what is set as a normal network behaviour and are typically carried out by an outsider who wants to gain access to the network resources illegally or unethically [5], [17]. Intrusion detection systems (IDSs) are a common measure to detect external attacks [38], [39]. IDSs can be divided into two basic categories: anomaly-based IDS (AIDS) and signature-based IDS (SIDS) [17], [38]. SIDS pre-defines specific abnormal patterns (i.e. signatures) of a system in advance, and any new matched incoming traffic triggers an alarm for attention. The critical limitation of an AIDS is its reliance on up-to-date signatures, which means it cannot detect zero-day attacks. An AIDS, on the other hand, operates on a set of rules that define normal behaviour, and any network traffic that deviates from the rules is treated as an anomaly. AIDSs can thus detect novel abnormal patterns, including zero-day attacks. Many artificial intelligence techniques, such as fuzzy interpolation [40], [41], AIS [17], and artificial neural networks [42], have been employed to develop AIDSs.

The insider threat, which refers to malicious actions performed by insiders within an organisation through their authorised accounts with the intention of causing information theft, electronic fraud, or system sabotage, is gaining more attention these days. Insider threats can take many forms, including disgruntled employees, consultants, or officers within organisations [5]. IDSs and firewalls are typically used to protect organisation networks from external threats, but because insider threats usually originate from trusted accounts, they cannot be detected by these externally-facing security measures [5]. The more money an insider threat incident usually costs the organisation, the longer it goes unnoticed [6].

Insider threat identification is regarded as a highly difficult undertaking in organisations of all sizes due to the nature and sophistication of such threats. However, insiders' suspicious behaviour is frequently used as a precursor to potential insider threats [5], [6], such as downloading or accessing huge amounts of private data over the company network, or copying files from private folders over the company network. Support vector machines and deep learning, amongst other machine learning and artificial intelligence techniques, have been created as a possible alternative to these conventional precautionary measures that successfully identify, contain, and discourage insider threats [5].

2) IDS USING ARTIFICIAL IMMUNE SYSTEMS

AIS is a subset of computational intelligence methods that takes its cues from the biological immune system's innate defences and is intended to address engineering issues related categorisation, optimisation, and anomaly detection [15]. Dendritic cells (*DCs*) in the natural immune system are in charge of gathering antigens (such as viruses and bacteria) and signals (such as contextual information), which results in a specialised immune response (i.e. deletion or tolerance) [43]. Because of their functionality, *DCs* are regarded

TABLE 1. Blockchain-based e-Government projects.

Country Name	Project Description	Status
Estonia [32]	Implemented blockchain technology in electronic identification (eID), e-health, and e-residency..	Running;Initiated in 2014.
Dubai [33]	By 2020, use blockchain technology to power all public transactions.	Ongoing; Initiated in 2016.
Switzerland [31]	In the Swiss city of Zug, create an Ethereum-based, uPort-powered e-residency ID.	Running, Initiated in 2019.
USA [34]	Create new rules that will allow the government to use blockchain for security and collaboration.	Ongoing; Announced in 2016.
Luxembourg [31]	Create a public framework that will enable blockchain applications to be integrated into all industries.	Ongoing; Announced in 2019.
Canada [35]	Create an e-Government information system using the Ethereum blockchain.	Ongoing, Initiated in 2018.
Mexico [36]	Embrace blockchain technology in finance, agriculture, and public procurement.	Ongoing, Announced in 2017.
China [31]	Add blockchain integration to e-health, e-ID, and e-voting systems.	Ongoing; Initiated in 2016.
France [30]	Support the development of blockchain systems by banks and other businesses to enable secure business transactions.	Ongoing, Announced in 2016.
Russia [30]	Investigate the use of blockchain to manage government records, e-health services, and land and property register.	Ongoing; Announced in 2017.
Africa [31]	Ghana, Kenya, South Africa, Ethiopia, Liberia, and Nigeria are just a few of the nations looking at the potential applications of blockchain in the agricultural sector, land registry, finance, transit, and e-services.	Ongoing; Announced in 2017.
Argentina [37]	Integrate blockchain-based eID to improve citizens' access to public services.	Ongoing; Announced in 2019.
Singapore [31]	ESince 2019, educational institutions have used the Ethereum blockchain to provide digital certificates.	Running; Initiated in 2018.
Sweden [31]	Investigate the use of blockchain in the current land registry system.	A proof-of-concept since 2016.
New Zealand [31]	In 2018, blockchain was used in electronic voting.	Running; Initiated in 2018.
India [31]	Investigate the use of blockchain in e-voting, land registry, and e-Government.	Ongoing, Announced in 2018.

as the body's own intrusion detection agents. *DCs* express costimulatory molecules (*csm*) on their cell surface to limit the number of antigens they can sample while in tissue, such as the skin or lung.

Three biological signals are essential in tissues for *DCs* maturation and amplification [43]. Pathogenic associated molecular pattern (*PAMP*) signals are those produced by viruses or bacteria that activate immune responses. *PAMP* is a strong indicator of abnormal behaviour in the tissue. The disrupted host tissue or stressed cells emit danger signals (*DS*). *DS* indicates the possibility of an anomaly, but with a lower probability than *PAMP*. Safe signals (*SS*) are produced by the naturally programme cell decay process, and thus it is an indicator of normal tissue behaviour.

In their lifetime, each *DC* exists in three states [43]. Immature *DC* (*iDC*) are immune-free and do not contribute to any immune actions. They are responsible for the collection of antigens and the signals that go with them. When a *DC*'s *csm* concentration exceeds the migration threshold, it migrates to a fully mature or semi-mature state. Compared to *iDCs*, an *smDC* has sampled a higher concentration of *SS* than *PAMPs* and *DS*, and a *mDC* has sampled a higher concentration of either *PAMPs* or *DS* than *SS*.

The DCA was derived from its natural counterpart for computer network intrusion detection [15], [16]. First, feature selection is used to select the most informative features, which are then classified as *PAMP*, *DS*, or *SS* based on their biological metaphor definitions. Then, a population of artificial *DCs* responsible for signal sampling is created in a pool, and *DCs* that meet certain criteria will stop sampling and process the data for classification. It is worth noting that each *DC* is given a unique sampling threshold by simulating the function of costimulatory molecules.

Once ceasing sampling, the *DCs* then detect the contexts of *smDC* by:

$$Context[smDC] = \sum_{d=1}^{m} \frac{\sum_{j=1}^{3}(c_j * w_j)}{\sum_{j=1}^{3} w_j}, \tag{1}$$

where $c_j (j \in \{1, 2, 3\})$ represent the signal values of *PAMP*, *DS* and *SS*, respectively; w_i^j indicate the corresponding weights of *PAMP*, *DS* and *SS*, respectively; m expresses the number of data items sampled in the *DC* based on the threshold, i.e. *csm* value. The context value of *mDC* can be calculated in the same way, but with a different set of weights. Weights are usually pre-defined [15] or learned through optimisation approaches such as generic algorithms [39]. By comparing the context values of *mDC* and *smDC*, the *DC* is assigned a context of normal or anomaly, depending on which is greater. The context is then attached to the data items sampled by the *DC*. It should be noted that each data item is sampled by multiple *DCs*, resulting in multiple attachments of context values. The final classification label of the data item is generated using an aggregation approach based on this.

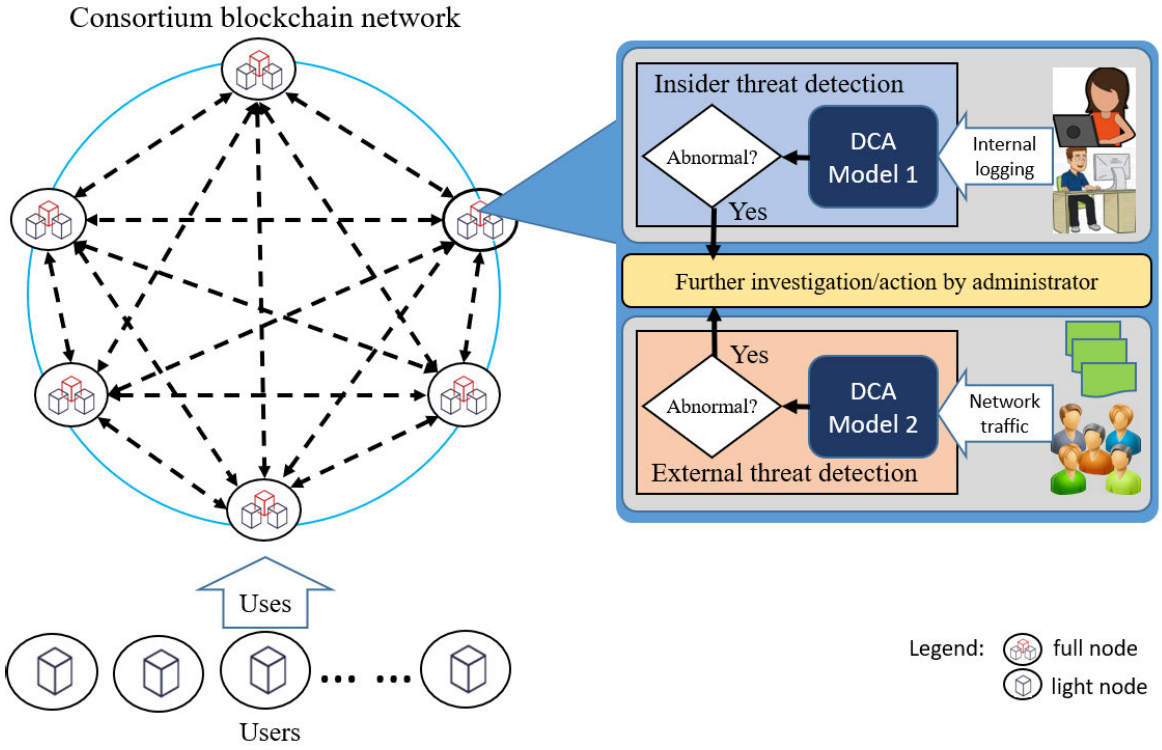


FIGURE 2. The proposed secure and privacy-preserving e-Government framework.

III. PROPOSED E-GOVERNMENT FRAMEWORK

The proposed secure and privacy-preserving e-Government framework is comprised of three modules as illustrated in Fig. 2. First, a peer-to-peer consortium blockchain network serves as the underlying computing and communication structure for various e-services, with each node representing a government agency. Second, an external attack detection module using an artificial immune system based on DCA detects suspicious traffic from the Internet for further investigation by network administrators. Third, an insider threat detection module using the DCA identifies internal anomalies generated by legitimate e-Government system accounts for further investigation. The internal and external threat detection modules jointly ensure their legitimacy before transactions are appended to the blockchain. Given that both the external attack detection module and the insider threat detection module are implemented using the DCA, they are discussed together in Sec. III-B.

A. BLOCKCHAIN NETWORK

The proposed e-Government framework adopts a consortium blockchain in order to eliminate a centralised control of data, to offer effective e-services at a reasonable computational cost, to safeguard private data against unauthorised access, and to preserve privacy. The consortium blockchain enables flexible information access control, allowing the accessibility of the information to be selectively limited to validators (such as e-Government departments), authorised users (such as registered citizens, businesses, and shareholders), or unregistered users, and facilitating cross-Government collaboration [9], [12].

1) BLOCKCHAIN STRUCTURE

In the proposed e-Government framework, two types of blockchain nodes are used: full nodes and light nodes. Each full node stores a copy of the entire blockchain; Government departments or their computing devices are configured as full nodes, which collectively form the backbone of the blockchain network. Light nodes are registered and configured for general e-Government users. Light nodes do not keep a copy of the whole blockchain on their servers. Instead, individuals connect to a complete node for authorised information access using their accounts and wallets. In a nutshell, a blockchain wallet is a digital repository that enables users to manage and save their login information, including IDs, passwords, private and public keys, and other account-related data. The unique wallet ID assigned to each user enables safe and secure information exchange and transfer, and wallets are accessible via mobile or web applications. All general e-Government users must register with one of the full nodes for authorisation and information access. Any new transaction from a general user is relayed to one of the full nodes and then propagated to other full nodes in the network. This means the full nodes are responsible for synchronising their local blockchain copy with the rest of the P2P network.

In order to authenticate transactions sent to the network, a number of devices representing various e-Government departments have been pre-selected as validators. Before a new transaction can be added to the blockchain ledger, it must first be approved by a predetermined number of validators. The proposed e-Government architecture incorporates an additional layer of protection to identify any disguised transactions proposed by hackers, by deploying an IDS on each full node in the consortium network, with full details described in Sec. II-D. An authorised organisation from an e-Government agency, such as the governance board, implements the pre-selection through the usage of an approved application programming interface (API) based on a predefined set of operational rules. Other e-Government agencies that are not validators are permitted to develop, examine, and submit new transactions to the blockchain but not to participate in the consensus and validation procedures. In other words, any government official has access to the data used by the blockchain to identify a specific user or organisation.

A Government node, i.e a full node, can be added to the consortium blockchain network using Algorithm 1. When a new node joins the network, its public and private keys, blockchain wallet, and address are generated, as indicated in lines 2, 3, and 4 in Algorithm 1. Because the generated keys and wallet are used to sign and validate transactions, they must be kept safe, which is accomplished through the use of the function *safelyStorePrivateKey*() in line 5. Following address generation, a node contacts validators in the blockchain network to transmit its registration request as stated in line 6. One of the validators will then authenticate this registration and transfer some e-Government tokens (registration record) to the node's blockchain address. A validator is selected to broadcast the new node information to other peers as shown in line 7. Following that, as expressed in lines 8, 9, and 10, the chosen validator broadcasts the new node's registration information to the network peers. This enables other network peers to get their wallet information in order to send transactions during the subsequent cycle. Here, N in line 8 represents the entire set of currently registered devices in the network, and n indicates each individual device that is receiving the broadcasted information about the new node. The process of adding a new node is completed when a full network node is successfully set up and broadcasted to the network. After registration, the newly added node will be able to sign up for and verify transactions using the blockchain address, wallet, and private and public key pair.

The procedure of adding e-Government nodes, i.e, full nodes, as depicted in Algorithm 1, effectively allows the construction of the backbone of the proposed e-Government framework. This also clearly differentiates the functionality between the service provider, i.e. government departments, and the service receiver, i.e. users, when registering new nodes in the blockchain network. The management of all transactions by the e-Government nodes via keeping a full complete copy of the blockchain ledger ensures the efficiency of the proposed e-Government system. This is ensured by

Algorithm 1 Full Node Registration

input: A new device m,
A set of N nodes in the current consortium network,
tokens
output: Registered node m

1: **if** (a full node request is valid from the Government) **then**
2: $(K_{pub}, K_{pr}) \leftarrow generateKeys()$
3: $Addr \leftarrow createBlockchainAddress() + (K_{pub}, K_{pr})$;
4: $Walt \leftarrow createBlockchainWallet() + (K_{pub}, K_{pr})$;
5: *safelyStorePrivateKey*();
6: $Addr \leftarrow Addr$+tokens;
7: $\beta \leftarrow selectMiner(N)$;
8: **for each** $n \in \{N - \beta\}$ **do**
9: *distributeRegistration*(n, m);
10: **end for**
11: $m \leftarrow verifiedNewNode()$;
12: **else**
13: create a lightweight node m using Algorithm 2;
14: **end if**

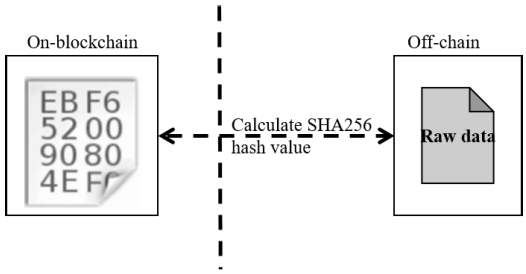


FIGURE 3. Example of an Off-chain storage for blockchain.

fewer full nodes and thus faster transaction processing due to the use of a consortium blockchain rather than a private blockchain. Additionally, the e-Government nodes in the proposed system are made to be able to store off-chain (sideDB) data like images, PDF, DOC, text documents, contracts, and other non-transactional data files that are either too big to be stored in the blockchain or are subject to future modification or deletion [11]. In the proposed decentralised consortium blockchain-based system, these files are encrypted on the user side and stored off-chain. In other words, the SHA256 algorithm is used to hash the raw data of the original file, such as a photo, video, or text file, producing a hash value that is saved in the blockchain with a reference to the original file kept in the off-chain database [44], as illustrated in Fig. 3. Of course, the original file or raw data in this instance is not available to the general public.

Off-chain transactions can be processed considerably more quickly than on-chain transactions as they do not require the consensus from all full nodes. When a transaction is carried out on off-chain documents, the responsibility of e-Government nodes is to ensure that the blockchain protocols have been followed in the proposed system and that the transaction execution has been verified. For instance,

the e-Government blockchain full node will carry out the transaction via the exchange of private keys when a user accepts the terms and conditions to share their off-chain documents with a third party organisation. It is only necessary to provide personal information when registering a new user, such as social security numbers, birth dates, and names. The blockchain ledger in the proposed framework only stores the hash of the personal data to provide maximum privacy because the data is hashed and cryptographically connected with the ledger.

2) USER REGISTRATION AND AUTHENTICATION

Citizens, businesses and other users can register to e-Government for various services with the registration process summarised in Algorithm 2. Note that, lines 2-11 in Algorithm 1 and lines 1-9 in Algorithm 2 are similar in that every device in a blockchain network needs to be created with cryptographic keys, an identity, an address, and a wallet. In Algorithm 2, in order to identify the user, a user ID is first produced as mentioned in line 2, followed by the creation of a new blockchain address for the user that contains both the public and private keys as shown in line 3. The generated user ID and private key are stored safely by the selected node, as illustrated in line 4. With this information, a blockchain wallet is generated for this new user as shown in line 5, and the generated wallet is broadcasted to all nodes in the blockchain network as expressed in lines 6 to 8. The created blockchain wallet will be used to send and receive pertinent transactions to this account. Through the wallet interface, the user may conveniently check their history as well as any fresh transactions that have been made available to them in their blockchain addresses.

Algorithm 2 New User Registration

input: User registration request
Nodes N in the current network
output: A newly registered user u
1: $(K_{pub}, K_{pr}) = generateKeys()$
2: $uID \leftarrow createUserID()$;
3: $Addr \leftarrow createBlockchainAddress() + (K_{pub}, K_{pr})$;
4: $(uID, K_{pr}) \leftarrow Safelystore(uID, K_{pr})$;
5: $Walt \leftarrow createBlockchainWallet() + (K_{pub}, K_{pr})$;
6: **for each** $n \in N$ **do**
7: $distributeWallet(n, Walt)$;
8: **end for**
9: $u \leftarrow verifiedNewUser()$;

The registration of new users not only allows the users of the e-Government system to implement transactions, but also facilitates the access control of user information and their transactions. The registration mechanism makes the proposed e-Government system work efficiently with a clearly defined responsibility for every node in the consortium blockchain. When a user submits a record to the e-Government network, the transaction will first go through the authentication process to facilitate the initialisation of the transaction. From this, the block is updated to a new version which is broadcasted across the network for validation and then transferred to the user's blockchain address stored in the e-Government consortium network. The transferred record in the user's blockchain address includes the following content: 1) the ID of the user, 2) the record value or the transaction, such as property registration, and 3) the record identification, such as property registration number. Each data instance in a blockchain represents a virtual asset.

When a third-party organisation requests to access a user's information, the user needs to provide their blockchain address for verification. The organisation can then use the blockchain web API to access the blockchain data stored in the user's address. All e-Government users are required to keep their private keys safe and secure, with the support of a backup. If any user lost their private key, the user will be required to create a new blockchain address and make a request to one of the e-Government department nodes to transfer the user's information and associated records from the old blockchain address to the newly created blockchain address.

The identity of a registered user will be validated and authenticated when the user wants to access the blockchain network through a registered device. Note that human errors remain a great cause of cyber security breaches in public and private organisations [45]; this access validation and authentication help to significantly reduce human errors which have always been considered as a main cause of failure and a weak link to access information stored in information systems [46]. As a result, Government information will flow securely and seemingly to the right individuals at the right time through a user's device, regardless of the user's location as long as the Internet is available.

3) E-GOVERNMENT SIMULATION

Due to the substantial hardware devices required to implement the proposed e-Government framework, this study employs the eVIBES simulator [18] to simulate the system. Briefly, eVIBES is a configurable and open source framework for simulating large-scale Ethereum networks in order to observe the empirical behaviours and dynamic properties of P2P nodes [18]. The eVIBES is broadcast-based, event-driven, scalable, message-oriented and concurrent. The eVIBES was adopted in this project due to its effective deployment mechanism and low running cost in comparison to alternative options [18]. In addition, its high scalability allows the network to accommodate a large number of nodes, without compromising the network speed or efficacy.

The proposed e-Government framework can be customised when it is simulated using eVIBES. The configurable parameters include the number of P2P nodes, the number of transactions, the rate of transaction generation, the initialisation of the genesis block, the sidechains or off-chain database, and the smart contract mode [18]. The sidechains were configured

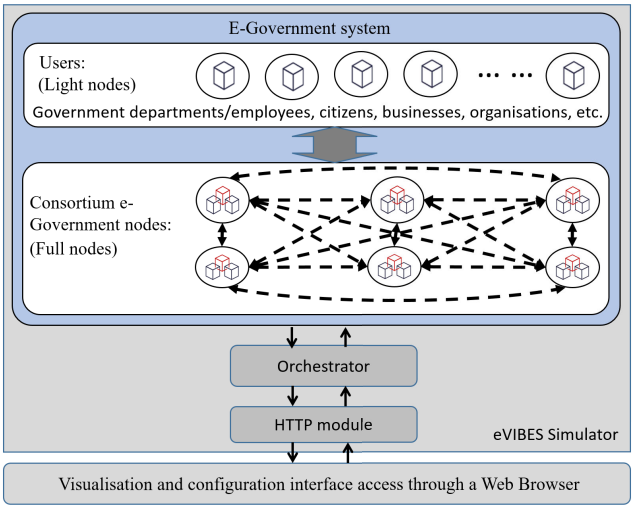


FIGURE 4. Simulated environment of the proposed e-Government system.

using the same implementation of database and servers as reported in [18]. A smart contract contains a set of rules under which the participating parties agree upon and run on the top of Ethereum blockchain to support the exchange of users' information, documents, property, etc., between P2P nodes without the involvement of a trusted third party. The smart contract mode of the simulation allows smart contracts to be uploaded for execution. The performance of the proposed e-Government system can be assessed by the outputs of the simulator on a number of metrics, such as the overall execution time, the total number of transactions processed, throughput or the number of transactions per second, and the block propagation delay.

The architecture of the blockchain network of the proposed e-Government framework is summarised in Fig. 4. In addition to the full nodes and the light nodes of the blockchain, the main component orchestrator manages the entire simulated network by sending the configuration, the parameter settings, and other messages to all nodes in the system. It enables system configuration, monitors and records system states, and communicates with users about the results and performance. In other words, the orchestrator regulates the entire simulated e-Government system including the generated blockchain. The HTTP module enables the interaction between the orchestrator and the system interface through a browser in an event driven manner.

B. INSIDER AND EXTERNAL THREAT DETECTION

Both insider and external threat detection functions are realised using an optimised DCA algorithm, which is deployed on all full nodes to detect abnormal behaviour and unwanted traffic. In particular, two DCA models are trained using two distinct datasets, one for external attack detection and one for insider threat detection. Rather than using the original DCA with a signal categorisation process as discussed in Sec. II-D, a recently proposed variant that takes all features as system inputs is adapted in this work to support the proposed e-Government framework by realising that a threat is frequently relevant to more than three features and that the importance of each feature can be determined using a general optimisation algorithm [47]. In other words, rather than employing the signal categorisation process in analogy to the natural biological process, the general optimisation algorithm directly maps features to threats using weights and the irrelevant features are weighted as 0.

The weight associated with each feature is computed using the genetic algorithm (GA) [47]. In fact, GA has been widely used to optimise weights, such as rule base optimisation in fuzzy inference systems [41]. Briefly, GA starts with the initialisation of a population of random individuals, with each representing a possible solution of weights. Then, the population evolves through a number of operations, typically including elitism, mutation and crossover; and more effective individuals are survived and evolved over time until a specified level of performance or the maximum number of iterations is reached.

An individual (I), in this work, is a vector comprising of all the weights used by the DCA algorithm. The size of the population ($\mathbb{P}$) is a problem-specific adjustable parameter depending on the scale of the e-Government system, typically in a range from tens to thousands, with 10 to 50 being widely used [48]. The objective function is simply defined as the accuracy of threat detection in the e-Government system. The fitness proportionate selection method is employed for individual selection and reproduction with the support of crossover and mutation. When the GA terminates by either reaching the maximum number of iterations or pre-defined optimal accuracy requirement, the fittest individual in the current population is taken as the optimal set of weights.

The accuracy of threat detection used by the objective function is achieved by applying the revised DCA to a training dataset. The training dateset is pre-processed through a feature selection procedure which selects the most informative features for implementing the DCA-based threat detection systems. Many feature selection approaches are readily available for this project, such as the work reported in [49] and [50], but the traditional information gain approach [51] is employed in this work due to its simplicity and efficiency. From this, the selected features are normalised using the min-max normalisation method before they are fed into the GA algorithm for optimal weights generation. The DCA with the optimal set of weights provides the artificial immune function to the proposed e-Government system to exclude fraud transitions in the blockchain and minimise inappropriate information access.

IV. EXPERIMENTAL EVALUATION

This section describes the procedures of the experiments, and analyses the findings. All experiments and simulations were carried out using an HP workstation equipped with an Intel processor ® Xeon™ E5-16030 v4 @3.70 GHz and 32GB RAM. Two experiments in particular were carried out

TABLE 2. Parameters for simulating the proposed e-Government blockchain system.

Parameter	Setting
Number of boot nodes	100
Number of transactions in the network	20000
Rate (in sec) of transaction generation	0.01
Initial number of peers per node	5
Transactions in a batch	10
Initial number of user accounts	100
Smart contracts mode	Independent execution
Genesis block	Override

and are thus reported here. Firstly, the performance of the proposed consortium blockchain-based e-Government system was assessed using a variety of widely used metrics. Secondly, the insider and intrusion detection functions were evaluated using two publicly available datasets, including the CERT for insider threat detection, and the UNSW_NB15 for external threat detection.

A. BLOCKCHAIN PERFORMANCE EVALUATION

This experiment was implemented through the eVIBES simulator. The parameters used to configure the e-Government blockchain network are presented in Table 2. The maximum number of nodes that need to boot with the genesis block was set to 100 in the initialisation stage, and adjusted to up to 200 during the simulation. Each initialised node is assigned with an initial account including a node ID, a blockchain wallet, private and public keys for account accessing and transaction validation. The number of transactions that an individual node can process in a batch and add into the blockchain was set to 10, and the rate of transaction in the network was set to 0.01 seconds for fast and efficient block creation and propagation in the network [18], [24], [25]. In this simulation, the maximum amount of transactions that the blockchain network can store across all nodes was set to 20,000.

Transactions were executed in the form of smart contracts in the eVIBES simulator, and the smart contracts mode was set as an independent execution. Briefly, each node is able to preserve its own database state whilst carrying out the assigned transactions before sharing and adding them to the blockchain thanks to the autonomous execution mechanism of smart contracts [18]. To preserve consistency and act as the foundation for the ledger on the consortium blockchain, the default genesis block is replaced with the customised Ethereum genesis block values before the simulation starts [18], [24]. All nodes periodically send their transactions to the consortium blockchain storage in every 10 seconds, in order to capture the throughput, the rate at which transactions were validated, the block propagation time, and the processing time for each transaction.

1) BLOCK PROPAGATION TIME

This experiment measured the average propagation time of blocks in seconds by linearly increasing the number of Ethereum nodes, or validators, with the result displayed in Fig. 5. It is worth noting that each Ethereum node represents a dedicated e-Government full node, so the number of Ethereum nodes equals the number of e-Government full nodes. This figure shows that the block propagation time increases linearly along with the increase of the number of Ethereum nodes in the network, demonstrating the scalability of the proposed e-Government system.

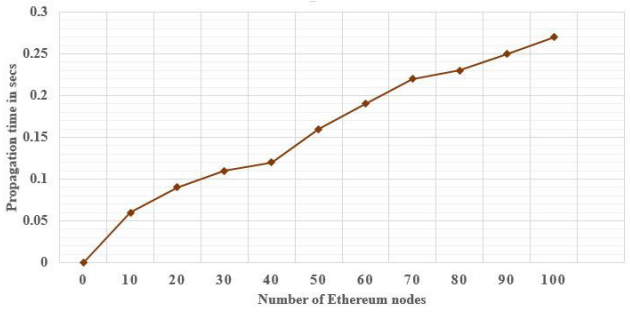


FIGURE 5. Block propagation time against the number of nodes.

2) TRANSACTION THROUGHPUT

This experiment investigated the transaction throughput of a consortium blockchain network with different numbers of nodes under the proposed e-Government framework. The performance of the consortium blockchain network as measured by the number of transactions processed per second (i.e. throughput) with an increasing number of Ethereum nodes (i.e. validators) is depicted in Fig. 6. It should be noted that the ideal situation occurs when all user transactions are validated in one second or less. The figure clearly shows that the number of average transactions validated per second decreases as the number of nodes in the consortium blockchain network grows. This is due to the communication overhead required to choose a node among the consortium nodes to validate new transactions, create a new block, and append it to the ledger. As a result, if a given blockchain network requires a large number of nodes, the transaction processing speed will decrease. This may not be the case in e-Government systems because all participating departments or agencies serve and share the same goal of delivering public services; in these systems, fewer validators will be sufficient to process and validate transactions whilst other Ethereum nodes communicate with the e-Government users.

From Fig. 6, a consortium network with up to 40 validators can validate up to 100 transactions in less than a second, which is close to the ideal case. However, for a consortium network with more than 50 validators, the network needs well over a second to validate 100 transactions, which deviates from the ideal situation. Consequently, as the number of nodes increases in the network, the network requires more time to validate transactions. Apparently, a trade-off must be made between the desired network performance, the number of transactions that can be processed in a second, and the number of nodes in the consortium blockchain network. This

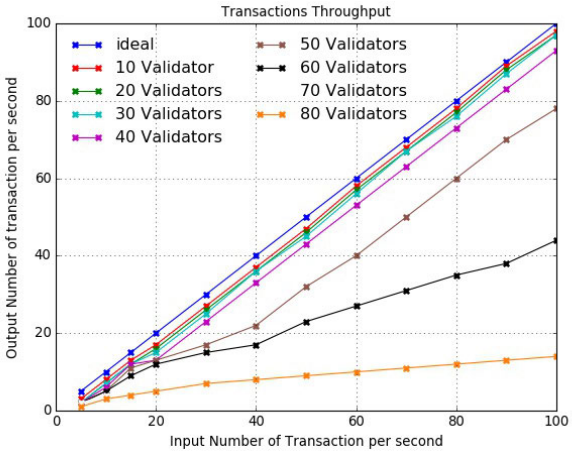


FIGURE 6. The number of transactions per second against the number of e-Government nodes.

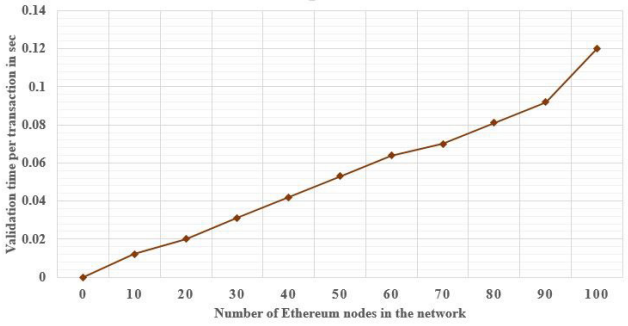


FIGURE 7. Validation time against the number of e-Government nodes.

trade-off should be carefully taken into account during the design phase of an e-Government system. More precisely, special consideration must be given to choosing the right number of validating Ethereum nodes for the consortium network in order to strike a balance between security and transaction throughput.

3) SINGLE TRANSACTION VALIDATION TIME

It has been observed in this experiment that more time was required to validate a transaction as the number of Ethereum nodes was increased, as shown in Fig. 7. In the experimental setting, the validation time grew up to 0.12 seconds when the number of Ethereum nodes was increased to 90, but it was less than 0.1 seconds on average if the network consisted of fewer than 90 validators. This performance is generally commensurate with a typical consortium blockchain network regarding the average validation time for a transaction [9], [12]. This implies a relationship between the desired network performance, the number of transactions per second, the number of Ethereum nodes, and the network resources, which frequently requires careful design and, at times, compromise before the full realisation and deployment of an e-Government system.

4) TIME REQUIREMENT FOR ADDING NEW BLOCKS

The goal of this experiment was to determine how long it takes a validator to add a new block of transactions to the

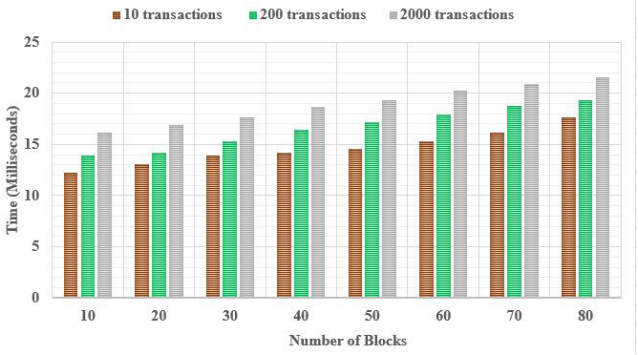


FIGURE 8. Time required in milliseconds to add a new block to the blockchain network.

blockchain. To add a new block, a validator must: i) receive a request from the initiating node, ii) validate the requesting node, iii) create the block from the associated transactions, and iv) update the rest of the e-Government nodes with the new block.

The experimental results are summarised in Fig. 8. Note that, as the number of blocks grows, the validator must validate more transactions and blocks, essentially increasing the time needed to add a new block to the blockchain network. It is also clear that the quantity of transactions that must be verified directly influences how long it takes to add a new block to the blockchain. This is because it takes longer to perform the validation of a block when it is comprised of more transactions.

B. SECURITY AND PRIVACY ANALYSIS

The functioning mechanism of the blockchain primarily benefits the anonymity of the planned e-Government system. It can successfully restrict unauthorised access to, misuse of, and abuse of the hardware and data infrastructure. Every blockchain network device has its own storage for maintaining data that needs to be kept private, including eIDs, private keys, wallet information, etc. Since each e-Government user depends on their key to communicate with the consortium network, the usage of unique private keys can also shield the network to some extent against an attacker who tries to deanonymise a user. In addition, because all of the blocks that contain transactions are hashed, the anonymity that blockchain offers enhances user privacy.

The adoption of blockchain technology greatly facilitates the security enforced by the proposed e-Government system. To maintain integrity, each transaction in the proposed consortium blockchain network consists of the hash of the information, or transactions. Public-key cryptography is used to encrypt each transaction in order to protect the secrecy of the shared data. In order to maintain information integrity and confidentiality, it must not be made available to unauthorised users. The proposed blockchain-based e-Government system uses encryption and digital signatures to offer security, privacy, and proper access control to the stored documents.

Attacks on authentication and authorisation take place when nefarious persons attempt to take over the blockchain

TABLE 3. Security services in the proposed e-Government system.

Security service	Realisation
Authentication	Blockchain address and digital signature
Access Control	Digital signature, encryption and IDS
Confidentiality	Encryption and IDS
Integrity	Encryption, digital signature and IDS
Non-Repudiation	Encryption and digital Signature
Availability	Distributed/decentralised services and IDS
Trust	Transparency, decentralisation, encryption and digital signature
Validation	Consensus algorithm, encryption, hashing and IDS
Authorisation	Blockchain wallet, blockchain ID, digital signature
Verification	Blockchain wallet, blockchain ID, and digital signature
Identification	Blockchain wallet, blockchain ID, and digital signature
Anonymity	Encryption and access control

network in order to approve themselves or introduce phony nodes that seem to be authorised nodes, which will eventually compromise the e-Government network. This is difficult to happen in the proposed e-Government framework because every full node of the consortium e-Government network is required to be pre-selected by authorised entities from an e-Government agency of a specific Government beforehand and every full node is closely monitored by other departments, which is typical in all e-Government systems.

Anonymity in blockchain-based networks could make it easier for opponents to engage in unlawful activities [12]. The blockchain information is only accessible to consortium nodes and authorised users in the proposed e-Government framework. As a result, any adversary attempting to establish an anonymous connection can be detected instantly because all user devices attempting to access the blockchain network must first be validated. Furthermore, the proposed blockchain-based e-Government framework improves information security and privacy by incorporating immutability (i.e. unchangeable records once added), resiliency (i.e. avoiding a single point of failure), verifiability (i.e. only validated transactions), and distributed consensus by validators, among other features. The security services and corresponding implementation approaches provided by the proposed framework is summarised in Table 3.

The use of consortium blockchain in e-Government systems can help to prevent common threats to information security and privacy. DoS and DDoS attacks, for example, are common attacks launched by cyber criminals against web servers in order to consume a significant amount of bandwidth and resources until the service is rendered inoperable. The lack of attack detection functions in decentralised blockchain-based services may make the system vulnerable to DoS or DDoS attacks. In the most extreme case, when the target of the attack is the majority of the nodes in the network, the decentralisation nature of the proposed e-Government system can enable the allocation of data and bandwidth to the least overloaded nodes in the blockchain network to absorb DoS and DDoS attacks should they occur. As a result, the proposed e-Government framework is, to some extent, equipped with 'self-healing' functionality when attacked.

Other common security threats are also taken into account in the proposed e-Government framework. These attacks have the potential to make the proposed e-Government framework vulnerable, but defence strategies are readily available thanks to the use of blockchain, as shown in Table 4. These security attacks and their corresponding defence strategies are detailed in the referred work, such as those described in [52], [53], [54], [55], [56], and [57], thus extra details beyond that listed in Table 4 are not duplicated here.

C. EVALUATION OF THREAT DETECTION

The proposed e-Government system incorporates both insider threat detection and external threat detection functionality. Two datasets were used in this study to train attack detection models and evaluate the system. In particular, the CERT insider threat dataset V4.2 was used to validate the performance of insider threat detection [5]. Briefly, the CERT dataset is a synthetic dataset that details the everyday computer activities of insiders over the course of 17 months. Of the 1000 user accounts used to collect the data, 70 were engaged in harmful activities within the organisation. There were five actions that insiders took during this time period that were recorded in the dataset, including logging in and off of the computers, sending and receiving emails, connecting and unplugging external devices, the type of file accessed, and HTTP URLs visited. After being pre-processed, 80% of the data was used for training and the rest for testing.

The UNSW_NB15 dataset is a publicly available external threat detection dataset [19]. Reconnaissance, Shellcode, Exploit, and Fuzzers are examples of modern attack types included in this dataset but not usually found in other datasets. The class label is represented as the last feature of 49 features in the UNSW_NB15 dataset. The 49 features can be clustered in six groups, including flow features, basic features, content features, time features, additional generated features, and labelled features. Flow features include the traffic flow captured between a client and a server. The attributes that characterise the protocols for connections are referred to as basic features. Content features are characteristics of TCP/IP and HTTP services. Time features are timing attributes such as TCP protocol arrival round trip time and time between packets. The additional generated features are synthetic features generated randomly. This dataset has been pre-processed and is ready for training and testing. The training dataset contains 175,341 records comprising of 56,000 normal activities and 119,341 anomalous activities, whilst the testing dataset has 82,332 data instances including 37,000 normal activities and 45,332 anomalous activities.

1) PARAMETER SETTINGS

The information gain method was employed for feature selection. As commonly employed by other projects [16], a population of 100 *DCs* was initialised in the sampling pool and the size of the mature pool was set to 10 *DCs*. With a mean of 5.0 and a standard deviation of 1, a Gaussian distribution was used to determine the migration criteria for DCs. The percentage of anomalies in the datasets was used to calculate

TABLE 4. Common security attacks and countermeasures.

Attack	Defence
DoS and DDoS	Typically, this happens when attackers flood online services with massive amounts of fake traffic in order to make the service unavailable. The proposed system is a decentralised P2P system in which user data is stored in multiple nodes, ensuring system availability by avoiding any single point of failure.
Nodes injection attack	This might happen if hostile nodes attempt to join and authenticate the consortium blockchain network and seize control of it. Since each peer in the consortium network is pre-selected by an authorised entity from an e-Government agency, this is challenging under the proposed architecture.
Modification attack	Due to the fact that the hash value of the current block header (parent) is linked to and stored in the following block (child), this is challenging because if the content of any block changes, its hash will also change, and the change will be propagated throughout the network to invalidate that block.
51% attack	This will happen if an attack is successful in changing every instance of a record by controlling at least 51% of the network peers. This is difficult because an attacker would need to alter each copy of a block in the network in order to alter any block in the blockchain, and then they would need to persuade all nodes that the modified block is the legitimate one. Typically, it is impossible to accomplish this.
False validators	This is impossible in the proposed framework because all consortium network peers are pre-selected by an authorised entity from e-Government agencies.
Eavesdropping and Traffic analysis	This attack cannot succeed since all user blocks in the proposed e-Government system are hashed, and the unreadable hashes of the transactions are kept in the blockchain.
Man in the Middle Attack	This is not conceivable under the proposed system since each peer in the consortium network has been pre-selected from e-Government agencies by an authorised organisation.
Sybil attack	Due to the requirement that all full nodes of the consortium e-Government network be preselected by an authorised body from among e-Government agencies in advance, this is challenging under the proposed e-Government architecture.
Impersonation and Non-traceability	In order to protect the security and privacy of shared transactions among its unreliable participants, the proposed system is built as an immutable and distributed database.
Replay Attack	Since any transactions received from a network node are verified by witnesses, it is difficult for rogue nodes to start harmful transactions.
Masquerading	This is difficult in the proposed system because when a person wants to access the blockchain network, their identity must be verified and authorised using a registered device.
Transaction Tampering	This is impossible since there is no possibility of deletion in the future because the blockchain database is append-only and immutable. Each node in the network has a copy of the same ledger, and it uses a consensus mechanism to decide whether to add new transactions.

TABLE 5. Parameter settings for the GA used in the experimentation.

Number of Individuals	50
Number of Iterations	250
Crossover Rate	0.95
Mutation Rate	0.1

the anomaly threshold for both inside and external attack detection. The parameter values of GA for the generation of the optimal weights of DCA is summarised in Table 5, following the typical settings in the literature as reported in [41] and [48].

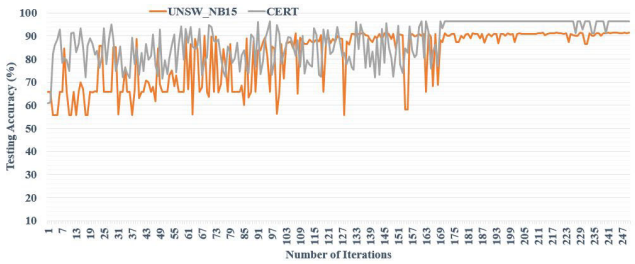


FIGURE 9. Optimised GA-DCA fine-tuning the testing accuracies.

2) RESULTS AND ANALYSIS

The training processes using the two datasets over 250 iterations are illustrated in Fig. 9. The intrusion detection models stabilised after about 200 training iterations. In this experiment, the performance of the proposed insider and external threat detection was measured using accuracy and detection rate (i.e. sensitivity). Briefly, the detection rate is defined as the percentage of successfully detected positive cases amongst all positive cases in the dataset. The performance of the models was further evaluated via precision and F-Score metrics to measure the effectiveness of the models on datasets with uneven class distribution (i.e. class imbalance). It is important to note that high accuracy indicates that the model is performing better only when the dataset contains balanced data samples between classes (symmetric). The F-score is more efficient than accuracy when the dataset has uneven class distributions.

The experimental results are summarised in Table 6, which clearly demonstrates the ability of DCA to support the proposed e-Government system. The high detection rate to inside threats can help identify insiders at an early stage and thus ensure the integrity of the ledger as the inclusion of any block must be supported by over half of the full nodes. Similarly, the high precision of detecting external treats indicates the reliability of the proposed system for the detection of external intrusions. The competitive accuracies and F-scores for both insider and external threats show that the proposed system can detect threats associated with uneven class distributions

TABLE 6. Results on classification accuracy, detection rate, F-score and precision.

Dataset	Accuracy (%)	Detection Rate (%)	F-Score (%)	Precision (%)
CERT-Insider	96.52	97.65	93.19	89.12
UNSW_NB15	91.64	92.72	95.21	97.84

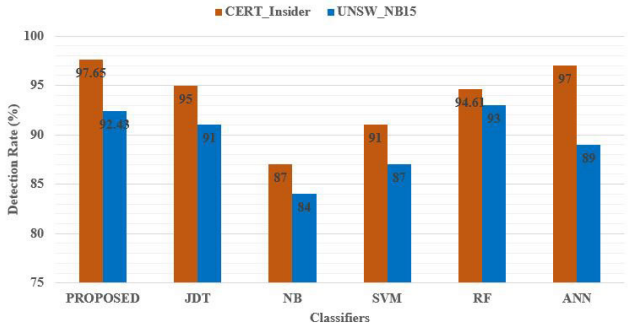


FIGURE 10. Insider and external threat detection rates.

between different types of attacks and normal data traffic, which is common in e-Government systems.

The detection rates were additionally compared with those led by five commonly used classifiers, including the J48 Decision Tree (JDT), the Naive Bayes (NB), the Support Vector Machine (SVM), the Random Forest (RF), and the Artificial Neural Network (ANN). The results of the comparison are shown in Fig. 10. The findings of this comparative study imply that the proposed system is able to competitively detect outside attackers as well as insiders in an e-Government system in a manner akin to the functioning mechanism of human immune systems protecting against infections and internal threats.

V. CONCLUSION

This paper presents a decentralised, secure, and privacy-preserving e-Government framework using consortium blockchain and artificial immune systems. The decentralised structure and encryption/validation mechanism provided by blockchain technology ensure the security, privacy, and integrity of information, which is further enhanced by the insider and external threats detection functionalities realised through an artificial immune system. The proposed framework was implemented using the eVIBES simulator. The experimental results show that the proposed e-Government framework can provide e-services to users in an effective and secure manner, with the potential of increasing trust in public sectors. A direct piece of future work following the experimentation will be to investigate the innovative application of advances in artificial intelligence with the goal of speeding up block creation when there is a spike in transactions in the e-Government network so as to make the system more scalable and robust. In addition, it is worthwhile to study the application of other artificial immune systems to provide a security shield to the proposed e-Government system.

REFERENCES

[1] L. Carter and V. Weerakkody, "E-government adoption: A cultural comparison," *Inf. Syst. Frontiers*, vol. 10, no. 4, pp. 473–482, Sep. 2008.

[2] L. Yang, N. Elisa, and N. Eliot, "Privacy and security aspects of E-government in smart cities," in *Smart Cities Cybersecurity and Privacy*. Amsterdam, The Netherlands: Elsevier, 2019, pp. 89–102.

[3] R. Palanisamy and B. Mukerji, "Security and privacy issues in E-government," in *Cyber Behavior: Concepts, Methodologies, Tools, and Applications*. Pennsylvania, PA, USA: IGI Global, pp. 880–892, 2014.

[4] N. Elisa, L. Yang, F. Chao, and Y. Cao, "A framework of blockchain-based secure and privacy-preserving E-government system," *Wireless Netw.*, vol. 24, pp. 1–11, Dec. 2018.

[5] J. Glasser and B. Lindauer, "Bridging the gap: A pragmatic approach to generating insider threat data," in *Proc. IEEE Secur. Privacy Workshops*, May 2013, pp. 98–104.

[6] (2019). *Verizon Insider Threat Report*. Accessed: Mar. 22, 2020. [Online]. Available: https://www.verizon.com/about/news/verizon-refocuses-cyber-investigations-spotlight-world-insider-threats/

[7] N. Elisa, L. Yang, H. Li, F. Chao, and N. Naik, "Consortium blockchain for security and privacy-preserving in E-government systems," 2020, *arXiv:2006.14234*.

[8] N. E. Nnko, *A Decentralised Secure and Privacy-Preserving E-Government System*. Tyne, U.K.: University of Northumbria at Newcastle, 2020.

[9] Z. Li, J. Kang, R. Yu, D. Ye, Q. Deng, and Y. Zhang, "Consortium blockchain for secure energy trading in industrial Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 14, no. 8, pp. 3690–3700, Aug. 2017.

[10] S. Underwood, "Blockchain beyond bitcoin," *Commun. ACM*, vol. 59, no. 11, pp. 15–17, 2016.

[11] N. Elisa, L. Yang, H. Li, F. Chao, and N. Naik, "Consortium blockchain for security and privacy-preserving in E-government systems," in *Proc. ICEB*, 2019, pp. 99–107.

[12] O. Dib, K.-L. Brousmiche, A. Durand, E. Thea, and E. B. Hamida, "Consortium blockchains: Overview, applications and challenges," *Int. J. Adv. Telecommun.*, vol. 11, nos. 1–2, pp. 1–14, 2018.

[13] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," Decentralized Bus. Rev., Manubot, Tech. Rep. 21260, 2008.

[14] A. M. Antonopoulos, *Mastering Bitcoin: Unlocking Digital Cryptocurrencies*, Sebastopol, CA, USA: O'Reilly Media, 2014.

[15] J. Greensmith, U. Aickelin, and S. Cayzer, "Introducing dendritic cells as a novel immune-inspired algorithm for anomaly detection," in *Proc. Int. Conf. Artif. Immune Syst.* Springer, 2005, pp. 153–167.

[16] Z. Chelly and Z. Elouedi, "A survey of the dendritic cell algorithm," *Knowl. Inf. Syst.*, vol. 48, no. 3, pp. 505–535, Sep. 2016.

[17] N. Elisa, L. Yang, X. Fu, and N. Naik, "Dendritic cell algorithm enhancement using fuzzy inference system for network intrusion detection," in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ-IEEE)*, Jun. 2019, pp. 1–6.

[18] A. Deshpande, P. Nasirifard, and H.-A. Jacobsen, "EVIBES: Configurable and interactive ethereum blockchain simulation framework," in *Proc. 19th Int. Middleware Conf. (Posters)*, Dec. 2018, pp. 11–12.

[19] N. Moustafa and J. Slay, "UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," in *Proc. Mil. Commun. Inf. Syst. Conf. (MilCIS)*, Nov. 2015, pp. 1–6.

[20] J. R. Gil-Garcia, S. S. Dawes, and T. A. Pardo, "Digital government and public management research: Finding the crossroads," *Public Manage. Rev.*, vol. 20, no. 5, pp. 633–646, May 2018.

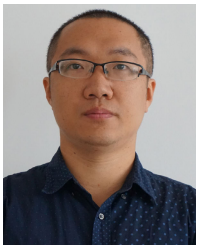
[21] *E-Government in Support of Sustainable Development/UN Department of Economic and Social Affairs*, United Nations E-Government Survey 2014, New York, NY, USA, 2016.

[22] M. Stefanova, S. Stefanov, and O. Asenov, "Identity protection accessing E-government through the biometric authentication methods," in *Proc. 6th IEEE Int. Conf. Intell. Syst.*, Sep. 2012, pp. 403–408.

[23] V. Ndou, "E–government for developing countries: Opportunities and challenges," *Electron. J. Inf. Syst. Developing Countries*, vol. 18, no. 1, pp. 1–24, 2004.
[24] V. Buterin, "A next-generation smart contract and decentralized application platform," *White Paper*, vol. 3, no. 37, pp. 1–2, 2014.
[25] C. Cachin, "Architecture of the hyperledger blockchain fabric," in *Proc. Workshop Distrib. Cryptocurrencies Consensus Ledgers*, vol. 310, 2016, pp. 1–4.
[26] Z. Zheng, S. Xie, H. Dai, X. Chen, and H. Wang, "An overview of blockchain technology: Architecture, consensus, and future trends," in *Proc. IEEE Int. Congr. Big Data (BigData Congress)*, Jun. 2017, pp. 557–564.
[27] U. Bodkhe, S. Tanwar, K. Parekh, P. Khanpara, S. Tyagi, N. Kumar, and M. Alazab, "Blockchain for industry 4.0: A comprehensive review," *IEEE Access*, vol. 8, pp. 79764–79800, 2020.
[28] D. Di Francesco Maesa and P. Mori, "Blockchain 3.0 applications survey," *J. Parallel Distrib. Comput.*, vol. 138, pp. 99–114, Apr. 2020.
[29] A. A. Monrat, O. Schelén, and K. Andersson, "A survey of blockchain from the perspectives of applications, challenges, and opportunities," *IEEE Access*, vol. 7, pp. 117134–117151, 2019.
[30] M. Jun, "Blockchain government—A next form of infrastructure for the twenty-first century," *J. Open Innov., Technol., Market, Complex.*, vol. 4, no. 1, p. 7, Dec. 2018.
[31] M. Kuperberg, S. Kemper, and C. Durak, "Blockchain usage for government-issued electronic IDS: A survey," in *Proc. Int. Conf. Adv. Inf. Syst. Eng.* Springer, pp. 155–167, 2019.
[32] C. Sullivan and E. Burger, "E-residency and blockchain," *Comput. Law Secur. Rev.*, vol. 33, no. 4, pp. 470–481, 2017.
[33] *Dubai Blockchain Strategy*, Smart Dubai, Dubai Government, Dubai, United Arab Emirates, Dec. 2016.
[34] (2016). *Blockchain Project in USA*. Accessed: May 27, 2020. [Online]. Available: https://consensys.net/blog/enterprise-blockchain/which-governments-are-using-blockchain-right-now//
[35] (2018). *Blockchain Project in Canada*. Accessed: Mar. 22, 2020. [Online]. Available: https://bitaccess.ca/blog/government-of-canada-ipfs///
[36] (2017). *Blockchain Project in Mexico*. Accessed: May 22, 2020. [Online]. Available: https://www.gob.mx/cidge/acciones-y-programas/blockchain-hackmx/
[37] (2019). *Blockchain Project in Argentina*. Accessed: May 22, 2020. [Online]. Available: https://www.bloomberg.com/press-releases/2019-08-26/nec-idb-lab-and-ngo-bitcoin-argentina-to-deploy-a-blockchain-ba/
[38] H.-J. Liao, C.-H. R. Lin, Y.-C. Lin, and K.-Y. Tung, "Intrusion detection system: A comprehensive review," *J. Netw. Comput. Appl.*, vol. 36, no. 1, pp. 16–24, 2013.
[39] N. Elisa, L. Yang, and N. Naik, "Dendritic cell algorithm with optimised parameters using genetic algorithm," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Jul. 2018, pp. 1–8.
[40] L. Yang, J. Li, G. Fehringer, P. Barraclough, G. Sexton, and Y. Cao, "Intrusion detection system by fuzzy interpolation," in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ-IEEE)*, Jul. 2017, pp. 1–6.
[41] N. Naik, R. Diao, and Q. Shen, "Dynamic fuzzy rule interpolation and its application to intrusion detection," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 4, pp. 1878–1892, Aug. 2018.
[42] G. Wang, J. Hao, J. Ma, and L. Huang, "A new approach to intrusion detection using artificial neural networks and fuzzy clustering," *Exp. Syst. Appl.*, vol. 37, pp. 6225–6232, Sep. 2010.
[43] P. Matzinger, "Essay 1: The danger model in its historical context," *Scandin. J. Immunology*, vol. 54, nos. 1–2, pp. 4–9, Jul. 2001.
[44] U. Mukhopadhyay, A. Skjellum, O. Hambolu, J. Oakley, L. Yu, and R. Brooks, "A brief survey of cryptocurrency systems," in *Proc. 14th Annu. Conf. Privacy, Secur. Trust (PST)*, Dec. 2016, pp. 745–752.
[45] M. Evans, L. A. Maglaras, Y. He, and H. Janicke, "Human behaviour as an aspect of cybersecurity assurance," *Secur. Commun. Netw.*, vol. 9, no. 17, pp. 4667–4679, Nov. 2016.
[46] E. Luiijf, "Understanding cyber threats and vulnerabilities," in *Critical Infrastructure Protection* (Lecture Notes in Computer Science), vol. 7130, J. Lopez, R. Setola, and S. D. Wolthusen, Eds. Berlin, Germany: Springer, 2012, pp. 52–67, doi: 10.1007/978-3-642-28920-0_4.
[47] N. Elisa, F. Chao, and L. Yang, "A study of the necessity of signal categorisation in dendritic cell algorithm," in *Proc. U.K. Workshop Comput. Intell.* Cham, Switzerland: Springer, 2019, pp. 210–222.
[48] J. Li, L. Yang, Y. Qu, and G. Sexton, "An extended Takagi–Sugeno–Kang inference system (TSK+) with fuzzy interpolation and its rule base generation," *Soft Comput.*, vol. 22, no. 10, pp. 3155–3170, May 2018.
[49] Y. Qu, G. Yue, C. Shang, L. Yang, R. Zwiggelaar, and Q. Shen, "Multi-criterion mammographic risk analysis supported with multi-label fuzzy-rough feature selection," *Artif. Intell. Med.*, vol. 100, Sep. 2019, Art. no. 101722.
[50] Q. Zhang, Y. Qu, A. Deng, and L. Yang, "Hierarchical quotient spaces-based feature selection," in *Proc. 10th Int. Conf. Adv. Comput. Intell. (ICACI)*, Mar. 2018, pp. 770–775.
[51] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*, San Mateo, CA, USA: Morgan Kaufmann, 2016.
[52] M. Azees, P. Vijayakumar, and L. J. Deborah, "Comprehensive survey on security services in vehicular ad-hoc networks," *IET Intell. Transp. Syst.*, vol. 10, no. 6, pp. 379–388, 2016.
[53] M. Azees, P. Vijayakumar, M. Karuppiah, and A. Nayyar, "An efficient anonymous authentication and confidentiality preservation schemes for secure communications in wireless body area networks," *Wireless Netw.*, vol. 27, no. 3, pp. 2119–2130, Apr. 2021.
[54] A. Maria, V. Pandi, J. D. Lazarus, M. Karuppiah, and M. S. Christo, "BBAAS: Blockchain-based anonymous authentication scheme for providing secure communication in VANETs," *Secur. Commun. Netw.*, vol. 2021, pp. 1–11, Feb. 2021.
[55] P. Vijayakumar, M. Azees, A. Kannan, and L. J. Deborah, "Dual authentication and key management techniques for secure data transmission in vehicular ad hoc networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 4, pp. 1015–1028, Apr. 2015.
[56] M. Azees, P. Vijayakumar, and L. J. Deboarh, "EAAP: Efficient anonymous authentication with conditional privacy-preserving scheme for vehicular ad hoc networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 9, pp. 2467–2476, Feb. 2017.
[57] A. Begum Shakeel Ahamed, N. Kanagaraj, and M. Azees, "EMBA: An efficient anonymous mutual and batch authentication schemes for vanets," in *Proc. 2nd Int. Conf. Inventive Commun. Comput. Technol. (ICICCT)*, Apr. 2018, pp. 1320–1326.

NOE ELISA (Student Member, IEEE) received the B.Sc. degree in telecommunication engineering from the University of Dar es Salaam, Tanzania, in 2010, the M.Tech. degree in computer networks and information security from Jawaharlal Nehru Technological University, Hyderabad, India, in 2014, and the Ph.D. degree in cyber security from Northumbria University, U.K., in 2021, under the support from the Commonwealth Scholarships Commission. He has more than ten years of teaching experience with The University of Dodoma, Tanzania, where he worked as a Lecturer. He has about 13 publications in the domain of computational intelligence and cyber security. His research interests include information security, privacy assurance, e-Government systems and smart cities, blockchain technology, computational intelligence, machine learning, human–computer interaction, and cloud computing.

LONGZHI YANG (Senior Member, IEEE) received the B.Sc. degree from the Nanjing University of Science and Technology, Nanjing, China, in 2003, the M.Sc. degree from Coventry University, Coventry, U.K., in 2006, and the Ph.D. degree from University of Wales, Aberystwyth, U.K., in 2011, all in computer science. He is currently a Professor of computer science and artificial intelligence with the Department of Computer and Information Sciences, Northumbria University, Newcastle upon Tyne, U.K. His research interests include computational intelligence, machine learning, big data, cyber security and privacy, computer vision, intelligent control systems, robotics, and the application of such techniques in real-world uncertain environments. He is a Senior Fellow of the Higher Education Academy and the Founding Chair of the IEEE Special Interest Group on Big Data for Cyber Security and Privacy.



FEI CHAO (Member, IEEE) received the B.Sc. degree in mechanical engineering from Fuzhou University, China, in 2004, and the M.Sc. degree (Hons.) in computer science and the Ph.D. degree in robotics from the University of Wales, Aberystwyth, U.K., in 2005 and 2009, respectively. He is currently an Associate Professor with the School of Informatics, Xiamen University, Xiamen, China. He has published more than 100 peer-reviewed journals and conference papers. His research interests include developmental robotics, machine learning, and optimization algorithms.

NITIN NAIK (Member, IEEE) received the Polytechnic, B.Sc., M.Sc., M.Tech., M.B.A., and M.S.W. degrees in electrical engineering and the Ph.D. degree in computer science from Aberystwyth University, Aberystwyth, U.K. He is currently a Senior Lecturer with the School of Informatics and Digital Engineering, Aston University, Birmingham, U.K. He has authored more than 100 peer-reviewed articles in the areas of artificial intelligence, cybersecurity, big data, cloud computing, the Internet of Things, and game-based learning.

TOSSAPON BOONGOEN received the Ph.D. degree in computer science from Cranfield University, U.K., in 2003. From 2007 to 2011, he was a PDRA and a Visiting Research Fellow at the Department of Computer Science, Aberystwyth University, U.K. He is currently an Associate Professor with the School of Information Technology, Mae Fah Luang University. He has published articles in well-known venues like IEEE TRANSACTIONS OF KNOWLEDGE AND DATA ENGINEERING, IEEE TRANSACTIONS OF CYBERNETICS, *Machine Learning*, *Expert Systems with Applications*, and *Artificial Intelligence and Law*. His research interests include AI, machine learning, data science, uncertainty, and fuzzy systems. He serves as an Associate Editor for IEEE ACCESS, *PeerJ Computer Science*, *International Journal of Artificial Intelligence*, and *ICT Express*.