Enhancing Resilience in IoT Water Systems Using Data-Intelligence and Decentralisation

Haitham Mahmoud, Wenyan Wu, Mohamed Medhat Gaber and Yonghao Wang

***Abstract*—In recent years, concerns regarding the security of water networks have escalated due to the increasing integration of water assets (actuators and sensors) with the Internet, combining Information Technology (IT) and Operation Technology (OT). This integration promises improved services for water networks but also introduces the risk of cyber-attacks and physical threats. As a result, there is a growing need for novel security measures to protect integrated Cyber-Physical Systems (CPS) in water distribution systems (WDSs). This article assesses actual incidents and potential Cyber-Physical (CP) attacks on water systems, explores their operational impacts, and suggests mitigating measures. It introduces a secure architecture for an integrated CPS in WDS. The study incorporates attack detection and data validation models to enhance system robustness and reduce risks, adhering to the security criteria of Water 4.0. First, the attack detection model utilizes a two-stage architecture employing six Machine-Learning (ML) algorithms, resulting in developing a simulation model with the best-suited configuration. Second, the data validation model uses blockchain technology on transmitted data, creating a simulation model for water consumption data with various input types, consensus mechanisms, and data output conversion methods. Finally, this article provides a foundation for researchers, professionals, and operators in the water sector to experiment with, evaluate, and further develop this secure architecture for their water systems. Simulating their networks using the proposed architecture allows them to identify the most suitable configurations and parameters for their specific implementations.**

***Index Terms*—Water Systems, Provisioning IoT Security, Data Intelligence, Blockchain.**

I. Introduction

The rapid development of smart water networks, also known as Water 4.0, which seamlessly integrate IT and OT, has created an immediate demand for improved security in WDS. While this integration can potentially optimize water infrastructure and services, there is also an increased risk to safety. Water supply, water treatment, water distribution, and water sanitary removal (wastewater) are the four vital components of the water supply and distribution systems. The water supply system called the supply-side water distribution system, provides households with treated water transported from various sources, including reservoirs, dams, aquifers, wells, and aqueducts. The main goal of water treatment is to remove biological contaminants using filtration processes. The water distribution system uses smart meters, water tanks, and pumps to make transporting water through pipelines easier. Finally, sewers and subsystems are used in the sanitary and wastewater removal system to move untreated water to treatment facilities.

The authors work in the Birmingham City University, UK.
Manuscript received Month XX, 2023; revised Month day, 2023.

Since many of these systems date back to the late 1800s, they have aged and may require replacement or repair, potentially making them susceptible to disruptions and security threats. Protecting modern smart water networks necessitates controlling public access and implementing advanced security measures. In the context of the Smart Water Network (SWN), consisting of five layers - physical, sensing and control, collection and transmission, data management and presentation, and data fusion and analysis, cyber-physical systems are integrated into water infrastructure, enhancing automation, data analysis, and safeguarding against both cyber and physical threats. The risk is further highlighted by a 2019 RiskIQ report, revealing that cybercrime costs over $1.1 million every minute and affects more than 1,800 individuals by causing infrastructure damage and service disruptions [1]. The ICS security market, exhibiting a Compound Annual Growth Rate (CAGR) of 7.6% from 2014 to 2019 and witnessing an increase in investments for infrastructure and asset protection from $7.82 billion to $11.29 billion, underscores these vulnerabilities [2].

A secure smart water management system is needed to defend against ever-evolving cyber threats in contemporary water distribution networks. However, studies continue to focus on attack detection, accuracy and false positives due to a lack of existing simulators. Creating a method that minimizes false positives in an integrated threat detection system is essential to improving infrastructure security. Moreover, reliable data verification is necessary for smart water management. The literature on this subject lacks in-depth implementation. Moreover, a comprehensive investigation into data verification is lacking, which exposes a security flaw. Implementing an effective data verification system is essential to provide an additional line of defence against potential attacks.

A resilient smart water management system is proposed to detect and localize intrusions and enhance infrastructure resilience. Data intelligence and decentralization are key components of the system. This system uses blockchain technology for secure measurement verification inside sensor devices and ML for attack detection and localization.

- An attack detection system is proposed that utilises six state-of-the-art ML algorithms: Isolation Forest (IF), Bag of SFA-SAX Symbols (Boss), Random Forest (RF), Support Vector Machine (SVM), k-Nearest Neighbors (KNN), and Extreme Gradient Boosting (XGBoost). Moreover, four different-size WDS are included in the evaluation: Net1, Mini-town, C-town, and D-town. This assessment examines characteristics including accuracy, detection time, and time-to-detect while thoroughly testing the system's functionality under four attack types.

- The data verification system is proposed using blockchain technology. Two alternatives are investigated for data entry: 1) time series data relevance and 2) integration with one timestamp using a water modelling called EPANET. Moreover, four different consensus mechanisms working within the same water distribution systems are included in a thorough analysis. The system is evaluated based on the number of transactions, generation and mining time, throughput, time per transaction and transaction Efficiency, systematically assessing the performance of each mechanism.

This article is organised as follows: Section II presents the threat model and the existing study. Section III proposes the architecture by discussing the key components, attack detection and visualisation, simulation and GUI in the system. Section IV discusses the results and the future research directions. Section VI aims to address the future research directions. Finally, section VI concludes the work.

TABLE I
EVALUATION OF THE EXISTING STUDIES THAT USED TECHNOLOGIES (I.E., BLOCKCHAIN, SECURE IOT OR TRAFFIC ANALYSIS) FOR SECURITY, PRIVACY OR RESILIENCE IMPACT.

	System	Purpose	Tech. used			Impact		
			Blockchain	Secu. IoT	Traffic Analysis	Security	Privacy	Resilience
[7], [8]	Watering System	Watering Management	√	×	√	×	×	√
[5]	Watering System	Encrypted Comm.	×	√	×	√	×	√
[6]	WDS	Secure WDS	√	√	×	√	√	√
[9], [10]	WDS	Attack Detection	×	×	√	√	×	√
[11], [12]	WDS	Data verification	√	×	×	√	×	√
This paper	WDS	Data verification and Attack Detection	√	√	√	√	√	√

II. BACKGROUND

A. Threat Model

We aim to implement a secure smart water network with two primary objectives: 1) to mitigate CPS attacks on WDS, and 2) to validate shared data on the WDS within decentralized connected networks. We have developed a threat model that considers honest network peers to achieve these objectives. Effectively addressing potential threats and vulnerabilities is crucial for securing smart water networks. Robust access controls and authentication are essential to prevent unauthorized data access and safeguard sensitive information. Data integrity checks and encryption are critical as data tampering can compromise information integrity. Intrusion detection and device authentication are necessary to counter the infiltration of malicious nodes, including compromised Internet of Things devices. Real-time monitoring and intrusion detection systems are indispensable to avoid prolonged breaches due to delayed attack detection. Redundancy and proactive maintenance are vital to minimize the impact of service interruptions on customers. Considering centralization risks and ensuring transaction transparency are also important factors. In addition, verifying the data within a smart water network is crucial to ensuring its reliability and accuracy.

B. Related Works

This section provides an overview of studies in water systems employing blockchain, ML, or IoT security to leverage security (as shown in Table I). Blockchain has gained significant attention in watering systems and agriculture due to its ability to establish transparent, immutable water usage records. This ensures equitable water rights allocation and elevates the overall system security by safeguarding against data manipulation and unauthorized access. In parallel, other water systems use blockchain to promote water rights and transparency.

ML takes the lead in smart water systems, primarily dedicated to attack detection. These smart water systems encompass diverse domains, including distribution, supply, wastewater, drinking water, and irrigation. Some studies use ML and blockchain for multifaceted applications like seed monitoring, leakage detection, and predictive maintenance [3], [4]. However, it is crucial to highlight that the paramount focus in these integrated systems remains on enhancing security, with particular emphasis on identifying and thwarting potential attacks, thus creating resilient water infrastructure. Furthermore, blockchain and ML are important water distribution systems and irrigation technologies. Some studies have ventured into encryption for securing data transmission and enhancing data privacy and integrity [5], [6]. However, it is noteworthy that the primary objective of these investigations centres on leveraging the security and enhancing the resilience of water infrastructure, with relatively limited attention given to addressing privacy concerns.

This highlights the insufficiency of existing studies that comprehensively leverage the security potential presented by the convergence of IoT security, ML, and Blockchain.

III. PROPOSED ARCHITECTURE

The proposed WDS architecture comprises four core components: the WDS model on EPANET, data validation system (using blockchain technology), cloud/storage, and system services (see Figure 3). The EPANET WDS model allows users to design water distribution systems and is known for its adaptability and integration capabilities. The data validation process validates transmitted measurements from WDS assets like water tanks, using various methods, forming a chain of hashed data. Metadata is stored in the cloud for reference, and both data chains and metadata are compared using a fundamental hash function. Finally, data analysis and processing occur through system services applications, including SCADA, ICS, and attack detection.

A. Key Components

1) Use-cases: Four use cases cover a variety of WDS configurations and characteristics. With a granularity of one

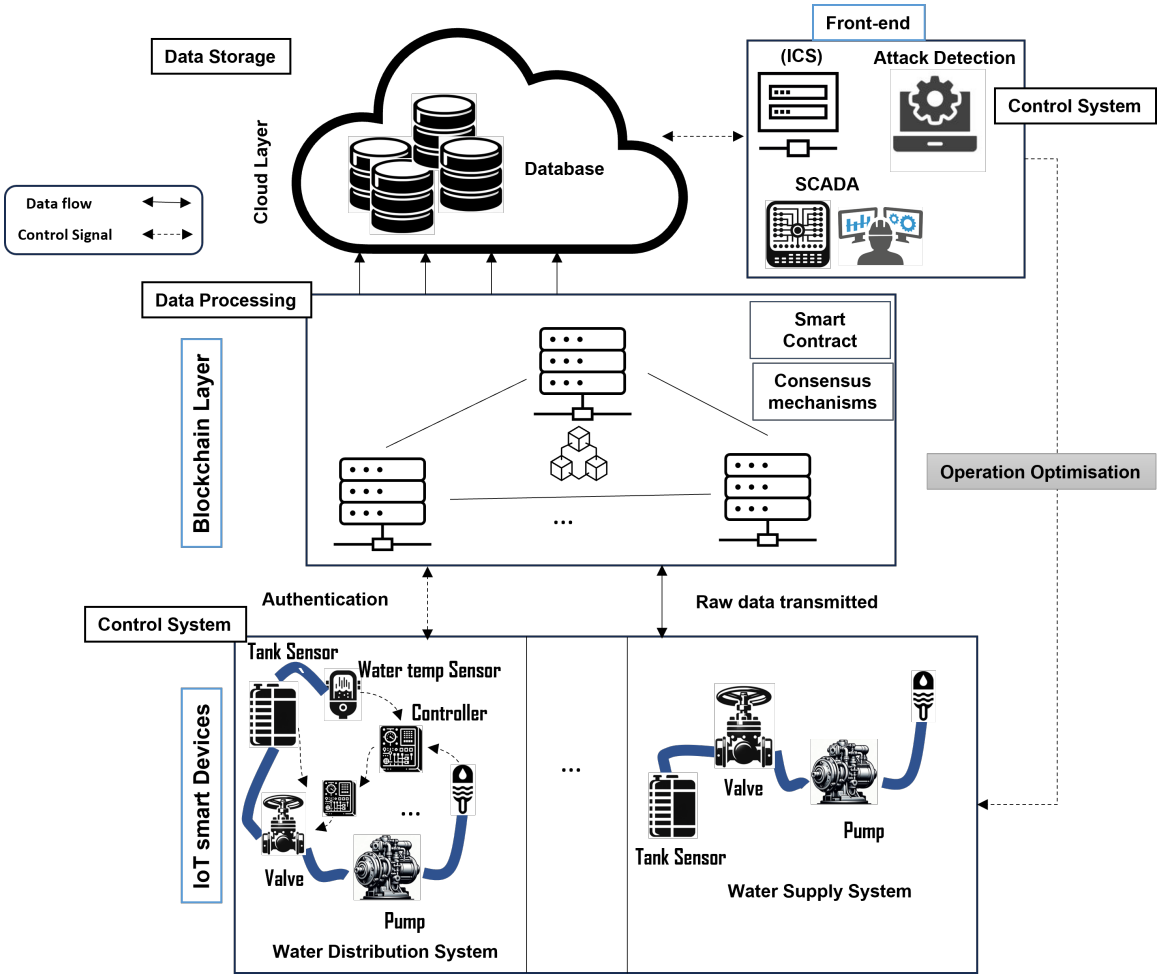

Fig. 1. The suggested architecture of secure IoT-WDS to enhance water resilience using Data intelligence and Decentralization. The architecture focuses on two key components which are data validation using blockchain and attack detection using ML.

hour, Net1 comprises 11 nodes, 12 pipelines, one pump, one tank, and one reservoir. Its duration is 336 hours. Mini-town has ten nodes, ten pipelines, two pumps, one tank, and one reservoir. It runs for 168 hours at a granularity of 20 minutes. C-Town is a 72-hour operation with a 6-minute granularity that consists of 429 pipelines, 11 pumps, seven tanks, one reservoir, and 396 nodes. Finally, D-Town uses 407 nodes, 443 pipelines, 11 pumps, seven tanks, and one reservoir over 72 hours, albeit with a 20-minute granularity.

2) IoT devices and Authentication: Cryptographic techniques are used for authentication between blockchain networks and IoT devices. Every IoT device is equipped with an individual public-private key pair ($DeviceKey\_public$, $DeviceKey\_private$), in which the private key is safely stored ($DeviceKey\_private$). The blockchain network creates a cryptographic challenge ($Challenge$), essentially a hashed value of particular data, and sends it to the device in response to an authentication request. The machine uses its private key to answer the challenge with a response (Response): Response = Hash($Challenge$, $DeviceKey\_private$). The blockchain network verifies the response by hashing the challenge using the public key linked to the device's identity and ensuring that Response = Hash($Challenge$, $DeviceKey\_public$). When verification is successful, the device is verified, access is granted via smart contracts, and an immutable audit trail is created for security and transparency.

3) Encryption and Data signature: A combination of encryption, hashing functions, and digital signatures is used to ensure messages are secure during transmission. The sender encrypts consumption data using the recipient's public key, ensuring that only the intended recipient can decrypt it with the private key. The blockchain generates digital signatures that can be used to verify the authenticity and source of data. Altering the data changes the digital signature, preventing potential attackers from manipulating consumption amounts. A bloom filter and zero-knowledge proof verify the blockchain's integrity. Zero-knowledge proofs verify data consumption without revealing customer identity, while bloom filters match pseudonyms with block data. Blockchain security and trustworthiness are ensured through the verification process overseen by a mining node.

4) Data Validation and Blockchain: These are pivotal elements of any blockchain system, overseeing communication and verification processes through consensus mechanisms. This study introduces five consensus mechanisms—Proof of Work (PoW), Proof of Trust (PoT), Proof of Assignment

(POA), and Proof of Vote (PoV) to accommodate fast-action, voting-based, and real-time verification in the context of water distribution systems. PoA is recommended due to the WDS phenomena' fast-paced and voting-based nature. PoW, the initial consensus mechanism derived from Bitcoin, requires all network peers to validate transactions before rewarding the first verifier. PoT relies on the peer with the highest reputation for data verification, employing a trust matrix with reputation increments for approved data. PoV uses a voting system for data verification, requiring over 3/4 of the feedback to indicate genuine data. PoA offers low-processing and fast verification, suitable for IoT systems with less stringent security needs.

5) Cloud Layer and Front-end APIs: When storing data on a blockchain, it is usually not saved directly but rather as a hash to the original data. A cloud-based storage system occupies the actual data. The hash value is a link to the cloud data and is used for verification. Access to the cloud, where the data is stored, is necessary to recover the original data. The hash value and the reference to the original data are all on the blockchain; the actual data is not there. When data is accessed on the blockchain, the data is retrieved and matched from the cloud using the recorded information on the blockchain.

Establishing seamless connectivity with front-end applications, such as SCADA and ICS, is essential for improving the resilience of intelligent water management. One of the primary concerns of this article is the attack detection system. By achieving seamless integration and prioritizing attack detection, we aim to enhance the system's ability to monitor, control, and optimize water-related processes, ensuring the security and reliability of our critical water infrastructure.

6) Attack Detection and Visualisation: This study comprehensively evaluates attack detection systems and their resilience against malicious intrusions. Due to the prevalence of sensor tampering in critical infrastructure sectors like WDS, this article explores four different WDS scenarios and generates and assesses attacks against them. The effectiveness of SVM, KNN, RF, Boss, XGBoost, and IF algorithms in countering these attacks is assessed. The SVM specializes in identifying patterns, assisting in the detection of abnormalities that may indicate a possible attack. KNN assesses data point proximity, which aids in the discovery of anomalies in water system behaviour. RF improves attack detection by combining multiple decision trees with its ensemble learning technique. Boss, a symbolic representation-based program, detects and analyzes complicated patterns in water system data, revealing possible risks. As a strong gradient-boosting algorithm, XGBoost iteratively refines its understanding of system dynamics, resulting in better attack detection. Finally, IF's ability to separate anomalies is useful in identifying odd occurrences or attacks by differentiating them from the majority of typical system behaviour. Together, these classifiers offer a robust protecting water system from potential security threats.

Assessment time, accuracy, precision, and F1 score are evaluation metrics used in the performance assessment of attack detection systems. Assessment time measures the time necessary to identify potential attacks in water systems. Accuracy indicates the system's overall correctness in detecting both attacks and non-attacks. Precision quantifies the system's ability to identify attacks while minimizing false alarms correctly. The F1 score provides a balanced assessment of precision and recall, considering both false positives and false negatives in attack detection.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

B. Simulation and GUI

Data entry in the tool operates in two distinct modes: direct integration with EPANET, a prominent water modelling software, and the importation of time series data from CSV files. The direct integration is designed for real-time data processing, involving the transmission of individual timestamps. This differs from the importation of time series data, which handles pre-extracted time series data from EPANET, encompassing a wealth of datasets for comprehensive analysis.

Time series data is used, either directly from CSV files in MATLAB. The attack detection API extensively studies two detection techniques—self-supervised and unsupervised. Users select their preferred method during parameter selection in the model's graphical user interface (GUI). Data training and testing in different training and testing approaches—each with unique benefits—is determined by the selected method. Once time-series water data has been analyzed, attack origins can be identified via an attack localization technique. Users give information about the consensus method and data source (water distribution model or Excel file). Hashing-based validation (such as PoW, PoA, PoT, and PoV). Recognition of the chain or rejection for security-related reasons is established by data authorization. While the data validation model shows chained and dropped data and evaluates consensus process performance using multiple coefficients (e.g., latency, number of blocks per transaction), the attack detection model visualizes attacked assets and evaluates performance across data output coefficients (e.g., accuracy, precision, time-to-detect).

An intuitive GUI is developed using MATLAB, featuring user-friendly drop-down menus for dataset selection and user-activated buttons for ML algorithm selection. Moreover, the water distribution systems are visualised, emphasising the precise localisation of the attacks through highlighted sections.

IV. Discussion

A. Security Resilience and Attack Detection

Using EPANETCPA in MATLAB, data manipulation attacks are created within water distribution systems [13]. Four attacks are generated for security assessment as the following

- #Attks1 and 2 involves manipulate water tank sensor (T7) control signal before sending it to the SCADA. Through signal alteration, a potentially drought-causing false HIGH signal is falsely indicated when it is LOW.
- #Attks3 and 4 display inaccurate water level information, leading to operational failures like leaving the pump running and creating flooding.

TABLE II
EVALUATION OF ATTACKS DETECTION TOWARDS FOUR WATER MODELLING SYSTEMS (NET1, MINI-TOWN, C-TOWN, AND D-TOWN), IN WHICH 🏴‍☠️: FAILED TO DETECT AND 🛡️: DETECTED THE ATTACK .

WDS	#	Detection Algorithms					
		RF	SVM	KNN	Boss	XGboost	IF
Net1	1	🛡️	🛡️	🛡️	🛡️	🛡️	🛡️
Mini-town	2	🛡️	🛡️	🛡️	🏴‍☠️	🛡️	🛡️
C-Town	2	🛡️	🏴‍☠️	🏴‍☠️	🛡️	🛡️	🛡️
	3	🛡️	🏴‍☠️	🏴‍☠️	🏴‍☠️	🛡️	🛡️
	4	🛡️	🏴‍☠️	🏴‍☠️	🏴‍☠️	🛡️	🛡️
D-Town	2	🛡️	🏴‍☠️	🏴‍☠️	🏴‍☠️	🛡️	🛡️
	3	🛡️	🏴‍☠️	🏴‍☠️	🏴‍☠️	🛡️	🛡️
	4	🛡️	🏴‍☠️	🏴‍☠️	🏴‍☠️	🛡️	🛡️

Figures 4 present the assessment results, including evaluating assessment time, accuracy, precision, and F1 score across six algorithms for four datasets. With increasing dataset size, XGBoost, Boss, and SVM show the longest assessment times. Boss, XGBoost, RF and IF algorithms reach 5.7, 5.2, 4.9 and 4.5 seconds for the C-Town Dataset, indicating higher computational costs associated with more complex models. Regarding accuracy, all algorithms achieve 100% or 0% based on detected attacks, with Net1, Minitown and D-Town datasets showing nearly 100% accuracy, except when using SVM for Minitown, while D-town exhibits lower accuracy. Boss performs 35% accuracy for the C-Town because of its limitation in detecting attacks in the C-Town dataset. SVM performs the least effectively in precision attributed to its inherent limitations in handling high-dimensional data or datasets with significant overlap between classes, resulting in more false positives. KNN achieves 100% precision for Net1, demonstrating the effectiveness of true positives with zero false positives. RF, XGBoost, and IF achieve 100% precision for C-town and D-, highlighting the effectiveness of the ensemble methods for detecting attacks. 100% precision indicates all detected attacks are accurate, while 0% precision means none of the detected attacks are correct. RF, XGBoost, and IF exhibit the highest F1 scores across datasets, showcasing a strong balance between precision and recall, effectively measuring a model's attack detection accuracy while considering false positives and negatives. To enhance the system's accuracy and robustness, a combination of diverse learning techniques—namely Deep Neural Networks (DNN), RF, and IF is employed in a voting mechanism. The accuracy rates for Net1, Minitown, C-town, and D-town are 100%, 100%, 100%, and 80% respectively. This approach ensures that the detection process is dependable, as it is validated by multiple learning methods.

B. Data Validation

This system assesses system complexity using latency and throughput. Latency is the time it takes for a transaction to be verified and irreversible, determined by block generation time (t_G) and mining time (t_m). Throughput, measured in transactions per second ($TPS = \frac{N_T}{t_L}$), depends on the number of transactions (N_T) and latency ($t_L = t_G + t_{m,\delta}$) in which δ is the consensus mechanism. Security is paramount, and consensus mechanisms like PoW and PoT aim to ensure security by preventing majority power control (over 51% and 33.3% mining power, respectively). The formula $\nu_M \leq \nu_{T,\delta}$ helps guarantee security against malicious verifiers (ν_M), considering the number of true validators (ν_T) and total nodes (N). While PoA has the least behaviour, PoW allocates all peers in parallel, resulting in longer mining times. Other water distribution systems exhibit similar behaviour. The analysis presented in Table III investigates the Net1, Mini-town, C-town, and D-town datasets, focusing on consensus mechanisms including PoW, PoT, PoA, and PoV. Each block is constructed with one measurement from each sensor, and the number of transactions varies from 336 to 953. The study calculates these scenarios' generation, mining, total time, throughput, time per transaction and transaction Efficiency. Notably, the results reveal that PoT and PoA achieve the highest throughput, followed by PoV and PoW, as they require only one selected validator compared to PoV and PoW, which mandate the participation of all nodes in the validation process. The throughput spans 1.35 to 8.07, depending on the dataset and the consensus mechanism employed. The time per transaction varies from 0.13240 to 0.927615. Furthermore, the transactional efficiency ranges from 0.00534 to 0.127

Implementing blockchain at the water tank level safeguards sensing and control signals, preventing unauthorized alterations that could lead to potential flooding or drought. Some implementations in the network have failed to incorporate consensus mechanisms with an authorization point for data approval. Regarding data validation, water systems require strong security and rapid verification. To meet these criteria, PoT and PoV are proposed consensus mechanisms. PoV offers stronger security but consumes more energy and experiences delays due to data processing by all chosen blockchain peers. Given the limited number of blockchain nodes in water systems, real-time processing remains feasible. If stringent security measures are not necessary, PoT can be used to reduce resource consumption.

Using water tanks as authorized blockchain nodes addresses the significant processing requirements for transmitting and verifying water measurement data, especially when other components of the EPANET system lack computational ability. The mining time for verifying transactions varies according to the consensus techniques used. PoA has the shortest mining time, while PoW has the longest due to its intensive parallel data verification process. The PoT and PoV consensus processes are recommended in the context of water systems due to their ability to provide comprehensive security while allowing real-time processing. These measures successfully protect the integrity and security of data. Mining, complexity, and throughput parameters were used to analyze and compare the efficiency and performance of several consensus techniques, including PoA, PoT, PoW, and PoV. These measurements provide useful information about their unique capabilities and performance characteristics.

Blockchain technology provides several security benefits in water distribution systems, such as ensuring service availability, data immutability, transparency, and preventing single points of failure. It uses mining problems to train detection models and improves data integrity by disseminating datasets

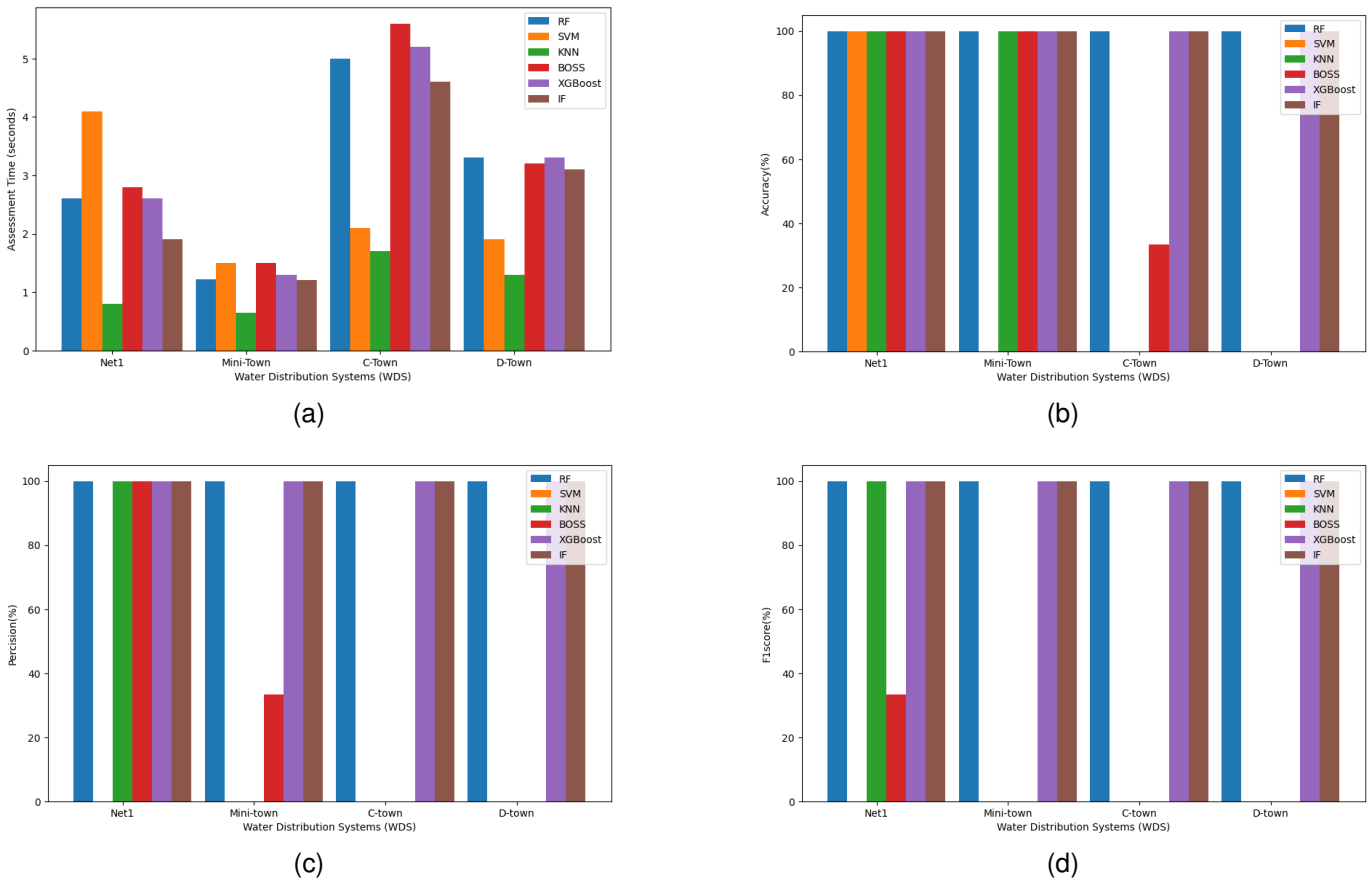

Fig. 2. Performance evaluation of attack detection system using six classifiers (RF, SVM, KNN, Boss, XGboost and IF) on four datasets (NEt1, Mini-town, C-town and D-Town) in terms of (a) Assessment time, (b) accuracy; (c) precision; and (d) F1score.

and learning algorithms across many blockchain nodes. However, key differences include the lack of communication protocol simulation, the exclusion of WDS controllers (PLCs), the disregard for broadcasting delays and data transfer processing, the requirement for direct communication among selected blockchain nodes, and vulnerabilities to Man-in-the-Middle (MiTM) attacks in certain consensus methods such as PoA. These distinctions emphasize the system's unique features and limits compared to real-world blockchain implementations in water distribution networks.

V. Future Research Directions

The extensive research on this topic offers numerous opportunities for future work within water applications and related IoT domains. This research includes:

- Integrating communication protocol simulations with WDS controllers is crucial. Simulating communication protocols allows for the testing and optimization of consensus mechanisms under various conditions, identifying potential issues like bottlenecks, latency, and security vulnerabilities. Including WDS controllers in these simulations adds a layer of realism, as these controllers manage physical components like pumps and valves, which significantly affect system performance and reliability.
- Extending the model to include hardware implementation is another point to explore. Implementing the best-suited model specifications, selected based on this model's outcomes, can be applied to real water distribution systems. This implementation can be investigated on various platforms, including Amazon Web Services, Google Cloud, IBM Cloud, Ethereum, Hyperledger Fabric, and IBM Watson.
- Investigating methods to optimize assessment times could significantly enhance the system's overall efficiency and responsiveness. This can enhance the scalability of this system to consider multiple subsystems in the water domain. For instance, incorporating adaptive algorithms that dynamically adjust based on network load and transaction volume could mitigate delays. Moreover, exploring parallel processing techniques or decentralized validation approaches may help distribute the computational load more evenly, thereby reducing bottlenecks and improving assessment times.
- The architecture's capabilities can be broadened to advance decision-making processes through optimization. Decision-making actions range from isolating infected components to network shutdowns for recovery. Integrating other blockchain use-cases, such as peer-to-peer trading models, can be explored to promote robust and precise decision-making, contributing to the shift towards autonomous decision-making. This system can dynamically adapt to changing conditions in the network.

TABLE III
PERFORMANCE EVALUATION OF DATA VALIDATION TOWARDS FOUR DATASETS (NET1, MINI-TOWN, C-TOWN AND D-TOWN) USING FOUR CONSENSUS MECHANISMS (POW, POT, POA AND POV) IN TERMS OF NUMBER OF TRANSACTION, GENERATION, MINING AND TOTAL TIME, THROUGHPUT, TIME PER TRANSACTION AND TRANSACTION EFFICIENCY

WDS	Cons.	Transac	Gener. time (sec)	Mining time (sec)	Total time (sec)	Throughput (TPS)	Time per Transaction (Sec/Trans.)	Trans. Efficiency (T^2/Trans.)
Net1	PoW	336	3.36	247.30	250.6643	1.34	0.746	0.00534
	PoT			48.59	51.95	6.46	0.1546	0.12435
	PoA			53.13	56.49	5.94	0.1681	0.1051
	PoV			99.50	102.86	3.26	0.30619	0.03169
Mini-Town	PoW	504	5.04	462.478	467.518	1.07	0.927615	0.002
	PoT			57.87	62.91	8.01	0.12482	0.127
	PoA			61.69	66.73	7.55	0.13240	0.1131
	PoV			260.38	265.42	1.89	0.5266	0.0071
C-Town	PoW	953	9.53	872.6	882.13	1.08	0.9256	0.00122
	PoT			109.2	118.73	8.07	0.12458	0.067969
	PoA			116.4	125.93	7.56	0.1321	0.0600
	PoV			491.3	500.83	1.90	0.5255	0.00379
D-Town	PoW	381	3.81	277.87	281.68	1.35	0.7393	0.004792
	PoT			54.6	58.41	6.52	0.1533	0.111624
	PoA			59.7	63.51	6.0	0.1666	0.09447
	PoV			111.8	115.61	3.29	0.303438	0.02845

Optimizing node selection for blockchain validators is a critical area to explore, as it can significantly impact the efficiency and security of the system. Implementing adaptive system management strategies will ensure that the network can respond to various challenges in real-time, enhancing overall resilience and performance.

- Extending the data validation model involves constructing a system that can select the most suitable nodes to serve as blockchain nodes. This selection process should be based on specific criteria like distance and computational capabilities. Since certain sensors and IoT devices may not support validation computations, introducing a novel approach for node selection becomes crucial to keep networks of varying sizes.
- Expanding the range of attack scenarios within the EPANETCPA to significantly mitigate complex advanced attacks. For instance, pump manipulation attacks can cause incorrect water pressure levels, leading to either system damage or inadequate water supply. Sensor spoofing can introduce false data into the network, resulting in incorrect system responses such as unnecessary water treatment processes or false contamination alerts. Communication interception can allow attackers to hijack control signals, causing operational disruptions and potentially endangering public health. Moreover, attacks targeting the SCADA system can disrupt the entire water distribution process, leading to severe service outages. Malware attacks on system controllers can corrupt essential software, compromising the integrity and functionality of the water distribution system. Unauthorized access to data logs can lead to the exposure of sensitive information, which can be exploited for further attacks or to undermine public trust.
- Implementing advanced detection techniques (i.e., federated learning) can significantly enhance the ability to identify and respond to threats. These methods analyze network patterns and behaviours to spot anomalies that might indicate an attack. Federated learning uses decentralized data from various sources to train models while maintaining data privacy.
- Extending on the consensus mechanism in data validation can involve ML. ML adds to alternate mining algorithms in PoW blockchains, improving security and efficiency. GENECI uses evolutionary ML to help in consensus inference by acting as an organizer for creating ensembles [14]. This includes refining data validation processes to ensure data integrity and security. Implementing multi-layered validation mechanisms can add extra layers of protection. Moreover, ML plays an essential role in safeguarding federated learning with blockchain, acting as the foundation for authenticating transactions and guaranteeing a secure and transparent federated learning environment [15].
- Enhancing the GUI is another essential point for user-friendly and adaptive system management. A well-designed GUI can simplify complex processes, making the system more accessible to users and allowing for easier monitoring and control. Integrating these elements can lead to a more robust and intuitive system, capable of adapting to various scenarios and maintaining optimal performance.

VI. CONCLUSION

In this paper, a proposed WDS architecture is a comprehensive framework comprising four basic components: an EPANET-based WDS model, a data validation system based on blockchain technology, cloud storage, and system services. This architecture integrates sophisticated technology to satisfy the requirement for secure and robust water distribution systems. The EPANET paradigm simplifies the construction of water distribution networks, while the blockchain-based data validation mechanism protects the confidentiality and integrity of data transmission. Cloud storage is used for data reference and system services such as SCADA, ICS, and

threat detection. The design exhibits its ability to improve the security of water distribution systems by effectively identifying attacks and safeguarding data integrity through a thorough review of ML methods. Furthermore, the study provides various consensus processes, providing insights into their effectiveness in data security and timely verification.

The proposed WDS design uses advanced technology to improve the security and resilience of water distribution systems. It combines an EPANET-based architecture, blockchain-based data validation, cloud storage, and system services to form a comprehensive framework for water infrastructure management. Examining ML algorithms and consensus methods provides vital insights into enhancing attack detection and data validation. This research helps the development of reliable and secure water distribution systems by tackling critical infrastructure concerns in the context of constantly changing technologies and threats.

FUNDING

This research has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie Training Networks (ITN)-IoT4Win grant agreement No. [765921]

REFERENCES

[1] Francis Bignell. Riskiq find cybercrime cost organisations just under $1.8million per minute, Jul 2021.

[2] Evgeny Goncharov, Kirill Kruglov, and Yuliya Dashchenko. Five ics cybersecurity myths based on kaspersky lab ics cert experience. *Automatisierungstechnik*, 67(5):372 to 382, 2019.

[3] Qi Peng, Lifen Tu, Yunyun Wu, Zhenyu Yu, Gerui Tang, Wei Song, et al. Automatic monitoring system for seed germination test based on deep learning. *Journal of Electrical and Computer Engineering*, 2022, 2022.

[4] Abdel Bayoumi and Rhea McCaslin. Internet of things–a predictive maintenance tool for general machinery, petrochemicals and water treatment. In *Advanced Technologies for Sustainable Systems: Selected Contributions from the International Conference on Sustainable Vital Technologies in Engineering and Informatics, BUE ACE1 2016, 7-9 November 2016, Cairo, Egypt*, pages 137–146. Springer, 2017.

[5] Cherine Fathy and Hassan M Ali. A secure iot-based irrigation system for precision agriculture using the expeditious cipher. *Sensors*, 23(4):2091, 2023.

[6] Haitham Hassan M Mahmoud, Wenyan Wu, and Yonghao Wang. Secure data aggregation mechanism for water distribution system using blockchain. In *2019 25th International Conference on Automation and Computing (ICAC)*, pages 1–6. IEEE, 2019.

[7] Hui Zeng, Gaurav Dhiman, Ashutosh Sharma, Amit Sharma, and Alexey Tselykh. An iot and blockchain-based approach for the smart water management system in agriculture. *Expert Systems*, 40(4):e12892, 2023.

[8] Bhabendu Kumar Mohanta, Sangay Chedup, and Mohan Kumar Dehury. Secure trust model based on blockchain for internet of things enable smart agriculture. In *2021 19th OITS international conference on information technology (OCIT)*, pages 410–415. IEEE, 2021.

[9] Haitham Mahmoud, Wenyan Wu, and Mohamed Medhat Gaber. A time-series self-supervised learning approach to detection of cyber-physical attacks in water distribution systems. *Energies*, 15(3):914, 2022.

[10] Riccardo Taormina, Stefano Galelli, Nils Ole Tippenhauer, Elad Salomons, Avi Ostfeld, Demetrios G Eliades, Mohsen Aghashahi, Raanju Sundararajan, Mohsen Pourahmadi, M Katherine Banks, et al. Battle of the attack detection algorithms: Disclosing cyber attacks on water distribution networks. *Journal of Water Resources Planning and Management*, 144(8):04018048, 2018.

[11] Haitham H Mahmoud, Wenyan Wu, and Yonghao Wang. Wdschain: A toolbox for enhancing the security using blockchain technology in water distribution system. *Water*, 13(14):1944, 2021.

[12] Thandile Nododile and Clement Nyirenda. A blockchain-based secure data collection mechanism for smart water meters. In *2023 IST-Africa Conference (IST-Africa)*, pages 1–8. IEEE, 2023.

[13] Riccardo Taormina, Stefano Galelli, HC Douglas, Nils Ole Tippenhauer, Elad Salomons, and Avi Ostfeld. A toolbox for assessing the impacts of cyber-physical attacks on water distribution systems. *Environmental modelling & software*, 112:46–51, 2019.

[14] Adrián Segura-Ortiz, José García-Nieto, José F Aldana-Montes, and Ismael Navas-Delgado. Geneci: A novel evolutionary machine learning consensus-based approach for the inference of gene regulatory networks. *Computers in Biology and Medicine*, 155:106653, 2023.

[15] Haitham HM Mahmoud, Wenyan Wu, and Yonghao Wang. Proof of learning: Two novel consensus mechanisms for data validation using blockchain technology in water distribution system. In *2022 27th International Conference on Automation and Computing (ICAC)*, pages 1–5. IEEE, 2022.

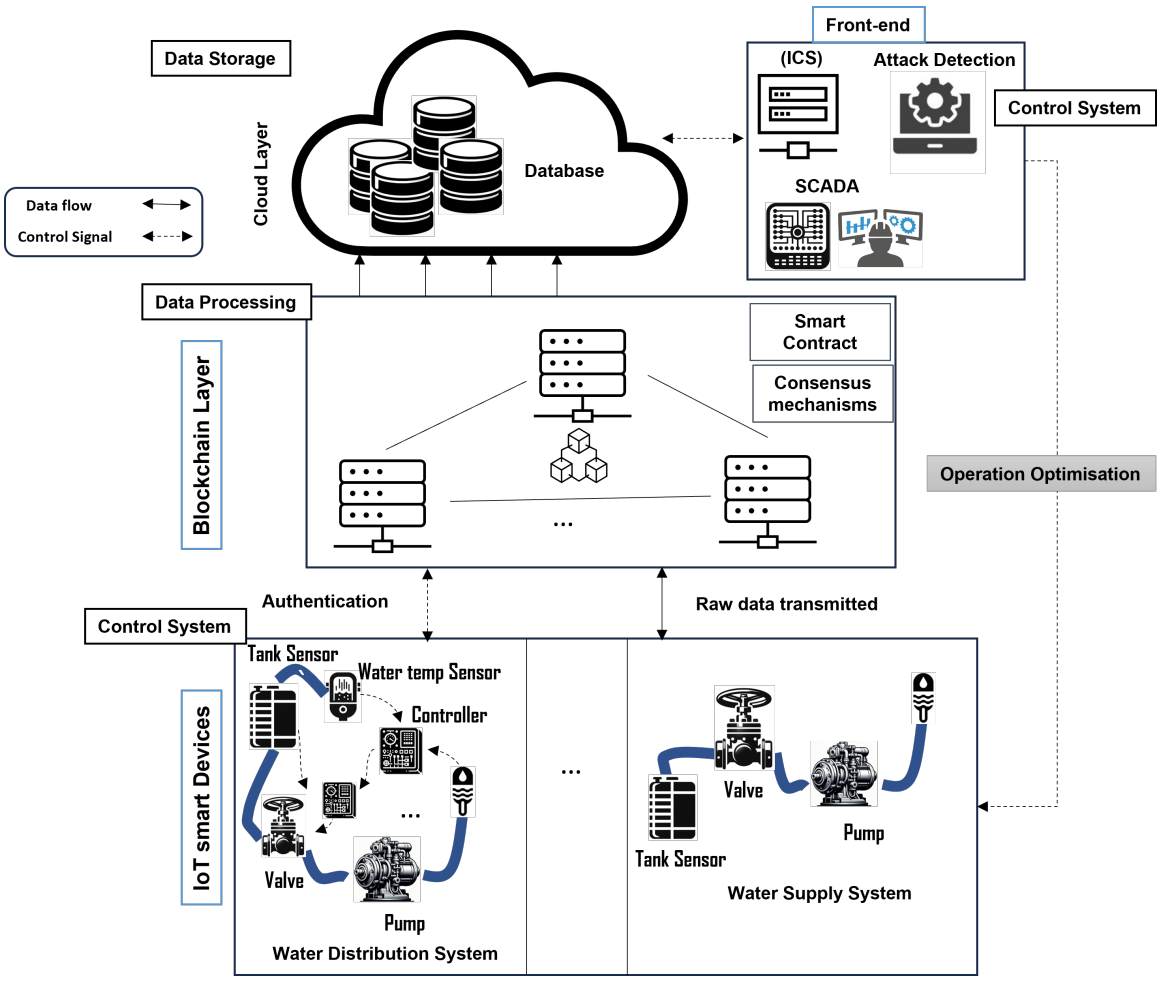

Fig. 3. The suggested architecture of secure IoT-WDS to enhance water resilience using Data intelligence and Decentralization. The architecture focuses on two key components which are data validation using blockchain and attack detection using ML.

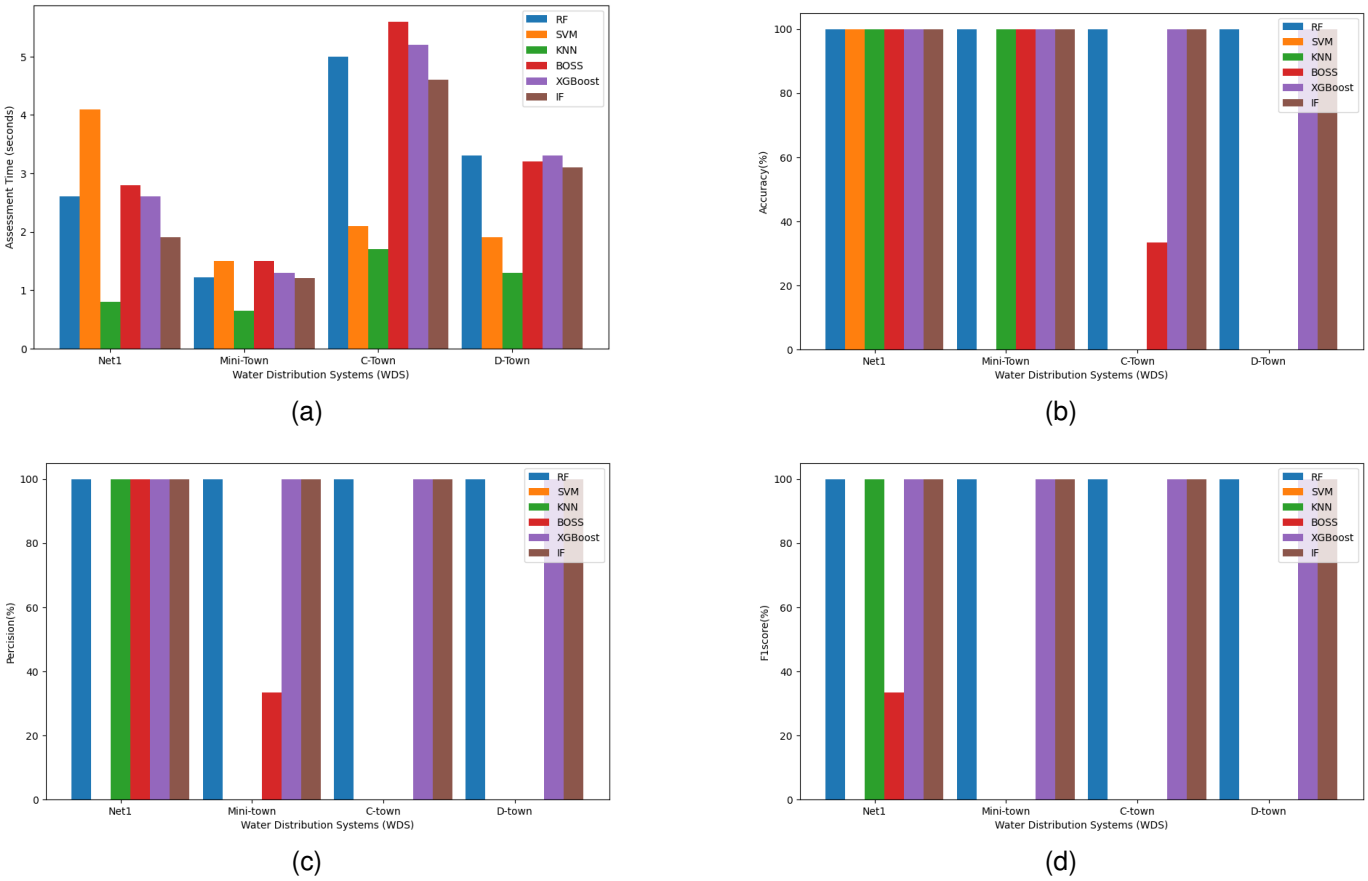

Fig. 4. Performance evaluation of attack detection system using six classifiers (RF, SVM, KNN, Boss, XGboost and IF) on four datasets (NEt1, Mini-town, C-town and D-Town) in terms of (a) Assessment time, (b) accuracy; (c) precision; and (d) F1score.