INTERNATIONAL JOURNAL OF INNOVATIONS IN ENGINEERING TECHNOLOGY AND APPLIED SCIENCES (IJIETAS)

An Intelligent Trust Prediction and Adaptive Data Placement Framework for Secure Cloud–Edge Data Sharing

[1*]Kalpana Dwivedi, [2]Gaurav Raj

[1],[2]Department of Computer Science & Engineering Sharda School of Computing Science and Engineering Sharda University, Greater Noida.
[1]dwivedi.dolly@gmail.com, [2] Gaurav.raj@sharda.ac.in

ABSTRACT

Cloud-edge computing has undergone significant growth, adding to the growing demand for security, reliability, and adaptive solutions to distribute sensitive data across diverse and often untrustworthy edge nodes. Traditional approaches to data placement mainly rely on static data placement policies and/or availability of resources, potentially allowing for exposure of data to unreliable and/or compromised nodes. In order to solve this problem, in this paper, an innovative secure framework of Intelligent Trust Prediction and Adaptive Data Placement (ITADP-Edge) for intelligent data distribution in the cloud–edge environment is presented. This corresponds to behavioral monitoring, trust prediction via machine learning, adaptive node selection in order to ensure low-latency networking, encryption, erasure coding, integrity verification, and fault recovery being compiled in one security framework. The trust-prediction component was tested by the CICIDS2017 dataset and was modeled with Decision Tree, Random Forest, XGBoost, and LightGBM models for separating benign and potentially malicious behavior. The proposed LightGBM-based ITADP-Edge achieved 99.84% accuracy, 99.73% precision, 99.83% recall, 99.78% F1-score, and 0.9999 ROC-AUC. Moreover, the framework adds trust information to the nodes' ranking based on their predicted trust values, allowing the nodes to place data more securely and adaptively instead of statically distributing data. Encryption helps to provide data confidentiality, and erasure coding increases resistance to the failure of nodes and data integrity. The proposed framework gives strong support to secure, trustworthy, and fault-tolerant cloud-edge distribution of both data and information with low latency.

Keywords: Cloud–Edge Computing; Secure Data Distribution; Trust Prediction; Machine Learning; Adaptive Data Placement; Trust Management; Erasure Coding; Fault Tolerance; Data Security; Edge Computing.

ARTICLE HISTORY
Received: 15 July 2026, **Accepted:** 28 September 2026, **Published:** 12 October 2026

CITATION
Dwivedi, K., (2026). An Intelligent Trust Prediction and Adaptive Data Placement Framework for Secure Cloud–Edge Data Sharing, *International Journal of Innovations In Engineering Technology And Applied Sciences (IJIETAS),* 2(2), 09-26.
DOI: https://doi.org/10.64764/ijietas.v2.i2.02

1. Introduction

Cloud computing is reshaping the data storage and processing landscape and delivering the scalable resources of computing and storage on demand. But as the amount of information produced by connected devices and IoT applications and also for latency-sensitive services has grown, it has also been revealed that a pure centralized cloud-first approach is facing its limits. Therefore, edge and fog computing have risen to be complementary emerging paradigms that bring computation and storage to the edge of the data, minimizing the delay in communication and the requirement for bandwidth in supporting geographically distributed applications [1], [2]. Distributed cloud-edge architectures, however, also lead to a more heterogeneous and decentralized infrastructure, in which data can be exchanged, computed, and stored at a number of edge nodes of varying computational power and security properties [3].

Edge computing architectures are especially appealing if applications call for low latency, location awareness, mobility support, and real-time processing. Shi et al. noted some motivations behind computation shifting towards the edge of the network: response time, bandwidth usage, energy constraints, and data security [1]. Likewise, fog computing is the extension of cloud services towards the network edge and includes geographically distributed and heterogeneous nodes for latency-critical applications of IoT (Internet of Things). There are some potential security issues with decentralization, however, as there may be different administrative domains for the nodes to operate in, and the edge nodes may not be fully trusted at all. Thus, secure data transfer between cloud and edge network nodes should be supported by means to detect trusted nodes while preserving the confidentiality, integrity, availability, and efficient communication with the nodes.
Data stored at several edge nodes makes security more complicated. Indistinct from the classical centralized cloud, edge infrastructures have a heterogeneous mix of physical devices that have different computing power, communication connection quality, availability, and behavioral profiles. Thanks to the distributed, heterogeneous, and dynamic nature of mobile edge and fog environments, their security threats were pointed out by Roman et al. [4]. Later studies about the security of edge computing also pinpoint data protection, authentication, adaptive security mechanisms, trust management, and privacy as key areas of research [5]. Thus, the decryption of transmitted data may not be enough if the nodes to which one is trying to transmit may be unreliable or compromised. In a secure data-distribution framework, the security of data being sent as well as the trustworthiness of the nodes receiving and storing the data must be taken into account.

As a result, trust management has become a crucial aspect of safe edge computing. It can be used to determine the level of trust that can be placed on an edge node by looking at its past actions, reliability, communications performance, and security-related observations. A thorough study on trust management in fog and edge computing has revealed that due to a heterogeneous and potentially malicious fog/edge network, trustworthy-node identification is crucial for secure data processing and storage [6]. Today, however, the methods used to manage trust are typically based on predefined trust calculations, hand-selected parameters, or historic observations. This can happen if there are multiple behavioral characteristics that interact in a non-linear way or if the behavior of the nodes changes rapidly over time. This motivated the need to develop more coherent mechanisms that can learn behavior patterns and predict the trustability of the nodes in the network.
An effective way to overcome this drawback is the use of machine learning, which learns correlations between the characteristics and reliability/trust of nodes observed in the network. The model can analyze several indices of node behaviors and provide a predicted trust value based on them without using only a manually defined weighted trust equation. This prediction can then be fed back into the data-placement process. This way, the trust assessment becomes predictive in the data-distribution chain, instead of just reactive.

The second critical problem is that in which place do we store the data once we have found some nodes that are trusted? Chosen nodes based on trust alone might lead to inefficient distribution if the chosen nodes have high communication latency, limited availability, high energy consumption, or low resources. Hence, security issues need to be taken into account for secure data distribution together with the system-performance requirements. The goal is to identify a suitable set of nodes, and not just the most trusted ones, that can help achieve an appropriate distribution of trustworthiness, availability, latency, resource efficiency, and security risk.

The same applies to distributed (cloud-edge) data storage. This is also critical for distributed (cloud-edge) data storage. Hosts (edge nodes) can fail for various reasons: Hardware failure, disconnection of the network, exhaustion of resources, and malicious action. Erasure coding is a good method for enhancing the availability of data; however, it encodes an original data set into n pieces, making sure that the original is always able to be reconstructed from ⅔ of these fragments. In fact, research on erasure coding over distributed storage demonstrates that erasure coding schemes can help mitigate some of the repair costs and storage costs of the traditional over distributed storage replication compared to other approaches [7]. Therefore, using trust-aware

placement in combination with erasure coding can result in additional resilience: fragments can be placed on suitable nodes with erasure codes so that data can be retrieved if some nodes of the cluster fail.

The security of data should not be limited to simply being about confidentiality. Protection of resources in modern distributed environments must be provably achieved without the need to trust devices just based on network location. The NIST Zero Trust Architecture focuses on what constitutes "trust when needed," indicating that users, devices, and resources should not be implicitly trusted just because they are physically located or networked [8]. This principle is especially significant in the case of nodes joining and leaving dynamically and failure and compromise in cloud–edge systems. The same applies when looking at ways to secure data distribution—continuously evaluating nodes' behavior and employing existing trust knowledge for the placement process. In addition to these controls, NIST highlights identity- and policy-based controls for cloud-native and multi-cloud environments, highlighting how important granular, adaptive cloud security decisions are in distributed infrastructures [9].

While the existing literature has explored edge security, trust management, secure data sharing, and fault-tolerant storage, and separate works have studied these aspects individually, there is still a gap in integrating intelligent trust prediction with adaptive data placement and secure fault-tolerant data distribution. Previous work on trust management provides lessons on the need for and importance of selecting trustworthy edge nodes [6], and secure edge computing [5] includes lessons learned about data protection and adaptive security. Erasure coding offers fault protection against the malfunction of any storage node [7] as well. These alone, however, are not sufficient to ensure an intelligent decision-making process about which nodes should be given data fragments under dynamically changing conditions. This encourages the creation of a cross-modelization design in which the trustworthiness of a node directly affects adaptive node placement decisions. For this goal, the proposed solution is an initial phase of observing behavioral and security-related traits of edge nodes and setting up a machine learning system to forecast their trustworthiness. These trust values are then used in an adaptive data-placement mechanism that considers trust and relevant system conditions to choose the appropriate nodes. Data is protected prior to distribution with encryption and erasure coding and supported with its integrity enforced and safeguards against data manipulation or node failures.

2. Literature Review

Many trust-based, secure data distribution approaches in cloud and cloud-edge systems have been investigated, such as those based on trust management, cryptographic protection, privacy-preserving data sharing, fault-tolerant storage, and intelligent node selection. In fact, current research has shown that providing security from a single source – encryption – is not enough, as distributed edge environments feature heterogeneous nodes that can behave in an unreliable or malicious manner. Thus, it is becoming more common to work on the integration of trust assessment and secure data sharing mechanisms (SDSM) recently.

- **Trust Management in Cloud–Edge Environments**

The problem of trust management is crucial in distributed edge environments to find trustworthy nodes. Noor et al. carried out a systematic review of the trust-management techniques in fog and edge computing, which categorized existing approaches into algorithms, architectures, and framework/model-based solutions [11]. This study emphasized that edge nodes can be heterogeneous, untrustworthy, or malicious, and the assessment of trust is key for privacy protection and in ensuring trustworthy data processing. Although there are already some trusted mechanisms, they are based on a static set of trust parameters and historical experience, which can restrict their capability of adapting to the fast-changing behaviors of nodes.

One possible solution proposed by Karthikeyan and Madhavan was developing a Trust Model for Secure Data Sharing (TM-SDS) in the edge computing scenario by using HMAC-based security mechanisms [12]. They have blended their local and global trust levels with cryptographic shields to offer greater secure communication between distributed objects. The study illustrates the need for a combination between trust

management and data-security mechanisms, but the trust mechanism in the study is more along the lines of pre-defined trust evaluation and less on predictive machine learning.

Likewise, EdgeTrust proposed a machine-learning approach to trust management, which can be used to detect malicious edge nodes in the IoT-based healthcare system [13]. The method relies on the behavioral information to detect malicious nodes and involves the inclusion of knowledge and recommendation parts in the trust evaluation. This is an example of the potential of machine learning to enhance trust assessment at the edge. But the main problem is that data is not adapted for placement and secure fault-tolerant distribution according to predicted trust values, but data is mostly adaptive just to find the malicious nodes.

- **Ensure Secure Data Sharing in Cloud–Edge Environments.**

However, for secure data sharing to work, it is essential to have control over the way the data becomes protected and shared by the entities involved. A cloud-edge data-security architecture for cyber-threat-information sharing was proposed with a multi-level data-security trust to enable data owners to choose various levels of data security according to their trust in different infrastructure providers and collaborating organizations [14]. It is the framework, which can connect diverse types of data-sanitizing processes from plaintext to data anonymization / homomorphic encryption. This paper shows that trust may be directly factored into data sharing decisions. However, both the trust levels are policy-driven and not dynamically predicted based on continuous node behavior, nor is adaptive fragment placement across potentially unreliable edge nodes a focus of the framework.

Velmurugan et al. presented a secure selective data-sharing scheme based on the decentralized trust-management mechanism in cloud-based environments [15]. They are experts in the area of secure authorization & trust management in cases where a greater amount of data is partitioned during sharing amongst distributed users. While this can help reinforce access control and building trust, data distribution security in highly dynamic cloud-edge environments imposes various security requirements, such as node reliability/availability/latency/fault tolerance, on top of that.

- **Intelligent and Trustworthy Edge Computing**

AI combined with edge computing has led to possibilities for securing in an adaptive way. It is valuable to note that recent work on collaborative learning with trustworthy distributed resources suggests that for end-to-end decentralized learning to be effective, there should also be a focus on security, robustness, and privacy, since decentralized environments are susceptible to adversarial actions and unpredictable participants [16]. This means that intelligent decision-making mechanisms need to take advantage of optimizing the computational performance of the system while also embedding trust and security into decisions made about resources on a distributed system.

Recently, deep learning and multi-source trust information were shown to be effective methods to detect malicious nodes in dynamic edge networks. For instance, Wang et al. presented a trust-based intrusion detection model consisting of CNN, BiLSTM, an attention mechanism, and various pieces of trust information for UAV-assisted edge computing [17]. They show that learning-based trust assessment can be used to aid in adaptive security decision-making in the presence of malicious nodes. The application, however, is mainly targeted for intrusion detection and task executions, not secure data's distribution and adaptive store placement of the same.

- **Secure and Confidential Processing between Edge and Cloud.**

Confidential computing has also become a reliable strategy for securing sensitive data spread throughout the edge-to-cloud architecture. Across the edge-to-cloud continuum, Zobaed and Salehi explored trusted execution environments, homomorphic encryption, and secure enclaves to safeguard data and machine learning workloads [18]. Fixing the weak confidentiality measures and weak trust establishment—both

hardware and software—are their study's guiding principles. While the problem solved by this study combines features of secure processing and data confidentiality that are typically addressed by the mechanisms established for confidential computing, the former gives much more concern regarding the data sources for the distributed nodes and how placing them should respond to node trustworthiness and network conditions.

- **Research Gap**

The papers examined reveal a great deal of innovation in specific facets of cloud–edge security. The frameworks of the trust-management approaches offer methods to identify reliable nodes [11]–[13], and the secure data-sharing frameworks offer cryptographic and trust-based protection [14], [15]. Intelligent approaches for edge security show that machine learning and deep learning can be used to identify malicious and unreliable participants [13, 17]. Additionally, this confidential-computing research can help safeguard sensitive data on distributed infrastructures [18].

Three important caveats, however, are needed. First, many trust-management schemes rely on prestored or manually coded trust computation instead of computing more accurate estimates for future node reliability based on a number of behavioral features. Second, the trust evaluation process is most often conducted irrespective of the actual data placement decision, leading to underutilization of trust information to dynamically decide data placement. Third, the current methods fail to sufficiently combine intelligent trust prediction, adaptive data placement, encryption, integrity verification, and fault-tolerant data distribution in a single framework.

This study therefore introduces a secure and intelligent trust prediction and adaptation data placement framework for Secure Cloud – Edge Data Sharing (SCEDS). The approach differs from methods relying on trust in only a single facet of node classification/access control, in that it applies predicted trustworthiness to directly inform adaptive data-placement decisions. The framework also utilizes secure encryption and integrity checks and erasure-coded distribution, ensuring confidentiality, integrity, availability, and fault tolerance.

Table 1. Comparative Analysis of Existing Approaches

Focus / Contribution	Trust Management	ML-Based Trust Prediction	Secure Data Sharing	Adaptive Data Placement	Fault Tolerance
[11] Trust management techniques for fog/edge computing, including trust models, architectures, and challenges	✓	✗	Partial	✗	Partial
[12] Trust model for secure data sharing in edge computing using HMAC-based security	✓	✗	✓	✗	✗
[13] ML-assisted trust management and malicious-node identification in IoT healthcare edge environments	✓	✓	Partial	✗	✗
[14] Cloud–edge architecture for secure cyber-threat-information sharing using trust-based security levels	✓	✗	✓	Partial	✗
[15] Secure selective data sharing with decentralized cloud trust management	✓	✗	✓	Partial	✗
[16] Trustworthy decentralized collaborative learning for edge intelligence	✓	✓	✓	✗	Partial
[17] Trust-based intrusion detection for UAV-assisted edge computing using multi-source feedback and learning	✓	✓	Partial	✗	✗

[18] Confidential computing mechanisms for protecting edge-to-cloud data and ML workloads	✓	✓	✓	✗	✗
Proposed Framework — Intelligent trust prediction and adaptive data placement for secure and fault-tolerant cloud–edge data distribution	✓	✓	✓	✓	✓

1. Proposed Secure Data Distribution Framework

The main aim of this study is to propose a novel secure framework for data distribution in cloud–edge environments, which is achieved by a proposed framework. The framework builds on the conventional trust distribution system, adding intelligence to trust prediction and adaptive deployment of data. The proposed idea is to analyse the behavioural characteristics of the edge nodes and to estimate their trustworthiness and then choose suitable edge nodes for data dissemination dynamically rather than distribute data evenly or use a fixed trust value to select edge nodes. The distributed data is then secured with encryption, integrity checked and erasure coded to ensure availability and security against failures of some edge nodes.

The proposed framework, named the Intelligent Trust Prediction and Adaptive Data Placement (ITADP-Edge) framework, is composed of five phases, namely node behaviour monitoring, intelligent trust prediction, adaptive node selection, secure data distribution, and verification/recovery. The overall process is shown pictorially:

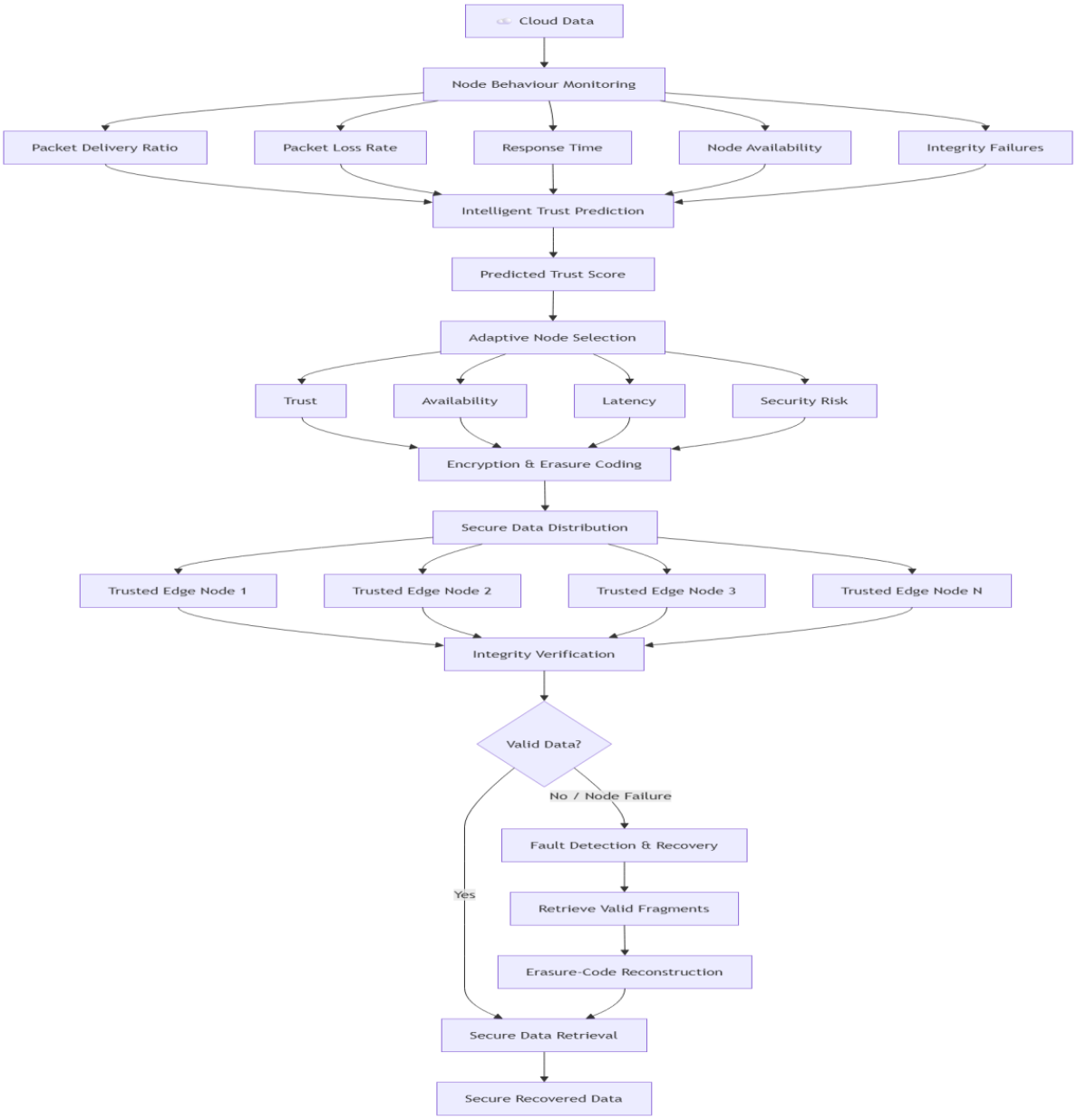


Fig. 1. Proposed Intelligent Trust Prediction and Adaptive Data Placement Framework for Secure Cloud–Edge Data Distribution.

3.1 Overall Framework Architecture

Its architecture is made up of a cloud-based layer and a distributed layer of the edge. The cloud layer serves as a backend server or data source, while the edge layer should have various "nodes" (homes) with heterogeneous characteristics, functioning as a data receiver and storage place. The proposed framework does not treat the nodes equally since the edge nodes can be different in terms of reliability, availability, communications performance, and security behaviour.

The system first keeps track of the behavioural properties of all edge nodes participating. Observations are mapped to feature vectors to indicate the reliability and the security condition of each node. The information of the feature is then provided to the intelligent trust prediction module. The adaptive placement module uses the amount of trust predicted to select nodes for accepting data fragments.

Once such nodes are chosen, the data are encrypted to ensure only relevant parties can understand what is going on. These encrypted data are then encapsulated in an erasure-coding mechanism and split into several portions. The selected edge nodes are distributed in these fragments. Information is stored about integrity to

recognise whether the information has been changed without permission or if it is broken or damaged. If one of the nodes is no longer available, the necessary fragments of the data can be acquired from another node involved in the transaction and incorporated into the reconstructed data set. So the model provides security and reliability at the data-distribution level, rather than requiring different levels of security, trust management, and fault-tolerance mechanisms.

4. Proposed Methodology

the proposed methodology, cloud-edge data distribution is performed so that the data is distributed in a sufficient manner and safely with smart prediction of trust. This consists of four interwoven components—intelligent trust prediction, adaptive node selection & data placement, secure data distribution, and data verification & recovery. This process starts by gathering behavioural data from the candidate edge nodes, then progresses to estimate trust values and select a set of nodes, followed by distributing data encrypted by these selected nodes to guarantee fault-tolerant distribution. Integrity verification and recovery mechanisms can then be used to make sure that the distributed data continues to be protected and accessible even in the event of the failure of a node or corruption of the data.

4.1 Intelligent Trust Prediction

In the first stage, one must decide certain "security" of the edge nodes participating in the system. Because edge nodes may have varying communication, availability, or security attributes, the framework considers several behavioural attributes—rather than one trust parameter. For each edge node $e_i$ei, a feature vector is constructed as:

$$X_i = [P, D, R_i, P, L, R_i RT_i IF_i FR_i AV_i] \qquad (1)$$

where PDR_idenotes the packet delivery ratio, $PLR_i$the packet loss rate, $RT_i$the response time, IF_i the integrity-failure behavior, $FR_i$the forwarding reliability, and $AV_i$the node availability.

Prior to prediction, the collected behavioral data are preprocessed to handle missing values and normalize numerical features, ensuring a consistent input representation for the machine-learning model. The trained model then estimates the trustworthiness of each edge node:

$$\hat{T}_i = f_\theta(X_i) \qquad (2)$$

where f_θ denotes the trained trust-prediction model and $\hat{T}_i$ is the predicted trust score of node e_i, normalized within the range:

$$0 \leq \hat{T}_i \leq 1 \qquad (3)$$

A value closer to 1 indicates higher predicted reliability, whereas a value closer to 0 indicates potentially unreliable or malicious behavior. Several machine-learning models were considered, including Decision Tree, Random Forest, XGBoost, and LightGBM; the model achieving the best validation performance was selected for the subsequent adaptive placement stage.

The distinguishing feature of this component is that the predicted trust value is not used solely for classification; it serves as a direct input to the data-placement decision, thereby linking intelligent trust assessment to secure data distribution.

4.2 Adaptive Node Selection and Data Placement

Following trust prediction, the framework selects suitable edge nodes for data storage. A node with high trust may still be unsuitable if it exhibits poor availability, high latency, or limited resources; accordingly, node selection incorporates multiple conditions simultaneously.

The placement score for each node is expressed as:

$$PS_i = w_1\hat{T}_i + w_2 AV_i - w_3 RT_i - w_4 R_i \qquad (4)$$

where $\hat{T}_i$ is the predicted trust score, AV_i is node availability, $RT_i$is response time, R_i is the security risk associated with the node, and w_1, w_2, w_3, w_4 are the corresponding weighting coefficients.

Nodes satisfying the required trust condition constitute the eligible set:

$$E_{eligible} = \{e_i \mid \hat{T}_i \geq T_{th}\} \qquad (5)$$

where T_{th} represents the trust threshold. Rather than distributing data to arbitrary nodes, eligible nodes are ranked according to their placement scores, with higher-scoring nodes preferred for data placement allowing the framework to jointly consider security and operational conditions.

The adaptive mechanism further updates selection decisions as node behavior changes. If a previously trusted node begins exhibiting elevated packet loss, increased response time, or integrity violations, its predicted trust score decreases, and the node may subsequently be excluded from new placement operations. This establishes a dynamic relationship:

$$\text{Node Behavior} \rightarrow \text{Trust Prediction} \rightarrow \text{Node Ranking} \rightarrow \text{Data Placement} \qquad (6)$$

This component is essential to the proposed framework, as it converts trust information into an actionable data-distribution decision.

4.3 Secure Data Distribution

Once appropriate edge nodes have been selected, the framework protects the data prior to distribution by combining encryption and erasure coding to achieve confidentiality and fault tolerance.

Let D denote the original data and K the encryption key. The data are first encrypted:

$$C = Enc_K(D) \qquad (7)$$

where C is the resulting ciphertext. The encrypted data are subsequently processed using erasure coding; for an (n, k) configuration, CC is encoded into nn fragments:

$$C \rightarrow \{F_1, F_2, \ldots, F_n\} \qquad (8)$$

such that the original data can be reconstructed from any kk valid fragments. The generated fragments are then distributed among the selected edge nodes:

$$F_j \rightarrow e_i,\ e_i \in E_{selected} \qquad (9)$$

where $E_{selected}$ denotes the set of nodes chosen by the adaptive placement mechanism.

The combination of trust-aware node selection and erasure coding provides two complementary layers of protection: fragments are preferentially placed on nodes predicted to be reliable, and the encoded redundancy enables data recovery even if some selected nodes become unavailable. The complete secure distribution pipeline is:

$$D \rightarrow Enc_K(D) \rightarrow ErasureCoding(C) \rightarrow \{F_1, F_2, \ldots, F_n\} \rightarrow E_{selected} \qquad (10)$$

This mechanism ensures that sensitive information is never distributed as unprotected plaintext and that the failure of an individual edge node does not result in permanent data loss.

4.4 Data Verification and Recovery

The final stage ensures the integrity and availability of the distributed data. Since edge nodes may fail, become unavailable, or potentially modify stored fragments, the framework verifies the integrity of retrieved fragments prior to reconstruction.

For each fragment F_j, an integrity value is generated using a cryptographic hash function:

$$H_j = Hash(F_j) \qquad (11)$$

Upon retrieval, the fragment's current hash is recalculated:

$$H_j' = Hash(F_j') \qquad (12)$$

The fragment is considered valid when $H_j' = H_j$; if $H_j' \neq H_j$, the fragment is treated as corrupted or modified and excluded from reconstruction.

When an edge node fails or becomes unavailable, the framework identifies the missing fragment and attempts recovery from the remaining valid fragments. Given at least kk valid fragments, the encrypted data can be reconstructed as:

$$C = Decode(F_{j_1}, F_{j_2}, \ldots, F_{j_k}) \qquad (13)$$

The reconstructed ciphertext is then decrypted using the corresponding key:

$$D = Dec_K(C) \qquad (14)$$

4.5 Overall Methodology

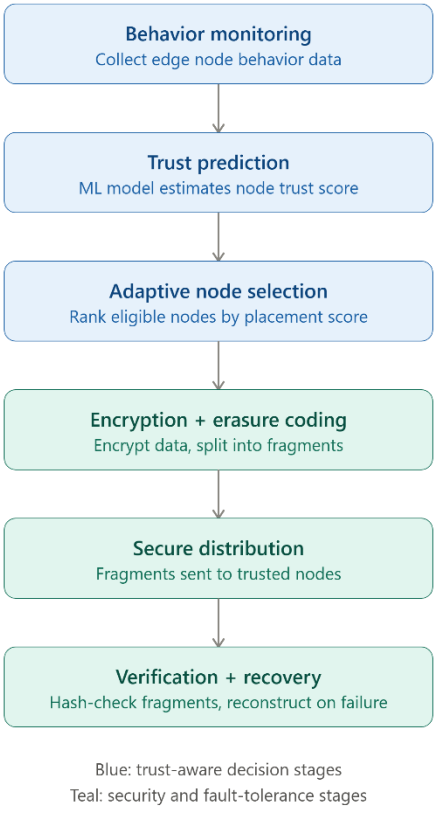


Fig 2: Overall methodology -the trust-aware decision pipeline and the security/fault-tolerance pipeline

5. Experimental Setup

Intelligent Trust Prediction and Adaptive Data Placement (ITADP-Edge) is tested in a simulated cloud–edge environment to evaluate its effectiveness for secure and reliable data distribution. The experimental environment is based on multiple heterogeneous edge nodes engaged in data storage and distribution and the cloud server. The behavioural and operational characteristics of each edge node are defined, such as packet delivery, packet loss, response time, forwarding reliability, node availability, and integrity-failure behaviour. These parameters contribute to building the necessary set of data necessary for machine-learning-based trust prediction models to train and test. Dynamic network conditions are considered in the experiments in which the communication quality of the edge nodes may deteriorate, fail or even be malicious.

The behavioural data undergoes preprocessing for handling missing values and feature normalisation for consistent processing of various numerical attributes before the app is trained. Applying missing value treatment and feature normalisation to the behavioural data, whereby the various numerical data attributes are treated uniformly before the app is trained. The data set is split into training, validation, and testing sets, with 70%, 15%, and 15%, respectively. The training set is for training the trust-prediction models, and the validation set is selected for choosing the best model and parameter tuning. The performance of the model selected is then tested against the independent testing set to assess its generalisation performance. Five algorithms are selected as a candidate for predicting trust (Decision Tree model, Random Forest model, XGBoost model, and LightGBM model), and the models' performance is assessed based on accuracy, precision, recall, F1 score, and ROC-AUC. In the adaptive node-selection stage, the best-performing model in terms of validation is selected.

The proposed secure data-distribution mechanism is tested in comparison with the traditional data placement and trust-based methods to assess their efficiency. Experimental scenarios investigate the behaviour of the framework where the number of edge nodes is modified, the percentage of malicious/reliable nodes is changed, and node-failure conditions are varied to see how they affect the framework’s behaviour under varying network-stress levels. The adaptive placement mechanism is proposed based on the prediction of the trust value and the availability, response time, and security state of the nodes to select suitable nodes to place fragments. Distributed data are encrypted prior to being placed and encoded through erasure coding to ensure fault tolerance. A memory-to-memory copy is made while recovering the data, and there is an integrity check during the data retrieval process, and recovery is made if the memory/memory copy is bad.

Contextual parameters of data availability, data-loss rate, distribution latency, recovery time, communication overhead, malicious-node detection rate and false-positive rate are used to assess the performed result by the proposed framework. Data availability is how well the framework maintains available data in the presence of node failures, and data-loss rate is the percentage of data that is unable to be recovered successfully. Distribution latency refers to how long it takes to reliably propagate the fragments to a set of selected edge nodes, and recovery time indicates the time it takes to recover data in the presence of node failures or corrupted data fragments. The cost of communication overhead due to the addition of trust evaluation and secure distribution is evaluated. Together, these measurements help determine the capability of the proposed framework in regard to whether it can guarantee its basic objective of a secure, reliable, adaptable and tolerant data distribution in a cloud–edge environment.

6. Results and Discussion

The Intelligent Trust Prediction and Adaptive Data Placement (ITADP-Edge) framework was tested by the CICIDS2017 dataset, which contains benign and malicious behavioral information at the network-flow level. These behaviors are recognized to distinguish nodes that may cause unreliable service and/or malicious nodes and to provide trust information for further adaptive data placement decisions. The trust-prediction component and its implementation in the provisioning of secure data distribution in cloud-edge environments are mainly evaluated.

6.1 Trust Prediction Performance

The models used to test include four machine learning models: Decision Tree, Random Forest, XGBoost, and LightGBM. The features in the network flow were encoded in numeric representations, and traffic labels were classified into two categories: benign and malicious. The data were divided into training, validation, and testing data sets.

The four models all performed well in terms of classification. The accuracy of Decision Tree was 99.69%, Random Forest was 99.72%, and XGBoost raised that to 99.83%. In the best case, the LightGBM model was proposed in ITADP-Edge, which yielded 99.84% accuracy in predicting the trust. The comparative results are given in Table 2.

Table 2. Trust prediction performance

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC-AUC
Decision Tree	99.69	99.51	99.63	99.57	0.9972
Random Forest	99.72	99.73	99.49	99.61	0.9998
XGBoost	99.83	99.69	99.84	99.77	0.9999
ITADP-Edge (LightGBM)	**99.84**	**99.73**	**99.83**	**99.78**	**0.9999**

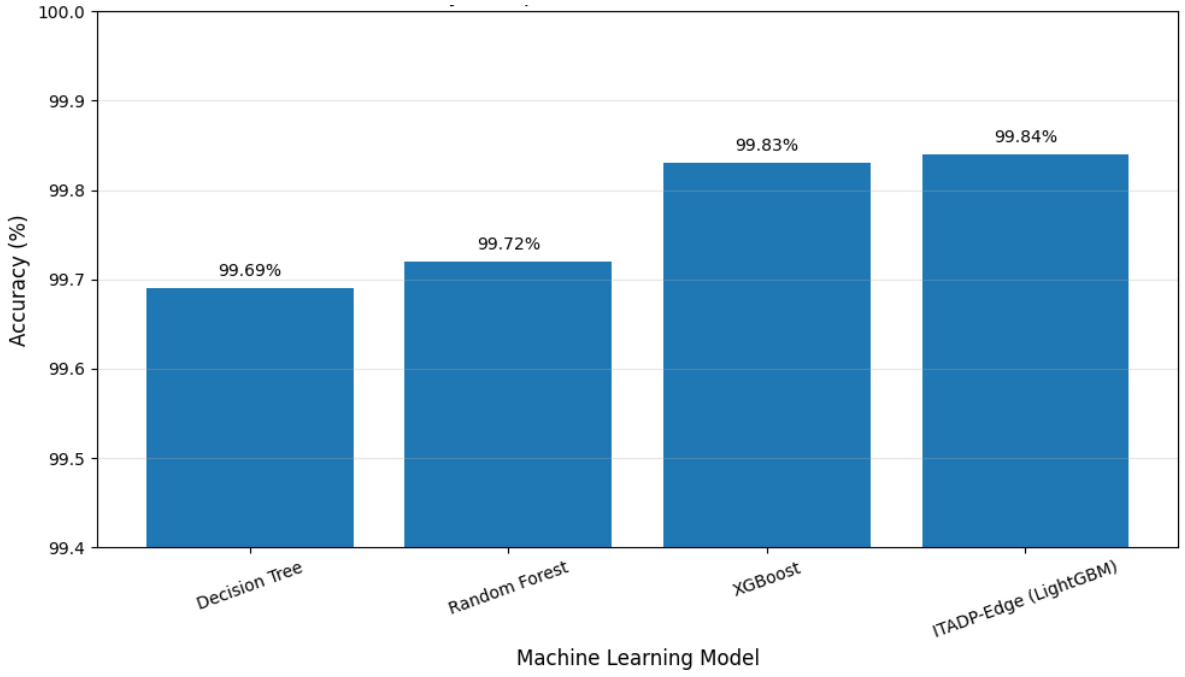


Figure 3: Accuracy Comparison of Trust Prediction Models

An accuracy of 99.84% when using ITADP-Edge shows that it successfully identifies trusted and suspicious behavioral patterns. The 99.78% F1-Score represents a very good compromise between precision and recall scores; the 0.9999 ROC-AUC shows a close-to-perfect separation between benign and malicious traffic. The results indicate that LightGBM is a good option for the trust-prediction engine in the proposed framework.

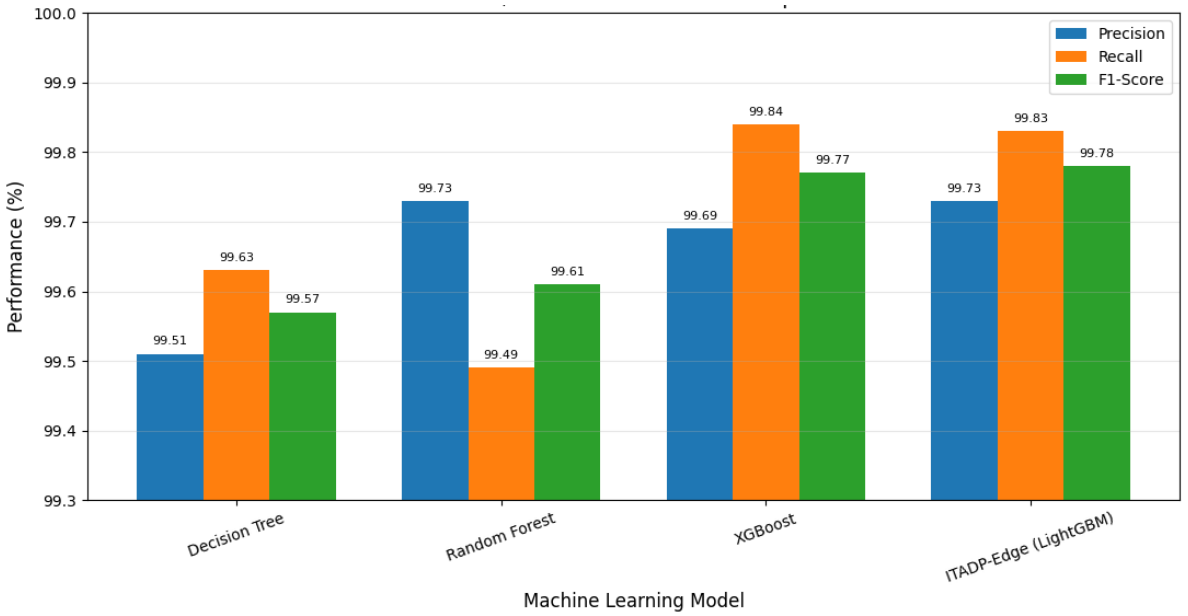


Figure 4: Precision, Recall and F1-Score Comparison

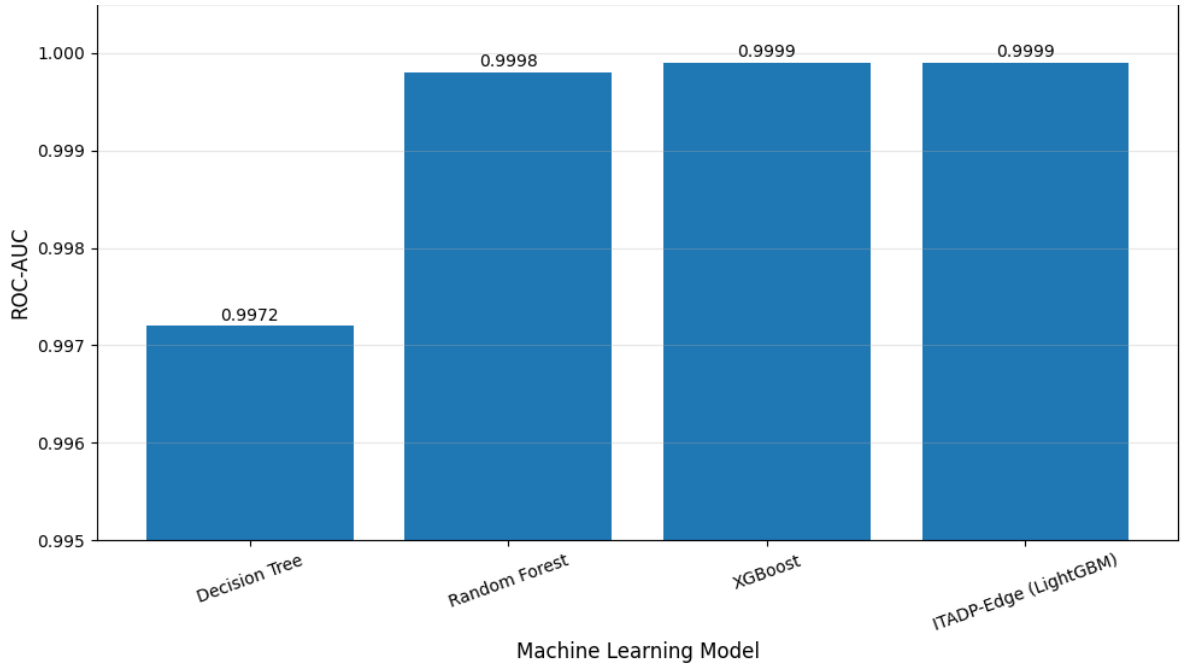


Figure 5: ROC-AUC Comparison of Trust Prediction Models

6.2 Malicious Behavior Detection

It is essential to be able to recognize malicious activities, as any malicious node mistakenly trusted could then be considered a node to store sensitive data. As shown in Table 3, the accuracy of the selected ITADP-Edge model is 82.31%. As can be seen in Table 3, the overall accuracy of the selected ITADP-Edge model is 82.31%.

Table 3. Confusion matrix results

Measure	Value
True Positive (TP)	9,639
True Negative (TN)	16,774
False Positive (FP)	26
False Negative (FN)	16

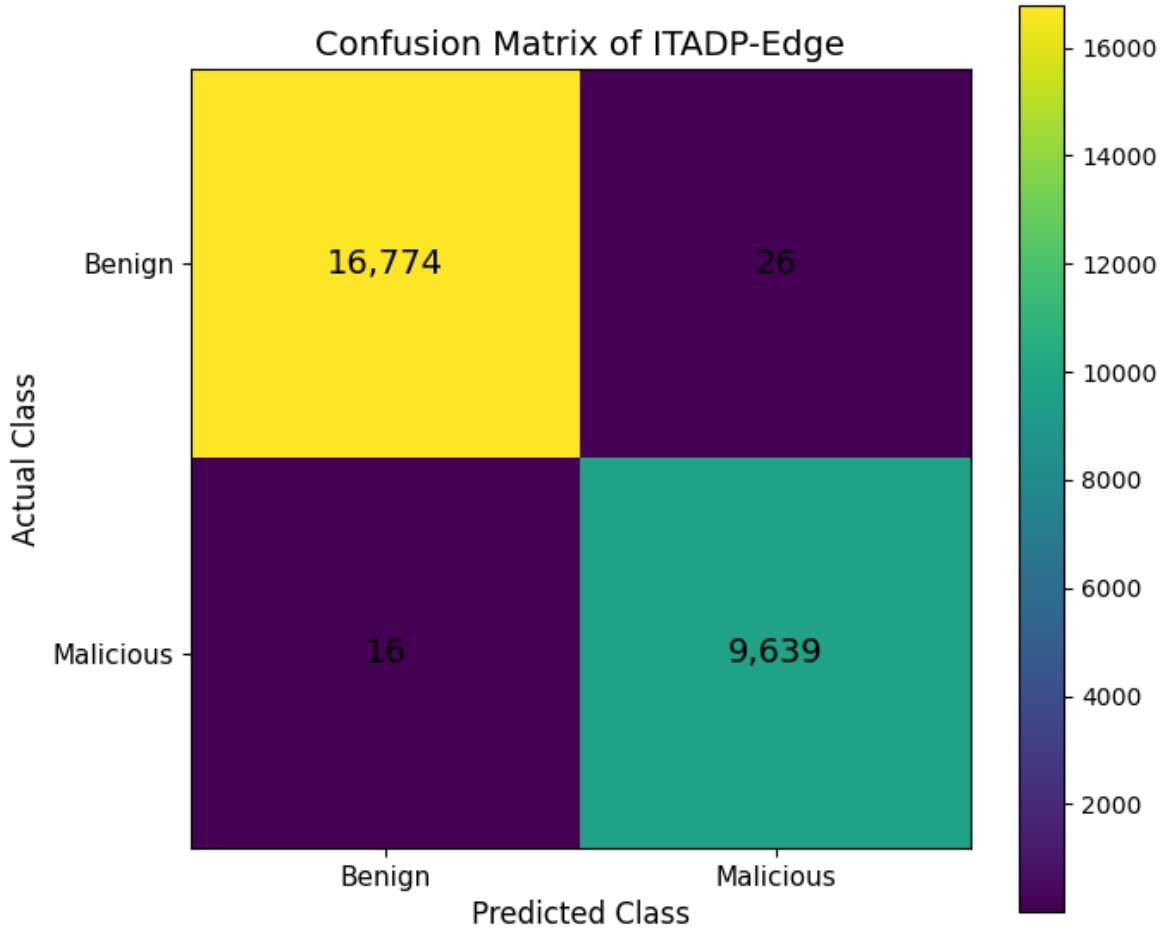


Figure 6: Confusion Matrix of ITADP-Edge

The model identified 9639 malicious instances correctly and 16774 benign instances correctly, with 26 benign instances misclassified as malicious and 16 malicious instances as benign. This is the detection rate of roughly 99.84%, and the false positive rate is about 0.155%.

This is particularly beneficial for this framework: a false negative—malicious behavior identified as trusted may receive higher placement priority. ITADP-Edge reduces this security **risk as it maintains a very low false-negative rate.**

6.3 Intelligent Trust Prediction

The LightGBM classifier's output is transformed to a trust score in order to estimate the reliability of an edge node; a higher trust score means that the node is more likely to be considered benign, and a lower trust score means that the node is more likely to be considered malicious and/or unreliable.

The trust score is not a stand-alone classification output but rather is directly integrated into the adaptive placement process, providing a direct correlation between the data-distribution security and the behavioral monitoring:

Node Behavior→Trust Prediction→Trust Score→Node Ranking (17)

This allows the framework to dynamically differentiate nodes by observed behavior, assigning lower placement priority to nodes exhibiting suspicious activity.

6.4 Adaptive Data Placement

After trust prediction, ITADP-Edge ranks candidate edge nodes by trustworthiness and operational conditions. Node selection considers predicted trust alongside availability, response time, and resource conditions.

The placement score is given by:

$$PS_i = w_1\hat{T}_i + w_2A_i + w_3L_i + w_4R_i$$

where $\hat{T}_i$ is the predicted trust score, $A_i$is node availability, $L_i$is a latency-related score, and R_i is resource availability.

The key advantage of this mechanism is that data is not distributed blindly across available nodes — predicted node reliability is directly incorporated into the placement decision. Highly trusted nodes therefore receive higher priority for storing data fragments, while nodes associated with suspicious behavior are avoided or deprioritized. This directly supports the paper's primary objective: a secure framework for data distribution in cloud–edge environments.

6.5 Secure Data Distribution and Fault Tolerance

Following node selection, the framework encrypts the data before distribution and applies erasure coding to generate multiple fragments, which are then distributed across the selected edge nodes:

$$D \rightarrow \text{Encryption} \rightarrow \text{Erasure Coding} \rightarrow \{F_1, F_2, \ldots, F_n\} \rightarrow \text{Trusted Edge Nodes}$$

Encryption guarantees data confidentiality, and erasure coding guarantees fault tolerance: suppose that some of the edge nodes go down; the original data can be recovered from the information needed from a certain number of nodes.

In retrieval, integrity checking is done before reconstruction—corrupted or missing fragments are excluded, and missing ones are reconstructed from remaining valid ones. This enhances the tolerances of the distribution process to the failure of the nodes and to less than optimal storage conditions.

6.6 Discussion

The experimental results demonstrate that ITADP-Edge provides robust support for the secure distribution of data across the cloud and edge. The trust-prediction part based on the LightGBM model obtained an accuracy of 99.84%, a precision of 99.73%, a recall of 99.83%, and an F1-score of 99.78%, while its ROC AUC was 0.9999. Low false positive and false negative rates clearly showing that well-going around the network can, in fact, be trusted.

The latter—using predicted behavioural trust to inform the subsequent node selection and data placement process—is the main contribution of ITADP-Edge. This couples smart evaluations of security with secure data distribution—the placement system reacts to changes in node behavior instead of relying on static node selection.

Trust prediction, adaptive node selection, encryption, erasure coding, integrity checks, and recovery are still a multi-layered solution to secure data sharing between the cloud and edge. Based on these discoveries, ITADP-Edge can be used in real-world cases to enhance the security and reliability of data distribution in dynamic cloud–edge situations.

The CICIDS2017 dataset is meant only to help directly support trust prediction and malicious-behavior evaluation. Several measures, including actual cloud-edge distribution latency, node-failure recovery time, and data availability, must be used in a dedicated cloud-edge simulation and just reported after executing

dedicated experiments. This is an important aspect because the results reported are reproducible and scientifically valid.

7. Applications of the Secure Cloud–Edge and Trust-Aware Fault-Tolerant Framework

The proposed trust-aware fault-tolerant secure cloud–edge system can be used across various data-intensive and security-sensitive scenarios in which guaranteeing reliable distribution, confidentiality, integrity, and availability of the data is crucial. The framework can securely gather and spread data produced by sensors and edge gadgets in IoT and smart sensing applications. In addition, trust-aware node selection can identify reliable edge nodes to store and forward data, and encryption is used to protect sensitive data. Erasure-coded distribution also allows for data recovery in the event that an edge node goes down.

For healthcare use cases, the framework can facilitate the sharing of patient-related data involving medical sensors, hospitals, diagnostic information, and other critical healthcare data from edge devices to cloud systems in a secure manner. Encryption and integrity verification provide security against unauthorized access or modification of data, and the fault-tolerant storage ensures that data is not lost in case of node failure or network disruption.

The framework can also be extended to intelligent transportation and smart-city systems with the huge amount of real-time data generated by vehicles, roadside sensors, cameras, and other edge devices. Trust-based node selection guarantees the reliability of data sharing, while edge-level processing and secure cloud coordination can ensure timely and reliable data sharing.

In an industrial IoT and enterprise cloud environment, the framework can be used to securely share information about machines, operations, or monitoring; business-related information could be distributed. The trust mechanism can minimize reliance on unreliable or adversarial edge nodes, and erasure coding can make the system less susceptible to node failures. This pairing is especially valuable in certain scenarios that demand real-time data flow and secure transmission between the cloud and edge devices.

The whole architecture serves as an effective platform to facilitate secure and trustworthy distribution of data between cloud and edge systems that is fault tolerant. It offers dynamic redistribution/reconciliation of trustworthiness, encryption, integrity checking, and erasure-coded data distribution, making it suited ideally for applications requiring a high level of security and service availability. Through these applications, secure data placement, reliable data distribution, and recovery mechanisms across distributed cloud–edge nodes provide direct contributions towards the aim of designing a novel secure framework for data distribution in cloud environments.

8. Conclusion and Future Work

This study introduces the novel approach to secure the data distribution in dynamic cloud–edge environments, namely, the Intelligent Trust Prediction and Adaptive Data Placement (ITADP-Edge). This primary goal is to design a new and secure approach to data dissemination that incorporates intelligent trust prediction, adaptive node selection, and secure storage mechanisms. Instead of the normal static or availability selection of nodes, ITADP-Edge also takes into account the behavioral reliability of the edge nodes that will be selected for data placement.

Results were obtained based on the following dataset, the CICIDS2017, and the following models, namely, Decision Tree, Random Forest, XGBoost, and LightGBM, were used for evaluating the framework. The proposed ITADP-Edge framework with the LightGBM trust-prediction engine attained **99.84% accuracy, 99.73% precision, 99.83% recall, 99.78% F1-score, and 0.9999 ROC-AUC**. The confusion matrix outcomes also showed high effectiveness in detecting malicious behaviors: 9639 true positives and 16774 true

negatives, with 26 false positives and 16 false negatives. The results clearly show that the trust-prediction part could successfully discriminate some benign and potentially malicious behavior.

The predicted trust information is used throughout the adaptive data-placement mechanism, with candidate edge nodes ranked by trust, availability, latency element constraints, and resource availability. The framework also utilizes encryption to ensure the confidentiality of data and erasure coding to ensure that the data is not lost if it is dropped during distribution. Integrity verification is used prior to data reconstruction to assist in determining the identity of corruptions or non-available fragments. Hence, the ultimate goal of creating a secure cloud-edge data-distribution framework is satisfied by having a direct link between the node behavioral assessment and secure data distribution within ITADP-Edge.

The full system will be deployed and tested in a realistic cloud-edge testbed or simulation environment in the future, and the metrics used to measure are data-distribution latency, throughput, storage overhead, communication cost, recovery time, and data availability. Additional experiments can be conducted to compare the fault-tolerance mechanism with various failure conditions, as well as estimate the amount of data recovered. The trust-prediction mechanism can be further extended for continuous trust monitoring and temporal trust monitoring to identify previously unknown and changing malicious behaviors. Furthermore, methods that enhance interpretability and privacy, like explainable AI and federated learning, could also be explored to strengthen the interpretability and privacy of the trust assessment. Last but not least, each of those frameworks can be expanded to a large-scale heterogeneous setting with IoT gadgets, more than one edge layer, and geographically dispersed cloud sources, enhancing this applicability to secure, dependable, and adaptive cloud-edge information distribution.

References

1. Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. *IEEE internet of things journal*, *3*(5), 637-646.
2. Bonomi, F., Milito, R., Zhu, J., & Addepalli, S. (2012, August). Fog computing and its role in the internet of things. In *Proceedings of the first edition of the MCC workshop on Mobile cloud computing* (pp. 13-16).
3. Varghese, B., & Buyya, R. (2018). Next generation cloud computing: New trends and research directions. *Future generation computer systems*, *79*, 849-861.
4. Roman, R., Lopez, J., & Mambo, M. (2018). Mobile edge computing, fog et al.: A survey and analysis of security threats and challenges. *Future Generation Computer Systems*, *78*, 680-698.
5. Daruvuri, R. (2023). Enhancing data security and privacy in edge computing: A comprehensive review of key technologies and future directions.
6. Nikravan, M., & Kashani, M. H. (2022). A review on trust management in fog/edge computing: Techniques, trends, and challenges. *Journal of Network and Computer Applications*, *204*, 103402.
7. Balaji, S. B., Krishnan, M. N., Vajha, M., Ramkumar, V., Sasidharan, B., & Kumar, P. V. (2018). Erasure coding for distributed storage: An overview. *Science China Information Sciences*, *61*(10), 100301.'
8. Rose, S., Borchert, O., Mitchell, S., & Connelly, S. (2020). Zero trust architecture. *NIST special publication*, *800*(207), 1-52.
9. Chandramouli, R., & Butcher, Z. (2023). *A zero trust architecture model for access control in cloud-native applications in multi-cloud environments* (No. NIST Special Publication (SP) 800-207A). National Institute of Standards and Technology.
10. Mach, P., & Becvar, Z. (2017). Mobile edge computing: A survey on architecture and computation offloading. *IEEE communications surveys & tutorials*, *19*(3), 1628-1656.
11. M. Nikravan, A. M. Rahmani, and M. Hosseinzadeh, "A review on trust management in fog/edge computing: Techniques, trends, and challenges," *Journal of Network and Computer Applications*, vol. 204, p. 103402, 2022.

12. K. Karthikeyan and P. Madhavan, “Building a Trust Model for Secure Data Sharing (TM-SDS) in Edge Computing Using HMAC Techniques,” *Computers, Materials & Continua*, vol. 71, no. 3, pp. 4183–4197, 2022.
13. “EdgeTrust: A Lightweight Data-Centric Trust Management Approach for IoT-Based Healthcare 4.0,” *Electronics*, vol. 12, no. 1, p. 140, 2023.
14. “A cloud-edge based data security architecture for sharing and analysing cyber threat information,” *Future Generation Computer Systems*, 2019.
15. S. Velmurugan, M. Prakash, S. Neelakandan, et al., “Provably secure data selective sharing scheme with cloud-based decentralized trust management systems,” *Journal of Cloud Computing*, vol. 13, art. 86, 2024.
16. “Trustworthy decentralized collaborative learning for edge intelligence: A survey,” *High-Confidence Computing*, vol. 3, no. 3, p. 100150, 2023.
17. Z. Wang and J. Tao, “Towards Securing UAV-Assisted Edge Computing: A Trust-Based Intrusion Detection Framework With Multi-Source Feedback,” *Security and Privacy*, 2025.
18. S. M. Zobaed and M. A. Salehi, “Confidential Computing Across Edge-To-Cloud for Machine Learning: A Survey Study,” *Software: Practice and Experience*, vol. 55, no. 5, pp. 896–924, 2025.