Taylor & Francis
Taylor & Francis Group

Applied Artificial Intelligence
An International Journal

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/uaai20

Privacy-Preserving with Zero Trust Computational Intelligent Hybrid Technique to English Education Model

Yitian Zhang

To cite this article: Yitian Zhang (2023) Privacy-Preserving with Zero Trust Computational Intelligent Hybrid Technique to English Education Model, Applied Artificial Intelligence, 37:1, 2219560, DOI: 10.1080/08839514.2023.2219560

To link to this article: https://doi.org/10.1080/08839514.2023.2219560

Published online: 04 Jun 2023.

APPLIED ARTIFICIAL INTELLIGENCE
2023, VOL. 37, NO. 1, e2219560 (1853 pages)
https://doi.org/10.1080/08839514.2023.2219560

Taylor & Francis
Taylor & Francis Group

Privacy-Preserving with Zero Trust Computational Intelligent Hybrid Technique to English Education Model

Yitian Zhang

College of Foreign Language, Zhengzhou Preschool Education College, Zhengzhou, China

ABSTRACT

The urgent and indispensable requirement for secure and efficient remote work compels organizations to reconsider their approach to safeguarding themselves against cyber threats. The shift toward remote work amplifies the need to redirect more network traffic toward cloud-based applications rather than relying solely on the internal network. The growing adoption of the hybrid work model necessitates system administrators to increasingly provide access to applications and services beyond the conventional boundaries of enterprise networks. Ensuring privacy goes beyond mere compliance with regulations; it is crucial for demonstrating transparency and accountability, which are essential in building trust with stakeholders. Employing a zero-trust approach can proactively enhance privacy by implementing access controls based on the principle of least privilege and predefined purposes. Such an approach helps to limit potential damages and enhances the resilience of complex information systems. This work proposes an innovative privacy-preserving and zero-trust computational intelligent hybrid system. Building upon the zero-trust architecture, this system ensures a protected environment while preserving privacy. It achieves this by employing multi-level trust fields within a corporate network, where every access request undergoes comprehensive authentication, authorization, and encryption before being granted access. The system's efficacy is validated within a sports training application environment, with stringent authorization requirements and the corresponding need to safeguard personal privacy. By implementing the proposed system, the application environment can effectively mitigate privacy risks while providing secure access only to authorized individuals. The hybrid system's computational intelligence further enhances its ability to adapt to evolving threats and maintain the confidentiality and integrity of sensitive data. In summary, the current landscape of remote work necessitates organizations to prioritize cybersecurity and privacy. By embracing a zero-trust approach and implementing the privacy-preserving and zero-trust computational intelligent hybrid system, organizations can ensure robust protection, maintain privacy compliance, and establish a trusted foundation for remote work environments.

ARTICLE HISTORY
Received 8 May 2023
Revised 24 May 2023
Accepted 25 May 2023

CONTACT Yitian Zhang zhangyitian@zzpec.edu.cn Zhengzhou Preschool Education College, Zhengzhou 450000, China

Introduction

The relentless advancement and constant evolution of highly sophisticated cybercrime techniques and the growing intricacy of modern IT systems have given rise to a novel security architecture model known as "zero trust." Zero trust represents a comprehensive security framework encompassing meticulously crafted system design principles and a coordinated strategy. This model acknowledges that threats can originate externally and internally to the traditional network perimeters.

Zero trust arises from the imperative need to establish an approach that does not solely rely on conventional security measures such as firewalls and network boundaries. While crucial, these traditional security mechanisms are no longer sufficient to protect modern IT infrastructures against the ever-evolving threat landscape adequately. On the other hand, the zero trust model challenges the long-standing assumption of trust within the network perimeter, instead adopting a more cautious and meticulous approach to security. At its core, zero trust operates on the fundamental principle that no user, device, or network component should be automatically trusted, regardless of location or past authentication. This approach ensures that every access attempt, regardless of its source, is subjected to thorough scrutiny and verification before being granted any level of privileged access. The zero trust model employs a variety of advanced security controls, including multi-factor authentication, microsegmentation, and continuous monitoring, to rigorously authenticate and authorize every access request. Furthermore, zero trust emphasizes the importance of continuously monitoring and evaluating network traffic, user behavior, and system activities to identify suspicious or anomalous activities swiftly. This proactive monitoring approach enables organizations to promptly detect and mitigate potential security breaches, minimizing the impact of cyberattacks and reducing the risk of data loss or system compromise.

By implementing the zero trust model, organizations can achieve several notable benefits. Firstly, it reduces the likelihood of successful cyberattacks by significantly limiting the potential attack surface and minimizing the impact of security vulnerabilities. Secondly, zero trust provides enhanced visibility into network traffic, facilitating improved threat detection and response capabilities. Moreover, the zero trust architecture model enables organizations to adopt a more granular and precise approach to access controls, ensuring that only authorized entities gain access to sensitive resources.

In summary, the zero trust model represents a significant paradigm shift in cybersecurity. It recognizes cyber threats' ever-evolving and multifaceted nature and acknowledges that perimeter-based security measures are no longer sufficient. By adopting a zero-trust approach, organizations can

fortify their IT systems against emerging cyber threats and establish a robust security posture that aligns with the complexities of the modern digital landscape.

In particular, the fundamental principles of the model are (Shan, Iqbal, and Saxena 2022):

(1) *never trust, always verify*: every user, device, application, and the data stream is considered untrusted. Each of the above must be authenticated and then explicitly authorized with the minimum required privileges,
(2) *assume breach*: it is assumed that the operator's devices and network may have already been breached by some malicious group. The deny by default principle is applied to every user, device, application, and data stream access request. Access is granted after multiple parameters are thoroughly examined (e.g., user name, device name, location, time, previously recorded user behavior, etc.).

The zero-trust approach incorporates comprehensive tracking of movements (Janicke, Abuadbba, and Nepal 2020). All-access requests, configuration changes, and network traffic are recorded in log files, which are automatically checked continuously for suspicious activity. The model accepts that any authorization of access to critical resources carries risks and requires immediate preparedness to respond to incidents, assess the damage, and recover business operations (Karunarathne, Saxena, and Khurram Khan 2021).

Figure 1 shows an example of an application of zero trust, where the attacker has compromised the passwords of a legitimate user and is attempting to gain access to the organization's systems.

Although the considered architecture appears to be the optimal solution for complex environments, particularly when implementing the zero-trust model, numerous challenges emerge that undermine the solution's effectiveness and raise concerns about the reliability of the provided product. The first potential challenge stems from the lack of comprehensive support from all members of an organization's staff. It is imperative to wholeheartedly embrace the principles of zero fault tolerance and acknowledge the existence of threats both outside and within traditional network perimeters for any proposed solution to achieve success.

To ensure the success of a zero-trust architecture, it is crucial to obtain complete consent and expertise from administrators. Without their full understanding and commitment to the underlying principles, the implementation may suffer from inefficiencies and vulnerabilities. Additionally, if users are granted the ability to override established security policies, it undermines the entire purpose of a zero-trust approach. Allowing such overrides weakens the security posture and negates the benefits that a zero-trust architecture can provide.

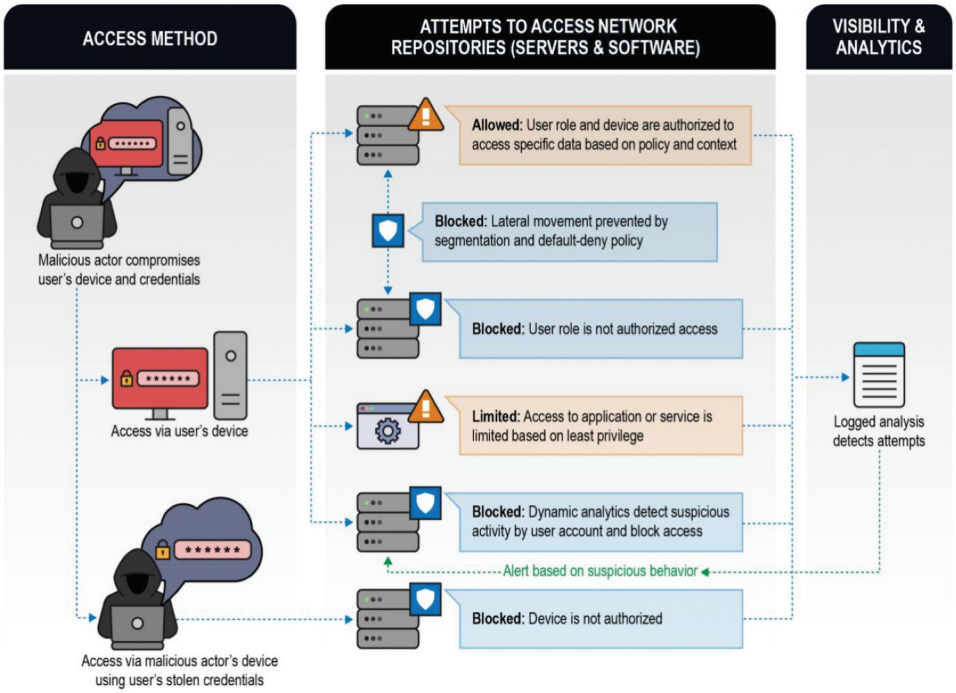

Figure 1. Zero trust remote exploitation scenario (https://media.defense.gov/).

Moreover, the reliability of the zero-trust model hinges on the organization's ability to accurately identify and authenticate users, devices, and applications. Inadequate identity verification mechanisms or loopholes in the authentication process can jeopardize the overall security of the architecture. Furthermore, the zero-trust model requires constant monitoring and analysis of network traffic, user behavior, and system logs to detect anomalous activities promptly. Failure to adequately monitor and analyze these factors can result in missed or delayed identification of potential threats, rendering the architecture less effective.

Another significant challenge lies in the complexity of implementing and managing a zero-trust architecture. It often requires a comprehensive reconfiguration of existing network infrastructure and the deployment of specialized security technologies. Organizations may face difficulties in integrating these new components with their legacy systems or ensuring seamless interoperability. Additionally, ongoing maintenance and regular updates to address emerging vulnerabilities and evolving security requirements are critical. Without proper management and continuous attention, the architecture's effectiveness may diminish over time.

In conclusion, while the considered architecture holds promise for addressing security challenges in complex environments, the successful implementation of a zero-trust model necessitates overcoming various challenges. These

challenges include obtaining full support and expertise from all stakeholders, eliminating user overrides, establishing robust identity verification mechanisms, ensuring continuous monitoring and analysis, managing the complexity of implementation, and providing ongoing maintenance and updates. By addressing these challenges, organizations can maximize the benefits of a zero-trust architecture and enhance their overall security posture (Alromaihi, Elmedany, and Balakrishna 2018).

Also, it should be understood that once even the most basic or intermediate zero trust capabilities are integrated into the network, monitoring is necessary to mature the implementation and achieve full benefits. Accordingly, the pervasive need for zero-trust applications across the enterprise environment mandates and drives the scalability of capabilities for the solution to be essential (Rodigari et al. 2021). An additional consideration is given to access control decisions that may have been made in the past are overridden and will be continually redefined as access to a sensitive resource is requested. This fact requires a robust infrastructure to create, enforce and record these access decisions. In addition, network elements that were not previously part of access control decisions may base on the architecture's policy implementation, become critical elements that require reliability and consistent use of their management modes (Samaniego and Deters 2018).

Persistence in the mind-set and implementation of the zero-trust security model over time is also an essential requirement. It should be ensured that administrators and architecture defenders, in general, provide continuous and uninterrupted enforcement of ever-tightening and updated security policies of deny-by-default and always assuming a breach occurs (Dzogovic et al. 2022).

It should be emphasized that the zero-trust architecture approach preserves and adheres to the observance of the above conditions so that its benefits are not degraded or eliminated with the evolution of the information system being implemented. However, because the implementation of such a strict policy implies its permanent dependence on human supervision and knowledge, which degrades its absolute success, this work presents an innovative privacy-preserving with zero trust computational intelligent hybrid system, which based on advanced artificial intelligence systems ensures the continuous application of zero trust architecture preserving privacy and providing the provision of a protected digital environment.

The motivation behind this research stems from the increasing need for robust privacy protection and secure environments in corporate networks, particularly in sensitive applications such as sports training. As organizations face evolving cyber threats and stringent privacy regulations, there is a demand for innovative approaches that combine privacy-preserving techniques with the zero-trust architecture to ensure comprehensive security.

The primary contribution of this research is the development and evaluation of a privacy-preserving and zero-trust computational intelligent hybrid

system. By integrating multi-level trust fields, authentication, authorization, and encryption processes, the system establishes a protected environment within corporate networks. The incorporation of k-anonymization and a mix of local and global recoding techniques further enhances privacy and data protection.

This research introduces several novel aspects to address the challenges of privacy preservation and security in corporate networks. The novelty lies in the following key elements:

(1) Hybrid System: The proposed system combines computational intelligence techniques, specifically the utilization of maximum likelihood estimates and fuzzy k-means algorithms. This hybrid approach leverages the strengths of both algorithms to enhance privacy preservation, data security, and decision-making processes within the zero-trust architecture.
(2) Multi-Level Trust Fields: The system incorporates multi-level trust fields, enabling a fine-grained control over access and ensuring that each request undergoes thorough authentication, authorization, and encryption. This approach provides an additional layer of security within the network.
(3) Application to Sports Training: While privacy-preserving systems have been explored in various domains, this research specifically focuses on the context of sports training applications. This domain introduces stringent authorization requirements and emphasizes the need for privacy protection, making it an important and challenging area to apply privacy-preserving techniques.
(4) Incorporation of Privacy Regulations: The research emphasizes the application of privacy regulations within the system. It explores methods to enforce privacy rules and compliance without explicit training, contributing to the development of systems that can adapt to evolving privacy regulations and policies.

By combining these novel elements, this research offers a comprehensive and innovative approach to privacy preservation, security, and trust within corporate networks. The integration of computational intelligence techniques, multi-level trust fields, and domain-specific application considerations makes this research unique and valuable in addressing the privacy and security challenges faced by organizations today.

Privacy-Preserving Zero Trust Intelligent System

The proposed system is based on a cryptographic protocol that provides anonymous access to trust information and is implemented in a network,

where each peer stores information and, at the same time, answers the trust queries of others. Chatters form anonymous groups and generate responses within the group. The answer to a trust query has k-anonymity protection against an adversary who can illegally obtain all communication on the network. In addition, the encryption system ensures that the initiator of a trust query can check the validity of an anonymous response (Gao et al. 2021; Vanickis et al. 2018).

Let D be a data set. The data is organized by rows and columns such that each row refers to a different person and each column relates to a different attribute of the person. The set of data D is also called a data table, and each line of the table will be called a tuple from now on. Modifying a data table to satisfy k-anonymity is inextricably linked to the generalization of the tuples that compose it (Bayardo and Agrawal 2005). This process of generalizing the tuples of a data table can be represented with the help of a generalization transformation. This transformation is called recoding. Mathematically, any generalization transformation can be expressed in the following way (Samaniego and Deters 2018):

$$\phi : T \rightarrow T' \tag{1}$$

where T is the space of tuples and applies $T = X_1 \times X_2 \times \ldots X_n$, where X_i is the field of values of the i-th attribute and T' is the space of generalized tuples for which $T' = X_1^{'} \times X_2^{'} \times \ldots X_n^{'}$ applies respectively.

To generalize a tuple, one can begin by generalizing its associated attributes. This abstract representation of data allows for identifying broader patterns and trends. Generalizing attributes independently allows for a more specific focus on each attribute's characteristics and determining an appropriate level of abstraction. One transforms their values into more generalized or less specific representations to generalize attributes. This can be done by substituting specific values with ranges or categories or applying mathematical functions to aggregate or summarize attribute values. A nuanced analysis can be performed by generalizing attributes separately, allowing for a deeper understanding of the data.

Different attributes may require different levels of generalization based on their inherent characteristics and analysis goals. For example, numerical attributes can be generalized by converting specific values into intervals or ranges, while categorical attributes may be generalized by grouping distinct values into broader categories. This approach to generalization provides a flexible and customizable framework for data analysis. It allows for identifying common patterns and trends across a diverse set of tuples. Furthermore, it facilitates the preservation of privacy and confidentiality by reducing the risk of exposing sensitive or personally identifiable information.

In summary, generalizing a tuple involves separately generalizing its attributes, enabling the identification of common patterns and trends through higher-level abstractions. This approach offers flexibility, customization, and privacy preservation in data analysis. More specifically, the value of each attribute of a tuple is assigned to a wider interval that contains it (e.g., the age attribute with a value of 23 can be set to the value range [20 — 25]). This process is repeated for all characteristics of all tuples. Finally, we have created a rule by which the attribute of each tuple is generalized to a specific field of values based only on its value (and not, for example, the values of other characteristics of the same tuple). For this reason, we say that this particular type of recoding works globally since, regardless of the tuple, the values of the same attributes are generalized in the same way, and for this reason, it is called global recoding. The mathematical expression of the transformation, in this case, will have the following form (Zhang et al. 2021):

$$\begin{gathered} \phi_i^g : X_i \to X_i^{'} i = 1..n \\ \phi^g(t) = \left(\phi_1^g(x_1), \phi_2^g(x_2), \ldots, \phi_n^g(x_n)\right) t = (x_1, x_2, \ldots, x_n) \in T \end{gathered} \tag{2}$$

where with φ^g we denote the value obtained by the transformation in its i-th dimension, each tuple has n attributes (so many are obviously the dimensions of the transformation), X_i is the field of values of the i-th attribute, and X_i is the generalized range of values and $t = (x_1, x_2, \ldots, x_n)$.

In the most general case, the determination of the generalized value of an attribute is not solely reliant on the initial value of the corresponding tuple. Instead, it can also be influenced by the values of other attributes within the same tuple. Consequently, it is possible for two tuples to possess identical attribute values, yet be generalized to different intervals. For instance, consider two tuples that share the value of 23 for the "age" attribute. In the process of generalization, these tuples may be transformed into the intervals [21 — 24] and [22 — 25], respectively.

During this transformation, the decision regarding the appropriate generalization to apply is contingent upon the location of the tuple. By analyzing the specific characteristics of the tuple, such as its attributes and values, a determination can be made as to which generalization method is most suitable. This allows for the creation of intervals or ranges that encapsulate a broader scope of data points, facilitating a more comprehensive representation of the underlying information. This transformation is called local recoding and is described as follows (Shan, Iqbal, and Saxena 2022; Wang et al. 2022):

$$\begin{gathered} \phi_i^l : T \to X_i^{'} i = 1..n \\ \phi^l(t) = \left(\phi_1^l(t), \phi_2^l(t), \ldots, \phi_n^l(t)\right) t \in T \end{gathered} \tag{3}$$

where all symbols used, have a similar meaning as before.

If we compare global and local recording techniques, we can observe notable differences in their approach. Global recoding involves mapping the value field of an attribute to a generalized value field, whereas local recoding maps a tuple to a generalized tuple. It is important to note that global recoding can be considered a specific case of local recoding. When considering performance, global recoding tends to result in a generalization that is, at best, comparable to the generalization achieved through local recoding. However, in general, local recording can potentially achieve less information loss compared to a global recording. Consequently, local recoding is generally preferred over global recoding due to its superior efficiency in preserving information.

To illustrate the advantages of local recording, let's consider an example. Suppose we have a dataset with attributes representing different countries and their corresponding populations. With global recoding, the values in the population attribute would be grouped into broad ranges or categories. For instance, populations could be categorized as "small," "medium," or "large" based on predetermined thresholds. While this approach provides a high-level understanding, it leads to information loss by losing the specific population values.

On the other hand, local recording allows for a more nuanced representation of the data. Instead of grouping the populations into predefined categories, local recoding retains the individual population values but abstracts them in a generalized manner. For example, the populations might be transformed into percentages representing the country's population size compared to the global population. By preserving the original population values while capturing the relative proportions, the local recording offers a more detailed and informative representation of the data than global recoding. It enables deeper analysis and potentially more accurate conclusions.

In summary, local recoding surpasses global recoding regarding information preservation and efficiency. Its ability to retain specific values while abstracting them in a generalized manner makes it preferable for data transformation tasks (Gao et al. 2021).

Another way to categorize the data generalization transformation is based on the dimensions we consider each time for generalization. In the simple case, the generalization is made for each dimension separately, and the recoding is characterized as single-dimensional recoding. In the general case, the generalization depicts not one attribute (or equivalently one dimension at a time) but represents the Cartesian product of several attributes or dimensions. In this case we have multi-dimensional recoding.

Each partition of the single dimensional space creates two separate groups of data, but the next partition affects both groups equally. Conversely, in the multi-dimensional approach, a partition of the space leaves two independent subspaces, each of which can be re-partitioned regardless of whether a possible intersection is allowed to the other or not. We understand that multi-

dimensional recoding is much more accurate while single-dimensional recoding is much simpler. Now combining the above categories, we create some variations on the above transformations. First, local recoding is performed multidimensionally by definition. Global recoding, however, can be performed in either a one-dimensional or multi-dimensional manner (Bayardo and Agrawal 2005; Wang et al. 2022).

In conclusion, we see why local recoding is considered more efficient than global recoding. Visually, we can say that while with global recoding this goal is to create as small partitions of the space as possible, with local recoding we can rename some tuples so that they participate in different groups. Since this goal is to create many (numerically) and small (in terms of the elements they contain) groups, we understand that local recoding is preferred to global recoding. At the same time, the multi-dimensional approach is also selected for the single-dimensional one.

To determine the distances between tuples, various methods can be employed. One such approach involves examining the distance of a point relative to two points that are significantly distant from each other (point1 and point2). In essence, these points can be regarded as the extreme points of the passage under examination. Subsequently, for every intermediate point, its distance from these two extreme points is evaluated. The point that is closest to it is then assigned to the same group. This process yields two distinct sets: one set comprising points that are closest to the extremity point1 (referred to as set part1), and another set consisting of points that are closest to point2 (referred to as set part2). This iterative procedure is continued until the resulting partitions can no longer be subdivided further, ensuring that all relevant distances between tuples have been accounted for (Tang et al. 2021).

A key point in the above analysis is the distance. We can use many heuristic functions that give us the distance between two points (for example, we could use the Euclidean distance or the Manhattan distance). However, since tuples other than points in some n-dimensional space have physical information, we want to use an expression directly related to the amount of information we lose when we group two individual tuples. For this reason, we use the Normalized Certainty Penalty (NCP). The NCP is calculated for a group of tuples (equivalently, this group can also be called an equivalence class) and is a comparison expression between the value range of the attributes of the group of tuples and the value range of the attributes of the data table. For a dimension (an attribute) the NCP is calculated for the equivalence class G as follows (Vanickis et al. 2018; Zeng et al. 2021):

$$NCP^{(i)}(G) = \frac{\overset{(i)}{max}_{G} - \overset{(i)}{min}_{G}}{\overset{(i)}{max} - \overset{(i)}{min}} \tag{4}$$

that is, it is the quotient between the range of values of the equivalence class in dimension i to the range of values of the data array in dimension *i*. In all dimensions, the NCP is calculated as follows (Hatakeyama, Kotani, and Okabe 2021; Lijun et al. 2021):

$$NCP(G) = \sum_{i \in QID} w_i NCP^{(i)}(G) \tag{5}$$

To ensure the correct transmission of a message, we encode the message, i.e., describe it in a suitable language, in which it will not be easy to mistake one word for another. For integer $b \geq 2$, the set $A = \{0, 1, \ldots, b-1\}$, consisting of b symbols, is an alphabet. An ordered sequence of k symbols from the set forms a word of length k. The coding distance is defined as follows (Kaihua et al. 2021; Zhang et al. 2021):

$$d(C) = \min_{w_1, w_2} \{d(w_1, w_2) \mid w_1, w_2 \in C\} \tag{6}$$

that is, the minimum is calculated for each pair of different code-words w_1, w_2 *of* C.

Ideally, the encoding should be such that the message can be understood even if there are a small number of errors. An encoding is k-error correcting if, for every codeword w of C, w is the unique codeword closest to any corruption of it in at most k symbols. Thus, if such a word is transmitted, we can locate the correct word even if errors occur.

The proposed encoding C is fault-tolerant as long as $d(C) \geq 2k + 1$. Specifically, considering such an encoding and even if at most k errors are noted during the transmission of each code word. Let w be a code word sent and let w' be the word received. Since at most k errors occur during the transmission of w, we conclude that (Magazzù, Ciarpi, and Saponara 2018):

$$d(w, w') \leq k \tag{7}$$

Let z be any codeword in the C encoding other than w. Then:

$$d(w, z) \geq 2k + 1 \tag{8}$$

and by the triangular inequality:

$$d(w, z) \leq d(w, w') + d(w', z) \tag{9}$$

We have:

$$d(w', z) \geq d(w, z) - d(w, w') \geq 2k + 1 - k = k + 1 \tag{10}$$

Thus, w is the only codeword with a shorter distance than the word w', which is therefore correctly interpreted as w.

But there is an upper bound on the number of codewords of C. Specifically, any codeword (of length n) of the encoding C with $S(w)$ the set of all words w

that differ k from w in at most n of the symbols of, is calculated as (Ohnishi et al. 2018; Rodigari et al. 2021):

$$|S(w)| = 1 + \sum_{i=1}^{k} \binom{n}{i} (b-1)^i = \sum_{i=0}^{k} \binom{n}{i} (b-1)^i = s \tag{11}$$

Since the encoding C is k error-tolerant, only w of these words can belong to the codewords of C. Moreover, for two different codewords w and w' of C, it holds that $S(w) \cap^S (w') = \emptyset$. If there was a word $y \in S(w) \cap^S (w')$, then due to the triangular distance inequality (Bayardo and Agrawal 2005; Berger 2013):

$$d(w, w') \leq d(w, y) + d(y, w') \leq k + k = 2k \tag{12}$$

But this contradicts the fact that $d(C) \geq 2k + 1$. Therefore, the number of codewords of the C encoding is at most b^n/s considering that the number of words of length n from an alphabet of b symbols is b^n.

This approach utilizes the concept of homomorphic encryption to ensure a robust level of privacy and adheres to the principles of zero trust architecture. This is achieved by incorporating a noise component into the original message. Homomorphic encryption is a cryptographic technique that allows computations to be performed on encrypted data without requiring decryption. By applying this method, sensitive information remains protected throughout the entire computation process, minimizing the risk of unauthorized access or data leakage. To bolster privacy, we introduce an additional layer of security by introducing noise to the original message. This noise is designed to obfuscate the underlying data and prevent any meaningful information from being discerned by unauthorized parties. By employing this technique, we ensure that even if an attacker gains access to the encrypted data, they will be unable to extract valuable insights or sensitive details without the necessary decryption keys.

By leveraging homomorphic encryption in conjunction with noise injection, this system guarantees end-to-end privacy and a stringent zero trust framework. This means that at no point in the computation or data handling process can any entity be granted implicit trust. Instead, every transaction, access request, or computation is rigorously verified, authenticated, and encrypted, thereby mitigating potential vulnerabilities and ensuring the utmost privacy and security of the involved data. This noise is controllable and can be corrected by decryption (by the owner of the corresponding key), analogous to an error-correcting code. The fully homomorphic CS cryptosystem is defined by the algorithms (KeyGen; Encrypt; Decrypt; Eval) and supports the evaluation of functions $f \in \mathcal{F}_{\mathcal{CS}}$.

The original CS cryptosystem has a security parameter λ and consists of the following algorithms (Behera and Rani Prathuri 2020; Bellare and Goldreich 2006):

$$\text{KeyGen}\left(1^{\lambda}\right) = p \tag{13}$$

$$\text{Encrypt}(p, m) = m' + pq \tag{14}$$

$$\text{Decrypt}(p, c) = (cmodp)mod2 \tag{15}$$

$$\text{Decrypt}(p, c) = (cmod2) \oplus (c \div pmod2) \tag{16}$$

$$\text{Decrypt}(p, c) = LSB(c) \oplus LSB(c \div p) \tag{17}$$

We observe that all ciphertexts are m' away from multiples of p. This distance is the noise of the ciphertext. Despite its existence, the decryption works correctly as the original message maintains parity.

Finally, and to ensure the continued application of the zero-trust architecture, preserving privacy, and ensuring the provision of a protected digital environment, a hybrid computational intelligence system is used based on the maximum likelihood estimation and the fuzzy k-means algorithm. Specifically, let the set $\mathrm{D} = \{\mathrm{x}_1, \ldots, \mathrm{x}_\mathrm{n}\}$ with n unlabeled samples chosen randomly and independently of the density mix (Bien et al. 2019; Cavazos-Cadena and Montes-de-Oca 2003):

$$p(\mathrm{x} \mid \theta) = \sum_{j=1}^{c} p\left(\mathrm{x} \mid w_\mathrm{j}, \theta_\mathrm{j}\right) P\left(w_\mathrm{j}\right) \tag{18}$$

where the full vector θ is assumed to be constant and unknown. The likelihood function of the observed samples is defined as (Burgin and Rocchi 2019):

$$p(\mathrm{D} \mid \theta) = \prod_{k=1}^{n} p(\mathrm{x}_\mathrm{k} \mid \theta) \tag{19}$$

The maximum likelihood estimate $\hat{\theta}$ is that value of θ that maximizes the likelihood function $p(D|\theta)$. In particular, if we assume that $p(D|\theta)$ is a differentiable function of θ, then necessary conditions for $\hat{\theta}$ arise. Let l be the logarithm of the likelihood function and let $_{\theta_\mathrm{i}}l$ be the rate of change of l for θ, then (Iezzi 2020):

$$l = \ln\, p(\mathrm{D} \mid \theta) = \sum_{k=1}^{n} \ln\, p(\mathrm{x}_\mathrm{k} \mid \theta) \tag{20}$$

Therefore:

$$_{\theta_\mathrm{i}}l = \sum_{k=1}^{n} \frac{1}{p(\mathrm{x}_k \mid \theta)} \quad _{\theta_\mathrm{i}}\left[\sum_{j=1}^{c} p\left(\mathrm{x}_k \mid w_j, \theta_\mathrm{j}\right) P\left(w_j\right)\right] \tag{21}$$

If we consider that θ_i and θ_j are mutually independent and the bound probability:

$$P(w_i \mid x_k, \theta) = \frac{p(x_k \mid w_i, \theta_i)P(w_i)}{p(x_k \mid \theta)} \tag{22}$$

Then we have:

$$\begin{aligned} {}_{\theta_i} l &= \sum_{k=1}^{n} \frac{1}{p(x_k \mid \theta)} \quad {}_{\theta_i} p(x_k \mid \theta) \\ &= \sum_{k=1}^{n} P(w_i \mid x_k, \theta) \frac{1}{p(x_k \mid w_i, \theta_i)P(w_i)} \quad {}_{\theta_i} p(x_k \mid \theta_i) \\ &= \sum_{k=1}^{n} P(w_i \mid x_k, \theta) \quad {}_{\theta_i} \ln\, p(x_k \mid \theta_i) \end{aligned} \tag{23}$$

Accordingly, the fuzzy k-means algorithm seeks to minimize the following cost function:

$$J_{fuz} = \sum_{i=1}^{c} \sum_{j=1}^{n} \left[\hat{P}\left(w_i \mid x_j, \hat{\theta}\right)\right]^b \| x_j - \mu_i \|^2 \tag{24}$$

where b is a free parameter that regulates the mixing of the various groups. The probabilities of participation in the groups are normalized in such a way that the total sum is zero:

$$\sum_{i=1}^{c} \hat{P}(w_i \mid x_j) = 1, j = 1, \ldots, n \tag{25}$$

Then, at a (local) minimum of the function J_{fuz} will hold:

$$\frac{\partial J_{fuz}}{\partial \mu_i} = 0 \text{ and } \frac{\partial J_{fuz}}{\partial \hat{P}_j} = 0 \tag{26}$$

so finally, the following solutions emerge:

$$\begin{aligned} \mu_i &= \frac{\sum_{j=1}^{n} \left[\hat{P}(w_i|x_j)\right]^b x_j}{\sum_{j=1}^{n} \left[\hat{P}(w_i|x_j)\right]^b}, \\ \hat{P}(w_i \mid x_j) &= \frac{\left(1/d_{ij}\right)^{1/(b-1)}}{\sum_{r=1}^{c} \left(1/d_{rj}\right)^{1/(b-1)}} \kappa\alpha 1 d_{ij} = \| x_j - \mu_i \|^2 \end{aligned} \tag{27}$$

Due to the infrequent existence of analytical solutions for the aforementioned equations, the computation of cluster centers and prior probabilities necessitates an iterative calculation process, as outlined by the above algorithm. This algorithmic approach enables the determination of these values through successive iterations, ensuring a convergence toward accurate results. The iterative calculation algorithm is designed to resolve the equations involved by employing a systematic step-by-step procedure. It begins by initializing the cluster centers and prior probabilities to initial values, typically set beforehand

or chosen randomly. These initial values provide a starting point for the iterative process.

Subsequently, the algorithm enters a loop where it performs a series of calculations in each iteration. The calculations involve updating the cluster centers and the prior probabilities based on specific rules and criteria. The specific method employed for updating these values varies depending on the chosen algorithm or technique for solving the problem. Throughout the iterations, the algorithm adjusts the cluster centers, refining their positions to better represent the underlying data distribution. Simultaneously, the prior probabilities, which indicate the likelihood of each data point belonging to a particular cluster, are also iteratively updated to improve their accuracy. The iterative process continues until a predefined termination condition is met. This condition can be based on several factors, such as the number of iterations performed, the convergence of the values, or the achievement of a desired level of accuracy. Once the termination condition is satisfied, the algorithm halts, and the final values for the cluster centers and prior probabilities are obtained.

It is important to note that the success and efficiency of the algorithm heavily rely on the initial values chosen, the termination condition defined, and the convergence criteria employed. Careful consideration and fine-tuning of these parameters are crucial for achieving satisfactory results.

In summary, due to the limited availability of analytical solutions for the equations involved, the iterative calculation algorithm presents an effective approach to determine cluster centers and prior probabilities. By iteratively refining these values through a well-defined step-by-step procedure, the algorithm ensures convergence toward accurate solutions for the given problem.

Literature Review

Zero trust techniques are becoming more popular, and the research community is proposing methods and techniques to utilize them further to take advantage of their privacy-oriented nature. This chapter presents recent research from various domains to highlight the applicability of Zero trust principles.

Li et al (Shan, Iqbal, and Saxena 2022) aimed to give a collection of new research directions and concepts for 5 G/6 G-enabled IoT and new IoT technology trends. They researched emerging technologies, including zero-trust and blockchain. They presented a zero-trust security framework for Future IoT and a blockchain-based device validation in the IoT ecosystem that may allow safe device admin rights and authentication. Finally, they highlighted the primary obstacles and possible research trends.

Wang et al (Wang et al. 2022) suggested an accurate locality-sensitive hashing-based traffic low prediction method with the capacity to preserve

privacy and handle and mine collected massive traffic data quickly. By deploying a series of tests on an actual traffic dataset, they proved the viability of their concept in terms of forecast accuracy, efficiency, and sensor data privacy. The fundamental weakness of their study is that they neglected to account for some background aspects, such as weather, occurrences, etc. Incorporating such variables into their forecasting model might significantly increase forecast accuracy and efficacy.

Ghimire and Ravat (Ghimire and Rawat 2021) provided a practical application of blockchain-enabled federated learning to accomplish completely safe, privacy-preserving, and verifiable collaborative training for the Internet of Vehicles which can provide trustworthy and secure services. The Internet of Vehicles is susceptible to several cyberattacks and privacy problems. Still susceptible to contamination and reverse engineering threats is Federated Learning. Blockchain has already shown a zero-trust, completely secure, decentralized, and auditable method for recording and distributing data. Their way minimized any anomalies that would have otherwise occurred on the Collaborative Training server or during conversations.

THEMIS was presented by Pestana et al (Pestana et al. 2021) as a decentralized, scalable, privacy-by-design advertising network that needs no user confidence. THEMIS offered traceability to its participants, compensated users for watching advertisements and let advertisers check the success and billing detailed reports of their advertising campaigns. Using smart contracts and zero-knowledge techniques, they develop a prototype of THEMIS, and preliminary performance evaluations indicate that it can grow linearly in a multi-sidechain environment. Even though other firms and initiatives have recommended using cryptocurrency for internet ads, they think their approach is the initial solution to deliver on this promise. Given the technique's applicability and the mix of protection, privacy, and performance, THEMIS may serve as the basis for an entirely new attitude to digital advertising.

Through a mix of blockchain and zero-trust concepts, Sultana et al (Sultana et al. 2020) intended to improve the privacy of medical records and picture transfer. Several papers dealing with these two ideas were examined, and a distributed, and trustless architecture for safe medical data and picture transport and storage was presented by merging these two notions. Blockchain was used to maintain an audit log of medical/health data transfers for future investigation. Zero trust principles were implemented to keep medical data secure during transmission and to increase user security. They want to put it on the Ethereum platform to test its scalability and efficacy in the real world and discover methods to make the system quicker and more user-friendly.

Jinila et al (Bevish Jinila, Prayla Shyry, and Christy 2022) established a Zero Trust Model to manage the security of Internet of Medical Things info, reduce

the incidence of risks, and provide additional benefits to consumers. Due to the COVID-19 pandemic, the Internet of Medical Things is seeing a firmer boom, and remote patient monitoring is crucial for senior patients. However, patient records are more sensitive, and improper handling might result in undesirable circumstances. Their suggested approach is a multi-component architecture using a multi-factor authentication process, and user confidence is considered. They want to examine and select the gathered data in the future so that it may be improved.

Experiment: English Education Model

Let D be a data array with a finite number of tuples. The finite set $[A_1, A_2, \ldots, A_n\}$ is the set of attributes of D. The data table shown in Table 1, depicts the medical data of some patients who participate in a sports training program. In this table, based on the above definition, the set of attributes are {*Age, Sex, Zipcode, Disease*} while the tuples are the rows of the table.

As is clear from the above definition and the above table, each attribute of the table is unique since it assigns a semantic property to each tuple of the table. Beyond the property, each attribute is directly linked to a value field. For example, the {*Sex*} attribute of the table has as a value field the set {*Male, Female*}. Conversely, each tuple in the array is not necessarily unique (Hamer 2012).

We now define the concept of table view or table view on another table that contains the same tuples but a subset of the attributes of the original table. That is, the data matrix D with a set of attributes A and let $A' \subseteq A$. The table D′, which contains the same tuples as D, with a set of attributes A′ is called the projection or view of the table D for A′ and is denoted by D[A′]. For example, the view of Table for the set $A' = \{Age, Zipcode\}$ is denoted as *D[Age, Zipcode]* and is the Table 2 below.

It appears that both tables have identical tuples, but the view table has fewer attributes.

Continuing, we define the quasi-identifier which plays a critical role in the k-anonymity model. In particular, for table D with a set of attributes A, we define as Quasi Identifier or QI and denote by QID a minimal subset of A,

Table 1. English Education data.

Age	Sex	Zipcode	Disease
25	Male	53711	Flu
25	Female	53712	Hepatitis
26	Male	53711	Brochitis
27	Male	53710	Broken Arm
27	Female	53712	AIDS
28	Male	53711	Hang Nail

Table 2. Age and Zipcode data.

Age	Zipcode
25	53711
25	53712
26	53711
27	53710
27	53712
28	53711

such that the table D[QI_d] if merged with other data, can lead to de-anonymization of one or more people.

It is clear that the QI is not unique, and we can say that the choice of attributes that create the QI in each case depends on the type of external data we have available for linking. For example, if we join the table *D* with $QI_D = \{Age, Zipcode\}$ to the following Table 3 (let *V*) and it concerns data contained in a list of athletes who will take part in qualifying matches, then we can find that Wang has a virus (Flu).

This can be found by concatenating the arrays $D[QI_d]$ and $V[QI_d]$. The two table views show that the tuple (25, 53711) is common. Table D and V shows that the person representing the tuple is called Wang, suffers from a virus (Flu), and is a man (Male). So we see that with an appropriate choice of QI we can remove anonymity with the help of linking. Based on the above definitions, we can now express the k-anonymity property (Bayardo and Agrawal 2005).

Let D be a matrix and QI_d the Quasi Identifier of this matrix. We say that the matrix D satisfies k-anonymity if and only if every tuple of the matrix D [QI_d] appears in the matrix at least k times. Let's see how we can change the Table English Education data so that it satisfies k-anonymity for various values of k for $QI_d = \{Age, Zipcode\}$ (Table 4 for k-anonymity (k = 2) and Table 5 for k-anonymity (k = 3).

Table 3. Preliminary games data.

Name	Age	Sex	Zipcode
Wang	25	Male	53711
Li	28	Female	55410
Chen	31	Female	90210
Liu	19	Male	02174
Zhang	40	Female	02237

Table 4. English Education data with k-anonymity (k = 2).

Age	Sex	Zipcode	Disease
25	Male	[53711–53712]	Flu
25	Female	[53711–53712]	Hepatitis
[26–28]	Male	53711	Brochitis
27	Male	[53710–53712]	Broken Arm
27	Female	[53710–53712]	AIDS
[26–28]	Male	53711	Hang Nail

Table 5. English Education data with k-anonymity (k = 3).

Age	Sex	Zipcode	Disease
[25–26]	Male	[53711–53712]	Flu
[25–26]	Female	[53711–53712]	Hepatitis
[25–26]	Male	[53711–53712]	Brochitis
[27–28]	Male	[53710–53712]	Broken Arm
[27–28]	Female	[53710–53712]	AIDS
[27–28]	Male	[53710–53712]	Hang Nail

With the help of Tables 4 and 5 we can understand why satisfying k-anonymity means maintaining the anonymity of individuals. If we attempt linking as before, we will see that the tuple from $V[QI_d]$ is not the same as a tuple from $D[QI_d]$ but lies within the intervals showing 2 tuples of $D[QI_d]$ in Tables 4 and 3 tuples of $D[QI_d]$ of Table 5. That is, if table D satisfies k-anonymity, then every random tuple whose attributes are $\mathrm{QI_d}$ will either be inside the intervals defining at least k (identical) tuples of D or will be outside of the intervals representing all the tuples of D. In other words, any linking attempt will result either in k tuples which we cannot distinguish or none.

Upon analysis of the aforementioned information, it becomes evident that the level of anonymization provided by this proposal is directly proportional to the value of k. In practical scenarios, the objective is to achieve the highest degree of anonymization while minimizing the loss of information. It is widely acknowledged that the problem of optimal k-anonymization falls within the category of NP-complete problems. Consequently, the pursuit of an ideal solution is not feasible within reasonable time constraints. Instead, the focus is on attaining a satisfactory solution that can be executed within a short timeframe (Goldreich 2011).

Regarding the efficiency of the intelligent system, let n possible events $A_1, \ldots, A_n$ occur with probabilities $P_1, \ldots, P_n$, respectively. Then entropy, symbolically $\mathrm{info}(A_1, \ldots, A_n)$, is defined as the average of the information of all events, i.e.:

$$\mathrm{info}(A_1, \ldots, A_n) = -\sum_{i=1}^{n} P_i \log_2(P_i) \tag{28}$$

The priors are then used to define a heuristic profit function. Let be a set T of multiplicity data $|\mathrm{T}|$ and let $\mathrm{freq}(C_\mathrm{j}, T)$ be the total number of data of the category $C_\mathrm{j}, \mathrm{j} \in \{1, \ldots, k\}$. According to the previous ones, we consider that the class C_j carries information equal to (Albarakati et al. 2018; Jaynes 2003):

$$-\log_2\left(\frac{\mathrm{freq}(C_\mathrm{j}, T)}{|T|}\right) \tag{29}$$

So, the entropy of the set of categories in the set T is:

$$\text{info}(T) = -\sum_{j=1}^{k} \frac{\text{freq}(C_j, T)}{|T|} \log_2 \left(\frac{\text{freq}(C_j, T)}{|T|} \right) \tag{30}$$

Then we calculate the average information in the partition of the set T into n subsets $T_1, \ldots, T_n$ as a result of applying a test X:

$$\text{info}_X(T) = -\sum_{i=1}^{n} \frac{|T_i|}{|T|} \text{info}(T_i) \tag{31}$$

Considering two categories, the «Disease» category with 9 data and the «no Disease» category with 5 data. The entropy of the set T is:

$$\text{info}(T) = -\frac{9}{14} \log_2 \left(\frac{9}{14} \right) - \frac{5}{14} \log_2 \left(\frac{5}{14} \right) = 0.940\text{bits} \tag{32}$$

Using the feature «age» as a test X we partition the set T into three subsets with entropy:

$$\begin{aligned} \text{infox}_X(T) = & \frac{5}{14} \left[-\frac{2}{5} \log_2 \left(\frac{2}{5} \right) - \frac{3}{5} \log_2 \left(\frac{3}{5} \right) \right] \\ & + \frac{4}{14} \left[-\frac{4}{4} \log_2 \left(\frac{4}{4} \right) - \frac{0}{4} \log_2 \left(\frac{0}{4} \right) \right] \\ & + \frac{5}{14} \left[-\frac{3}{5} \log_2 \left(\frac{3}{5} \right) - \frac{2}{5} \log_2 \left(\frac{2}{5} \right) \right] \\ = & \ 0.694\text{bits} \end{aligned} \tag{33}$$

So, the information gain in this case is 0.940–0.694 = 0.246 bits.

The aforementioned bias of the gain criterion can be removed by normalization which is achieved by dividing the information gain by the amount of split information defined as:

$$\text{split info(X)} = -\sum_{i=1}^{n} \frac{|T_i|}{|T|} \log_2 \left(\frac{|T_i|}{|T|} \right) \tag{34}$$

Finally, the heuristic function is derived:

$$\text{gain ratio(X)} = \frac{\text{gain(X)}}{\text{split info(X)}} \tag{35}$$

which is effective for selecting a test.

In conclusion, it is important to consider that when dealing with a feature that exhibits distinct values, there are various approaches one can adopt to determine the desired outcome. This can involve assigning a unique result for each additional value that the specific component receives or associating a specific outcome for each group within a partition of the values that the particular feature encompasses. On the

contrary, when working with a trait that assumes continuous numerical values, a binary test can be employed to assess its characteristics. This test involves establishing a specific threshold value, beyond which the trait is deemed to possess a certain attribute or characteristic, while values below the threshold indicate the absence of that attribute. By considering these different approaches, we can effectively analyze and handle features and traits within a technical context, allowing us to derive meaningful insights and make informed decisions.

Conclusion

This paper presents a comprehensive examination of an innovative privacy-preserving and zero-trust computational intelligent hybrid system. Built upon the zero-trust architecture, this system aims to establish a secure environment by incorporating multi-level trust fields within a corporate network. Each access request undergoes thorough authentication, authorization, and encryption processes before being granted access. The proposed system is evaluated in the context of a sports training application environment, which entails stringent authorization requirements and the need for robust privacy protection.

To enhance privacy, the suggested method incorporates k-anonymization techniques along with a combination of local and global recoding. These mechanisms help safeguard sensitive information and ensure that privacy-preserving measures are effectively implemented. In order to maintain the integrity of the zero-trust architecture and create a secure digital environment, a hybrid computational intelligence system is employed, leveraging the maximum likelihood estimates and fuzzy k-means algorithms.

Future directions for this research focus primarily on analyzing and enhancing the model by incorporating features that emphasize the application of privacy regulations, without requiring explicit training. To fully leverage the potential of interconnected learning systems, the automated system needs to explore various hyperparameter coordination strategies to improve precision and efficiency.

In summary, this paper presents a detailed exploration of a cutting-edge privacy-preserving and zero-trust computational intelligent hybrid system. By incorporating advanced techniques such as k-anonymization and a hybrid computational intelligence approach, the proposed system ensures a secure environment for sensitive applications, such as sports training. The research also outlines future avenues for further analysis and development, emphasizing the incorporation of privacy regulations and optimization of hyperparameter coordination strategies to enhance overall system performance.

Despite the promising aspects of the proposed privacy-preserving and zero-trust computational intelligent hybrid system, there are several limitations to

consider. First, the evaluation of the system is primarily focused on a sports training application environment. The effectiveness of the system in other domains or real-world scenarios remains to be explored. Additionally, the evaluation may not capture all possible attack vectors or scenarios, leading to potential vulnerabilities that require further investigation.

Another limitation is the reliance on the maximum likelihood estimates and fuzzy k-means algorithms as the basis for the hybrid computational intelligence system. While these algorithms have demonstrated effectiveness, there might be alternative algorithms or approaches that could yield even better results. Exploring and comparing different computational intelligence techniques could provide valuable insights for improving the system's performance.

Furthermore, the current study primarily focuses on the application of privacy regulations without explicitly addressing regulatory compliance and legal considerations. It is crucial to take into account the legal and regulatory frameworks that govern privacy protection and ensure compliance with relevant laws and policies. Future research should incorporate a legal and regulatory perspective to provide a more comprehensive understanding of the system's implementation in different jurisdictions.

In order to address the limitations mentioned above and further advance the field, future research directions can be pursued.

(1) Extension to diverse application domains: Conducting evaluations and case studies in various domains beyond sports training would validate the effectiveness and versatility of the proposed system. This could include healthcare, finance, or other industries where privacy and security are paramount.
(2) Exploration of alternative computational intelligence techniques: Investigating different computational intelligence algorithms and methodologies can contribute to enhancing the performance and efficiency of the hybrid system. Comparative studies that evaluate the pros and cons of various techniques would provide valuable insights for system optimization.
(3) Integration of legal and regulatory considerations: Future research should incorporate legal and regulatory aspects related to privacy protection. This involves analyzing and aligning the system with applicable privacy laws, data protection regulations, and industry standards. Such integration would ensure compliance and facilitate the system's adoption in real-world settings.
(4) Scalability and performance optimization: As the system operates within a corporate network, future research could focus on improving scalability and performance. Exploring hyperparameter coordination strategies, optimizing resource allocation, and considering distributed

architectures can enhance the system's efficiency and ensure its effectiveness in larger and more complex network environments.

(5) User-centric evaluation: Conducting user studies and gathering feedback from stakeholders can provide valuable insights into the usability and user experience of the system. User-centric evaluations would help identify potential challenges, usability issues, and areas for improvement from the end-user perspective.

By addressing these areas of future research, the proposed privacy-preserving and zero-trust computational intelligent hybrid system can be further refined and expanded, ultimately contributing to the development of more secure and privacy-conscious environments across various domains.

Disclosure statement

No potential conflict of interest was reported by the author.

References

Albarakati, A., B. Moussa, M. Debbabi, A. Youssef, B. L. Agba, and M. Kassouf. 2018. "Openstack-based evaluation framework for smart grid cyber security." *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, Aalborg, Denmark.

Alromaihi, S., W. Elmedany, and C. Balakrishna. 2018. "Cyber security challenges of deploying IoT in smart cities for healthcare applications." *6th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW)*, Barcelona, Spain.

Bayardo, R. J., and R. Agrawal. 2005. "Data privacy through optimal k-anonymization." *21st International conference on data engineering*, Tokyo, Japan. (ICDE'05).

Behera, S., and J. R. Prathuri (2020). Application of homomorphic encryption in machine learning. In *2020 2nd PhD Colloquium on Ethically Driven Innovation and Technology for Society (PhD EDITS)*, 1–2. Bangalore, India: IEEE.

Bellare, M., and O. Goldreich. 2006. On probabilistic versus deterministic provers in the definition of proofs of knowledge. *IACR Cryptol ePrint Arch* 2006: 359.

Berger, J. O. 2013. *Statistical decision theory and Bayesian analysis*. Berlin / Heidelberg, Germany: Springer Science & Business Media.

Bevish Jinila, Y., S. Prayla Shyry, and A. Christy. 2022. "A multi-component-based zero trust model to mitigate the threats in internet of medical things." *Data Engineering for Smart Systems: Proceedings of SSIC 2021*, Manipal University Jaipur, India.

Bien, J., I. Gaynanova, J. Lederer, and C. L. Müller. 2019. Prediction error bounds for linear regression with the TREX. *Test* 28:451–74. doi:10.1007/s11749-018-0584-4.

Burgin, M., and P. Rocchi (2019). Ample probability in cognition. In *2019 IEEE 18th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC)*, 62–65. Milan, Italy: IEEE.

Cavazos-Cadena, R., and R. Montes-de-Oca. 2003. The value iteration algorithm in risk-sensitive average Markov decision chains with finite state space. *Mathematics of Operations Research* 28 (4):752–76. doi:10.1287/moor.28.4.752.20515.

Dzogovic, B., B. Santos, I. Hassan, B. Feng, N. Jacot, and D. Thanh Van 2022. "Zero-Trust cybersecurity approach for dynamic 5g network slicing with network service mesh and segment-routing over IPv6." *2022 International Conference on Development and Application Systems (DAS)*, Suceava, Romania.

Gao, P., L. Yan, Z. Chen, X. Wei, L. Guo, and R. Shi. 2021. "Research on zero-trust based network security protection for power internet of things." *2021 IEEE 4th International Conference on Automation, Electronics and Electrical Engineering (AUTEEE)*, Shenyang, China.

Ghimire, B., and D. B. Rawat. 2021. Secure, privacy preserving, and verifiable federating learning using blockchain for internet of vehicles. *IEEE Consumer Electronics Magazine* 11 (6):67–74. doi:10.1109/MCE.2021.3097705.

Goldreich, O. 2011. *Studies in complexity and cryptography: Miscellanea on the interplay between randomness and computation*. vol. 6650. Springer. 10.1007/978-3-642-22670-0

Hamer, D. 2012. Probability, anti-resilience, and the weight of expectation. *Law, Probability and Risk* 11 (2–3):135–58. doi:10.1093/lpr/mgs004.

Hatakeyama, K., D. Kotani, and Y. Okabe. 2021. "Zero trust federation: Sharing context under user control towards zero trust in identity federation." *2021 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*, Kassel, Germany.

Iezzi, M. 2020. "Practical privacy-preserving data science with homomorphic encryption: An overview." *2020 IEEE International Conference on Big Data (Big Data)*, Atlanta, GA, USA.

Janicke, H., S. Abuadbba, and S. Nepal. 2020. "Security and privacy for a sustainable internet of things." *2020 Second IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications*, Atlanta, GA, USA. (TPS–ISA).

Jaynes, E. T. 2003. *Probability theory: The logic of science*. Cambridge university press. doi:10.1017/CBO9780511790423.

Kaihua, F., W. Zhang, Q. Chen, D. Zeng, X. Peng, W. Zheng, and M. Guo. 2021. "Qos-aware and resource efficient microservice deployment in cloud-edge continuum." *2021 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, Portland, OR, USA.

Karunarathne, S. M., N. Saxena, and M. Khurram Khan. 2021. Security and privacy in IoT smart healthcare. *IEEE Internet Computing* 25 (4):37–48. doi:10.1109/MIC.2021.3051675.

Lijun, Z., H. Guiqiu, L. Qingsheng, and D. Guanhua. 2021. An intuitionistic calculus to complex abnormal event recognition on data streams. *Security and Communication Networks* 2021:1–14. doi:10.1155/2021/3573753.

Magazzù, G., G. Ciarpi, and S. Saponara. 2018. "Design of a radiation-tolerant high-speed driver for mach Zehnder modulators in high energy physics." *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, Florence, Italy.

Ohnishi, R., W. Di, T. Yamaguchi, and S. Ohnuki. 2018. "Numerical accuracy of finite-difference methods." *2018 International Symposium on Antennas and Propagation (ISAP)*, Busan, Korea (South).

Pestana, G., I. Querejeta-Azurmendi, P. Papadopoulos, and B. Livshits. 2021. "Themis: A decentralized privacy-preserving ad platform with reporting integrity." *arXiv preprint arXiv:2106.01940*.

Rodigari, S., D. O'Shea, P. McCarthy, M. McCarry, and S. McSweeney. 2021. "Performance analysis of zero-trust multi-cloud." *2021 IEEE 14th International Conference on Cloud Computing (CLOUD)*, Chicago, IL, USA.

Samaniego, M., and R. Deters. 2018. "Zero-trust hierarchical management in IoT." *2018 IEEE international congress on Internet of Things (ICIOT)*, San Francisco, CA, USA.

Shan, L., M. Iqbal, and N. Saxena. 2022. Future industry internet of things with zero-trust security. *Information Systems Frontiers* 1–14. doi:10.1007/s10796-021-10199-5.

Sultana, M., A. Hossain, F. Laila, K. Abu Taher, and M. Nazrul Islam. 2020. Towards developing a secure medical image sharing system based on zero trust principles and blockchain technology. *BMC Medical Informatics & Decision Making* 20 (1):1–10. doi:10.1186/s12911-020-01275-y.

Tang, P., W. Wang, J. Lou, and L. Xiong. 2021. Generating adversarial examples with distance constrained adversarial imitation networks. *IEEE Transactions on Dependable and Secure Computing* 19 (6):4145–55. doi:10.1109/TDSC.2021.3123586.

Vanickis, R., P. Jacob, S. Dehghanzadeh, and B. Lee. 2018. "Access control policy enforcement for zero-trust-networking." *2018 29th Irish Signals and Systems Conference (ISSC)*, Belfast, UK.

Wang, F., L. Guangshun, Y. Wang, W. Rafique, M. R. Khosravi, G. Liu, Y. Liu, and Q. Lianyong. 2022. Privacy-aware traffic flow prediction based on multi-party sensor data with zero trust in smart city. *ACM Transactions on Internet Technology (TOIT)*, . doi: 10.1145/3511904.

Zeng, R., N. Li, X. Zhou, and Y. Ma. (2021). Building a zero-trust security protection system in the environment of the power Internet of Things. In *2021 2nd International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT)*, Shanghai, China, 557–60. IEEE.

Zhang, P., C. Tian, T. Shang, L. Liu, L. Lei, W. Wang, and Y. Zhao. 2021. "Dynamic access control technology based on zero-trust light verification network model." *2021 International Conference on Communications, Information System and Computer Engineering (CISCE)*, Beijing, China.