Journal of Algorithms and Computation

journal homepage: http://jac.ut.ac.ir

Privacy-preserving Data Fabric with Attribute-based Authentication and Access control (DF3A)

Sadra Fathenojavan[*1], Ali Moeini[†1], Ali Kamandi[‡1]

[1] School of Engineering Sciences, College of Engineering, University of Tehran

ABSTRACT

A data fabric layer guarantees that different kinds of data may be quickly and easily integrated, accessed, transformed, and managed. As the volume of data and access to it increases, the concern about its security also increases. Recently, special attention has been devoted to proposing new efficient architectures for a data fabric to provide secure communication between the data storage and their users. In this paper, using Shamir's Secret Sharing (SSS) and secure Attribute-Based Encryption (ABE), a new efficient data fabric architecture is proposed named Data Fabric with Attribute-based Authentication and Access Control (DF3A). DF3A offers security features such as perfect forward secrecy, untraceability, anonymity, mutual authentication, and availability. We verify these features using both informal and formal techniques. The verification results based on formal and informal techniques demonstrate that DF3A provides greater security than other architectures.

Keywords: Data Fabric, Elliptic Curve Cryptography, Attribute-based Authentication, Anonymity, Real-or-Random

AMS subject classification: 93C05.

ARTICLE INFO

Article history:
Research paper
Received 16, October 2024
Accepted 09, November 2024
Available online 20, December 2024

[*] sfathenojavan@ut.ac.ir
[†] Corresponding author: Ali Moeini, Email: moeini@ut
[‡] kamandi@ut.ac.ir

1 Introduction

Recent technological advancements, combined with the Data Fabric architecture, have paved the way for intelligent environments powered by sensors and actuators, enabling the delivery of services that address societal needs across a wide range of applications. By leveraging Data Fabric's ability to integrate and manage complex, heterogeneous data sources, valuable insights can be extracted from data collected by multiple sensors, transforming raw information into actionable intelligence. Effective data fusion, a critical component of Data Fabric, ensures seamless integration, unification, and data accessibility, which are essential for informed decision-making [3,12].

Data Fabric is a framework designed to enable seamless integration, management, and data accessibility from various sources and environments. Its goal is to offer a unified view of data, regardless of its location—whether on-premises, in the cloud, or at the edge. This framework addresses several challenges inherent in traditional data management systems, such as data silos, complexity, and the necessity for real-time data access [1]. A data-centric and security-focused data fabric intended for digital health applications is presented in this study. The increasing amount of Internet of Things (IoT) data from wearables, smartphones, and ambient sensors, driven by the growing interest in digital health research, has led to a significant rise in data volumes. Managing this vast amount of data, including various data types and time spans, is crucial [26]. Moreover, it is essential to comply with contractual and legal requirements. Data fabric consists of an architecture and a set of tools that facilitate the integration of disparate data sources from various contexts, presenting the data cohesively in dashboards. Additionally, the data fabric supports the development of modular and flexible data integration components that can be released as open-source or internal software. These components are used to build data pipelines that can be scheduled and deployed on-premises or in the cloud [2,3]. Therefore, an efficient data fabric architecture that ensures privacy-preserving data delivery to customers is critical [4].

In recent years, several attribute-based authentication (ABA) protocols [5,6] have been proposed to ensure secure communications. Attribute-based authentication is a method of authenticating parties based on their attributes rather than their identities, allowing them to be authenticated anonymously and preserving their privacy [7]. Due to its greater flexibility compared to traditional identity-based authentication schemes, ABA has been widely applied in various domains, such as multi-agent systems [8], eHealth systems [9], cloud computing [10], and more. However, it is challenging to implement existing ABA schemes in resource-constrained RFID systems, as they rely on computationally expensive operations [11]. In this paper, we address this issue by introducing DF3A, a lightweight elliptic curve attribute-based authentication protocol for the data fabric architecture. DF3A is based on elliptic curve cryptography and leverages the complexity of the elliptic curve discrete logarithm problem. The entities involved in DF3A include a trusted registration authority, the data store, and multiple customers. Customers initiate the protocol by specifying their policy in the form of a policy tree for the data store. In this tree structure, leaf nodes represent attributes, while internal nodes are threshold gates. The trusted registration authority then uses a multilevel threshold secret sharing scheme, called the hierarchical attribute-based secret sharing (HASS) scheme, to generate a set of attribute tokens for each customer based

on the policy tree. During the authentication phase, the data store verifies a customer's authenticity by checking the correctness of its blinded attribute tokens. After customer authentication, the data store authenticates itself to the customer. We provide an informal security analysis of DF3A and demonstrate that it can withstand common attacks. We also compare DF3A with several popular ABA schemes in terms of security properties and performance. The main contributions of this paper are as follows: DF3A is a novel lightweight elliptic curve attribute-based authentication and access control layer for Data Fabric, ensuring customer anonymity and untraceability against attackers. The proposed architecture provides distributed authentication and access control within the data fabric. In security analysis, it is shown that DF3A satisfies critical security properties relevant to Data Fabric systems. The structure of the paper is as follows: in Section 2 related works are reviewed. In Section 3 the technical foundations used throughout the study are presented. Details of the hierarchical attribute-based secret-sharing scheme are outlined. In Sections 5 and 6, common security properties and potential threats to Data Fabric systems are discussed. Our main contribution in-depth explanation of DF3A is presented in Section 7. Section 8 is devoted to evaluating the security features of DF3A. In Section 9, DF3A's performance and security characteristics with other well-known ABA schemes are compared. Finally, Section 10 presents the paper's conclusion.

2 Related work

Kuftinova et al. explore the principles and applications of Data Fabric in managing large-scale data for traffic and road systems. Their work highlights how platforms like IBM, Microsoft, and Cloudera enable the integration of big data, cloud computing, and blockchain to enhance Intelligent Transport Systems (ITS). This approach addresses challenges in data collection, analysis, and security using machine learning, microservices, and virtualization. Blockchain is employed to securely manage traffic events, accidents, and road system equipment data, facilitating efficient use of legacy ITS infrastructure and supporting crowdsourcing [12].

Kang Liu et al. developed a metadata-based Data Fabric architecture called M-Data-Fabric, which integrates metadata detection, analysis, and knowledge map creation to maximize the utility of diverse data sources. Their approach involves examining relationships between data sources to aggregate metadata, ultimately constructing a comprehensive business knowledge map based on correlations [3].

Italo Buleje et al. proposed a data architecture for a home-based telemonitoring study involving senior citizens in collaboration with the University of California, San Diego (UCSD). The study demonstrates the efficient integration of information from various IoT sensors and mobile apps to provide a comprehensive view of the health status of elderly individuals for further analysis [13].

Aurora Macías et al. created a comprehensive framework to enhance data fusion in pervasive systems. This framework includes two phases: the first focuses on system architecture design and data management, while the second evaluates data fusion systems based on international standards. The framework also includes guidance for building an evaluation subsystem to assess data fusion during runtime. They illustrated this approach through a mechanism designed to curb the spread of COVID-19 in assisted living facilities [15].

Ana-Maria Ghiran and Robert Andrei Buchmann proposed a method to eliminate the need for extensive data transformation knowledge by offering a diagrammatic integration management approach. Their technique uses agile modeling language to assist the front-end interface in combining relevant data, while RDF (Resource Description Framework) serves as a common representation to capture integration outputs in an Enterprise Knowledge Graph [16].

Seok-Jae et al. introduced a novel platform built on Data Fabric designed for cloud environments to overcome the limitations of legacy systems. This platform enables distributed data interoperability by physically linking data sources using knowledge graph databases. It also incorporates Holochain technology to integrate data into a scalable platform across multi-cloud environments, ensuring secure and authorized access regardless of data location. The use of a knowledge graph database mitigates issues related to heterogeneous data in decentralized settings, while Holochain improves memory and security processes in comparison to conventional blockchains [17].

Attribute-based authentication (ABA) is a promising security mechanism that allows users to authenticate anonymously based on their attributes, proving they own the required attributes without revealing their identity. Khader [19] introduced the first systematic description of ABA schemes and constructed an ABA protocol allowing verifiers to determine a set of credentials for the signing members of a particular group. Zhang et al. [20] expanded this by proposing a secure multi-agent protocol, where each agent holds a set of policy attributes to access protected datasets and ensure legitimate communication between agents. Guo et al. [21] proposed a decentralized system that uses verifiable attributes for mutual authentication while preserving privacy. Yang et al. [22] enhanced the group signature protocol from [23] by integrating it with dynamic policy trees to construct a more flexible ABA scheme. Ibrahim et al. [24] introduced two authentication schemes for cloud services supporting private attribute-based authentication. In the context of Data Fabric, we propose an authentication protocol that ensures users can trust that their data will reach the intended recipient, while the recipient can confirm the sender's authenticity without knowing their identity (anonymously), provided the intersection of their attribute sets meets a predefined threshold. In other words, a message can be exchanged only if the policy attribute in the message matches the recipient's required attribute.

3 Preliminaries

In this section, we give the basic definitions and preliminaries that will be used in DF3A. For the sake of description, the main notations used in this paper are listed in Table 1.

Table 1. Notations

Notations	Meaning
$\mathbf{s_0}, \mathbf{s_1}$	two secret values of the data store
$\mathbf{S_0}, \mathbf{S_1}$	the elliptic curve points of customer
$\boldsymbol{\Gamma} = (\mathbf{V}, \mathbf{E}, \boldsymbol{\vartheta})$	The policy tree
$\mathbf{x_R}$	the secret value of the customer
$\mathbf{ID_R}$	Identity of the customer
$\mathbf{r_1}, \mathbf{r_2}, \mathbf{r_3}$	The random values
$\mathbf{C_R}$	the elliptic curve point of customer

Λ_R	the set of attribute tokens
$\mathcal{I}_R$	the universe index set of attribute tokens

3.1 Data Fabric Concept

The new requirements and features are shown in Figure [1], a high-level diagram of DF3A architecture.

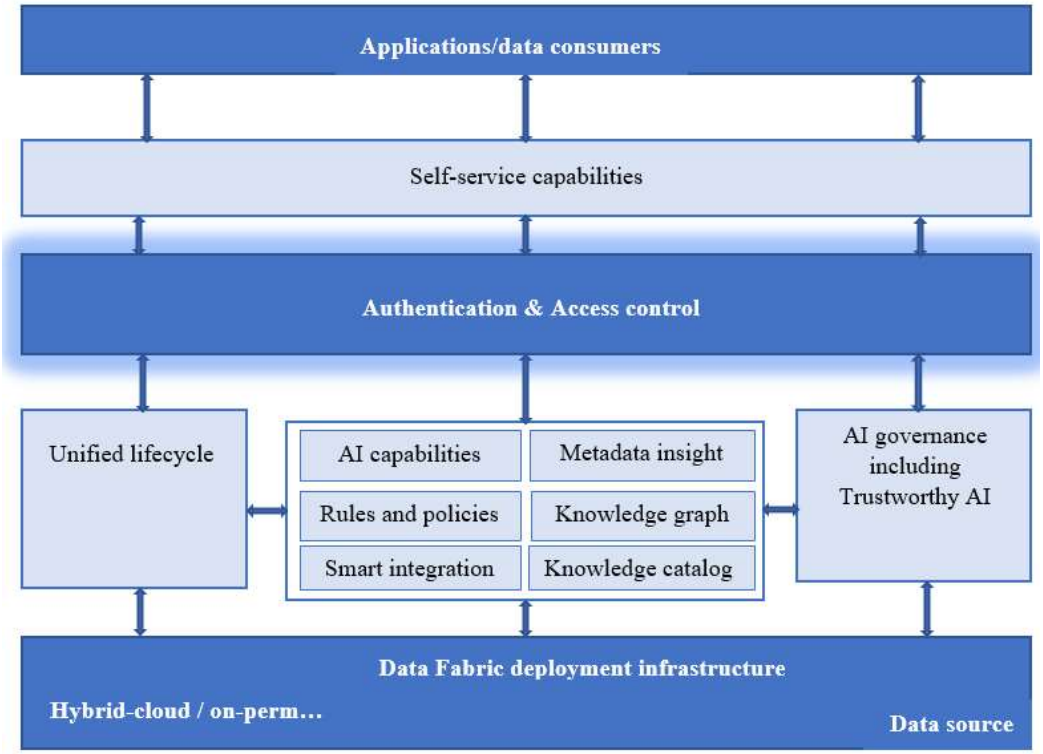

Figure 1. DF3A Framework.

The key functions and duties of the DF3A are briefly described here. It is simple to transfer these responsibilities and competencies to this architecture, as shown in Figure 1.

Catalog all your data: Business lexicon and metadata (business, technical, and operational) at design time and run-time are included.

Enable self-service capabilities: including data exploration, discovery, profiling, quality evaluation, and data consumption as a product, among other things.

Provide a knowledge graph: Visualizing the relationships between data, people, systems, processes, etc. to get more practical understanding.

Provide intelligent (smart) information integration: helping business users and IT personnel with data virtualization, federation, and integration and transformation projects.

Derive insight from metadata: Organizing and automating tasks and assignments for data engineering, data integration, and data governance from end to end.

Ensure trustworthy AI: Making sure AI techniques and results are transparent, explainable, and trustworthy; taking appropriate action in the event of bias, drift, declining accuracy and precision, etc.

Access Control & Authentication: this any procedure used by a system to confirm a user's identity before granting them access. Authentication is critical to successful security since access control is usually predicated on the identity of the user requesting access to a resource.

Enforce local and global data rules/policies: This includes the automated creation, modification, and enforcement of rules and policies based on AI and ML.

Manage an end-to-end unified lifecycle: putting in place a comprehensive and unified lifecycle for all Data Fabric tasks that spans platforms, identities, and organizations.

Enforcing data and AI governance: entails expanding the purview of traditional data governance to encompass AI artifacts, such as pipelines and models, and considering newly developing rules and privacy laws pertaining to data governance.

3.2 Elliptic curve cryptography

Elliptic curve cryptography (ECC) [18] is a public-key cryptosystem based on the algebraic structure of elliptic curves over finite fields (or Galois fields). There are two types of finite fields over which elliptic curves are defined, namely, prime fields and binary extension fields. However, in this paper, for the sake of simplicity, we only consider prime fields.

Let E be an elliptic curve defined over a prime field $\mathbb{F}_p$ $(p > 3)$ by a Weierstrass equation:

$y^2 = x^3 + ax + b$ (mod p), where $a, b \in \mathbb{F}_p$ and $4a^3 + 27b^2 \neq 0$ (mod p).

All points on E with the point at infinity O form an additive cyclic group, which is denoted by $E(\mathbb{F}_p)$:

$$E(\mathbb{F}_p) = \{(x, y) \in \mathbb{F}_p \times \mathbb{F}_p \mid y^2 = x^3 + ax + b\} \cup O. \quad (1)$$

The number of points in $E(\mathbb{F}_p)$ is defined as the order of E over $\mathbb{F}_p$ and is denoted by q. The addition of two points $P, Q \in \mathrm{E}(\mathbb{F}_p)$ is another point R such that $R = P + Q$. The repeated addition of P to itself $k - 1$ times is denoted by kP and is referred to as the scalar point multiplication. The order of any point $P \in \mathrm{E}(\mathbb{F}_p)$ is defined as the smallest positive integer k such that $kP = O$ A point $P \in \mathrm{E}(\mathbb{F}_p)$ is called a base point of $\mathrm{E}(\mathbb{F}_p)$ iff its order is equal to the order of $\mathrm{E}(\mathbb{F}_p)$.

The security of ECC depends on the ability to compute the elliptic curve discrete logarithm problem (ECDLP), which is defined as follows. Given two points $P, Q \in \mathrm{E}(\mathbb{F}_p)$, find an integer $1 \leq d \leq q$, if it exists, such that $Q = dP$. This problem is computationally hard when the order of $\mathrm{E}(\mathbb{F}_p)$, has a large prime factor.

3.3 Policy tree

A policy tree is a Layered structure where attributes form the leaf nodes, and internal nodes are threshold gates. Each threshold gate has a set number of child nodes and a threshold value that

defines the minimum number of child nodes that must be satisfied for the gate to be considered valid. A threshold of 1 represents an OR gate, while an AND gate is indicated when the threshold equals the total number of child nodes.

4 Secure Data Fabric Architecture

In this section, we introduce DF3A, an attribute-based authentication protocol designed for the Data Fabric architecture. The main entities involved in DF3A are the trusted registration authority, the customer, and the data store. The protocol operates in four distinct phases: setup, registration, authentication and access control, and key agreement. Each phase plays a critical role in ensuring the security and efficiency of the system. In the following subsections, we provide a detailed explanation of these phases. Table 1 summarizes the notations used throughout the remainder of this paper for clarity and reference.

Data Fabric facilitates real-time data access, lowers data silos, and helps enterprises streamline and improve their data management procedures. In addition, it facilitates the development of a self-serve data marketplace where users may find, exchange, and utilize data in an authorized and safe way. In order to handle the rising amount, velocity, and diversity of data being produced by digital business activities, data fabric is becoming an increasingly important part of modern information infrastructures. In our research, we are utilizing ECC-based encryption to privacy-preserving customers of the data fabric. We have improved a Data Fabric architecture by add access control and authentication layer. The customers are run on customer PCs, and they were authenticated by their attribute tokens, resulting in encrypted data being saved in our data lake. Using ECC-based encryption, a server can perform light weight computations on the attribute tokens, allowing for the detection of a legal customer or not. The server can be certain this customer is legal without knowing the identity of the customer because DF3A is anonymous and untraceable. This approach benefits from the Data Fabric's ability to provide a unified and integrated platform that enables data discovery, integration, management, authentication, and access control. By utilizing ECC encryption and a Data Fabric architecture, we can guarantee privacy-preserving customers who connect with Data Fabric, contributing to the development of more secure and privacy preserving Data Fabric Architect.

4.1 Setup

In this phase, the registration authority RA constructs an elliptic curve $E(\mathbb{F}_p)$ of large prime order q over a prime field $\mathbb{F}_p$, where P is a primitive point, and publishes the elliptic curve parameters. Then, it defines the attribute universe $A = \{a_1, a_2, ..., a_{|A|}\}$. Each attribute $a_i \in A$ is assigned a unique universe index $\mathcal{I}_T \in \mathbb{Z}_q$.

4.2 Registration

The registration phase consists of two key steps: data store registration and customer registration.

Data store registration: The data store S initiates this step by picking two random values $s_0 \in_R \mathbb{Z}_q$ and $s_1 \in_R \mathbb{Z}_q$. Then, it computes the elliptic curve points $S_0 = s_0 P$ and $S_1 = s_1 P$. Next, RA correspond to policy in data security constructs a policy tree Γ. Finally, RA stores the triple

(S_0, S_1, Γ) in a secure store. The sequence diagram of the receiver registration step is shown in Figure 2.

Customer registration: The registration authority RA initiates this step by picking a unique random value $x_R \in_R \mathbb{Z}_q$ for a given the customer R having an attribute set $\Omega_R \subseteq A$ each one of data which the data store has requested. Then, it constructs a HASS scheme over Γ by considering x_R as the secret. After that, RA leverages the HASS scheme to generate an attribute token τ_i for each attribute $a_i \in \Omega_R$. To do so, it finds a leaf node $v \in V_L$ such that $\Gamma(v) = a_i$ and computes $\tau_i = f_{\mathrm{p(v)}}\eta_{l(v)}S_0$, where $\eta_{l(v)}$ is the local order number of v within its siblings. It also sets the universe index of τ_i to $\eta_{\mathrm{u}(\tau_i)} = \eta_{\mathrm{u}(a_i)}$. The set of attribute tokens of R is denoted by Λ_R. Finally, RA computes the elliptic curve points $ID_R = x_R P$ and $C_R = x_R S_1$, and securely stores the 4-tuple $(ID_R, C_R, \Lambda_R, \mathcal{I}_R)$ in the customer, where $\mathcal{I}_R$ is the universe index set of attribute tokens in Λ_R. The sequence diagram of the Data Store registration step is shown in Figure 3.

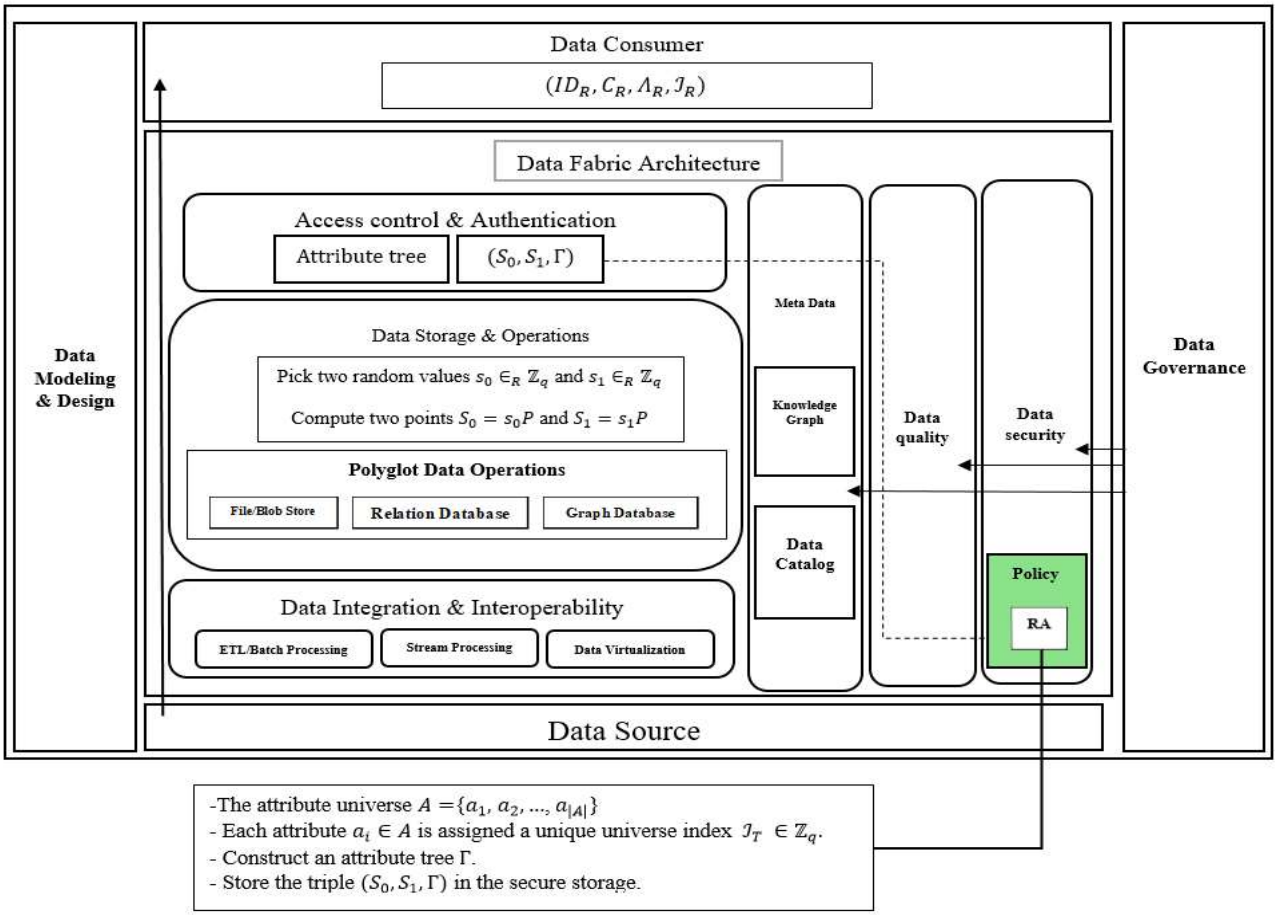

Figure 2. The sequence diagram of the data store registration step.

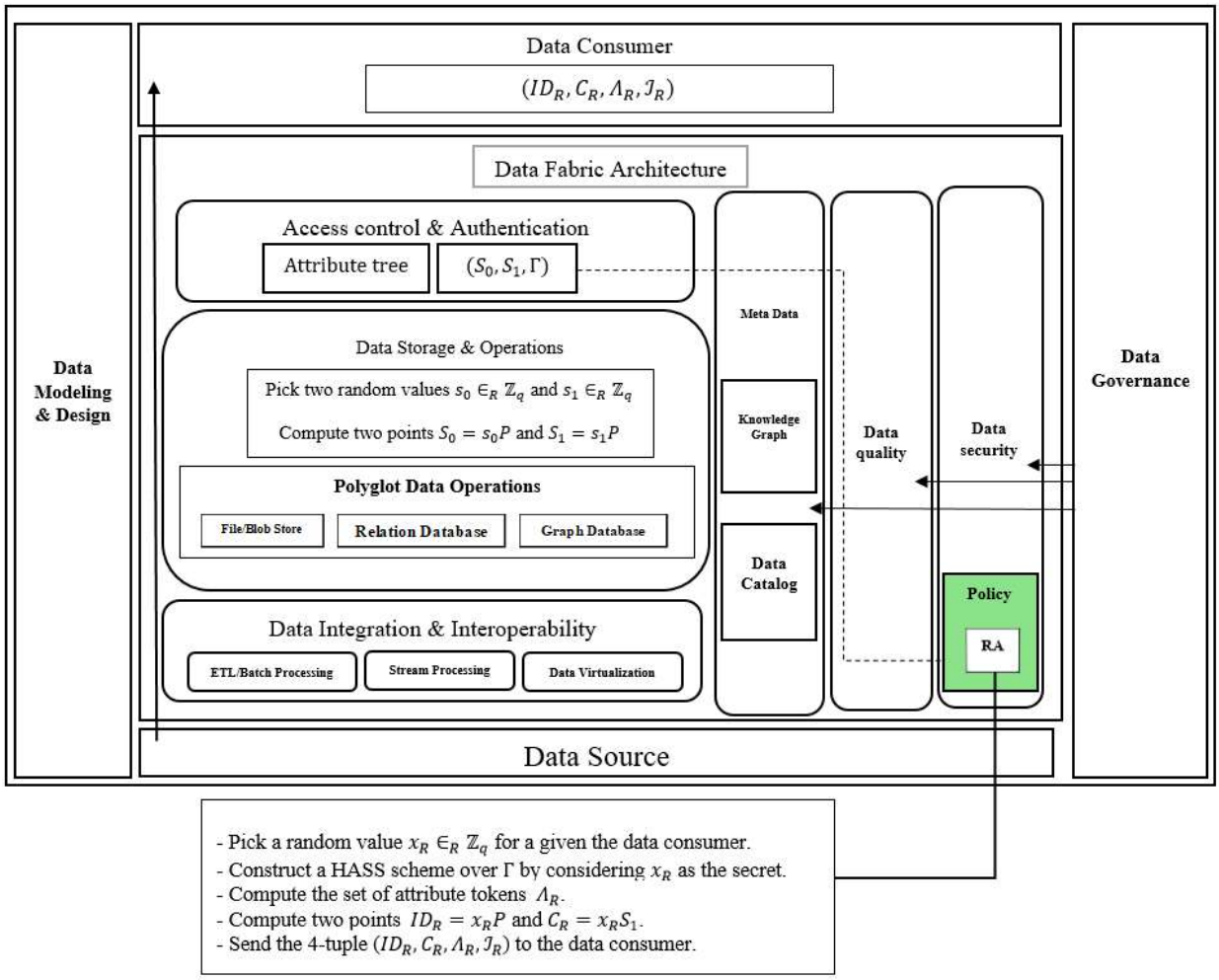

Figure 3. The sequence diagram of the Customer registration step.

4.3 Authentication & Access control

In this phase, the data store (δ) and the customer (R) exchange a series of messages to achieve mutual authentication. The interactions between δ and R are outlined as follows:

Step 1. δ picks a random value $r_1 \in_R \mathbb{Z}_q$ and sends it to R.

Step 2. R picks two random values $r_2, r_3 \in_R \mathbb{Z}_q$ and computes the elliptic curve points $ID'_{\mathrm{R}} = r_3 ID_{\mathrm{R}}$, $C'_{\mathrm{R}} = r_3 C_{\mathrm{R}}$, and $\Lambda'_{\mathrm{R}} = \bigcup_{\tau_i \in \Lambda_{\mathrm{R}}} \{r_3 \tau_i\}$. Then, R computes $M = H(ID'_{\mathrm{R}} || r_1)$ and sends the 4-tuple $(\Lambda'_{\mathrm{R}}, \mathcal{I}_{\mathrm{R}}, M, r_2)$ to δ.

Step 3. After receiving the message of R, δ recursively traverses Γ backward (i.e., starting from the leaf nodes) and checks $\Gamma_v(\mathcal{I}_{\mathrm{R}})$ for each node $v \in V$. If v is a leaf node and $\Gamma_v(\mathcal{I}_R)$ is True, then δ chooses an attribute token $\tau'_i \in \Lambda'_R$ having the same universe index as the attribute of v (namely, $\eta_u(\tau'_i) = \eta_u(\gamma(v))$) and sets the intermediate point $g(v)$ to τ'_i. More formally,

$$g(v) = \tau'_i = r_3 \tau_i = r_3 f_{p(v)}\big(\eta_l(v)\big) S_0 r_3 f_{p(v)}(\eta_l(v)) s_0 P. \qquad (2)$$

Otherwise, if v is an internal node and $\Gamma_v(\mathcal{I}_\mathrm{R})$ is True, then δ computes the intermediate point $g(v)$ as

$$g(v) = \sum_{u \in C'(v)} \ell_{\eta_l(u)}(0)\, g(u) \tag{3}$$

$$= r_3 \sum_{u \in C'(v)} \ell_{\eta_l(u)}(0)\, f_{p(u)}\big(\eta_l(u)\big) s_0 P$$

$$= r_3 \sum_{u \in C'(v)} \ell_{\eta_l(u)}(0)\, f_v\big(\eta_l(u)\big) s_0 P,$$

where $\ell_{\eta_l(u)}(0)$ is a Lagrange coefficient, namely,

$$\ell_{\eta_l(u)}(0) = \prod_{j \in L(p(u))} \frac{j}{j - \eta_l(u)} \bmod q$$

$$= \prod_{j \in L(v)} \frac{j}{j - \eta_l(u)} \bmod q \tag{4}$$

$C'(v)$ is the set of child nodes of v that are satisfied by the attributes of R, and $L(v)$ is the set of local indices of these child nodes, more formally,

$$C'(v) = \{u | u \in C(v) \wedge \Gamma_u(\mathcal{I}_\mathrm{R}) = \text{ True}\} \tag{5}$$

and

$$L(v) = \cup_{u \in C'(v)} \{\eta_l(u)\}. \tag{6}$$

Thus, the intermediate point of the root node r is computed as

$$g(r) = r_3 \sum_{u \in C'(r)} \ell_{\eta_l(u)}(0)\, f_r(\eta_l(u)) s_0 P = r_3 x_\delta s_0 P. \tag{9}$$

Then, δ multiplies $g(r)$ by $s_0^{-1} \bmod q$ to obtain the elliptic curve point $ID''_\mathrm{R} = s_0^{-1} g(r) = r_3 x_\mathrm{R} P$ and computes $M' = H(ID''_\mathrm{R} || r_1)$. If M' is equal to M, then δ authenticates R without being able to reveal its real identity and otherwise aborts it. Next, δ multiplies ID''_R by s_1 to obtain the elliptic curve point $C''_\mathrm{R} = s_1 ID''_\mathrm{R}$ and computes $M'' = H(C''_\mathrm{R} || r_2)$. Finally, it sends the message M'' to R.

Step 4. Upon receiving the message of δ, R computes $M''' = H(C'_R || r_2)$. If M''' is equal to M'', then R authenticates δ and otherwise aborts it.

Step 5. After receiving the message of R, δ recursively traverses Γ backward (i.e., starting from the leaf nodes) and checks $\Gamma_v(\mathcal{I}_\mathrm{R})$ for each node $v \in V$. If v is a leaf node and $\Gamma_v(\mathcal{I}_R)$ is True, then

δ chooses an attribute token $\tau_i' \in \Lambda_R'$ having the same universe index as the attribute of v (namely, $\eta_u(\tau_i') = \eta_u(\gamma(v))$) and sets the intermediate point $g(v)$ to τ_i'. More formally,

$$g(v) = \tau_i' = r_3\tau_i = r_3 f_{p(v)}\big(\eta_l(v)\big)S_0 = r_3 f_{p(v)}(\eta_l(v))s_0P. \tag{10}$$

Otherwise, if v is an internal node and $\Gamma_v(\mathcal{J}_\mathrm{R})$ is True, then δ computes the intermediate point $g(v)$ as

$$\begin{aligned} g(v) &= \sum_{u\in C'(v)} \ell_{\eta_l(u)}(0)\, g(u) \qquad (11)\\ &= r_3 \sum_{u\in C'(v)} \ell_{\eta_l(u)}(0)\, f_{p(u)}\big(\eta_l(u)\big)s_0P \\ &= r_3 \sum_{u\in C'(v)} \ell_{\eta_l(u)}(0)\, f_v\big(\eta_l(u)\big)s_0P, \end{aligned}$$

where $\ell_{\eta_l(u)}(0)$ is a Lagrange coefficient, namely,

$$\begin{aligned} \ell_{\eta_l(u)}(0) &= \prod_{j\in L(p(u))} \frac{j}{j-\eta_l(u)} \bmod q \\ &= \prod_{j\in L(v)} \frac{j}{j-\eta_l(u)} \bmod q \qquad (12) \end{aligned}$$

$C'(v)$ is the set of child nodes of v that are satisfied by the attributes of R, and $L(v)$ is the set of local indices of these child nodes, more formally,

$$C'(v) = \{u | u \in C(v) \wedge \Gamma_u(\mathcal{J}_\mathrm{R}) = \text{ True}\} \tag{13}$$

and

$$L(v) = \cup_{u\in C'(v)}\{\eta_l(u)\}. \tag{14}$$

Thus, the intermediate point of the root node r is computed as

$$g(r) = r_3 \sum_{u\in C'(r)} \ell_{\eta_l(u)}(0)\, f_r(\eta_l(u))s_0P = r_3 x_\delta s_0 P. \tag{15}$$

Then, δ multiplies $g(r)$ by $s_0^{-1} \bmod q$ to obtain the elliptic curve point $ID_\mathrm{R}'' = s_0^{-1}g(r) = r_3 x_\mathrm{R} P$ and computes $M' = H(ID_\mathrm{R}''||r_1)$. If M' is equal to M, then δ authenticates R without being able to reveal its real identity and otherwise aborts it. Next, δ multiplies ID_R'' by s_1 to obtain the elliptic curve point $C_\mathrm{R}'' = s_1 ID_\mathrm{R}''$ and computes $M'' = H(C_\mathrm{R}''||r_2)$. Finally, it sends the message M'' to R.

Step 6. Upon receiving the message of δ, R computes $M''' = H(C'_R||r_2)$. If M''' is equal to M'', then R authenticates δ and otherwise aborts it.

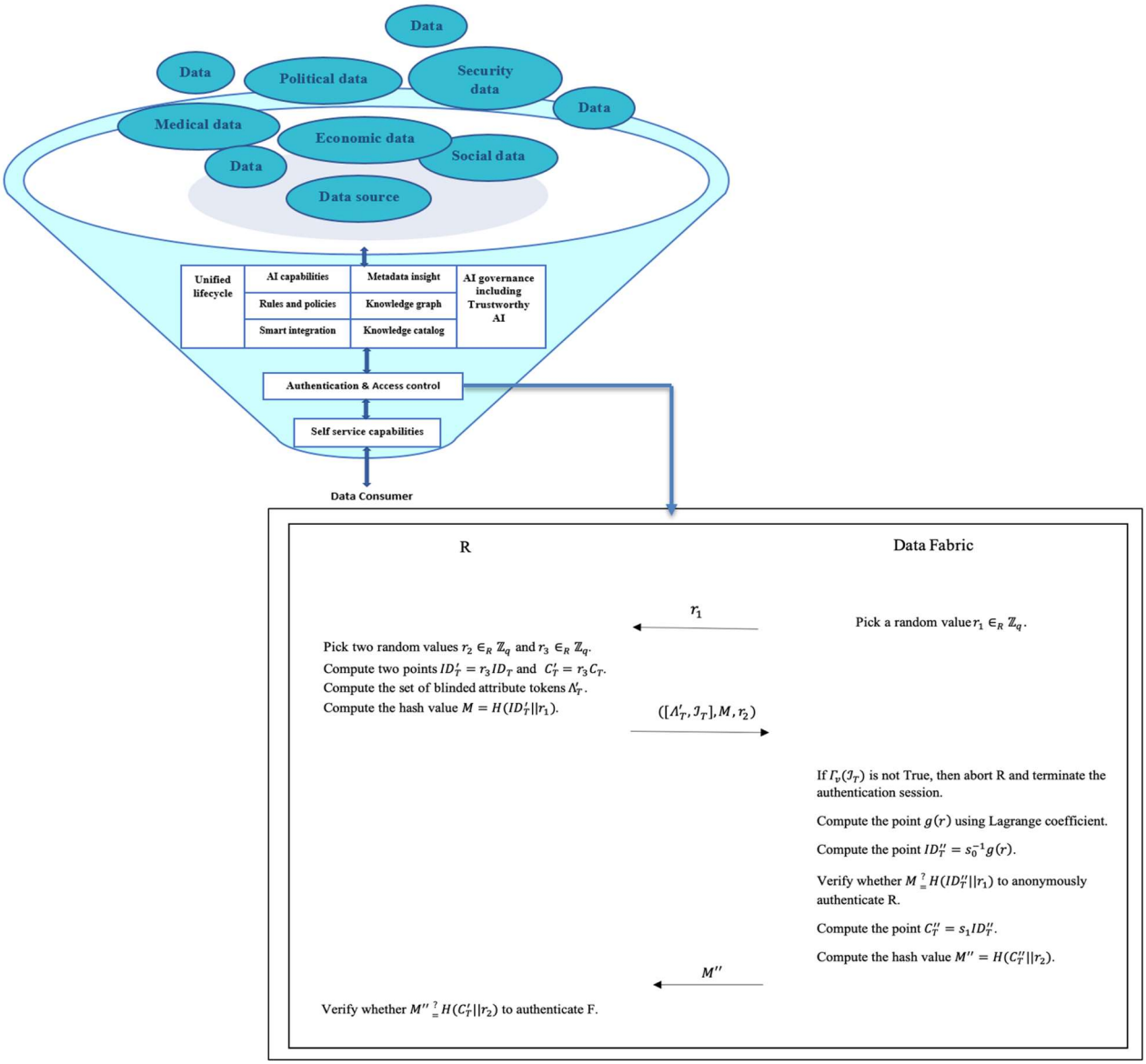

Figure 4. The sequence diagram of the mutual Authentication and generate key agreement step.

4.4 Key agreement

Note that C'_R and C''_R are equal elliptic curve points and are refreshed in each authentication session. Therefore, δ and R can respectively use C''_R and C'_R as the session key to encrypt the communication channel.

5 Security analysis

In this section, we first discuss the secure interaction between a research physician and the hospital's data store within the DF3A Data Fabric Architecture. We then provide both informal

and formal security analyses of DF3A, highlighting how it satisfies key security properties and mitigates various security threats.

5.1 Case study

In this scenario, patient medical data is stored in a hospital data store within DF3A architecture. A research physician, who needs access to specific data for their research, must securely interact with the data store using the DF3A architecture to ensure proper access control, privacy, and security.

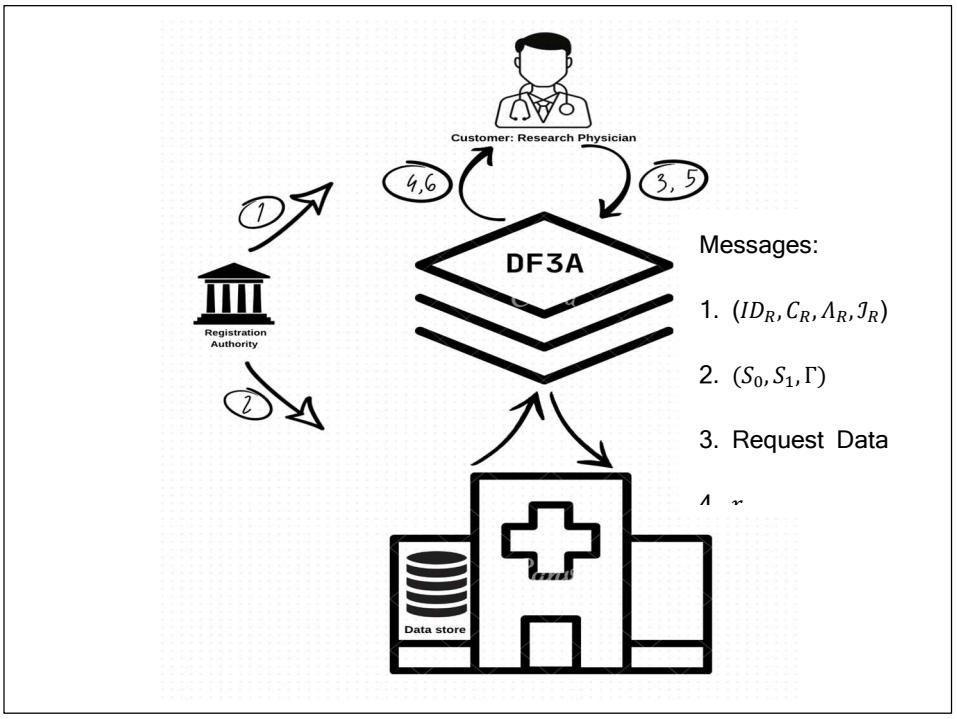

Figure 5. The interaction scenario between the research physician and the data store in the DF3A architecture.

Initial Setup

The Registration Authority (RA) configures the security parameters: An elliptic curve of large order is established, and indices for attributes (such as patient conditions, treatments, etc.) are defined.

Access policies are established for medical data. For example, access to patient data might be limited to researchers with a valid credential and specific attributes like a specialization in oncology.

Registration

Research Physician Registration: The process begins with the research physician (R) submitting their credentials, such as medical license number, specialization, and research domain, to the Registration Authority (RA). The RA generates unique attribute tokens corresponding to these credentials and securely provides them, along with public and private keys, to the physician. These attribute tokens enable the physician to access data relevant to their defined attributes while preserving their anonymity and ensuring their actual identity remains undisclosed.

b. Data Store Registration: The data store (S) is responsible for generating encryption keys and defining policies that govern data access control, ensuring robust security measures. These keys and policies are securely stored within the system to prevent unauthorized access. By leveraging these mechanisms, the data store ensures that only authorized users can interact with its resources, maintaining the confidentiality and integrity of the stored data.

Authentication & Access Control

Initiating the Interaction:

The research physician sends a request for access to specific patient data to the data store.

Mutual Authentication:

The data store sends a cryptographic challenge to the research physician.

The physician uses their attribute tokens and private key to generate a response, which is sent back to the data store.

The data store checks the response against the access policies and the physician’s attributes to ensure authorization.

Authentication Confirmation:

If the authentication is successful, the physician is granted access to the specific data that matches their attributes, without revealing their real identity. And the physician sure that the data store is legal.

Key Agreement: After successful mutual authentication:

The research physician and the data store generate a unique session key.

This session key is used to encrypt the data exchanged during the session, ensuring confidentiality and protecting against eavesdropping.

Benefits and Application in This Scenario

Patient Privacy: Patient identities are protected, and only the relevant data according to the physician's attributes is provided.

Secure Interactions: Two-way authentication and advanced encryption guard against attacks and eavesdropping.

Flexibility: The physician accesses only the data related to their research domain, with no unnecessary exposure to unrelated or sensitive information.

Fit for Sensitive Environments: The protocol is designed for environments with sensitive data, such as healthcare, ensuring compliance with privacy regulations.

5.2 Informal security analysis

Confidentiality: An adversary who intercepts messages cannot deduce the secret identifier x_R of the customer and the secret values s_0 and s_1 of the data store, which are included in attribute tokens, because of the difficulty of the elliptic curve discrete logarithm problem.

Mutual authentication: During the mutual authentication phase, the data store (δ) transmits the challenge r_1 to the customer, who then responds with the hash value M and a set of blinded attribute tokens Λ'_R. δ verifies the authenticity of R by validating the correctness of M. To do so, δ applies the Lagrange interpolation to the blinded attribute tokens in Λ'_R to obtain the elliptic curve point $g(r) = r_3 x_R s_0 P$, where r_3 is a random value and s_0 is a secret value only known to δ. Then, it multiplies $g(r)$ by $s_0^{-1} \bmod q$ to obtain the elliptic curve point $ID''_R = r_3 x_R P$ and computes $M' = H(ID''_R || r_1)$. If M' is equal to M, then δ ensures that the attribute tokens of R are valid and thus authenticates R. Similarly, R sends the challenge r_2 to δ and then receives $M'' = H(C''_R || r_2)$ from it, where $C''_R = s_1 ID''_R$. R believes δ is authentic by checking the correctness of M''. To do so, R computes $M''' = H(r_3 C_R || r_2)$, where C_R is a secret elliptic curve point stored in the memory of δ. If M''' is equal to M'', then δ authenticates R (because of the value of s_1 is only known to R).

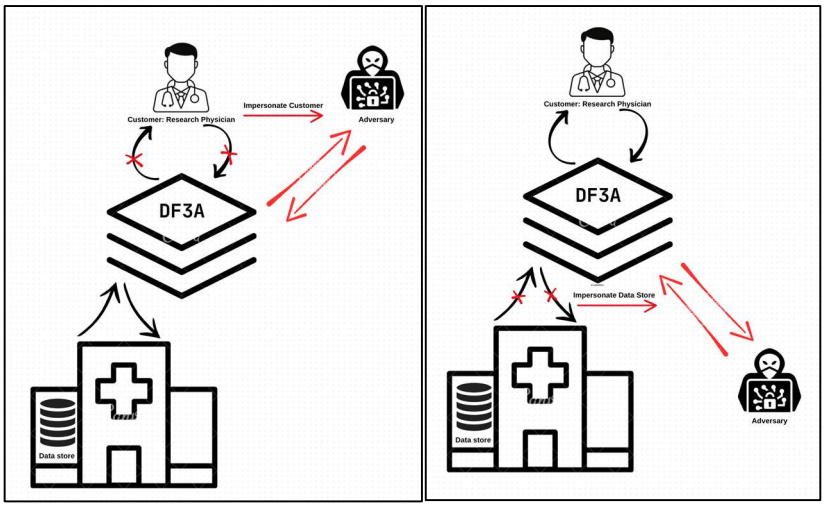

Figure 6. The DF3A architecture prevents impersonate attacks through mutual authentication.

Availability: Numerous authentication protocols require frequent updates to the secret values among the involved entities, increasing their susceptibility to desynchronization or denial of

service attacks. In contrast, DF3A operates continuously between the client and the data store without the need for synchronized updates to the secret data held in their memories.

Anonymity: The customer is authenticated based on its universal attributes and not on its identity. Thus, the honest-but-curious data store (or an adversary) cannot identify the real identity of δ.

Untraceability: The customer picks a random value r_3 in each authentication session then blinds its attribute tokens and other elliptic curve points stored in its memory. Therefore, the honest-but-curious data store (or an adversary) cannot trace the customer.

Perfect forward security: An adversary who compromises the customer and obtains the data stored in its memory, namely, ID_R, C_R, and Λ_R, cannot trace back any of the past communications of customer. This is because the adversary still does not know the blinding factor r_3, which is randomly generated in each authentication session. Thus, the adversary cannot identify messages transmitted in the past communications of the customer.

In the following, we show that our protocol prevents several known attacks.

Impersonation attack: An adversary who tries to masquerade the customer will not be able to generate valid attribute tokens without knowing the elliptic curve point S_0, which is only known to the registration authority and the data store. Besides, the adversary cannot masquerade the customer without knowing the secret values s_0 and s_1, which are only known to registration authority.

Replay attack: An adversary may attempt to pass the authentication by reusing the hash values M or M''. However, this attempt fails because of the freshness and unpredictability of these values in each authentication session.

Customers collaboration attack: Malicious customers in the system may attempt to collaborate and combine their attribute tokens to pass the authentication. Note that each customer is assigned a unique random secret x_R independent of other the customers in the system, which is shared among the attribute tokens of customer. Since x_R is only known to the registration authority, a malicious the receiver cannot combine its attribute tokens with other malicious customer to successfully bypass Step 3 of the mutual authentication phase.

Desynchronization attack: An adversary cannot block the communication between a receiver and the data store, because they do not need to update any secret data stored in their memory or database in each authentication session.

Track and trace attack: Both an adversary and an honest-but-curious data store cannot track and trace the customer by observing the messages sent by R. Since in each authentication session, a random value r_3 is used to refresh these messages.

5.3 Formal Security Analysis Using RoR Oracle Model

This section evaluates a session key (M'') security of our protocol from a passive/active adversary by using RoR oracle model [25]. We first introduce the RoR oracle model prior to demonstrating SK security for our protocol.

Table 2. Queries and descriptions

Queries	Description
Send ($\boldsymbol{\Gamma^R, message}$)	Based on Send(.), the adversary can transmit the message to Γ^R and receive the respond message via a public channel.
Corrupt $\boldsymbol{R}$ ($\boldsymbol{\Gamma_R^{t_1}}$)	This query is considered as the customer R stolen attacks, where the adversary can extract the secret parameters stored in R.
Corrupt $\boldsymbol{\delta}$ ($\boldsymbol{\Gamma_\delta^{t_2}}$)	This query is considered as the data store δ capture attacks, where the adversary can extract the secret parameters stored in δ.
Execute ($\boldsymbol{\Gamma_R^{t_1}, \Gamma_\delta^{t_2}}$)	Using this query, the adversary performs the passive/active security attacks by eavesdropping the transmitted messages between each participants via a public channel.
Test ($\boldsymbol{\Gamma^R}$)	An unbiased coin c is tossed prior to starting of the games. If adversary obtains the condition c=1 using Test(.) query, it denotes a SK between $\Gamma_R^{t_1}$ and $\Gamma_{DS}^{t_2}$ is fresh. If the adversary obtains the condition c=0, it denotes a SK is not fresh; otherwise, the adversary obtains a null value ($\perp$).
Reveal ($\boldsymbol{\Gamma^R}$)	Based on this query, the adversary reveals a SK generated between $\Gamma_D^{t_1}$ and $\Gamma_{DS}^{t_2}$ using Reveal(.) query.

Table 2 tabulates the necessary queries for the RoR oracle model such as Reveal(.), Corrupt(.), Send(.), Execute(.), and Test(.) to prove the mathematical (formal) security analysis. Furthermore, a "collision-resistant hash function Hash(.)" as random oracles are used.

Theorem. $Adv_{adversary}^{C_R}$ denoted the capabilities of adversary in violating SK security for the proposed DF3A scheme. Then, we can derive as follows:

$$Adv_{adversary}^{C_R} \leq \frac{q_h^2}{|hash|} + 2max\{\mathrm{C}.q_{send}^{DS}, \frac{q_{DS}}{2^{l_1}}, \frac{q_{DS}}{2^{l_2}}\} \quad (16)$$

q_{send}, q_h, $Hash$, and q_P, are the $send(.)$ query, the range space of hash function $h(.)$, the number of Hash query, and also, s,C, l_n, l_m are Zipf's parameters.

Proof. We present the sequence of four games namely GM_i ($i\epsilon[0,3]$). We denote that $Adv_{adversary,GM_i}^{C_{DS}}$. All GM_i are described as following:

Game $\boldsymbol{GM_0}$: This game is based on an actual security attack carried out by the adversary in our protocol. The bit c is randomly generated before the beginning of GM_0. Based on the GM_0, we obtain the following result:

$$Adv_{adversary}^{C_R} = |2.Adv_{adversary,GM_0}^{C_R} - 1| \quad (17)$$

Game $\boldsymbol{GM_1}$: GM_1 is considered that the adversary executes the eavesdropping attacks in which the exchangd messages are intercepted between $\Gamma_R^{t_1}, \Gamma_{DS}^{t_2}$ using $Execute(.)$ query. Moreover, the adversary performs $Reveal(.)$ and $Test(.)$ queries to derive a SK. The output of these queries determines whether the adversary obtains the sensitive knowledge and C_R between $\Gamma_{DS}^{t_1}, \Gamma_{\mathrm{R}}^{t_2}$. To derive a C_R, the adversary needs to the random nonces $\{RN_{C_R''}, RN_{r_2}\}$ and the secret credentials

$\{x_R, s_1, s_0\}$. Hence, the adversary's probability to win game GM_1 by performing the eavesdropping attacks is not at all increased. We derive the following result:

$$Adv^{C_R}_{adversary,GM_0} = Adv^{C_R}_{adversary,GM_1} \quad (18)$$

Game GM_2: This game is modeled as the passive/active attacks by using $send()$ and $Hash$ queries. In this game, the adversary can eavesdrop the exchanged messages $\{M,M''\}$ during the authentication and key agreement process. All messages are protected utilizing a $h(.)$. in addition, random nonce $\{\Lambda'_R, \mathcal{I}_R\}$ not compromised from the exchanged messages because Λ'_R is protected due to $ECDLP$ hard problem and $\mathcal{I}_R$ is policy of access structure to control access. we get the following result:

$$|Adv^{C_R}_{adversary,GM_2} - Adv^{C_R}_{adversary,GM_1}| \leq \frac{q_h^2}{2|hash|} + Adv^{DLP}_{adversary} \quad (19)$$

Game GM_3: In this final game, GM_3 is modeled the simulation of the Corrupt δ (.) and CorruptR(.) queries. The adversary cannot distinguish the secret values x_R, s_1, s_0 because of the difficulty of the elliptic curve discrete logarithm problem , therefore, the probability of guessing the secret key of R is l_1 and the secret keys of δ is l_2 and l_3 by the adversary approximately $\frac{1}{2^{l_1}}$, $\frac{1}{2^{l_2}}$ and $\frac{1}{2^{l_3}}$. Consequently, GM_2 and GM_3 are indistinguishable if the offline or online attacks are not successfully implemented. Hence, we obtain the following results:

$$\left|Adv^{C_R}_{adversary,GM_3} - Adv^{C_R}_{adversary,GM_2}\right| \leq \left\{C.q^{DS}_{send}, \frac{q_{DS}}{2^{l_1}}, \frac{q_{DS}}{2^{l_2}}, \frac{q_{Ds}}{2^{l_3}}\right\} + Adv^{DLP}_{adversary} \quad (20)$$

After $GM_0 - GM_3$ are performed successfully, the adversary tries to guess the valid c bit to win the games by using Test(.) query. Therefore, we obtain the following result:

$$Adv^{C_R}_{adversary,GM_3} = \frac{1}{2} \quad (21)$$

By combining (1),(2), and (5), we obtain the following result:

$$\frac{1}{2}Adv^{C_R}_{adversary,GM_3} = \left|Adv^{C_R}_{adversary,GM_0} - \frac{1}{2}\right| = \left|Adv^{C_R}_{adversary,GM_1} - \frac{1}{2}\right| \quad (22)$$

By combining the triangular inequality with Eqs.(3),(4), and (6), we obtain the following result:

$$\begin{aligned}\frac{1}{2}Adv^{C_R}_{adversary} &= \left|Adv^{C_R}_{adversaryGM_1} - Adv^{C_R}_{adversar\ \ GM_3}\right| \\ &\leq \left|Adv^{C_R}_{adversaryGM_1} - Adv^{C_R}_{adversary}GM_2\right| \\ &+ \left|Adv^{C_R}_{adversaryGM_2} - Adv^{C_R}_{adversary}GM_3\right| \\ &\leq \frac{q_h^2}{2|hash|} + \left\{C.q^{DS}_{send}, \frac{q_{DS}}{2^{l_1}}, \frac{q_{DS}}{2^{l_2}}, \frac{q_{DS}}{2^{l_3}}\right\} \\ &+ 2Adv^{DLP}_{adversary} \end{aligned} \quad (23)$$

We obtain the following result:

$$Adv^{C_R}_{adversary} \leq \frac{q_h^2}{|hash|} + \left\{C.q^{DS}_{send}, \frac{q_{DS}}{2^{l_1}}, \frac{q_{DS}}{2^{l_2}}, \frac{q_{DS}}{2^{l_3}}\right\} + \varepsilon(.) \quad (24)$$

Finally, we prove the semantic security of the proposed protocol using the RoR model. □

5.4 Comparative analysis

Most of the customers have limited resources. Hence, it is very important to have a realistic performance analysis in the security protocols. In this section, we give a performance analysis of DF3A in terms of store requirement, computational cost, and communication overhead to demonstrate its practicality in real-world scenarios. in this section, we demonstrate the detailed comparative analysis of DF3A with related scheme's M. Ibrahim et al. [24] in terms of "security functionalities", "communication overhead", "computation costs", and "Storage requirement".

Communication overhead: We calculate the total bits transmitted during the communication to compare the communication cost of M. Ibrahim et al. [24]. We assumed that $|A_s| = 10$, $|A_u| = 10$, $|K_u| = 10$, $n = 100$ and size of message=M=10B. These parameters are very large compared to elliptic curve private key of 160 bits for q and 512 bits for p in our basic and extended schemes. The total communication cost of the existing and proposed schemes is as presented in Table 3.

Table 3. Communication overhead.

DF3A	M. Ibrahim et al. [24]	Protocol
$2(\lvert A_u\rvert + 2)\lvert p\rvert = 12288\ bits$	$2(\lvert A_u\rvert + 1)\lvert p\rvert + \lvert q\rvert = 11424\ bits$	**RA → R**
$2(2\lvert p\rvert) = 2048\ bits$	_	$\boldsymbol{\delta}$ **→ RA**
_	$2\lvert p\rvert + \lvert q\rvert = 1184\ bits$	**R → RA**
_	$\lvert p\rvert\ per\ user = 51200\ bits$	**RA →** $\boldsymbol{\delta}$
$2(\lvert A_u\rvert)\lvert p\rvert + \lvert q\rvert = 10400\ bits$	$\lvert m\rvert + (2\lvert A_u\rvert + 5)\lvert p\rvert + \lvert q\rvert = m + 12960\ bits$	**R →** $\boldsymbol{\delta}$
$\lvert q\rvert = 160\ bits$	_	$\boldsymbol{\delta}$ **→ R**
$24{,}896\ bits$	$M + 76{,}768\ bits = 76{,}848\ bits$	Sum

Computation overhead: We assumed that $|A_s| = 10$, $|A_u| = 10$, $|K_u| = 10$ and $n = 100$.

Table 4. Computation cost of different cryptographic operations mapped to the computation time of one-way hashing operations.

Symbol	Description	Cost
T_h	Execution time of one hash invocation	$\mathbf{T_h}$
T_{mm}	Execution time of one modular multiplication	$\mathbf{2.5T_h}$
T_{ma}	Execution time of one modular addition	$\mathbf{0.3T_h}$
T_{exp}	Execution time of one modular exponentiation	$\mathbf{600T_h}$
T_{inv}	Execution time of one modular inverse	$\mathbf{200T_h}$
T_{ecsm}	Execution time of one scalar multiplication	$\mathbf{72.5T_h}$
T_{pa}	Execution time of one point addition	$\mathbf{13T_h}$
T_{pair}	Execution time of one elliptic curve pairing	$\mathbf{1550T_h}$

Table 5. Concrete evaluation of the computation cost mapping to the time taken by one SHA-1 hash operation as the time unit.

DF3A	Ibrahim et al [24]	Protocol
$(2+\|A_u\|)T_{ecsm} = (2+\|A_u\|)72.5T_h = 870T_h$	$2(T_{mm}) + 2(T_{ma}) + 5(T_{exp}) + \|A_u\|(T_{ecsm}) = 3005.5T_h + (\|A_u\|)72.5T_h = 3{,}731T_h$	**R**
$2(T_{ecsm}) = 145T_h$	$K_u(T_{mm}) + \big((4+K_u)(T_{exp})\big) + K_u(T_{pair}) = K_u(2.5T_h) + \big((4+K_u)(600T_h)\big) + K_u(1550T_h) = 23{,}925T_h$	$\boldsymbol{\delta}$
$(2+\|A_u\|)T_{ecsm} = (2+\|A_u\|)72.5T_h = 870T_h$	$\big((4+K_u)\|A_u\|)(T_{mm})\big) + ((\|A_u\|(K_u - 1)(T_{ma})) + 4(T_{exp}) + \|A_u\|(T_{ecsm}) = \big((4+K_u)\|A_u\|)(2.5T_h)\big) + ((\|A_u\|(K_u - 1)(0.3T_h)) + 4(600T_h) + \|A_u\|(72.5T_h) = 3{,}502T_h$	RA
$870T_h + 145T_h + 870T_h = 1{,}885T_h$	$3730.5T_h + 23925T_h + 3502T_h = 31{,}158T_h$	Sum

Storage overhead: The Storage required by the registration authority $RA = (|\,Au\,|)\,|\,Q\,| + n\,|\,p\,|$ bits, customer $S = (n + 2|AS|)|p|$ bits and Storage required by the customer or user $U = |q| + 2(|AU| + 1)|p|$ bits.

The customer DF3A protocol stores two series values s_0 , s_1 and two points S_0 , S_1 in its Storage, it stores three points of customer τ_{A_i} ، SID_R and C'_R in its Storage. The node and the registration authority keep the two points S_0 , S_1 sent by the customer. Detailed calculation of these values is given in Table 7.

Table 6. Storage requirement symbols.

Size of Customer's attribute	|q| bit
Customer's ID	|p| bit
output Mix-net Size of	|p| bit
The public key of each attribute	2|p| bit

Attribute of the customer	2|p| bit
The number of customer attributes	|Au|
The number of customer attributes	|As|

The following assumptions are used for non-parametric comparison of two protocols.

$$|A_s| = 10\ ;\ |A_u| = 10\ ;\ |K_u| = 10\ ;\ n = 100$$

Table 7. Storage overheads.

Protocol	δ	R	RA	Sum
Ibrahim et al [25]	$(n + 2\|AS\|)\|p\| = 120\|p\|$	$\|q\| + 2(\|AU\| + 1)\|p\| = \|q\| + 22\|p\|$	$\|AU\|\|q\| + n\|p\| = 10\|q\| + 100\|p\|$	61440+11424+52800=125,664
DF3A	$2(2\|p\|) + 2\|q\| = 4\|p\| + 2\|q\|$	$2(2\|p\|) + 2\|AU\|\|p\| = 24\|p\|$	$2(2\|p\|) = 4\|p\|$	2048+12288+4344=18,680

6 Conclusion

In this paper, we presented the design of a Data Fabric with Attribute-based Authentication and Access Control (DF3A), featuring a dedicated layer for access control and authentication. Inspired by the attribute-based authentication protocol of Ibrahim et al., our proposed architecture overcomes the security vulnerabilities identified in their scheme. DF3A ensures essential security requirements such as mutual authentication, anonymity, untraceability, and enhanced security capabilities. By relying exclusively on elliptic curve mathematics and the hardness of the elliptic curve discrete logarithm problem, the architecture achieves lightweight and efficient computational overhead, making it particularly well-suited for resource-constrained environments, such as the Internet of Things (IoT).

In DF3A, tokens are generated for each customer based on their attributes, enabling mutual authentication where the customer authenticates the data store, and the data store authenticates the customer. The proposed architecture demonstrates resilience against various security threats and has undergone thorough analysis. Formal verification using the Real-or-Random (RoR) model confirmed session key security, while informal analyses established resistance to attacks such as impersonation, customer collaboration, desynchronization, and tracking or tracing. Moreover, DF3A ensures confidentiality, availability, anonymity, untraceability, and perfect forward secrecy. With its strong security guarantees, low computational costs, energy efficiency, and economic feasibility, DF3A offers a sustainable and practical solution for modern data fabric applications.

References

[1] E. Hechler, M. Weihrauch, and Y. Wu, *Data Fabric and Data Mesh Research Areas,* Data Fabric and Data Mesh Approaches with AI: A Guide to AI-based Data Cataloging, Governance, Integration, Orchestration, and Consumption, Berkeley, CA: Apress, 2023, pp. 375-392.

[2] S. Hu, S. Jiang, Q. Miao, F. Yang, W. Zhou, and P. Duan, *Provably secure ECC-based anonymous authentication and key agreement for IoT,* Appl. Sci., vol. 14, no. 8, p. 3187, 2024.

[3] K. Liu, M. Yang, X. Li, K. Zhang, X. Xia, and H. Yan, *M-data-fabric: A data fabric system based on metadata,* 2022 IEEE 5th International Conference on Big Data and Artificial Intelligence (BDAI), 2022, pp. 57-62.

[4] M. Castelluccio, *Data fabric architecture,* Strategic Finance, vol. 103, no. 4, pp. 57-58, 2021.

[5] W. Yu and S. Wang, *Attribute-Based Authentication Scheme from Partial Encryption for Lattice with Short Key,* CMC-Comput. Mater. Continua, vol. 75, no. 1, pp. 67-80, 2023.

[6] S. Meel and S. Ulukus, *HetDAPAC: Distributed Attribute-Based Private Access Control with Heterogeneous Attributes,* arXiv preprint arXiv:2401.13653, 2024.

[7] H. Yang and V. A. Oleshchuk, *A dynamic attribute-based authentication scheme,* Codes, Cryptology, and Information Security: First International Conference, C2SI 2015, Rabat, Morocco, May 26-28, 2015, Proceedings-In Honor of Thierry Berger 1, Springer International Publishing, 2015, pp. 106-118.

[8] Q. Zhang, Y. Mu, and M. Zhang, *Attribute-based authentication for multi-agent systems with dynamic groups,* Comput. Commun., vol. 34, no. 3, pp. 436-446, 2011.

[9] S. Shukla and S. J. Patel, *A design of provably secure multi-factor ECC-based authentication protocol in multi-server cloud architecture,* Cluster Comput., vol. 27, no. 2, pp. 1559-1580, 2024.

[10] V. Sucasas, G. Mantas, M. Papaioannou, and J. Rodriguez, *Attribute-based pseudonymity for privacy-preserving authentication in cloud services,* IEEE Trans. Cloud Comput., 2021.

[11] W. Huang, *ECC-based three-factor authentication and key agreement scheme for wireless sensor networks,* Sci. Rep., vol. 14, no. 1, p. 1787, 2024.

[12] N. G. Kuftinova, O. I. Maksimychev, A. V. Ostroukh, A. V. Volosova, and E. N. Matukhina, *Data fabric as an effective method of data management in traffic and road systems,* in 2022 Systems of Signals Generating and Processing in the Field of On-Board Communications, 2022, pp. 1-4.

[13] I. Buleje, V. S. Siu, K. Y. Hsieh, N. Hinds, B. Dang, E. Bilal, T. Nguyen, E. E. Lee, C. A. Depp, and J. L. Rogers, *A versatile data fabric for advanced IoT-based remote health monitoring,* 2023 IEEE International Conference on Digital Health (ICDH), 2023, pp. 88-90.

[14] S. A. Rieyan, M. R. Kabir News, A. B. M. M. Rahman, S. A. Khan, S. T. J. Zaarif, M. G. R. Alam, M. M. Hassan, M. Ianni, and G. Fortino, *An advanced data fabric architecture leveraging homomorphic encryption and federated learning,* Inf. Fusion, vol. 102, p. 102004, 2024.

[15] A. Macías, D. Muñoz, E. Navarro, and P. González, *Data fabric and digital twins: An integrated approach for data fusion design and evaluation of pervasive systems,* Inf. Fusion, vol. 103, p. 102139, 2024.

[16] A.-M. Ghiran and R. A. Buchmann, *The model-driven enterprise data fabric: A proposal based on conceptual modelling and knowledge graphs,* International Conference on Knowledge Science, Engineering and Management, Cham: Springer International Publishing, 2019, pp. 572-583.

[17] S.-J. Moon, S.-B. Kang, and B.-J. Park, *A study on a distributed data fabric-based platform in a multi-cloud environment,* Int. J. Adv. Cult. Technol., vol. 9, no. 3, pp. 321-326, 2021.

[18] S. Rostampour, M. Safkhani, Y. Bendavid, and N. Bagheri, *ECCbAP: A secure ECC-based authentication protocol for IoT edge devices,* Pervasive Mobile Comput., vol. 67, p. 101194, 2020.

[19] D. D. Khader, *Attribute based authentication schemes,* PhD diss., University of Bath, 2009.

[20] Q. Zhang, Y. Mu, and M. Zhang, *Attribute-based authentication for multi-agent systems with dynamic groups,* Comput. Commun., vol. 34, no. 3, pp. 436-446, 2011.

[21] L. Guo, C. Zhang, J. Sun, and Y. Fang, *A privacy-preserving attribute-based authentication system for mobile health networks,* IEEE Trans. Mobile Comput., vol. 13, no. 9, pp. 1927-1941, 2013.

[22] H. Yang and V. A. Oleshchuk, *A dynamic attribute-based authentication scheme,* Codes, Cryptology, and Information Security: First International Conference, C2SI 2015, Rabat, Morocco, May 26-28, 2015, Proceedings-In Honor of Thierry Berger 1, Springer International Publishing, 2015, pp. 106-118.

[23] D. Boneh, X. Boyen, and H. Shacham, *Short group signatures,* Annual International Cryptology Conference, Springer Berlin Heidelberg, 2004, pp. 41-55.

[24] M. H. Ibrahim, S. Kumari, A. K. Das, and V. Odelu, *Attribute-based authentication on the cloud for thin customers,* J. Supercomput., vol. 74, pp. 5813-5845, 2018.

[25] S. Pape, *Sample or random security–a security model for segment-based visual cryptography,* in Financial Cryptography and Data Security: 18th International Conference, FC 2014, Christ Church, Barbados, March 3-7, 2014, Revised Selected Papers 18, Springer Berlin Heidelberg, 2014, pp. 291-303.

[26] A. Saini, V. Sharma, M. Kumar, A. Dey, and D. Tripathy, *10 Security, privacy, and authentication framework for web-driven data fabric,* Data Fabric Architectures: Web-Driven Applications, 2023, p. 183.