Passive User Authentication in Industrial Internet of Things

by

Guozhu Zhao

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(The School of Systems Information Science)
in Future University Hakodate
February 2023

To my family

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor Professor Xiaohong Jiang, who is intelligent, knowledgeable, visionary and approachable. It has been a great honor for me to be one of his Ph.D. students and the Ph.D. experience under his supervision is definitely life-changing for me in Future University Hakodate. I appreciate all his contributions of time and ideas that made my Ph.D. experience productive and exciting. He has been teaching me, both consciously and unconsciously, the important skills required to be a good researcher and the great personality traits that make a better man. I would also like to express my heartfelt gratitude to Professor Jiang's wife, Mrs. Li, for her countless care.

Besides my advisor, I would like to thank the rest of my thesis committee: Professor Yuichi Fujino, Professor Hiroshi Inamura, Professor Masaaki Wada, and Professor Shigemi Ishida. Their valuable advice and feedback on my dissertation research help me improve this thesis.

I would also like to give my sincere gratitude to Pinchang Zhang of Nanjing University of Posts and Telecommunications, China, who helped me a lot in improving both the quality and the clarity of dissertation research. He showed me the way to be an excellent researcher.

My sincere thanks also go to other members in our laboratory Wenhao Zhang, Yan Liu, Xinzhe Pi, He Zhu, Jiaqing Bai, and Zhen Jia for their contributions in some way to this thesis.

Last but not the least, I would like to thank my family: my wife Huanhuan Zhao,

eldest son Lingheng Zhao, newborn second son Yinheng Zhao, parents, brother and sister. Words cannot express how grateful I am to them for all of the sacrifices they have made for me.

ABSTRACT

Passive User Authentication in Industrial Internet of Things

by

Guozhu Zhao

Industrial Internet of Things (IIoT) serves as an important network architecture for information collection, exchange and analysis in the industrial platform. An IIoT system usually consists of a vast number of users with highly diverse authority rights and constantly generates/stores huge amounts of confidential information, so how to design flexible and cost-effective authentication approaches to ensure the security of IIoT systems becomes an increasingly urgent demand. Specially, the passive user authentication is of great importance for IIoT systems to implement continuous and non-intrusive user identity verification. The IIoT can be roughly divided into three layers according to the functions of IIoT, i.e., the Manufacturing Execution (ME) layer, Monitoring and Control (MC) layer, and Decision and Optimization (DO) layer. This dissertation develops user authentication schemes corresponding to these three layers to ensure the secure operation of IIoT systems. First, for user authentication of the ME layer, this dissertation explores the common behavioral biometrics from user sequential operation actions in IIoT systems to propose a passive authentication framework, which provides continuous/non-intrusive user authentica-

tion and poses good anti-interference capability in the interference-intensive environment of the ME layer. Second, for user authentication of the MC layer, we explore the user consecutive screen-touch actions during routine work processes and propose a passive authentication method based on both the time-varying characteristics and spatial image characteristics of the user touch trajectory sequences, which provides implicit/non-intrusive user identity verification and can meet the real-time authentication requirement of the MC layer. Finally, for user authentication of the DO layer, we develop a novel two-dimensional passive authentication framework by jointly utilizing both the time-varying characteristics of the user sequential operation actions and spatial variation characteristics of Channel State Information (CSI) caused by these actions, which applies to the authentication of the DO layer with high security requirement. It is expected that the new authentication methods proposed in this dissertation can significantly facilitate the applications of IIoT systems.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

CHAPTER I

Introduction

In this chapter, we first introduce IIoT system and its function architecture. Then we demonstrate the essence of IIoT and the difference between IIoT and IoT. Subsequently, we clarify new security challenges in IIoT and authentication challenges in IIoT scenarios. We further present the objective and main work of this thesis. Finally, we give the outline and main notations of this thesis.

1.1 Overview of IIoT

1.1.1 IIoT System

General Electric (GE) coined the name "Industrial Internet" as their term for the Industrial Internet of Things (IIoT), and others such as Cisco termed it the Internet of Everything and others called it Internet 4.0 or other variants [1]. Generally, IIoT refers to the extension and use of the Internet of Things (IoT) in industrial sectors and applications. With a strong focus on Machine-to-Machine (M2M) communication, big data, cloud computing, and machine learning, IIoT enables industries and enterprises to have better efficiency and reliability in their operations. IIoT encompasses industrial applications, including robotics, medical devices, and software-defined production processes. IIoT serves as an important network architecture for information

collection, exchange, and analysis in the industrial platform. With the rapid merging of Information Technology (IT) and Operational Technology (OT), IIoT becomes highly promising to significantly boost the automation, efficiency, and productivity in the global manufacturing industry. By now, IIoT has been widely employed in some critical industrial applications like the automotive industry, oil/gas industry, healthcare, energy production, and agriculture industry [1–3].

IIoT provides a way to get better visibility and insight into the company's operations and assets through the integration of machine sensors, middleware, software, backend cloud computing, and storage systems. Therefore, it provides a method of transforming operational business processes by using the results gained from interrogating large data sets through advanced analytics as feedback. The business gains are achieved through the improvement of operational efficiency and accelerated productivity, which results in reduced unplanned downtime and optimized efficiency, and thereby profits. Although the technologies and techniques used in today's industrial environments may look similar to IIoT, the scale of operation is vastly different. For example, huge data streams can be analyzed online using cloud-hosted advanced analytics at wire speed. Additionally, vast quantities of data can be stored in distributed cloud storage systems for future analytics performed in batch formats. These massive batch job analytics can glean information and statistics from data that would never have previously been possible because of the relatively tiny sampling pools. Process engineers can then use the results of the analytics to optimize operations and provide the information that executives can transform to knowledge, in order to boost productivity and efficiency, and to reduce operational costs [1, 4].

1.1.2 IIoT Function Layer Architecture

The core functional principle of IIoT is based on the comprehensive interconnection and deep collaboration between the physical system and the digital space driven

by data, as well as intelligent analysis and decision-making optimization in the process. As shown in Fig. 1.1, the function layer architecture of IIoT mainly includes three basic layers of Manufacturing Execution (ME) layer, Monitoring and Control (MC) layer, and Decision and Optimization (DO) layer, as well as a closed loop of industrial digital application optimization consisting of bottom-up information flow and top-down decision-making flow [5–10].

The ME layer usually covers industrial production, manufacturing, and assembly of IIoT systems. Common application cases of the ME layer are the production and manufacture of high-precision electronic components, automatic assembly and production workshops of automobiles, and production lines of smart home appliances. In the ME layer, due to the needs of product production and manufacturing, a large amount of alternating current, motor equipment with changing strong magnetic fields, and cross coverage of various wireless signals generate more electromagnetic interference. Therefore, the ME layer is usually accompanied by a large amount of electromagnetic interference. The MC layer often involves critical human-computer interaction, access control, command transmission, and data exchange in IIoT system, where there are a large number of important real-time instruction uploading and downloading, user access control, and manufacturing process monitoring. Examples of typical applications for the MC layer include industrial APP access and interaction, robot systems, manufacturing modeling, and identification analysis system. The DO layer mainly covers decision making, optimization, description, diagnosis, business operations, and management. In highly intelligent IIoT systems, artificial intelligence platforms, and intelligent decision-making systems are usually deployed on the DO layer. In the DO layer, a large amount of confidential information and sensitive data (such as finance, core technology, core algorithms, operation and sales strategies, crucial customer information, and key management technical services) are generated, stored, and exchanged.

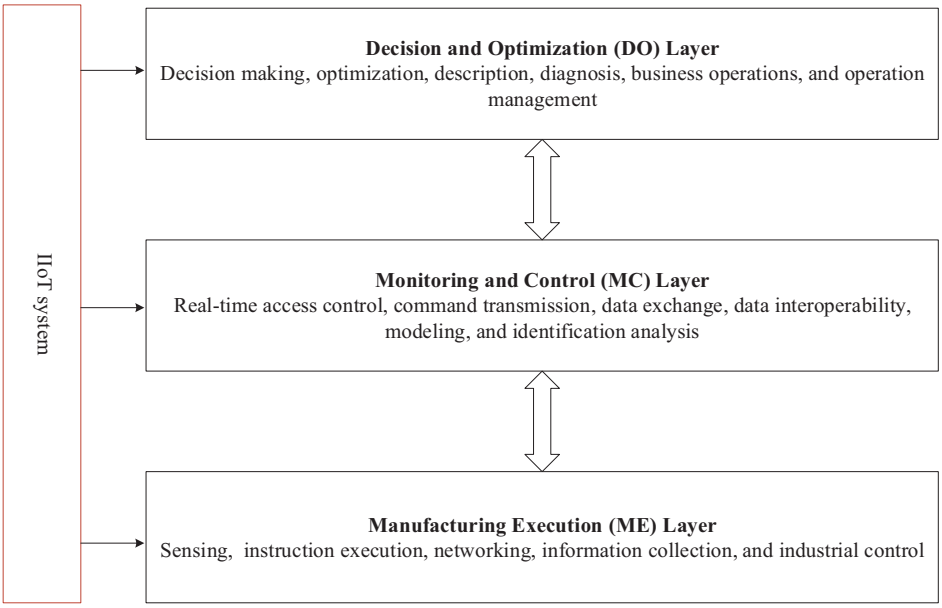


Figure 1.1: IIoT function layer architecture.

1.2 IIoT vs. IoT

To illustrate the difference between IoT and IIoT, we first introduce the definition of IoT. According to [11], IoT comprises large numbers of smart devices at the network edge that may have to collaborate and interact with each other in real time. In [12], the authors define IoT as an environment in which objects (devices) are given unique identifiers and the ability to transfer data over a network without having human-to-human or human-to-computer interaction. From another view in [13], IoT could be specified as a worldwide network of interconnected entities. As stated in [14], IoT is an ecosystem that interconnects physical objects with telecommunication networks, joining the real world with the cyberspace and enabling the development of new kinds of services and applications. Combining with the definition of IIoT mentioned in Section 1.1, we can see that IIoT is generally different from IoT in definition and function. Then, as shown in Table 1.1 we clarify the characteristics of IoT and IIoT in terms of components, applications, and security challenges. From the perspective

of components, IIoT is more complex than IoT, and consists of machine sensors, middleware, APP, backend cloud technology, SDN, IoT, and storage technologies. However, IoT generally serves as the foundation and supporting technology to provide a platform for other applications. In terms of application scenarios, IIoT puts more emphasis on industrial manufacturing, industrial chain, manufacturing ecosystem, and some critical industrial applications, while IoT focuses on the unique identity of the device, network interconnection, and human-to-human or human-to-computer interaction. Finally, we can see from Table 1.1 that there are new security challenges faced by IIoT compared with IoT.

Table 1.1: IIoT and IoT

Networks	**Components**	**Applications**	**Security challenges**
IoT	Sensors, devices, connectivity, data processing, user interfaces, software platforms	Transportation, medical care, smart grid, smart city, logistics, monitoring, unique identifiers	Poor visibility, limited security integration, open-source code vulnerabilities, overwhelming data volume, poor testing, unpatched vulnerabilities, vulnerable APIs, weak passwords
IIoT	IT (information technology), OT (operational technology), sensors, middleware, IoT, backend cloud, big data, SDN, industrial APP	Industrial manufacturing, industrial chain, manufacturing ecosystem, automotive industry, oil/gas industry, healthcare, energy production, agriculture industry	High security, real-time requirement, high anti-interference capability, passive authentication

Specifically, the current IIoT systems are facing various security challenges, both from their inherited IoT architecture and their own properties [15–18]. First, in industrial production sites, IIoT uses Industrial Control System (ICS) to collect, process, and analyze local data and resources. However, the current IIoT systems lack effective authentication and security mechanisms (e.g., authentication with high anti-interference capability) for the ICS. Second, in Cyber-Physical Systems (CPS), IIoT connects physical systems (hardware), software systems, and various types of systems through gateways. However, the current IIoT systems lack efficient gateways and real-time cross-layer security protocols. Third, IIoT systems process and store data in the cloud platform. However, the current IIoT systems lack effective protocols, frameworks, and algorithms to protect data security in heterogeneous cloud platform scenarios. Finally, IIoT often involves a large number of users, terminal devices, and industrial applications. The current IIoT lacks security mechanisms (e.g., to ensure the identity of a user interacting with IIoT is not impersonated) for satisfying different performance requirements across various IIoT scenarios to ensure the safe operation of the IIoT systems [15, 18].

1.3 User Authentication in IIoT

1.3.1 Authentication in IIoT

With the rapid merging of IoT, big data, and cloud computing technologies, IIoT becomes highly promising to boost the platform-based design, intelligent manufacturing, and networked collaboration in the global manufacturing industry [19–21]. Notice that the IIoT systems usually focus on critical industrial fields like the automotive industry, smart transport, medical care, and agriculture, so the security guarantee is of great importance for the secure operations of such systems [22–24]. Among these aforementioned security challenges, user authentication serves as a critical one since

such systems usually involve a large number of users (both for line operators and product designers) with highly diverse authority rights [25].

An IIoT system is generally a cloud-based system. The user authentication in the IIoT is usually encapsulated as a cloud service to verify the identities of cloud users who are attempting to access the system and thus to prevent unauthorized users from accessing to the sensitive information of the system [26, 27]. Depending on whether a user actively participates in the authentication process or not, the user authentication in IIoT systems can be roughly classified as active authentication and passive one. Active authentication usually requires some specified actions from a user to be authenticated (e.g., entering passwords and providing fingerprints), and it is commonly applied to the one-time authentication scenarios where the continuous monitoring of user legitimacy is not necessary once the user is successfully authenticated as a legitimate one [28–31]. In contrast, the passive authentication mainly explores the intrinsic properties related to the user inherent activities and behaviors to carry out user authentication, so it does not need additional actions from a user for authentication purpose and thus is highly appealing for the continuous user identity verification.

1.3.2 Why Passive Authentication in IIoT

Notice that IIoT systems usually involve a large number of users, where each user needs to manipulate multiple devices and perform many critical operations on these devices during his routine work process. In particular, to ensure the secure operation of an IIoT system, a user needs to frequently conduct user identity authentication whenever the user accesses his devices and whenever the user performs a critical operation. The conventional user authentication in IIoT systems usually adopts the active authentication methods like the pin-based or pattern-based authentication [32, 33]. Such active user authentication methods are suitable for entry-point or one-time au-

thentication since they require additional operation actions for authentication purpose (e.g., inputting password or ID information), but they are unrealistic for the frequent and non-intrusive authentication in IIoT systems where users are busy with their routine work operations and can hardly find time to frequently conduct additional operations for authentication. On the other hand, the passive authentication methods determine the identity of a user by exploiting his intrinsic behavioral traits during the routine work process, so they do not need additional operations from the user for authentication purpose and thus are highly appealing for frequent user authentication in the practical IIoT environments [34, 35]. Therefore, passive user authentication is of great importance for security guarantee in IIoT systems.

1.3.3 Challenges for Passive Authentication in IIoT

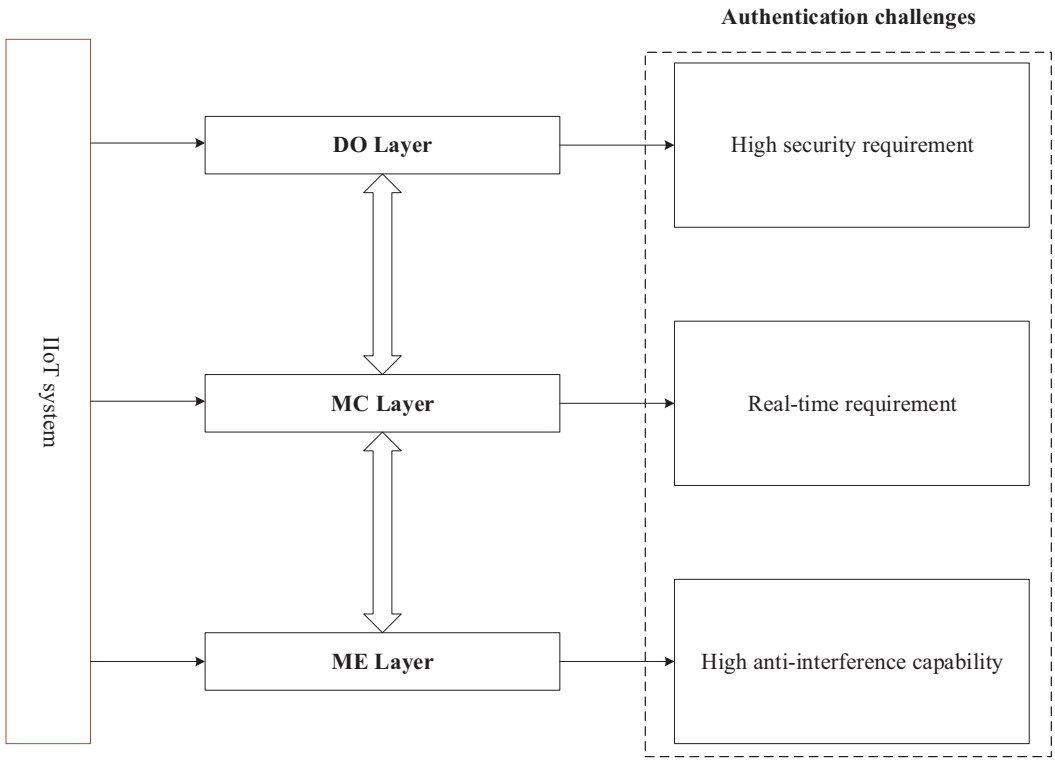


Figure 1.2: Challenges for passive authentication in IIoT.

In the actual operation process of IIoT, the ME layer, the MC layer, and the DO layer complement each other to form a bottom-up information flow and a top-down decision flow. As a result, different layers play their own unique roles while also facing different authentication challenges. As shown in Fig. 1.2, for the ME layer, a large amount of alternating current, motor equipment with changing strong magnetic fields, and cross coverage of various wireless signals generate more electromagnetic interference. So the user authentication in such layer requires the user authentication protocol to have better anti-interference ability. However, existing research efforts [36–41] lack effective interference removal or noise reduction methods, and the use of the motion characteristics of single sensor cannot effectively characterize users' identities in the presence of interference. For the MC layer, critical human-computer interaction, access control, command transmission, and data exchange are involved in IIoT systems. Therefore, the systems urgently need authentication protocols or frameworks with better real-time characteristics. Existing user identity authentication methods using human-computer interaction behaviors or behavioral biometric features (such as keystroke patterns-based authentication [42–44], gait-based authentication [45, 46], speaking-based authentication [47, 48], and touch-based authentication [35, 49]) have a large delay and consume more computing and storage resources in practical industrial application. For the DO layer, the core industrial applications or services related to the Industrial Internet generate, store, and exchange extensive sensitive data. Hence, an authentication protocol with high security performance is required to protect sensitive data in the DO layer. But, available authentication approaches [26, 32–35] are usually difficult to accurately depict the user identities and thus to achieve an acceptable user authentication performance based on only one-dimensional characteristics.

1.4 Objectives and Main Contributions

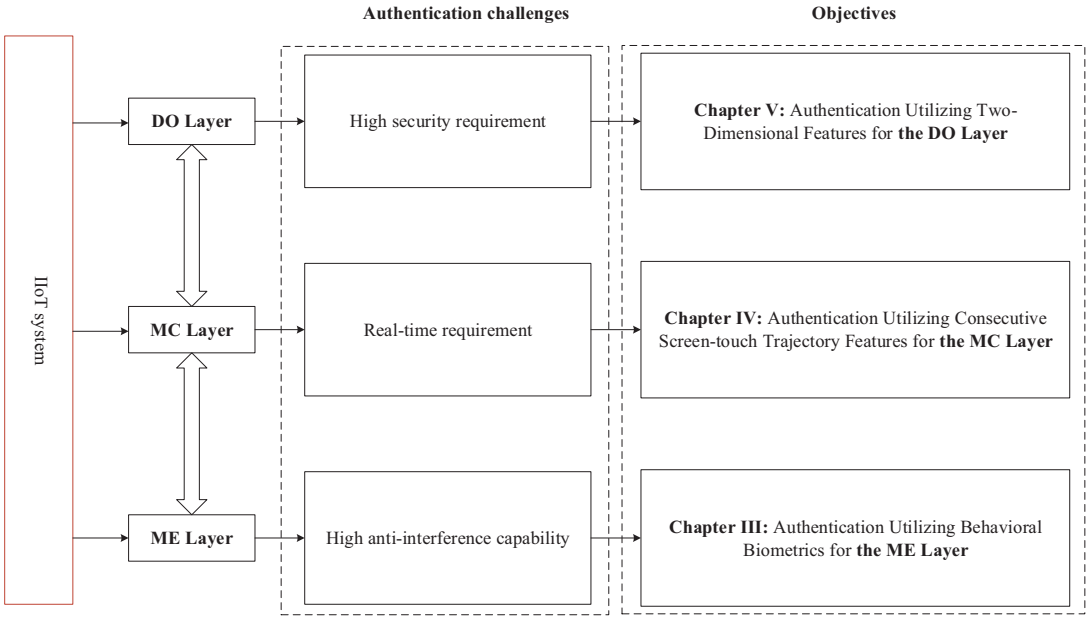


Figure 1.3: Objectives and main contributions of this thesis.

Corresponding to authentication challenges in the three layers of IIoT systems, we develop three schemes to satisfy these different authentication performance requirements across various IIoT scenarios, as shown in Fig. 1.3. For IIoT authentication requirements of high anti-interference capability in the ME layer, we first explore the common behavioral biometrics from sequential operation actions in IIoT systems to propose a passive authentication framework for continuous and non-intrusive user authentication against the impersonation attack. In the authentication scheme, we leverage the Kalman filtering and Wavelet techniques for noise elimination and the singular value decomposition method for the dimensionality reduction of feature space. For IIoT authentication requirements of high real-time performance in the MC layer, we propose a new user authentication framework based on the spatial-temporal features of screen-touch trajectories for continuous user authentication in practical IIoT scenarios by designing two classifiers corresponding to the spatial-temporal screen-

touch trajectory features and assigning each classifier an appropriate weight. In the authentication scheme, every time a user touches the screen, the IIoT authentication systems can verify the user's identity by analyzing the time-varying features of touch trajectory sequences and cumulative Screen-touch Trajectory Images (STTIs) characteristics, thus ensuring the real-time user authentication. For IIoT authentication requirements of high security performance in the DO layer, we develop a novel two-dimensional passive authentication framework by jointly utilizing both the time-varying characteristics of the user sequential operation actions and spatial variation characteristics of Channel State Information (CSI) caused by these actions. In the authentication scheme, by jointly exploiting the two-dimensional features of user sequential operation actions, we can not only provide a full spatial-temporal characterization of user identities but also significantly improve the security of the proposed passive user authentication.

Three commonly-used authentication performance metrics are of particular interest for authentication performance evaluation. They are false acceptance rate (FAR), false rejection rate (FRR), and equal-error rate (EER) [26, 50, 51]. The FAR denotes the ratio between the number of false acceptances and that of test samples from attackers, FRR denotes the ratio between the number of false rejections and that of test samples from legitimate users, and EER represents the sensitivity of the proposed authentication approach at the point where FAR = FRR [51]. We also adopt the authentication accuracy to evaluate the performance for resisting the impersonation attacks of the proposed framework, here the authentication accuracy is defined as the probability that the system successfully distinguishes between the legitimate users and impersonation attacks. The main contributions of this thesis are summarized in the following subsections.

1.4.1 Authentication Utilizing Behavioral Biometrics for the Manufacturing Execution (ME) Layer

It is demonstrated that the passive authentication mainly exploits the intrinsic behavioral traits of a user during the routine work process to determine its identity. By now, some research efforts have been devoted to the study of passive user authentication based on single user behavioral characteristic [36–41]. It is worth noting that the existing user authentication based on single action characteristic cannot be directly extended to the IIoT systems. On one hand, in industrial production process users are always required to operate mobile devices in a uniform and standardized manner, so relying on single behavioral characteristic can not ensure the uniqueness of user identity. On the other hand, in the IIoT environment there are often some specific user dress rules such as having to wear protective gloves or clothes, which results in a low discriminability of behavioral features and thus an inefficient authentication performance. In addition, the available single characteristic-based passive authentication solutions are usually sensitive to specific noise and interference [32], which makes them unsuitable for the complex IIoT systems suffering from intensive noise and interference (e.g., gravity components and non-stationary noise).

Based on this background, this work explores the common behavioral biometrics from sequential operation actions in IIoT systems and develops a multiple characteristics-based passive authentication framework for continuous and non-intrusive user identity verification. The main contributions of this work are summarized as follows:

- We first provide extensive experiment results to demonstrate that in IIoT systems the common behavioral biometrics from sequential user operation actions (i.e., walking, scanning, screen-touch, and photographing-uploading) exhibit good discriminability and stability in discriminating user identities.

- We then develop a theoretical framework for characterizing the intrinsic features

of sequential operation actions. In particular, we employ the Kalman filtering and Wavelet techniques to reduce the noise in sensor signals of user operation actions, and apply the singular value decomposition method to achieve the dimensionality reduction for the feature space of sequential operation actions.

- By modeling the transitions of operation actions as a Markov chain and applying the one-class classification technique for user classification, we develop a passive user authentication framework for continuous and non-intrusive user identity verification against the impersonation attack.

- Finally, experiment results are provided to illustrate the authentication performance of the proposed authentication framework in terms of the false acceptance, false rejection and equal-error rates. The related authentication efficiency issues, like the usability to the operation-action sequence length, the scalability to the number of features and user space, and the sensitivity to the operation action features, are also investigated.

1.4.2 Authentication Utilizing Consecutive Touch Trajectory Features for the Monitoring and Control (MC) Layer

Behavioral biometric user authentication solutions have been widely used for smart mobile terminals (MTs), such as keystroke patterns-based authentication [42–44], gait-based authentication [45, 46], speaking-based authentication [47, 48], and touch-based authentication [35, 49]. Among these behavioral biometric user authentication solutions, touch-based authentication is of particular interest for implementing the continuous and non-intrusive user identity verification. To the best of our knowledge, the existing literature only focuses on either screen-touch time domain characteristics of single trajectory [52], or spatial motion behavior characteristics of touch actions [26, 34]. To further improve the performance of touch-based authentication solutions for

the special IIoT environment, we develop the spatial-temporal screen-touch trajectory characteristics of touch trajectories in both time variation and space combinatorial distribution. The main contributions of this work are summarized as follows:

- By exploiting the time-varying nature of user screen-touch sequences from the routine work process of a user and applying the Hidden Markov Model (HMM) to characterize behavioral biometric characteristics of the user, we develop a new method to characterize the screen-touch-based behavioral biometric characteristics of users in IIoT scenarios.

- We then develop a theoretical framework based on Least Squares Polynomial Fit for characterizing spatial features of the STTIs from the user routine sequential touch trajectories. In particular, we successively reconstruct each touch trajectory of a screen-touch operation action sequence in an image to maintain the shape, relative position and length of the touch trajectory, and adopt average pressure, average curvature and average deviation degree of the trajectory to depict its RGB color in the image. We further apply the Speeded Up Robust Features (SURF) algorithm to extract user STTI features.

- We further design two classifiers based on HMM and eXtreme Gradient Boosting (XGBoost), respectively, corresponding to the above spatial-temporal screen-touch trajectory characteristics. By weighing outputs of these two classifiers, we thus develop a novel user authentication framework for continuous user authentication in various IIoT scenarios.

- We conduct extensive experiments to illustrate the authentication performance of the continuous authentication framework in terms of false acceptance rate, false rejection rate and equal-error rate. We also investigate the related authentication efficiency issues in terms of the usability to weights of two classifiers,

the operation-action sequence length, and the scalability to the number of user space.

1.4.3 Authentication Utilizing Two-Dimensional Features for the Decision and Optimization (DO) Layer

It is notable that when applying existing available passive authentication approaches in modern IIoT systems, it is usually difficult to accurately depict the user identities and thus to achieve an acceptable user authentication performance based on only one-dimensional characteristics. First, users in IIoT systems usually just follow the requirements of industrial production businesses to conduct some basic operations over their MTs in a standardized manner, making it difficult to accurately characterize the user identities with only the time-varying behavioral biometric features extracted from their operation actions. Second, the IIoT systems share a relatively uniform electromagnetic and space environment, so users there show a strong location correlation and thus a low discriminability in terms of the CSI spatial variation characteristics [53, 54]. However, our results in this work indicate that by jointly exploiting the two-dimensional features of the time-varying characteristics of user sequential operation actions and spatial variation characteristics of CSI caused by these actions, we can not only provide a full spatial-temporal characterization of user identities but also significantly improve the performance of passive user authentication. The main contributions of this work are summarized as follows:

- By constructing time-varying operation action sequences from the routine work process of a user and adopting the HMM to model these sequences, we develop a new method to characterize the behavioral biometric characteristics of users in IIoT scenarios.

- We then propose a new approach to depict the spatial-temporal variations of

CSI related to a user, in which the WiFi CSI data related to the user is first sliced to reduce the noise and interference from the random actions of the user, then the multi-domain features from the CSI data are extracted and the XGBoost model is applied to characterize these features.

- We further design two classifiers corresponding to the above two characteristics. By combining these two classifiers and assigning each classifier an appropriate weight, we thus develop a novel two-dimensional user authentication framework for passive, continuous and non-intrusive user authentication in IIoT scenarios.

- We conduct extensive experiments to evaluate the performance of the proposed authentication framework in terms of false acceptance rate, false rejection rate and equal error rate, and also examine the related authentication efficiency issues such as the sensitivity to the weights for classifiers, the sensitivity to authentication time and the capability of resisting against impersonation attacks.

1.5 Thesis Outline

The remainder of this thesis is outlined as follows. Chapter II introduces the related works of this thesis. In Chapter III, we focus on the common behavioral biometrics from sequential operation actions in IIoT systems and develop a multiple characteristics-based passive authentication framework for continuous and non-intrusive user identity verification. In Chapter IV we explore touch-based features of time-varying screen-touch trajectory sequences and cumulative consecutive screen-touch trajectory images from user touch actions during routine work processes in IIoT systems for passive authentication and in Chapter V we develop a novel two-dimensional user authentication framework for passive, continuous and non-intrusive user authentication by jointly exploiting the two-dimensional features of the time-varying characteristics of user sequential operation actions and spatial variation char-

acteristics of CSI caused by these actions. Finally, we conclude this thesis in Chapter VI.

1.6 Notations

The main notations of this thesis are summarized in Table 1.2.

Table 1.2: Main notations

Symbol	Definition
IIoT	Industrial Internet of Things
ROE/ME	R&D office environments/common manufacturing environments
Glo	Manufacturing environments requiring to wear protective gloves
FPC	Manufacturing environments requiring to wear full-body protective clothing
IT	Information technology
OT	Operational technology
MTs	Smart mobile terminals
SVD	Singular value decomposition
HMM	Hidden Markov Model
WIs	Work Instructions
STTIs	Cumulative Screen-touch Trajectory Images
XGBoost	eXtreme Gradient Boosting
SURF	Speeded Up Robust Features
CSI	Channel State Information
OAS	A sequence of successive operation actions collected from his routine work process

$\Upsilon_{Az}/\Upsilon_{Pi}/\Upsilon_{Ro}$	The orientation sensor values of Azimuth, Pitch, and Roll, respectively
B_{wal}	Matrix for characterizing the user's walking operation features
B_{sca}	Matrix for characterizing the user's scanning operation features
B_{tou}	Matrix for characterizing the user's screen-touch operation features
B_{pu}	Matrix for characterizing the user's photographing-uploading operation features
F	The set of a user's operation-action features
λ	The HMM
O	Observed operation-action feature sequence in HMM
X	An unknown user
U	The identity of the unknown user
FAR	False acceptance rate
FRR	False rejection rate
EER	Equal-error rate
§	The test sample set
$\mathbb{S}_{n\times m}$	The training sample space
$MMD(\cdot)$	Maximum Mean Discrepancy
χ	The screen-touch trajectory sequences dataset
ω_1, ω_2	Two weights of classifiers
φ	A preset threshold for authentication decision in the IIoT systems
ϖ_1	User claim is true (a legitimate user)
ϖ_2	User claim is false (an impostor)
O_L	An OSA of length L

R_{Bio}	A one-versus-all multi-class classifier based on behavioral biometric features of users
R_{CSI}	A one-versus-all multi-class classifier based on CSI features of users

CHAPTER II

Related Works

This chapter introduces the existing works related to our study of the thesis, including active user authentication and passive user authentication solutions. User authentication is the process of verifying user identity who is attempting to access the IIoT services, which can be used to prevent the unauthorized users from gaining access to sensitive information and thus to ensure the security of IIoT systems [55, 56]. Depending on whether a user actively participates in the authentication process or not, the user authentication in IIoT systems can be roughly classified as the active authentication or passive authentication [36, 57–60]. In the former case, some specific actions (like entering passwords or presenting credentials) from the user are required for authentication, so such a method is usually adopted for one-time authentication at entry point [52]. In contrast, the passive authentication mainly exploits the intrinsic behavioral traits of a user during the routine work process to determine its identity, so such a method can be used for continuous and non-intrusive authentication [57].

2.1 Active User Authentication

The main idea of active authentication is that the authentication system requires the users to actively participate in the authentication process. Active user authentication methods are suitable for entry-point or one-time authentication since they require

additional operation actions for authentication purpose (e.g., inputting password or ID information). Security and privacy protection are important for the widespread use of IoT in transportation and logistics, since many vehicle drivers are worried about information leakage and privacy invasion. Reasonable efforts in technology, law, and regulation are needed to prevent unauthorized access to or disclosure of the privacy data [22]. Behavioral biometric user authentication solutions have been widely used for MTs, such as keystroke patterns-based authentication [42–44], gait-based authentication [45, 46], speaking-based authentication [47, 48], and touch-based authentication [35, 49]. The authors in [29] propose a novel system for protecting fingerprint privacy by combining two different fingerprints into a new identity. A two-stage fingerprint matching process is proposed for matching the two query fingerprints against a combined minutiae template. By storing the combined minutiae template, the complete minutiae feature of a single fingerprint will not be compromised when the database is stolen. Furthermore, because of the similarity in topology, it is difficult for the attacker to distinguish a combined minutiae template from the original minutiae templates. The authors in [28] investigate the possibility of generating a 'MasterPrint', a synthetic or real partial fingerprint that serendipitously matches one or more of the stored templates for a significant number of users. Their preliminary results on an optical fingerprint data set and a capacitive fingerprint data set indicate that it is indeed possible to locate or generate partial fingerprints that can be used to impersonate a large number of users, which thus expose a potential vulnerability of partial fingerprint-based authentication systems, especially when multiple impressions are enrolled per finger. In [30], the authors propose an approach to counter replay attacks for face recognition on smart consumer devices using a noninvasive challenge and response technique. The image on the screen creates the challenge, and the dynamic reflection from the person's face as they look at the screen forms the response. The sequence of screen images and their associated reflections digitally

watermarks the video. By extracting the features from the reflection region, it is possible to determine if the reflection matches the sequence of images that were displayed on the screen. In [31], the authors adopt iris biometrics for unconstrained user authentication using hand-held devices such as smartphones. The analyses presented in their work indicate that a similar camera module with improved optics and sensors could combine iris biometrics with conventional front camera functions such as video calls and the capture of selfie images. The authors in [42] study the performance of Long Short-Term Memory (LSTM) networks for keystroke biometric authentication at a large scale in free-text scenarios. For this they explore the performance of LSTM networks trained with a moderate number of keystrokes per-identity and evaluated under different scenarios. The proposed approach in their work achieves state-of-the-art keystroke biometric authentication performance with an Equal Error Rate of 2.2% and 9.2% for physical and touchscreen keyboards, respectively, significantly outperforming previous approaches.

2.2 Passive User Authentication

Notice that IIoT systems usually involve a large number of users, where each user needs to manipulate multiple devices and perform many critical operations on these devices during his routine work process. In particular, to ensure the secure operation of an IIoT system, a user needs to frequently conduct user identity authentication whenever the user accesses his devices and whenever the user performs a critical operation. The conventional user authentication in IIoT systems usually adopts the active authentication methods like the pin-based or pattern-based authentication [32, 33]. Such active user authentication methods are suitable for entry-point or one-time authentication since they require additional operation actions for authentication purpose (e.g., inputting password or ID information), but they are unrealistic for the frequent and non-intrusive authentication in IIoT systems where users are busy with their

routine work operations and can hardly find time to frequently conduct additional operations for authentication. On the other hand, the passive authentication methods determine the identity of a user by exploiting his intrinsic behavioral traits during the routine work process, so they do not need additional operations from the user for authentication purpose and thus are highly appealing for frequent user authentication in the practical IIoT environments [34, 35].

By now, some research efforts have been devoted to the study of passive user authentication [36–41]. In [61–64] Single sign-on (SSO) mechanism enables users to be securely authenticated with multiple applications and websites by using just one set of credentials, which provides users with a new way of password management and user authentication. In [26], the authors utilize kinematic information sequences of multi-motion sensor behavior for passive user authentication when the user interacts with his smartphone, and also propose a decision procedure based on HMM to characterize the behavioral biometric feature space such that the continuous user identity verification can be implemented across various operational scenarios. The authors in [52] demonstrate the discriminability and robustness of features related to screen-touch behaviors, and then apply these features to develop a passive authentication solution for smartphone users. The authors in [53] show that it is possible to distinguish profiles of users by exploring CSI information even when they possess similar CSI fingerprints. They also design a practical user authentication approach based on the fine-grained CSI features to accurately determine the user identities in both lab and apartment environments. The literature [65] exploits the CSI of WiFi signals to extract the gesture features (like push, swing, and wave) and some identity-related imperceptible features, and then applies the HMM and Fresnel Model to develop a robust and efficient user authentication approach to determine user identities in IoT environments. In [32], the authors develop a non-intrusive and implicit authentication approach based on the accurate and fine-grained feature of mouse-interaction

behavior segments. The literature [45] focuses on exploiting both the three-dimension features (i.e., color, depth, and inertial) of dynamic gait and the multiclass support vector machine classifier to determine the user identity.

It is notable, however, that there are some problems for the above aforementioned authentication solutions.

1) The above passive authentication solutions cannot be directly extended to the IIoT systems. Although the above passive authentication solutions are effective for user identity verification in their concerned application scenarios, they cannot be directly extended to the IIoT systems. On one hand, in the industrial production process users are always required to operate mobile devices in a uniform and standardized manner, so relying on single behavioral characteristic can not ensure the uniqueness of user identity. On the other hand, in the IIoT environment there are often some specific user dress rules such as having to wear protective gloves or clothes, which results in a low discriminability of behavioral features and thus an inefficient authentication performance. In addition, the available single characteristic-based passive authentication solutions are usually sensitive to specific noise and interference [32], which makes them unsuitable for the complex IIoT systems suffering from intensive noise and interference (e.g., gravity components and non-stationary noise).

2) The existing literature only focuses on either screen-touch time domain characteristics of single trajectory [52], or spatial motion behavior characteristics of touch actions [26, 34]. It is worth noticing that for a given IIoT system, a user generally is engaged in specific work business according to work instructions (WIs), and interacts with the cloud platform by performing some common screen-touch operation actions (e.g., sliding up, sliding down, sliding left, and sliding right) on the touchscreens of (Mobile Terminals) MTs during routine work processes. To improve the performance of user authentication for the special IIoT environment, the characteristics of touch trajectories in both time variation and space combinatorial distribution should be

further explored. In practical IIoT scenarios it is urgent to explore touch-based features of time-varying screen-touch trajectory sequences and cumulative consecutive screen-touch trajectory images from user touch actions during routine work processes in IIoT systems and to develop touch-based passive authentication frameworks for continuous user identity verification.

3) It is notable, however, that when applying the above available passive authentication approaches in modern IIoT systems, it is usually difficult to accurately depict the user identities and thus to achieve an acceptable user authentication performance based on only one-dimensional characteristics. First, users in IIoT systems usually just follow the requirements of industrial production businesses to conduct some basic operations over their MTs in a standardized manner, making it difficult to accurately characterize the user identities with only the time-varying behavioral biometric features extracted from their operation actions. Second, the IIoT systems share a relatively uniform electromagnetic and space environment, so users there show a strong location correlation and thus a low discriminability in terms of the CSI spatial variation characteristics [53, 54]. However, our results in this work indicate that by jointly exploiting the two-dimensional features of the time-varying characteristics of user sequential operation actions and spatial variation characteristics of CSI caused by these actions, we can not only provide a full spatial-temporal characterization of user identities but also significantly improve the performance of passive user authentication.

CHAPTER III

Authentication Utilizing Behavioral Biometrics for the Manufacturing Execution (ME) Layer

3.1 Background and Related Work

The basic functions of the ME layer are sensing, instruction execution, networking, information collection, and industrial control. In the ME layer, IIoT receives specific manufacturing instructions, plans, and important parameters from the MC layer, and realizes manufacturing operations and execution through robots, industrial networks, robotic arms, controllers, automation systems, and instruction execution software. Due to the needs of product production and manufacturing, a large amount of alternating current, motor equipment with changing strong magnetic fields, and cross coverage of various wireless signals generate more electromagnetic interference. Therefore, the ME layer is usually accompanied by a large amount of electromagnetic interference. In order to ensure the robustness and reliability of the authentication, the authentication protocol at this layer needs to have good anti-interference performance.

By now, some research efforts have been devoted to the study of passive user authentication based on single user behavioral characteristic [33, 36, 36–41, 52, 66–68]. The authors in [26] show that it is possible to verify the identity of a smartphone user

by utilizing kinematic information sequences of multi-motion sensor behavior. They also investigate the reliability and applicability of using motion-sensor behavior for continuous smartphone user authentication across various operational scenarios. In [32], the authors develop a non-intrusive and implicit authentication approach based on the accurate and fine-grained feature of mouse-interaction behavior segments. The authors in [52] demonstrate the discriminability and stability of the feature for screen touch gestures, and then apply this feature to develop a continuous authentication solution for a user. The literature [45] focuses on exploiting both the three-dimension features of dynamic gait and the multiclass support vector machine classifier to determine the user identity.

3.2 Motivation

The existing passive authentication solutions are effective for user identity verification in their concerned application scenarios, but they cannot be directly extended to the IIoT systems. On one hand, in the industrial production process users are always required to operate mobile devices in a uniform and standardized manner, so relying on single behavioral characteristic can not ensure the uniqueness of user identity. On the other hand, in the IIoT environment there are often some specific user dress rules such as having to wear protective gloves or clothes, which results in a low discriminability of behavioral features and thus an inefficient authentication performance. In addition, the available single characteristic-based passive authentication solutions are usually sensitive to specific noise and interference [32], which makes them unsuitable for the complex IIoT systems suffering from intensive noise and interference (e.g., gravity components and non-stationary noise).

In this chapter, we explore the common behavioral biometrics from sequential operation actions in IIoT systems and develop a multiple characteristics-based passive authentication framework for continuous and non-intrusive user identity verification.

In particular, we first provide extensive experiment results to demonstrate that in IIoT systems the common behavioral biometrics from sequential user operation actions (i.e., walking, scanning, screen-touch, and photographing-uploading) exhibit good discriminability and stability in discriminating user identities. We then employ the Kalman filtering and Wavelet techniques to reduce the noise in sensor signals of user operation actions, and apply the singular value decomposition method to achieve the dimensionality reduction for the feature space of sequential operation actions. We further develop a passive user authentication framework for continuous and non-intrusive user identity verification against the impersonation attack. Finally, experiment results are provided to illustrate the authentication performance of the proposed authentication framework in terms of the false acceptance, false rejection and equal-error rates. The related authentication efficiency issues, like the usability to the operation-action sequence length, the scalability to the number of features and user space, and the sensitivity to the operation action features, are also investigated.

3.3 Threat Model and Overview of Our Approach

3.3.1 Threat Model

Consider an IIoT application scenario, where legitimate users interact with IIoT systems through mobile devices (industrial-level terminals) in the presence of a potential attacker. The attacker has access to physical mobile devices and thus can capture passcodes or proofs of identities to unlock mobile devices. Sensitive information on manufacturing core technologies might be exploited by the attacker to initiate malicious operations and to steal critical manufacturing data of IIoT systems. Therefore, the concerned IIoT system should be able to discriminate identities between legitimate users and attackers, based on the operation-action features constructed from sensor behaviors during the user routine work process. By utilizing the operation-

action features, one can achieve non-intrusive and continuous authentication when users implement routine operation actions with mobile devices.

3.3.2 Overview of Our Approach

A user holding a mobile device (information interaction terminal) always performs some common operation actions (e.g., walking, scanning, and screen-touch) during routine work processes. The mobile device is typically equipped with various sensors such as orientation, accelerometer, gyroscope, and touchscreen. It is noticed that industrial production processes are usually accompanied through the mutual switching of several user operation actions. The resulting sequences (for mutual switching) of the operation actions from different users can reflect different levels of posture and motion change of mobile devices. The switching sequences could represent unique behavioral characteristics of users. In other words, operation action behaviors can characterize the identities of users. In this work, we explore the discriminability and applicability of the operation-action features extracted from built-in sensors in mobile devices for passive and continuous authentication.

To examine the authentication performance under different IIoT scenarios, here we consider four scenarios: R&D office environments (ROE), common manufacturing environments (ME), manufacturing environments requiring to wear protective gloves (Glo), and manufacturing environments requiring to wear full-body protective clothing (FPC). Specifically, the ROE scenario refers to R&D office environments where there mainly exists diverse interferences from wireless communications; In the ME scenario, the interference generally comes from the various strong magnetic and electric fields induced by industrial devices; In the Glo scenario, users are required to wear gloves which follow the standard of HG/T 2584-2010; In the FPC scenario, users need to wear the full-body protective clothing which follows the standard of IEC/TR 61340-5-2-2007. These scenarios could roughly cover a user's routine operation-action

environments.

For these IIoT scenarios, we design an authentication framework consisting of four processes, as illustrated in Fig. 3.1: 1) Raw data collection and preprocessing; 2) Feature construction for operation actions; 3) Dimensionality reduction for operation-action features; 4) Passive authentication. In the raw data collection and

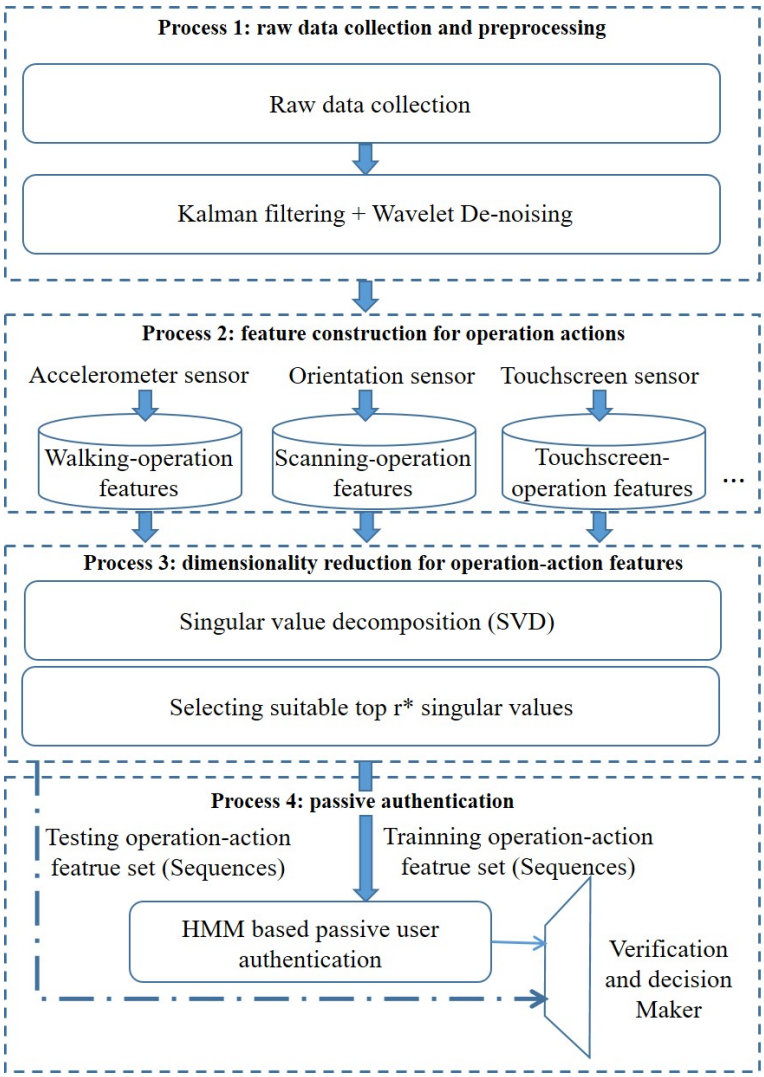


Figure 3.1: The four processes of the proposed authentication approach for IIoT scenarios.

preprocessing process, we collect real-time raw data from built-in sensors of mobile devices, and leverage a Kalman filter and Wavelet de-noising to reduce the noise along with the data. The process of feature construction for operation actions aims to an-

alyze the operation actions (e.g., walking, scanning, and screen-touch) of the user and then construct the corresponding operation-action feature space. The purpose of dimensionality reduction for operation-action features is to decrease the amount of data transmission and processing while enhancing the availability and practicality of the proposed authentication approach. In the passive authentication process, the identity of a user will be continuously validated.

3.4 Proposed Authentication Framework

3.4.1 Raw Data Collection and Preprocessing

We develop an application running in the background of mobile devices to obtain raw sensor data by accessing the APIs, which are provided by the SDK of the mobile device's operation system. For example, for a mobile device based on Android operation system, we can collect the accelerometer sensor data by employing the android.hardware.SensorManager package and listening to the event of $Sensor.Type\_Accelerometer$.

Figure 3.2: The process of raw sensor data collection during the user routine work process in IIoT scenarios.

For various application scenarios in IIoT systems, we utilize the raw data collected

from sensors of mobile devices to construct a user's unique operation-action characteristics over time, as shown in Fig. 3.2. In particular, we collect raw data by employing accelerometer, gyroscope, orientation, and touchscreen sensors. The four types of raw sensor data respectively correspond to walking, photographing-uploading, scanning, and screen-touch operation actions during the user routine work process in IIoT scenarios. The raw sensor data always exists several measurement errors due to diverse noise and interference (e.g., gravity components and non-stationary noise). Therefore, to reduce the impact of measurement errors on the authentication performance, it is necessary to conduct preprocessing for these raw sensor data.

1) Raw data filtering

Note that the interference from gravity components and invariable magnetic fields is commonly a constant, while sensor values triggered by a user's operation actions are usually time-varying variables. Similar to that in [26], we also adopt a Kalman filtering [69] to estimate the sensor value. Constant noise from IIoT scenarios attached to each sensor component can be reduced and even eliminated through the following steps.

Step 1 : If let $\hat{x}_{k-1}$ be sensor data state estimation at time $k-1$ and $\hat{x}_k^-$ be the a priori state estimation at time k, then the equation for the time update is expressed by

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1}, \tag{3.1}$$

where both A and B are coefficient matrices, related to the states at time $k-1$ and time k, respectively, and u_{k-1} is an optional control input.

Denoting by D_{k-1} the estimation of the error covariance at time $k-1$, the a priori estimation (i.e., predicted values) of the error covariance at time k denoted by D_k^- is calculated as

$$D_k^- = AD_{k-1}A^T + Q, \tag{3.2}$$

where Q is a process noise covariance in practical IIoT systems.

Step 2 : Based on (3.2), the Kalman gain denoted by K_k is expressed by

$$K_k = \frac{D_k^- H^T}{H D_k^- H^T + R}, \tag{3.3}$$

where H is the transformation matrix from state variables to measurements (observations), and R is the measurement noise covariance matrix. We use z_k to denote the measurement variable by actually measuring the process at time k, and then an a posteriori state estimation $\hat{x}_k$ is written as

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-). \tag{3.4}$$

Finally, we obtain the a posteriori error covariance estimation as

$$D_k = (I - K_k H) D_k^-. \tag{3.5}$$

In Fig. 3.3 we provide the comparison between the measured results and the filtered results for accelerometer sensor component values of Acc_x. We can see from Fig. 3.3 that constant noise (e.g., gravity components and invariable magnetic fields) in IIoT scenarios can be effectively reduced through Kalman filtering. Since Kalman filtering is recursive and can run in real time, it can meet the real-time and high-efficiency authentication requirements in IIoT systems.

2) Wavelet De-noising

Initial raw data directly collected from sensors always exists multiple peaks and other interference points due to the non-stationary noise in IIoT scenarios. The non-stationary noise is in general caused by electromagnetic interference (e.g., changing current and magnetic field, radio frequency signals) and man-made interference (e.g., collision, jitter, and accidental touch on mobile devices). In this work, we leverage

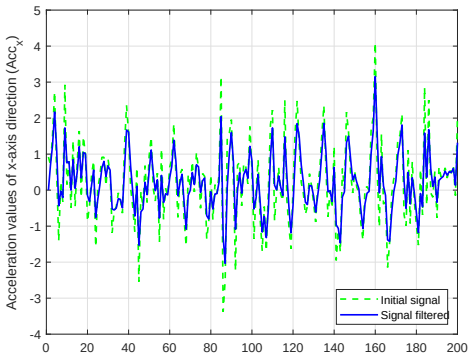
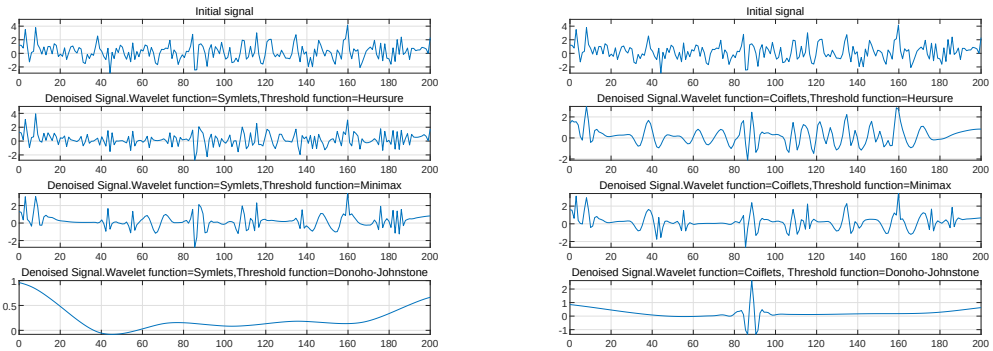


Figure 3.3: Kalman filtering for acceleration values of x-axis direction (Acc_x).



(a) De-noising wavelet function = '$Symlets$' (b) De-noising wavelet function = '$Coiflets$'

Figure 3.4: De-noising performance comparison of wavelet function $Symlets$ and $Coiflets$ in different threshold functions. (a) De-noising performance of wavelet function $Symlets$ under different threshold functions. (b) De-noising performance of wavelet function $Coiflets$ under different threshold functions.

wavelet de-noising to not only reduce the non-stationary noise and interference, but also retain the intrinsic feature of raw sensor data [26, 70].

To examine de-noising efficiency, we provide in Fig. 3.4 the de-noising performance comparison between the wavelet function $Symlets$ and $Coiflets$ under various threshold functions (e.g., $Heursure$, $Minmaxi$, and $Donoho - Johnstone$). We can see from Fig. 3.4 that the wavelet function $Symlets$ under the threshold function $Heursure$ is more efficient than $Coiflets$, since it reduces the non-stationary noise and interference while retaining the intrinsic feature of raw sensor data. Therefore, we adopt here the wavelet function $Symlets$ under the $Heursure$ threshold to achieve the optimal authentication performance of the proposed approach.

3.4.2 Feature Construction for Operation Actions

1) Walking Operation-action Feature

Due to the randomness of user motion, the posture of a mobile device is always time-varying. As a result, it is very difficult (if not impossible) to discriminate the values collected from each axis of the built-in accelerometer. To simplify the data processing and keep the original walking feature, we incorporate sensor component values collected from three axes into a robust compound variable to depict the user's robust gait characteristics.

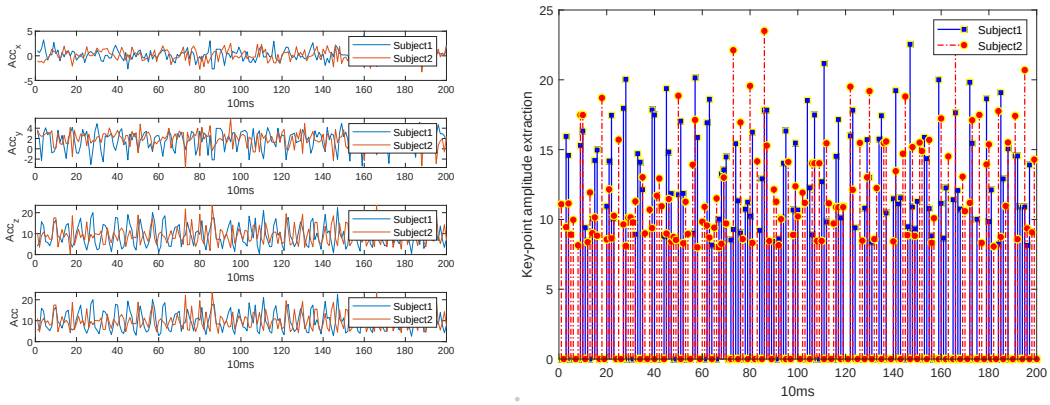
Let $Acc_x(t)$, $Acc_y(t)$, and $Acc_z(t)$ denote accelerometer component values at time t on the x, y, and z axes, respectively, and then the value of compound variable of accelerometer sensor denoted by $Acc(t)$ can be written as

$$Acc(t) = \sqrt{Acc_x^2(t) + Acc_y^2(t) + Acc_z^2(t)}. \tag{3.6}$$

If $Acc(t_1) \leq Acc(t_x) \leq Acc(t_2)$ and $|t_2 - t_1| \leq \partial$, $Acc(t_x)$ is referred to as a key-point amplitude of the gait curve, where ∂ is a threshold window specified by the

IIoT system. Then, $Acc(t_x) \in A$ and A is the vector of key-point amplitude.

Here, we collect M steps' key-point amplitudes of each user to characterize the user's walking actions. If there are K key-point amplitudes for the m-th step of the j-th user, then the key-point amplitude vector is denoted as $A^j_m = \{A^j_{m_1}, A^j_{m_2}, ..., A^j_{m_K}\}$. For the j-th user, the set of key-point amplitude data under the M steps is denoted by L^j_m. Note that the number of key-point amplitudes in a step is random. Hence, the dimensions of the elements in set L^j_m are not always the same. We use n to denote the dimension of the minimum dimensional element in L^j_m (i.e., $n = \min\{Length(L^j_m)\}$). We can obtain an $m \times n$ matrix B_{wal} for characterizing the user's walking-operation features.



(a) (b)

Figure 3.5: Examples of the accelerometer sensor data and Key-point amplitude from Subject1 and Subject2. (a) Accelerometer sensor components Acc_x, Acc_y, and Acc_z on the x, y, and z axes, and fused accelerometer sensor values Acc. (b) Key-point amplitude curves corresponding to Subject1 and Subject2.

As illustrated in Fig. 3.5, we plot the two-step curves of walking actions for two users (namely Subject1 and Subject2) and the corresponding performance comparison of the key-point amplitude discrimination. Fig. 3.5(a) shows the difference between accelerometer component curves Acc_x, Acc_y, Acc_z and the fused curves Acc of the two users, respectively. It is seen from Fig. 3.5(b) that the behaviors in terms of key-point amplitude for the two users are clearly different. This indicates that walking

operation action from a user has its own unique characteristics, and thus can be used for the identity of the user.

2) Scanning Operation-action Feature

Scanning two-dimensional code (i.e., QR code) through mobile devices is one of the most common operation actions for a user in IIoT scenarios to achieve the intelligent information management, such as automatic information processing, accurate and fast information collection, and information identification. The orientation sensor's spatial-temporal properties trigged by the user during the scanning operation process contain the robust and practical scanning behavior habits of the user. The user's scanning operation-action features can then be used to characterize the user's identity for authentication.

In general, an orientation sensor has three components Azimuth, Pitch, and Roll. We use Υ_{Az}, Υ_{Pi}, and Υ_{Ro} to denote the orientation sensor values of Azimuth, Pitch, and Roll, respectively. The values of Υ_{Az}, Υ_{Pi}, and Υ_{Ro} can be obtained from the mobile device rotation angle around z-axis, x-axis, and y-axis, respectively. Obvi-

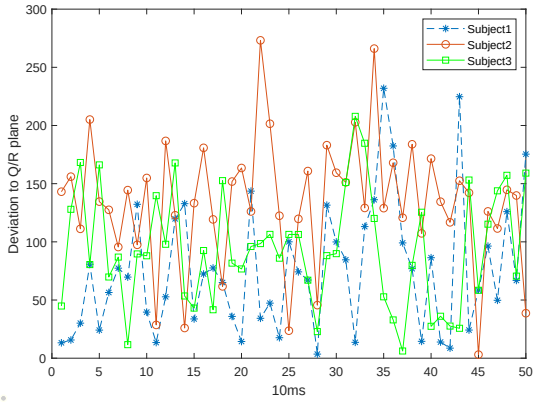


Figure 3.6: Comparison of the scanning operation-action features of three users (Subject1, Subject2, and Subject3).

ously, during the process of the user scanning the two-dimensional code, the spatial-temporal location of the mobile device at time t can be expressed as a vector $\boldsymbol{\beta}$. Then we have $\boldsymbol{\beta}_t = (\Upsilon_{Az,t}, \Upsilon_{Pi,t}, \Upsilon_{Ro,t})$. We use $\boldsymbol{\beta}_t^{j,i}$ to represent the orientation sensor vec-

$$D_t^{j\to i} = \sqrt{(\Upsilon_{Az,t}^{j,i} - \Upsilon_{Az,t_{par}}^{j,i})^2 + (\Upsilon_{Pi,t}^{j,i} - \Upsilon_{Pi,t_{par}}^{j,i})^2 + (\Upsilon_{Ro,t}^{j,i} - \Upsilon_{Ro,t_{par}}^{j,i})^2}. \quad (3.7)$$

tor consisting of the corresponding component values during the j-th user's scanning the i-th two-dimensional code at time t. Thus, we have $\boldsymbol{\beta}_t^{j,i} = (\Upsilon_{Az,t}^{j,i}, \Upsilon_{Pi,t}^{j,i}, \Upsilon_{Ro,t}^{j,i})$. Let $\boldsymbol{\beta}_{t_{par}}^{j,i}$ denote the orientation sensor vector at time t_{par} when the mobile device's plane is exactly parallel to the two-dimensional code plane. Thus, we have $\boldsymbol{\beta}_{t_{par}}^{j,i} = (\Upsilon_{Az,t_{par}}^{j,i}, \Upsilon_{Pi,t_{par}}^{j,i}, \Upsilon_{Ro,t_{par}}^{j,i})$. In order to quantitatively calculate the kinematic characteristics of the equipment plane relative to the two-dimensional code plane caused by the user's scanning action during the scanning two-dimensional code process. Let $D_t^{j\to i}$ denote the deviation degree of the mobile device plane relative to the two-dimensional code plane $\boldsymbol{\beta}_{t_{par}}^{j,i}$ from the j-th user's scanning operation actions at time t, and $D_t^{j\to i}$ is given in (3.7).

The moment when a user successfully scans a two-dimensional code is the origin of time denoted by t_0, and then we employ the orientation sensor value generated between T periods forward to t_0 and backward to t_0 to characterize the user's scanning operation actions. We use S to denote the number of samples collected from orientation sensor values and have $S = \frac{2T}{f}$, where f is the sampling frequency. We denote a scanning operation-action sequence of the j-th user by $\xi^j = \{D_{t_0-T}^{j\to i}, D_{t_0-T+1}^{j\to i}, \cdots, D_{t_0}^{j\to i}, D_{t_0+1}^{j\to i}, ..., D_{t_0+T}^{j\to i}\}$, and the dimension of ξ^j is S. We employ adjacent M times' scanning data to construct the j-th user's scanning feature matrix denoted by B_{sca}^j, and then we have $B_{sca}^j = [\xi_1^j; \xi_2^j; ...; \xi_M^j]$.

In Fig. 3.6, we construct three users' scanning operation-action features $\xi^{Subject1}$, ξ^{Subject2}, and $\xi^{Subject3}$, respectively. It indicates that each user's scanning operation action is discriminative and exhibits unique behavioral characteristics, which can be adopted to execute user authentication in IIoT scenarios.

3) Screen-touch Operation-action Feature

Touching screen is also one of the most common operation actions during the user

routine work process. The way (behavioral habits) of interaction with mobile devices from a user is proven to be applicable to determine the user's identity [52, 71].

To construct features of the screen-touch operation action, we study screen-touch behaviors based on 14 common feature parameters listed in Table 3.1. An important reason why we choose these features is that these screen-touch features are proven to be able to resist against the interference generated from the IIoT environment. More importantly, these features are easily extracted through information processing techniques [52]. These feature parameters are self-explanatory and can be determined through touchscreen sensors. In particular, using the screen-touch data sequences during a specific time interval ε, we first quantify the 14 feature values via statistical analysis for the touchscreen sensor data trigged and generated by the user's routine touch interactions. Then, the screen-touch feature vector denoted by u_{ε} can be obtained based on the resulting quantization values. Finally, we incorporate consecutive M screen-touch feature vectors into a matrix denoted by B_{tou} to characterize the user's robust and practical screen-touch behavioral features, and thus have $B_{tou} = [u_{\varepsilon_1}, u_{\varepsilon_2}, \ldots, u_{\varepsilon_M}]$.

Table 3.1: List of feature parameters of screen-touch

Dimensions	Definitions of feature parameters
1-4	Start x, y, stop x, y
5	Mid-stroke pressure
6	Average direction
7	Average velocity
8	Stroke duration
9	Direct end-to-end distance
10	Mobile device orientation
11	Ratio end-to-end dist and length of trajectory
12	Inter-stroke time
13	Largest deviation from end-to-end line
14	Mid-stroke area covered

4) Photographing-uploading Operation-action Feature

Photographing-uploading operation action through mobile devices is one of the

most common operation actions for a user to take and upload multimedia data (such as pictures and videos) in IIoT systems. The photographing-uploading operation action possesses the spatial-temporal properties which can be measured by accelerometer and gyroscope sensors [26, 72], and these properties can be leveraged to characterize the user's unique identity for authentication.

To extract the feature of the photographing-uploading operation, we first collect sensor data from accelerometer and gyroscope sensors. Specifically, in the duration of the photographing-uploading action, we collect 6 sensor component data sequences (2 sensors $\times$ 3 axes for each sensor) in the i-th photographing-uploading operation action denoted by vector $\Psi_i = (Acc_{(x,i)}, Acc_{(y,i)}, Acc_{(z,i)}, Gyr_{(x,i)}, Gyr_{(y,i)}, Gyr_{(z,i)})$. It is observed that each sensor sequence in Ψ_i represents a set of sensor events generated during the photographing-uploading operation action of a user. We then extract the spatial-temporal characteristics of the photographing-uploading action by analyzing user photographing-uploading action in time domain and frequency domain. Table 3.2 lists the time domain and frequency domain features of photographing-uploading operation action from each sensor sequence in Ψ_i. Specifically, most of time domain features are self-explanatory. For frequency domain features, we use MF to represent the mean of frequency f through FFT transform of each sensor sequence in Ψ_i. Let FC, RMSF, and RVF denote the barycenter frequency, root mean square frequency, and variance frequency, respectively. For a given frequency f, its frequency amplitude $s(f)$ can be obtained through FFT transform of Ψ_i component, and its FC, RMSF and RVF can be written as

$$FC = \frac{\sum fs(f)}{\sum s(f)}, \tag{3.8}$$

$$RMSF = \sqrt{\frac{\sum f^2 s(f)}{\sum s(f)}}, \tag{3.9}$$

$$RVF = \sqrt{\frac{(f - F_{FC})^2 s(f)}{\sum s(f)}}. \tag{3.10}$$

Finally, for each photographing-uploading operation action, the features for each data sequence in Ψ_i are concatenated together as a feature vector $\mathbb{P}_{\Psi_i}$, which can be used to characterize the i-th photographing-uploading operation action from a user. In practical applications, we need to get multiple samples to obtain a stable photographing-uploading operation feature. For m photographing-uploading operation actions, we denote the corresponding feature matrix as $B_{pu} = [\mathbb{P}_{\Psi_1}, \mathbb{P}_{\Psi_2}, \ldots, \mathbb{P}_{\Psi_m}]$.

Table 3.2: List of feature parameters of photographing-uploading operation actions

Category	Features	Definitions
Time domain	Maximum	Maximum sensor value of Ψ_i component
	Minimum	Minimum sensor value of Ψ_i component
	Mean	Mean sensor value of Ψ_i component
	Standard deviation	Standard deviation value of Ψ_i component
	Peak-to-peak	Difference between maximum and minimum value of Ψ_i component
	Kurtosis	Width of peak value of Ψ_i component
Frequency domain	Mean frequency (MF)	Mean frequency of Ψ_i component
	Barycenter frequency (FC)	Barycenter frequency of Ψ_i component
	Root mean square frequency($RMSF$)	Root mean square frequency of Ψ_i component
	Variance frequency (RVF)	Variance frequency of Ψ_i component

3.4.3 Dimensionality Reduction for Operation-action Features

To enhance the applicability of the approach developed in this work and to reduce the computational complexity, we consider a well-known singular value de-

composition (SVD) method to achieve dimensionality reduction for features' matrices. Suppose that the set of a user's operation-action features F is denoted as $F = [B_{wal}, B_{sca}, B_{tou}, B_{pu}]$ and F is an $m \times n$ matrix, which includes the three features we constructed in IIoT scenarios. Then, the singular value decomposition of F can be written as

$$F = U\Gamma V^T, \tag{3.11}$$

where U, V, and Γ denote the $m \times m$ left singular matrix, the $n \times n$ right singular matrix, and the $m \times n$ singular value matrix, respectively, which can be obtained as

$$V = F^T F, \tag{3.12}$$

$$U = F F^T, \tag{3.13}$$

$$\Gamma = \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix}, \tag{3.14}$$

where $\Lambda = \text{diag}(\delta_1, \delta_2, \cdots, \delta_\gamma)$, $\delta_1, \delta_2, \cdots, \delta_\gamma$ are singular values, they are arranged according to $\delta_1 > \delta_2 \cdots > \delta_\gamma > 0$, and $\gamma = \text{rank}(\Lambda)$. We use the first $r*$ singular values to obtain the approximate value of feature F, and F is given by

$$F = U_{m \times r*} \Lambda_{r* \times r*} V^T_{r* \times n}, r* < \min\{m, n\}. \tag{3.15}$$

In general, all singular values in Λ should be utilized for the feature construction and optimum authentication performance. Nevertheless, when $r*$ grows large, it may contain many extremely small singular values (even very approach to 0) in Λ. As a consequence, it might not be a good choice to set $r*$ as $r* = \gamma$. This is

because including these extremely small singular values in Λ causes a dimensionality bottleneck and an increase in computational complexity without much real benefit in terms of the authentication performance. In addition, extensive experiments in practical IIoT scenarios indicate that when the proportion of the sum of the first $r*$ singular values to that of all singular values is greater than a fixed threshold, we can not only reduce the dimension of F to lessen computational complexity and improve authentication efficiency, but also remain the features of user identities with tolerable feature loss. Hence, $r*$ is set as

$$\sum_{i=1}^{r*} \delta_i > \phi \sum_{i=1}^{\gamma} \delta_i, \tag{3.16}$$

where ϕ is a constant adjusted accordingly based on the requirements of the authentication performance in practical application scenarios.

3.4.4 Passive Authentication

1) Description of HMM-based authentication

Sequential operation actions exhibit the temporal nature of the problem of interest. Along with the systematical investigation of the operation-action events' time sequence distribution, the sequential operation actions appear to be stochastic and are viewed as a Markov process [73]. Here, we adopt the well-known HMM to describe the dynamic transferring processes along with time among operation actions and thereby characterize the user's identity during the routine work process.

We use $\lambda = (A, B, \pi)$ to denote the HMM in our approach, where A, B, π are the state transition probability matrix, the state observation probability matrix, and the vector of initial state probability, respectively. Herein, the hidden states refer to the real operation-action features in actual various IIoT production scenarios such as ROE, ME, Glo, and FPC. Thus, we construct the feature ma-

trices of hidden states leveraging the pre-enrolled (off-line) data in databases of IIoT systems. While the feature matrices of observed states are extracted and constructed by acquiring sensor data in real time during the user's routine work processes. Let $\{H_{pq}\}$ be the sets of hidden states and $\{O_{vw}\}$ be the set of observed states for $p, v \in \{\text{walking}, \text{scanning}, \text{screen}-\text{touch}, \text{photographing}-\text{uploading}\}$ and $q, w \in \{\text{ROE}, \text{ME}, \text{Glo}, \text{FPC}\}$. It is noticed that the degree of correlation between $F_{O_{vw}}$ and $F_{H_{pq}}$ can be used to show the connection of them and thereby present the state observation probability. The correlation coefficient between $F_{O_{vw}}$ and $F_{H_{pq}}$ is given by

$$R(F_{O_{vw}}, F_{H_{pq}}) = \frac{\sum_m \sum_n (F_{O_{vw}} - \bar{F}_{O_{vw}})(F_{H_{pq}} - \bar{F}_{H_{pq}})}{\sqrt{(\sum_m \sum_n (F_{O_{vw}} - \bar{F}_{O_{vw}}))^2 (\sum_m \sum_n (F_{H_{pq}} - \bar{F}_{H_{pq}})^2)}}, \tag{3.17}$$

where $\bar{F}_{O_{vw}}$ is the mean of all element values in $F_{O_{vw}}$ and $\bar{F}_{H_{pq}}$ is the mean in $F_{H_{pq}}$.

Based on (3.17), the probability that the hidden state H_{pq} generates the observed state O_{vw} is then written as

$$P(O_{vw}|H_{pq}) = \frac{R(F_{O_{vw}}, F_{H_{pq}})}{\sum R(F_O, F_{H_{pq}})}, \tag{3.18}$$

where O represents the set of all possible observed states corresponding to hidden state $F_{H_{pq}}$ and $\sum P(O|H_{pq}) = 1$. Therefore, the state observation probability matrix B connecting hidden states and observed states can be expressed as $B = P(O|H)$, which covers the probability of each hidden state generating all observed states. Employing the Baum–Welch method [74], the state transition probability matrix A and initial state probability vector π can be estimated.

For the i-th user to be authenticated, we denote the HMM associated with the time-varying operation actions as $\lambda^i = (A^i, B^i, \pi^i)$. The observed operation-action

feature sequence is $O^i = (o_{t_1}, o_{t_2}, \cdots, o_{t_n})$ with $o_{t_j} \in F$ representing the operation-action feature in time t_j, $j = 1, 2, \cdots, n$, under the length of the operation sequence n. Using the forward algorithm [75], the probability $P(O^i|\lambda^i)$ is then calculated to authenticate the user's identity under a given threshold.

2) Authentication decision

For each legitimate user, we develop his HMM-based one-class classifier and train the classifier by using its corresponding positive sample set in (3.26). Thus, each user is associated with one dedicated HMM-based one-class classifier to determine his identity in our authentication approach. Given an unknown user X who claims his identity is U, we use F^X to denote his corresponding real-time operation-action features, then the goal of our authentication is to determine whether the identity corresponding to F^X is U or not. Specifically, we set F^X as the input of the one-class classifier corresponding to user U, and then analyze the probability score of its output to determine the claimed identity of user X. If we use $S_u(F^X, U)$ to denote the similarity between operation-action features of user X and that of user U, then the authentication decision can be formulated as

$$(U, F^X) \in \begin{cases} ture, & S_u(F^X, U) \geq \varphi, \\ false, & \text{otherwise}, \end{cases} \tag{3.19}$$

where φ is a predefined threshold for the concerned IIoT system. From (3.19) we can see that when $S_u(F^X, U)$ is no less than φ, the claim of user X is true (a legitimate user), otherwise the claim is false (i.e., user X is an attacker).

3.5 Performance Modeling

1) Performance Metrics and Analysis

To evaluate the performance of the proposed approach, we investigate three typ-

ical performance metrics: FAR, FRR, and EER [26, 50, 51]. We also adopt the authentication accuracy to evaluate the performance for resisting the impersonation attacks of the proposed framework, here the authentication accuracy is defined as the probability that the system successfully distinguishes between the legitimate users and impersonation attacks.

Mathematically, the FAR and FRR can be formulated as follows according to classification results of our proposed authentication process,

$$FAR = \frac{FP}{FP + TN}, \tag{3.20}$$

$$FRR = \frac{FN}{TP + FN}, \tag{3.21}$$

where TP, FN, FP, and TN represent the number of yielding true-positive classification results when the samples actually are positive, the number of yielding false-negative classification results when the samples actually are positive, the number of yielding false-positive classification results when the samples actually are negative, and the number of yielding true-negative results when the samples actually are negative in a test. We use (F_i^X, U_i) to denote that the i-th unknown user X with operation action feature F_i^X claims that its identity is U_i, $i \in [1, n]$, and use $\S = \{(F_1^X, U_1), \ldots, (F_n^X, U_n)\}$ to denote the test sample set, here n is the number of test samples in $\S$. According to (3.19), we have

$$TP = \sum_{S(F_i^X, U_i) \geq \varphi} \S \in \omega_1, \tag{3.22}$$

$$FN = \sum_{S(F_i^X, U_i) \geq \varphi} \S \in \omega_2, \tag{3.23}$$

$$FP = \sum_{S(F_i^X, U_i) < \varphi} \S \in \omega_1, \tag{3.24}$$

$$TN = \sum_{S(F_i^X, U_i) < \varphi} \S \in \omega_2, \tag{3.25}$$

where ω_1 and ω_2 denote the set of the claim of user X is true and false, respectively.

We use $\mathbb{S}_{n\times m}$ to denote the training sample space with n samples and m labels, use $\mathbb{S}_{Lab_i}$ to denote all the samples with label Lab_i ($i \in [1, m]$) in $\mathbb{S}_{n\times m}$, and use $\bar{\mathbb{S}}_{Lab_i}$ to denote other samples with other labels except Lab_i. In our designed classifier based on HMM, we use $\mathbb{S}_{n\times m}$ to generate m sub-sample sets denoted by $\mathbb{S}_{sub}$,

$$\mathbb{S}_{sub} = \{ \begin{bmatrix} \mathbb{S}^+_{Lab_1} \\ \bar{\mathbb{S}}^-_{Lab_1} \end{bmatrix}, \begin{bmatrix} \mathbb{S}^+_{Lab_2} \\ \bar{\mathbb{S}}^-_{Lab_2} \end{bmatrix}, \ldots, \begin{bmatrix} \mathbb{S}^+_{Lab_m} \\ \bar{\mathbb{S}}^-_{Lab_m} \end{bmatrix} \}, \tag{3.26}$$

where '+' and '-' denote positive sample label and negative sample label, respectively.

According to the m sub-sample sets in $\mathbb{S}_{sub}$, we can obtain the set Ω_λ of m HMM models to create m one-class classifiers. From Section 3.4.4 we know that Ω_λ is given by

$$\Omega_\lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}. \tag{3.27}$$

By using the HMM models of the one-class classifiers in (3.27), (3.22)$\sim$(3.25) can be rewritten as

$$TP = \sum_{\max\{P(F_i^X|\Omega_\lambda)\} \geq \varphi} \S \in \omega_1, \tag{3.28}$$

$$FN = \sum_{\max\{P(F_i^X|\Omega_\lambda)\} \geq \varphi} \S \in \omega_2, \tag{3.29}$$

$$FP = \sum_{\max\{P(F_i^X|\Omega_\lambda)\} < \varphi} \S \in \omega_1, \tag{3.30}$$

$$TN = \sum_{\max\{P(F_i^X|\Omega_\lambda)\} < \varphi} \S \in \omega_2, \tag{3.31}$$

where

$$\max\{P(F_i^X|\Omega_\lambda)\} = \max\{P(F_i^X|\lambda_1), \ldots, P(F_i^X|\lambda_m)\}. \tag{3.32}$$

We can quantify the values of TP, FN, TP, and TN by calculating each element $P(F_i^X|\lambda_j)$ in (3.32), $j \in [1, m]$. Based on the parameter estimation method of λ_j and the forward algorithm in Section 3.4.4, $P(F_i^X|\lambda_j)$ can be determined through the following procedures [75].

Step 1, Let F^X denote the operation-action feature for the operation-action observation sequence from X, we have $F^X = O^X = \{o_1, \ldots, o_L\}$, here O_j denotes the j-th operation action feature for $j = 1, \ldots, L$, and L is the length of operation-action sequence. Given that HMM $\lambda = (A, B, \pi)$ with $\lambda \in \Omega_\lambda$, we use $\alpha_t(i)$ to denote the forward probability when the observation sequence is $(o_1, o_2, \ldots, o_t)$ and the state is $q_i \in A$. Thus, $\alpha_t(i)$ is determined as

$$\alpha_t(i) = P(o_1, o_2, \ldots, o_t, i_t = q_i|\lambda). \tag{3.33}$$

We calculate initial value $\alpha_1(i)$ as

$$\alpha_1(i) = \pi_i b_i(o_1), i = 1, 2, \ldots, l, \tag{3.34}$$

where o_1 is the first observation value in the operation-action observation sequence O^X, $\pi_i \in \pi$, and $b_i(\cdot) \in B$.

Step 2: Based on (3.34), the forward probability $\alpha_{t+1}(i)$ through $L-1$ iteration (i.e., $t = 1, 2 \ldots, L-1$) can be obtained as

$$\alpha_{t+1}(i) = [\sum_{j=1}^{l} \alpha_t(j) a_{ji}] b_i(o_{t+1}), i = 1, 2, \ldots, L. \tag{3.35}$$

Step 3: The probability $P(O^X|\lambda)$ of observation operation-action sequence O^X is

determined by

$$P(O^X|\lambda) = \sum_{i=1}^{l} \alpha_l(i). \tag{3.36}$$

It is observed from (3.33)∼(3.36) that FAR and FRR depend on the HMM model $\lambda(\cdot)$ and the operation-action length. The HMM model $\lambda(\cdot)$ captures the time-varying nature of the isolated operation-action features, while the length of the operation-action sequence determines the stability and distinguishability of user operation action characteristics.

2) User Space Upper Bound Analysis

Generally, for a larger number of users there is a higher risk that two users would have similar operation-action profiles [26]. To analyze user space upper bound, it needs to construct user space model for authentication. Since operation-action characteristics deviation from a user follows the normal distribution, the deviation between any two users' operation-action characteristics also approximately follows the normal distribution. Hence, for two user feature spaces $\mathbb{F}^a$ and $\mathbb{F}^b$ from users a and b, the corresponding deviation of the feature space is denoted as $\mathbb{D}(a,b) = \mathbb{F}^a - \mathbb{F}^b$ and we have $\mathbb{D}(a,b) \sim \mathcal{N}(\mu, \sigma^2)$, here $\mathcal{N}(\mu, \sigma^2)$ denotes a normal distribution random sequence with mean μ and variance σ^2. Thus, the feature space of user b is given by

$$\mathbb{F}^b = \mathbb{F}^a + \mathbb{D}(a,b). \tag{3.37}$$

Now we consider the more general case under given upper bounds of $\mu = \mu_{max}$ and $\sigma = \sigma_{max}$ in (3.37). For an arbitrary user a, we use $\mathbb{F}^{a^+_{max}}$ and $\mathbb{F}^{a^-_{max}}$ to denote the upper bound and lower bound of the feature space that has the minimum similarity with $\mathbb{F}^a$. Then $\mathbb{F}^{a^+_{max}}$ and $\mathbb{F}^{a^-_{max}}$ are given by

$$\mathbb{F}^{a^+_{max}} = \mathbb{F}^a + \mathbb{D}_{max}, \tag{3.38}$$

$$\mathbb{F}^{a^-_{max}} = \mathbb{F}^a - \mathbb{D}_{max}, \tag{3.39}$$

where $\mathbb{D}_{max}$ denotes a normal distribution random sequence with mean μ_{max} and variance σ^2_{max}.

We consider that the sample spaces are uniformly distributed between $\mathbb{F}^{a^{+}_{max}}$ and $\mathbb{F}^{a^{-}_{max}}$ with mean interval $\frac{\mu_{max}}{\mathbb{C}}$ and standard deviation $\dot{\sigma} = \frac{\sigma_{max}}{\mathbb{C}}$, where $\mathbb{C}$ is the number of user features distributed around $\mathbb{F}^a$. The i-th feature space between $\mathbb{F}^{a^{+}_{max}}$ and $\mathbb{F}^{a^{-}_{max}}$ can be written as

$$\ddot{\mathbb{F}}^a_i = \begin{cases} \mathbb{F}^a + \mathbb{D}_i, \mathbb{D}_i \sim \mathcal{N}(\mu_{max} * \dfrac{(\mathbb{C}-i)}{\mathbb{C}}, \dot{\sigma}^2), \\ \mathbb{F}^a - \mathbb{D}_i, \mathbb{D}_i \sim \mathcal{N}(\mu_{max} * \dfrac{(i)}{\mathbb{C}}, \dot{\sigma}^2), \end{cases} \tag{3.40}$$

where $i = 1, 2, \ldots, \mathbb{C}$.

Based on $\mathbb{F}^{a^{+}_{max}}$ and $\mathbb{F}^{a^{-}_{max}}$, $\ddot{\mathbb{F}}^a$ can be rewritten as

$$\ddot{\mathbb{F}}^a = \begin{bmatrix} \bar{\mathbb{F}}^a + \mathbb{D}_{max} + \hbar; \\ \bar{\mathbb{F}}^a + \mathbb{D}_{\mathbb{C}-1} + \hbar; \\ \bar{\mathbb{F}}^a + \mathbb{D}_{\mathbb{C}-2} + \hbar; \\ \cdots \\ \bar{\mathbb{F}}^a + \mathbb{D}_{\mathbb{C}-\mathbb{C}} + \hbar; \\ \bar{\mathbb{F}}^a - \mathbb{D}_1 + \hbar; \\ \bar{\mathbb{F}}^a - \mathbb{D}_2 + \hbar; \\ \cdots \\ \bar{\mathbb{F}}^a - \mathbb{D}_{\mathbb{C}} + \hbar; \\ \bar{\mathbb{F}}^a - \mathbb{D}_{max} + \hbar; \end{bmatrix}, \tag{3.41}$$

where $\bar{\mathbb{F}}^a$ denotes the mean of the feature space, and $\hbar$ is the noise associated with the IIoT scenario.

It is observed that as $\mathbb{C}$ increases, the similarity between elements in $\ddot{\mathbb{F}}^a$ becomes

$$MMD(\ddot{\mathbb{F}}_p^a, \ddot{\mathbb{F}}_q^a) = [\frac{1}{m^2}\sum_{i=1}^{m}\sum_{j=1}^{m}\mathbb{K}(\ddot{\mathbb{F}}_{p,i}^a, \ddot{\mathbb{F}}_{p,j}^a) - \sum_{i=2}^{m}\sum_{j=1}^{m}\mathbb{K}(\ddot{\mathbb{F}}_{p,i}^a, \ddot{\mathbb{F}}_{q,j}^a) + \sum_{i=1}^{m}\sum_{j=1}^{m}\mathbb{K}(\ddot{\mathbb{F}}_{q,i}^a, \ddot{\mathbb{F}}_{q,j}^a)]^{\frac{1}{2}}, \tag{3.42}$$

small (i.e., it is much more difficult to distinguish each other in $\mathbb{C}$). Therefore, it is necessary to quantitatively measure the amount of the similarity. We here use Maximum Mean Discrepancy $MMD(\ddot{\mathbb{F}}_p^a, \ddot{\mathbb{F}}_q^a)$ to quantify the similarity between any two elements $\ddot{\mathbb{F}}_p^a$ and $\ddot{\mathbb{F}}_q^a$ in $\ddot{\mathbb{F}}^a$, and $MMD(\ddot{\mathbb{F}}_p^a, \ddot{\mathbb{F}}_q^a)$ is written as the equation (3.42) [76], where $\mathbb{K}(\cdot)$ is the Gaussian kernel function for mapping the feature vectors of user operation actions to high-dimensional space.

Let $R(\ddot{\mathbb{F}}^a|\mathbb{C})$ denote the average of MMD similarity of all the adjacent elements in $\ddot{\mathbb{F}}^a$ with user space being equal to $\mathbb{C}$, we have

$$R(\ddot{\mathbb{F}}^a|\mathbb{C}) = \frac{1}{\mathbb{C}+1}\sum_{p\ adjacent\ q} MMD(\ddot{\mathbb{F}}_p^a, \ddot{\mathbb{F}}_q^a). \tag{3.43}$$

For a preset threshold ϕ_{EER}, the upper bound of user space $\mathbb{C}$ is denoted as $\mathbb{C}_{upper} = \max\{\mathbb{C}\}$, which satisfies

$$\epsilon(S(\ddot{\mathbb{F}}^a, R(\ddot{\mathbb{F}}^a|\mathbb{C}))) \leq \phi_{EER}, \tag{3.44}$$

where ϵ denotes the authentication performance determined by EER, S is the HMM-based one-class classifier, and R denotes the similarity of the user space.

Extensive simulations demonstrate that when the similarity R is no less than 2.76, the EER keeps under 10%. Fig. 3.7 shows that how the average of similarity R changes with the size of the user space $\mathbb{C}$, as well as the impact of operation-action length on the user space upper bound $\mathbb{C}_{upper}$. We can see that as the operation action length varies from 7 to 10, the upper bound of user space increases from 59 to 103. It indicates that the user space upper bound increases as the operation-action sequence

length increases. The reason is that a longer operation-action sequence yields a better stability and distinguishability of the user operation-action profiles, so increasing the length of operation actions serves as an effective way to improve the upper bound of user space in the practical applications of IIoT systems.

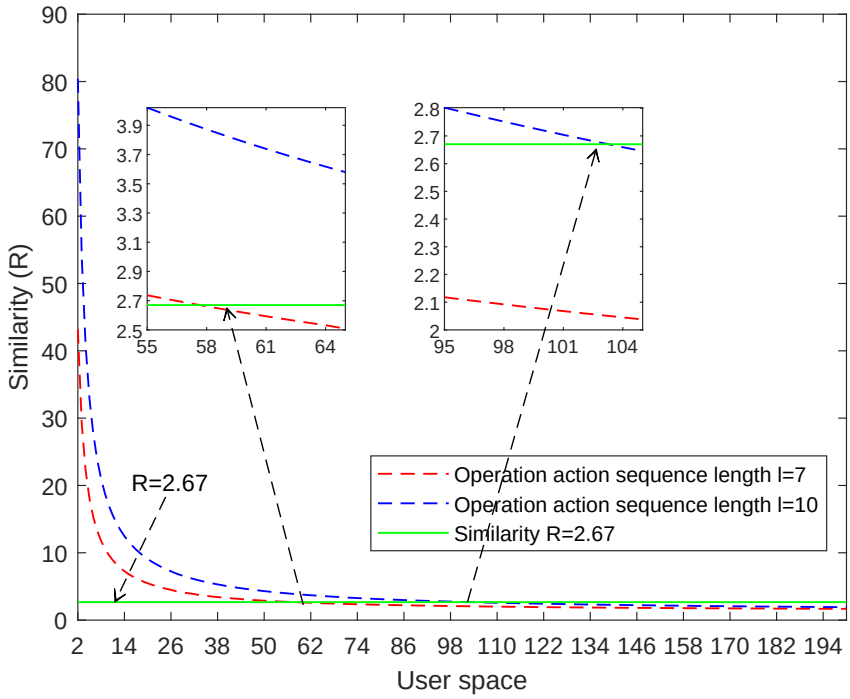


Figure 3.7: $R(\ddot{\mathbb{F}}^a|\mathbb{C})$ varies with user space $\mathbb{C}$ under operation-action sequence length being equal to 7 and 10.

3.6 Experiment and Analysis

Table 3.3: Main experiment datasets

Dataset Name	**Product Process**	**Subjects**	**Data Acquisition**	**Operational Method**
Dataset ♯1	Production-process: Aluminum electrolytic capacitor production process Product batch number: 20200506026, 20210508013,20210509009, 20210510015,20211101013	104 users Male/Female 78/26	Number of operation actions collected at one time (operation-action length): 9. ROE scenarios: 800 times (length=9, operation actions: random). ME scenarios: 800 times (length=9, operation actions: random). Glo scenarios: 800 times (length=9, operation actions: random). FPC scenarios: 800 times (length=9, operation actions: random).	Hand-hold
Dataset ♯2	Production-process: Aluminum electrolytic capacitor production process Product batch number: 20200610009, 20210508013,20210509009, 20210510015,20211101013	123 users Male/Female 87/36	Number of operation actions collected at one time (operation-action length): 2∼15. ROE scenarios: 600 times for each length (length=2∼15, operation actions:random). ME scenarios: 600 times for each length (length=2∼15, operation actions:random). Glo scenarios: 600 times for each length (length=2∼15, operation actions:random). FPC scenarios: 600 times for each length (length=2∼15, operation actions:random).	Hand-hold

3.6.1 Data Acquisition

To verify the effectiveness and stability of the proposed approach in practical applications for various IIoT scenarios, we develop an application which runs as a background and transparent process to collect sensor data during the user routine work process. We are allowed to run the designed application and assemble data from the manufacturing plant of Anhui YouKaiPu Electronics Co., Ltd (which is Industrial 4.0 oriented). Table 3.3 lists main experiment data collected from routine operation actions (i.e., walking, scanning, screen-touch, and photographing-uploading) for users (volunteers) who are working in the production process of 'aluminum electrolytic capacitor'. We collect the sensor data corresponding to each concerned operation action according to the timestamp when the action occurred. In dataset ♯1, we collect 800 sequences of operation actions in diverse interference scenarios (i.e., ROE, ME, Glo, and FPC) to construct operation-action features for 104 users. A sequence of operation actions is the random combination of the four typical operation actions and the length of that is 9. In dataset ♯2, we collect 600 (trials) sequences of operation actions for 123 users and take the length of the operation action sequence from 2 to 15.

3.6.2 Experimental Setting

Based on the HBuilderX development environment running in the background of Android mobile devices, we develop an APP to obtain raw sensor data from the APIs provided by the SDK of the Android operation system. The raw sensor data is transferred in real time to the virtual machine, which is running the Windows Server 2008 R2 operation system of the OpenStack private cloud platform. In the virtual machine, we use the Microsoft SQL Server 2008 R2 database to store the raw sensor data, and employ the Matlab R2019a and Microsoft Visual C++ to implement the proposed passive user authentication method.

For the ROE scenario, we adopt the HUAWEI Mate20 X (5G) mobile phone as an industrial mobile device, which has rich resources such as strong computing power and large storage to meet the demands of industrial data processing. For other three scenarios, we conduct data processing using the industrial customized Huawei Honor Play3 smartphone, which serves as the special equipment for workshop.

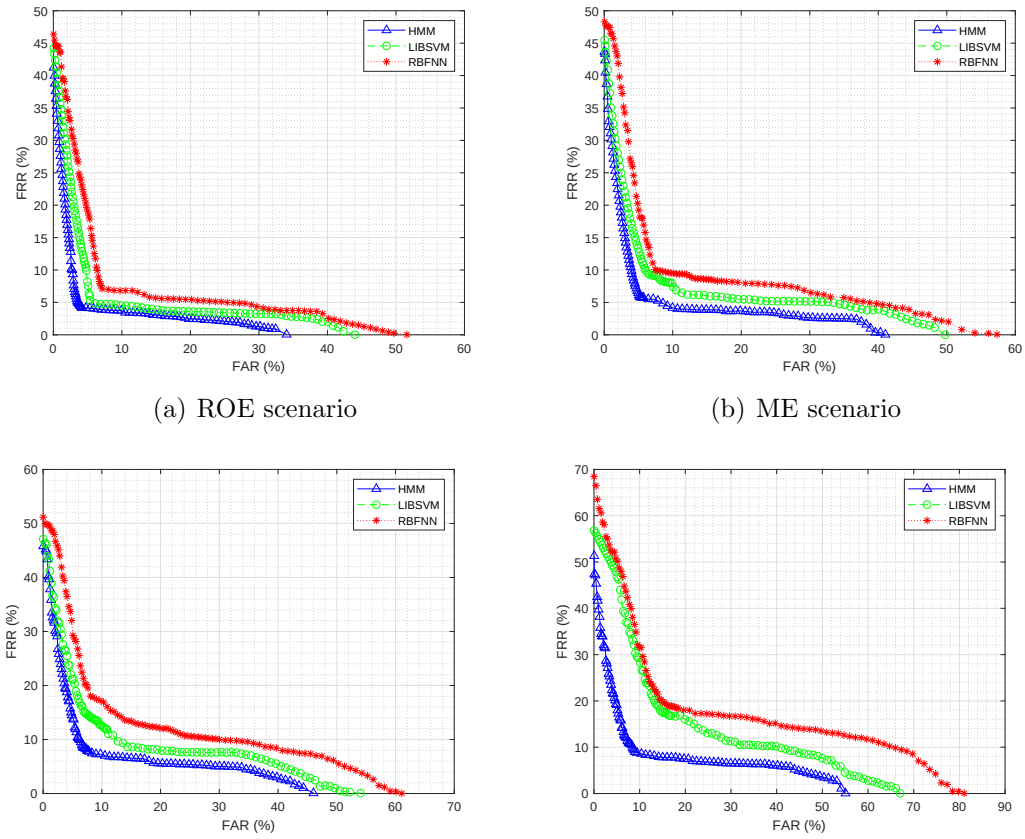
3.6.3 Authentication Performance

In order to show the authentication performance, the HMM-based classifier is compared with two other popular classifiers (i.e., LIBSVM [77] and RBFNN [78]) under the proposed passive authentication framework in Section 3.4. For the LIBSVM classifier, we set the SVM_type as 'one-class SVM' and the type of kernel function as 'radial basis function' [45] to satisfy the category requirements of a large number of users in the practical IIoT scenarios. For the RBFNN classifier, a three-layer neural network [26] is created with n input nodes, $2n$ hidden nodes, and n output nodes, where n is feature dimensionality. We combine multiple feature matrices (i.e., feature matrices of walking, scanning, screen-touch, and photographing-uploading) by concatenating them into one single feature matrix [45]. We then feed 60% random rows of the feature matrix into LIBSVM and RBFNN classifiers for training. Finally, the trained classifiers are used to verify new unknown data samples and thus determine identities of users.

Table 3.4: EER values of different classifiers

EER (%) Scenario / classifier	ROE	ME	Glo	FPC
HMM	4.16	5.74	7.79	8.97
LIBSVM	5.35	8.35	11.08	16.8
RBFNN	7.01	9.65	13.73	18.4

Extensive experiments have been conducted to verify authentication performance based on the dataset #1 in Table 3.3. By setting the number of random operation



(a) ROE scenario

(b) ME scenario

(c) Glo scenario

(d) FPC scenario

Figure 3.8: ROC curves of the proposed authentication approach for four operation-action scenarios under three types of classifiers. (a) ROE scenario. (b) ME scenario. (c) Glo scenario. (d) FPC scenario.

actions for authentication decision as 9, we plot in Fig. 3.8 the ROC curves of three classifiers (i.e., HMM, LIBSVM, and RBFNN) under four IIoT scenarios (i.e., ROE, ME, Glo, and FPC). It is observed from Fig. 3.8 that under the proposed passive authentication framework, the proposed HMM-based classifier outperforms the others in terms of ROC curves while the RBFNN classifier provides the worst authentication performance under the four IIoT scenarios. Specially, when FAR equals to 8%, the HMM-based classifier leads to an FRR of 3.8% and 5.27% for the ROE and ME scenarios, respectively. Even for the Glo and FPC scenarios with strong interference, the HMM classifier can still achieve an FRR of 7.46% and 10.7% when FAR=8%. These results indicate that the HMM-based classifier is promising to adapt to various complicated IIoT environments.

Another observation from Fig. 3.8 is that all the three classifiers achieve the best authentication performance in ROE scenarios, while the best FRRs provided by classifiers LIBSVM and RBFNN are 4.73% and 7.04% when FAR=8%. This is mainly due to the fact that the proposed HMM-based classifier not only utilizes the isolated operation-action features of user actions, but also captures the time-varying nature and dynamic properties of sequential operation actions during the user's routine industrial production process. Conversely, both LIBSVM and RBFNN classifiers ignore time-varying nature and dynamic properties of operation actions from the user's routine production process in IIoT scenarios, leading to the worse authentication performance. Thus, our proposed authentication approach exhibits a good applicability and robustness in IIoT environments.

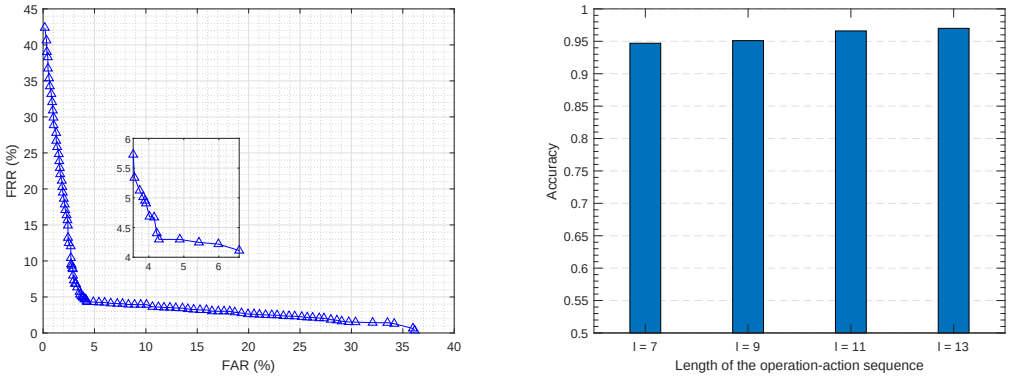
Table 3.4 illustrates that our proposed HMM-based classifier has the best authentication performance in terms of EER for all four scenarios, while the authentication performance based on LIBSVM is relatively better than that of RBFNN. Specifically, the EER value under the HMM-based approach is always less than 9% across four IIoT scenarios. Consequently, the proposed authentication framework under the

HMM-based classifier can provide more accurate and robust authentication services for the IIoT system. In addition, it is notable that the HMM-based classifier is competitive with the classifier in [26] for passive user authentication, where the best EER value even tends to 4% in the scenarios with less interference. This further demonstrates that the operation-action features generated from the routine work process of users are highly distinguishable, so the operation actions can be exploited to perform authentication in IIoT scenarios.

In summary, under all four scenarios considered above, our proposed HMM-based classifier outperforms the LIBSVM and RBFNN classifiers in terms of both the EER value and the ROC curve. Also, we can see from the Fig. 3.8 and Table 3.4 that the ROC curves are always distributed in the lower left corner of the two-dimensional coordinate system and the EER values are always less than 18.4% under each IIoT scenario, indicating that the concerned operation-action features can be used to efficiently characterize the identity of a user and thus can be leveraged to perform passive authentication in IIoT scenarios with different interferences. Therefore, the operation actions generated during the user's routine work might be explored to enhance the traditional authentication methods such as pin-based and pattern-based, and the proposed passive authentication method may serve as a promising alternative to the traditional ones for some special application scenarios like Industrial Internet.

3.6.4 Performance of Resisting Impersonation Attacks

To investigate the performance of resisting impersonation attacks for the proposed passive user authentication approach, we first randomly select 30 users from dataset ♯2 as impersonation attacker (IA) group and further randomly select 30 users from the remaining dataset ♯2 as legitimate user (LU) group. We then conduct the one-to-one randomly pairing between users in the IA group and LU group, and each user in the IA group will impersonate the operation actions of his corresponding



(a) (b)

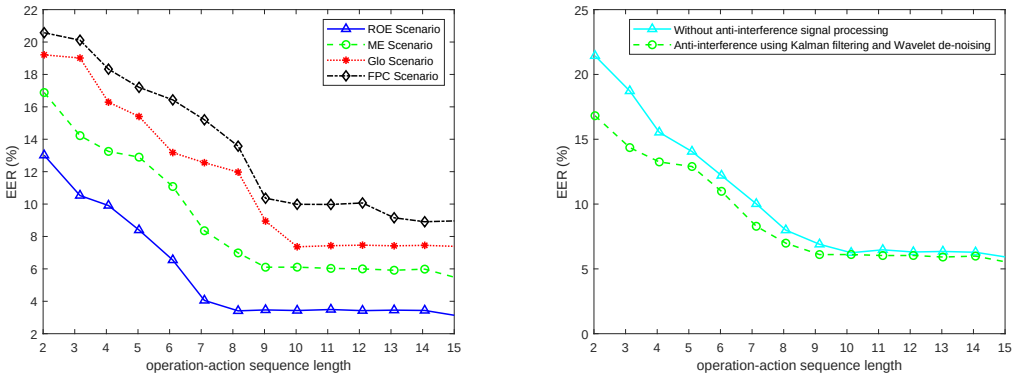
Figure 3.9: Performance of resisting impersonation attacks for the proposed passive user authentication approach. (a) ROC curves. (b) Accuracy.

pair in the LU group. To ensure that adversaries in the IA group are fully familiar with the legitimate users' action behavioral habits and the amplitude of the action space, the adversaries are arranged to work in the same production process as the legitimate users and repeatedly impersonate the legitimate users' operation actions at least 10 times one day for no less than 3 days. Specially, for the ROE scenario, we collect 10 sequential operation action sequences from each user of the two groups to obtain an operation action dataset with impersonation attacks. Finally, we apply the proposed passive authentication approach to perform user authentication based on the impersonation attack dataset.

To show the performance of resisting impersonation attacks for the proposed authentication approach, we present in Fig. 3.9(a) the impact of impersonation attacks on authentication performance in terms of ROC curves based on the impersonation attack dataset with operation action length $l = 9$. As observed from Fig. 3.9(a), the authentication performance in terms of ROC is always distributed in the lower left corner of the two-dimensional coordinate system and the EER value is less than 4.3%, thus the proposed authentication approach exhibits good discriminability and stability for resisting impersonation attacks.

We further present in Fig. 3.9(b) the authentication accuracy performance under the impersonation attack dataset with $l = \{7, 9, 11, 13\}$. It is observed from Fig. 3.9(b) that the accuracy of the proposed authentication approach is 94.7%, 95.1%, 96.6% and 97% under operation action lengths of 7, 9, 11 and 13, respectively. The results in Fig. 3.9(b) indicate that the proposed user authentication approach achieves an excellent performance for resisting impersonation spoofing attacks, and adopting a longer operation-action sequence length will lead to a better authentication accuracy.

3.6.5 Authentication Stability Analysis



(a) EER vs. length of operation-action sequence

(b) Anti-interference capability

Figure 3.10: EERs vary with the length of operation-action sequence. (a) EER vs. length of operation-action sequence for four operation-action scenarios. (b) Anti-interference capability of the proposed framework in the ME scenario.

Same as [26, 45, 52], we adopt the EER metric to characterize the authentication stability against spoofing attacks and show in Fig. 3.10(a) the corresponding results, which are obtained based on the dataset ♯2 in Table 3.3 with randomly selected 63 subjects from various IIoT scenarios. One can see from the Fig. 3.10(a) that for the four scenarios concerned, the EER of the proposed authentication approach monotonously decreases as the operation-action sequence length increases from 2 to 15, but such trend becomes less significant if we increase the sequence length further. It indicates

that when the operation-action sequence length is relatively small, we can get a significant improvement in the authentication performance in terms of EER by increasing the sequence length, but a too large sequence length might not be cost efficient since using more operation actions in the observed sequences will lead to a long authentication time without yielding a significant authentication performance enhancement. Therefore, it is wise to select a suitable operation-action sequence length for various IIoT applications with different authentication performance requirements.

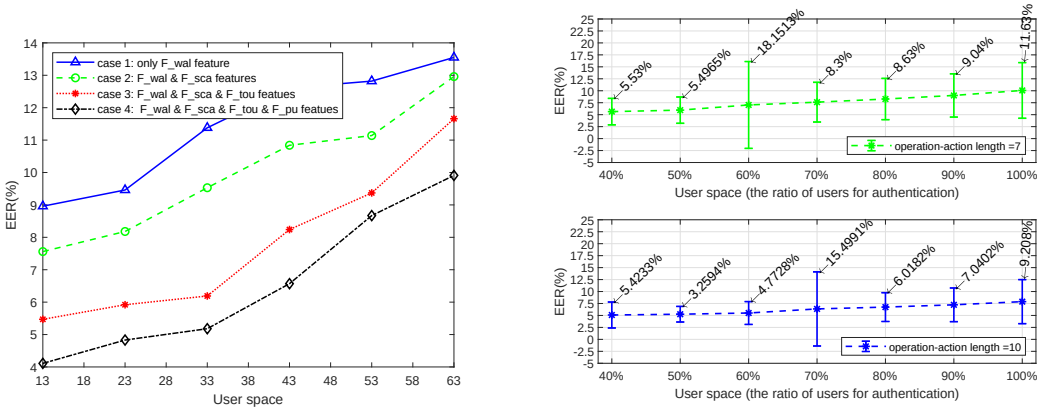
To show the anti-interference capability for the proposed authentication approach, we present in Fig. 3.10(b) the impact of interference on authentication performance in terms of ROC curves based on the dataset ♯2 in Table 3.3 in the ME scenario. It is observed from Fig. 3.10(b) that under the proposed passive authentication framework, the proposed HMM-based classifier using Kalman filtering and Wavelet de-noising outperforms the classifier without anti-interference in terms of ROC curves. Specially, when the length of operation-action sequence is less than 10, the EER value of the classifier without anti-interference fluctuates greatly, and when the length is greater than 10, the EER value of the classifier using Kalman filtering and Wavelet de-noising is always lower than that of the classifier that does not apply anti-interference. We can see that the passive authentication framework designed in this thesis can well weaken or eliminate various interference and noise in the industrial environment, so as to meet the requirements of the ME layer for anti-interference capability.

3.6.6 Scalability to the Number of Features and User Space

1) Scalability to the Number of Features

To evaluate the scalability to the number of features for the HMM-based authentication approach, we present in Fig. 3.11(a) the impacts of user space size and number of features on EER by varying user space size from 13 to 63 and considering four different combinations of features (case 1: utilizing only walking feature; case

2: jointly utilizing walking and scanning features; case 3: jointly utilizing walking, scanning and screen-touch features; and case 4: jointly utilizing walking, scanning, screen-touch and photographing-uploading features). For the sake of clarity, we consider here the ME scenario with operation-action sequence length 7 and randomly selected 63 users in dataset ♯2.



(a) EERs vary with the number of features

(b) EERs vary with the ratio of users for authentication

Figure 3.11: EERs vary with the number of features and user space. (a) EERs vary with the number of features under different user space. (b) EERs vary with the ratio of users for authentication under operation action length being equal to 7 and 10 in the ME scenario (The total number of users is 63).

As observed from Fig. 3.11(a), the authentication performance of case 4 significantly outperforms that of other cases, while the case 1 leads to the worst authentication performance. It indicates that utilizing more operation-action features usually leads to a more accurate characterization of user identities. Also, we can see from Fig. 3.11(a) that by setting user space size to be less than 63, the corresponding EER values of case 1, case 2 and case 3 are under 13.6%. This demonstrates that for a relatively small user space size, three-dimensional operation-action features might be enough to effectively discriminate user identities in IIoT scenarios.

In addition, it can be seen from Fig. 3.11(a) that the EER of each case increases as user space size increases. In particular, as user space size increases from 13 to 63, the

EER of case 1 increases from 8.97% to 13.55%, the EER of case 2 increases from 7.66% to 12.96%, the EER of case 3 increases from 5.47% to 11.66%, while the EER of case 4 only increases from 4.11% to 9.91%. Thus, as the user space size increases, jointly utilizing multiple dimensional operation-action features is helpful for maintaining a better authentication accuracy in practical applications of IIoT systems.

2) Scalability to User Space

To evaluate the scalability to user space for the HMM-based authentication approach, we present in Fig. 3.11(b) how EER varies with the user space size in the ME scenario by considering randomly selected 63 users in dataset ♯2 and the operation-action length of $\{7, 10\}$. The results show that in general EER increases as the size of the user space increases. Specifically, when the user space size increases from 40% to 100% the EER increases from 5.64% to 10.07% under the sequence length of 7. It indicates that the user space upper bound is close to 63 with the constraint of EER $< 10\%$ and sequence length of 7, which agrees with the mathematical analysis in Section 3.5. Another observation from the Fig. 3.11(b) is that the EER under the sequence length of 10 only increases from 5.09% to 7.89%, so increasing the operation action sequence length is an effective method to expand the upper bound of user space. Actually, we can see from the simulation result in Fig. 3.7 in Section 3.5 that the upper bound of user space would be 102 under the sequence length of 10 and the constraint of EER $< 10\%$.

To sum up, the EER increases as the user space becomes large under a given operation action length, and a longer operation-action sequence can lead to a larger upper bound of user space. Therefore, we can obtain a larger upper bound of user space by increasing the length of operation actions in practical applications for various IIoT scenarios.

3.6.7 Sensitivity to Operation-action Features

We explore how the proportion of a certain action feature (the ratio of the number of certain action feature to the total number of action feature in the observed sequence used for authentication) would affect the performance of the proposed HMM-based authentication approach. For brevity, we only show the impact of the screen-touch action proportion on the authentication performance across the four IIoT scenarios.

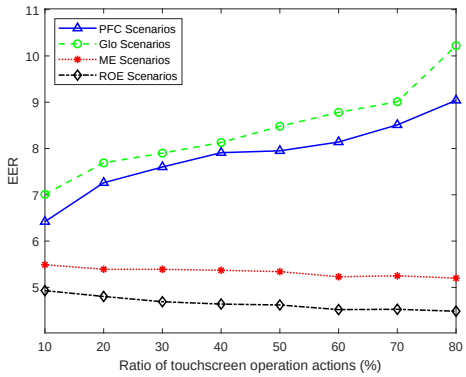


Figure 3.12: EERs vary with the percentage of screen-touch.

As shown in Fig. 3.12, the EER values in Glo and FPC scenarios quickly increase when varying the proportion of screen-touch actions from 10% to 80%. However, the EER values in ROE and ME scenarios tend to decrease slightly when the screen-touch action proportion increases. This may be due to the fact that in Glo and FPC environments, the user's touch behaviors are generally accompanied by interference from gloves and full-body protective clothing worn by the user, which causes touch behavior deformation (such as great amplitude and strong randomness). As a result, the originally stable screen-touch actions become difficult to be captured accurately, which degrades the authentication performance. Thus, by reducing the proportion of screen-touch actions we can effectively improve the performance of authentication in Glo and FPC scenarios. In contrast, in ROE and ME scenarios, increasing the proportion of screen-touch is beneficial to the improvement of the performance. Therefore,

it may be a good choice to enhance the accuracy and stability of the proposed approach by adaptively adjusting the proportion of some actions in various application scenarios of IIoT systems.

3.6.8 Sensitivity to Authentication Time

The authentication time is defined as the time required for the total authentication process, which can be calculated based on the time cost of each process involved in our authentication framework. In the process of raw data collection and preprocessing, our authentication framework first determines the operation-action sequence length l to collect the sensor data used for authentication. Since the data collection as well as its preprocessing are conducted in real time, an empirical setting of the corresponding time consumption $\mathcal{T}_{col}$ is $\mathcal{T}_{col} < 0.2$s. In the feature construction process for operation actions, the operation actions of walking, scanning, screen-touch, and photographing-uploading take about 0.9s, 1.2s, 1.4s, and 1.6s, respectively. Hence, we can obtain the average time consumption of an operation action in the feature construction process as $\mathcal{T}_{con} = 1.28$s. Regarding the time cost $\mathcal{T}_{dim}$ of the dimensionality reduction process for operation-action features and the time cost $\mathcal{T}_{pas}$ of the passive authentication process, an empirical setting under operation action length $l \leq 10$ is given as $\mathcal{T}_{dim} <$ 0.5s and $\mathcal{T}_{pas} < 0.8$s [26, 40].

To sum up, the time consumption of our proposed passive authentication approach is no larger than 10.46s and 14.3s under the operation-action sequence length of $l \leq 7$ and $l \leq 10$ respectively. Such time cost is acceptable to secure the sensitive information in various scenarios of the IIoT system.

3.7 Discussion

For IIoT authentication scenarios with high anti-interference capability requirements, we explore the common behavioral biometrics from sequential operation ac-

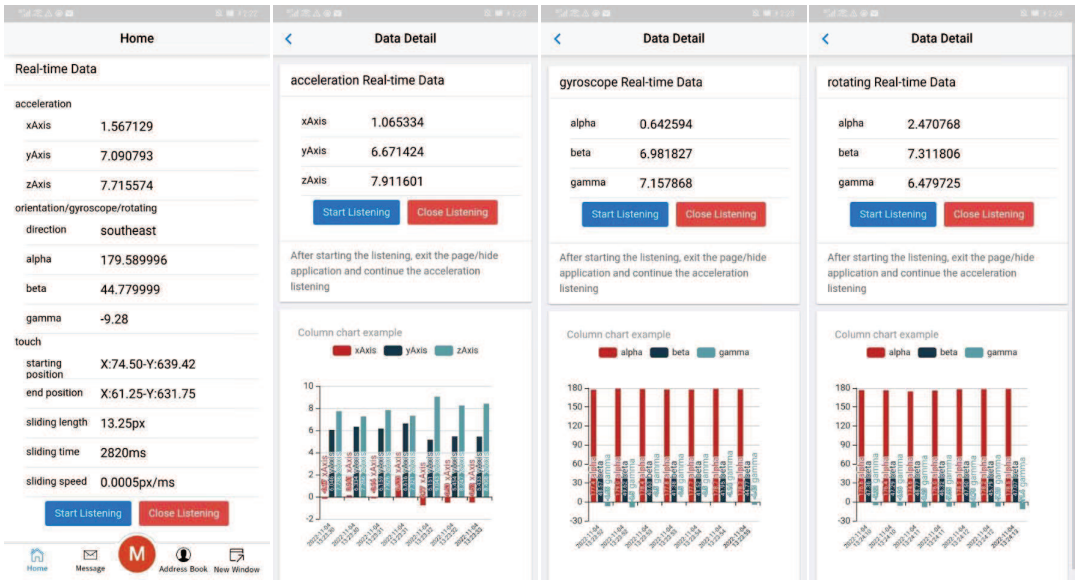


Figure 3.13: Various sensor data triggered by operation actions during user routine work processes through Android smartphone in IIoT systems.

tions in IIoT systems and leverage the Kalman filtering and Wavelet techniques for noise elimination to propose a passive authentication framework for continuous and non-intrusive user authentication against the impersonation attack. As shown in Fig. 3.13, we collect real-time raw data from built-in sensors of mobile devices to depict user operation action features during user routine work processes. Hence, whether sensor data can be accurately collected determines whether we can accurately depict the characteristics of user operation actions, and then determine the authentication performance of the entire user authentication protocol. However, during the practical application of IIoT, there is a lot of electromagnetic interference and noise generated by the interaction of electronic components on the IIoT site. This interference and noise will interfere with the collection of the sensor data, making a certain difference between the collected sensor data and the true value of the sensor.

In various practical IIoT scenarios, we leverage a Kalman filter and Wavelet denoising to reduce the noise along with the data. The interference removal performance is shown in Fig. 3.3 and Fig. 3.4. We can see from Fig. 3.3 that constant noise (i.e.,

gravity components and invariable magnetic fields) in IIoT scenarios can be effectively reduced through Kalman filtering. Since Kalman filtering is recursive and can run in real time, it can meet the real-time and high-efficiency authentication requirements in IIoT systems. We can see from Fig. 3.4 that the wavelet function *Symlets* under the threshold function $Heursure$ is more efficient than $Coiflets$, since it reduces the non-stationary noise and interference while retaining the intrinsic feature of raw sensor data. Therefore, we adopt here the wavelet function *Symlets* under the $Heursure$ threshold to achieve the optimal authentication performance of the proposed approach. More importantly, for the IIoT scenarios with high requirement on anti-interference capability we present in section 3.6 that the new authentication framework enables a flexible authentication performance control to be achieved by adjusting the system parameters like the length of operation sequence, number of features, size of user space, and the proportion of a certain action feature.

However, real IIoT scenarios are extremely complex. The user's actions are far more than the four described in this work, and there are many types of interference, smart devices, and sensors in IIoT scenarios. For simplification, we developed a passive user authentication prototype system where the IIoT scenarios exit some simple constant noise (e.g., gravity components and invariable magnetic fields) and non-stationary interference (e.g., radio frequency signals, changing current and magnetic field) that can be eliminated by simple filtering and denoising algorithms. Our experiments only use mobile smartphones and corresponding built-in sensors to collect sensor data triggered by 4 common operation actions in IIoT scenarios. In the subsequent research, we will consider more complex noise and interference scenarios in IIoT scenarios, and develop more diverse filtering and denoising algorithms to meet the IIoT authentication requirements with high anti-interference capability.

In the process of user operation action feature enrollment, when a new user joins the authentication system, we require the user to complete specific operation actions

according to certain operation procedures, thereby completing user behavioral biometric enrollment and storage. In future work, we will further research and develop the dynamic registration and update method of user operation action features to improve the usability of the authentication framework.

3.8 Summary

For IIoT authentication requirements of high anti-interference capability in the ME layer, this work proposed a novel passive user authentication framework by exploiting the behavioral biometrics from user sequential operations in IIoT systems. We demonstrated that the new framework enables a flexible authentication performance control to be achieved by adjusting the system parameters like the length of operation sequence, number of features, size of user space, and the proportion of a certain action feature. Thus, the proposed framework is promising for satisfying different performance requirements across various IIoT scenarios. Moreover, it is expected that the proposed authentication framework with the HMM-based classifier can serve as a good enhancement and complement to the traditional authentication solutions for IIoT systems.

CHAPTER IV

Authentication Utilizing Consecutive Touch Trajectory Features for the Monitoring and Control (MC) Layer

4.1 Background and Related Work

The basic functions of the MC layer are real-time access control, command transmission, data exchange, data interoperability, modeling, and identification analysis. In the MC layer, IIoT receives the product manufacturing and optimization solutions from the upper layer (i.e., the DO layer), and quickly generates product models. At the same time, this layer collects on-site production element information (people, materials, manufacturing methods, and environment) from the ME layer to form all instructions and data structures for product manufacturing. So we can see that the MC layer often involves critical human-computer interaction, access control, command transmission, and data exchange in IIoT systems, where there are a large number of important real-time instruction uploading and downloading, user access control, and manufacturing process monitoring. In order to ensure the timeliness and real-time of instruction transmission and information collection, the authentication protocol at this layer needs to have real-time performance guarantee.

By now, some research efforts have been devoted to the touch-based behavioral biometric authentication solutions with real-time performance guarantee, such as keystroke patterns-based authentication [42–44], gait-based authentication [45, 46], speaking-based authentication [47, 48] and touch-based authentication [35, 49, 79, 80]. The literature [52] investigates touch movements of user interacting with the touchscreen of a smartphone to obtain a set of 30 touch-based behavioral biometric features, and demonstrates that different users populate distinct subspaces of the touch-based feature space. Then, an SVM-based classification technology is applied to verify user identities continuously based on touch-based behavioral biometric features. The authors in [34] conduct sufficient experiments to illustrate the discriminability and robustness of the intrinsic features from consecutive operation actions (e.g., walking, scanning, screen-touch, and photographing-uploading), the Hidden Markov Model is adopted to determine user identities passively utilizing the one-class classification technique. In [26], smartphone internal sensors are used to measure the behavioral activity characteristics of users when they perform touch actions on the touchscreen of smartphones. Statistic-, frequency-, and wavelet-domain features are extracted from sensor values corresponding to these touch actions to realize continuous user authentication across various operational scenarios.

4.2 Motivation

It is notable that for the MC layer of IIoT, the touch-based behavioral biometric authentication is of particular interest for implementing the continuous and non-intrusive user identity verification. First, due to the increasing development of touchscreen-based devices, IIoT systems usually involve a large number of screen-touch operations during user interacting with their MTs, so user behavioral biometric features can be explored from these screen-touch behaviors to perform user authentication without requiring additional user actions or equipments for the purpose of

authentication. Second, an MT needs a continuous authentication mechanism that can protect a user throughout the entire working session, which complements the initial-login authentication to provide more comprehensive security protection. More importantly, our experiments show that touch-based behavioral biometric authentication methods can meet the requirements of the MC layer for real-time performance guarantee. Thus, we are motivated to design a flexible and cost-effective touch-based authentication approach for continuous and non-intrusive authentication in IIoT systems.

In this chapter, we explore the user consecutive screen-touch actions during routine work processes and propose a more advanced passive authentication method based on both the time-varying characteristics and spatial image characteristics of the user touch trajectory sequences for implicit and non-intrusive user identity verification. In particular, by exploiting touch trajectory sequences constructed through touch trajectories with different velocities from the user routine work process and adopting the HMM to model these sequences, we develop a new method to characterize the behavioral biometric characteristics of users in IIoT scenarios. We then reconstruct each touch trajectory in an image to maintain the shape, relative position and length of the touch trajectory, and adopt average pressure, average curvature and average deviation degree to depict its RGB color in the image. We further design two classifiers corresponding to the above two characteristics. By weighing the outputs of these two classifiers, we thus develop a novel user authentication framework for continuous user authentication in IIoT scenarios. Finally, we conduct extensive experiments to evaluate the performance of the proposed authentication framework in terms of false acceptance rate, false rejection rate and equal error rate, and also examine the related authentication efficiency issues such as the sensitivity to the weights for classifiers, the sensitivity to authentication time and the capability of resisting against impersonation attacks.

4.3 Problem Formulation

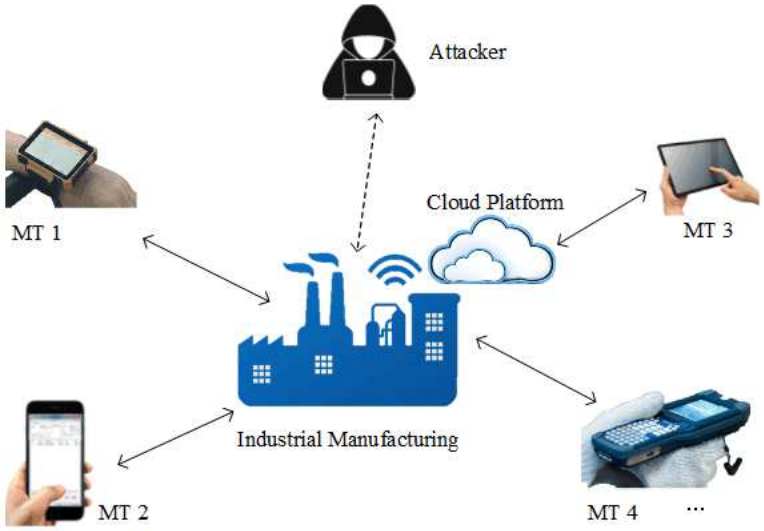


Figure 4.1: IIoT system, where a manufacturing cloud platform exchanges the information involved in industrial production business through MTs with a large number of users in the presence of one potential adversary.

4.3.1 Network Model

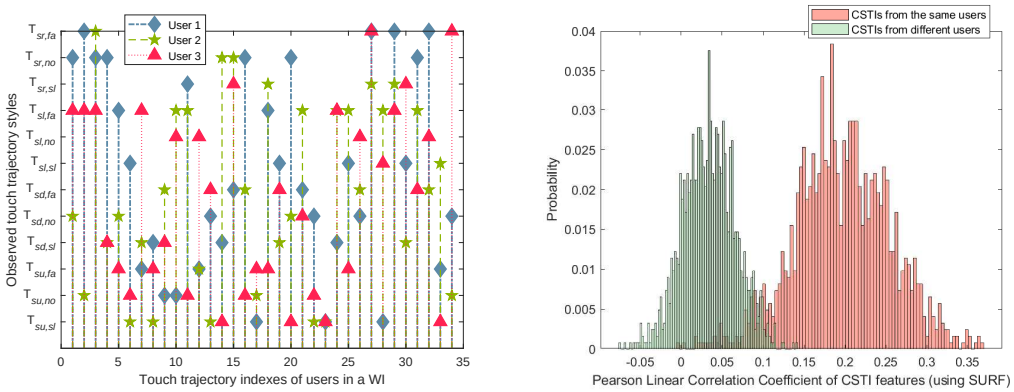
As illustrated in Fig. 4.1, we consider an IIoT system, where a large number of users use MTs to exchange the information involved industrial production business with a manufacturing cloud platform in the presence of one potential adversary. The manufacturing cloud platform provides diverse services (e.g., intelligent manufacturing) for these users simultaneously. The enormous confidential information (e.g., control instructions, business confidential data, and core technologies) is frequently exchanged between the cloud platform and users through the MTs. Due to the open nature of the shared transmission medium, a potential adversary attempts to masquerade as any legitimate user by using the identity of the user in order to inject harmful data and/or attain the information on business confidential data and core technologies, etc.

It is worth noticing that for a given IIoT system, a user generally is engaged

in specific work business according to WIs, and interacts with the cloud platform by performing some common screen-touch operation actions (e.g., sliding up, sliding down, sliding left, and sliding right) on the touchscreens of MTs during routine work processes. Unlike daily interaction with personal mobile smart phones, a user generally interacts with the industrial MT according to designated WIs, which are guiding documents specially prepared to complete a certain or the same type of work and generally specify work contents and work flow in IIoT systems. For example, a WI for starting a motor through the industrial APP is: opening the motor control page, starting the motor self-test, viewing the self-test messages, motor exception handling, abnormal diagnosis, confirming the motor is intact, and turning on the motor switch.

For a user, its consecutive screen-touch trajectory sequence is a series of consecutive touch trajectories generated by the user's interaction with the MTs in order to complete the contents specified in WIs. It is proven through extensive experiments with commercial MT devices that a consecutive screen-touch trajectory sequence under a specific WI from a user can reflect spatial-temporal information of behavioral biometric characteristics for the user. Specifically, we investigate 3675 screen-touch trajectory sequences with sequence lengths from 6 to 17 involving 50 users in IIoT scenarios, and demonstrate that the consecutive screen-touch trajectories are the state-of-the-art solution for user identity characterization in the complex IIoT systems. We provide in Fig. 4.2(a) the differences of time-varying properties from random 3 users, and present in Fig. 4.2(b) the Pearson Linear Correlation Coefficients of STTI features from the same users and different users [81, 82]. We can see from Fig. 4.2 that these consecutive screen-touch trajectories under the IIoT WIs during user routine work process pose unique spatial-temporal variation characteristics among users, and thus can be employed to authenticate users from accessing to the sensitive information of the IIoT system.

To comprehensively examine the availability and practicability of the proposed



(a) (b)

Figure 4.2: The differences of time-varying properties and STTI features from users under a specific WI. (a) Observed touch trajectory style sequences from random 3 users. (b) Pearson Linear Correlation Coefficient of STTI features from the same users and different users.

approach, we consider two common IIoT scenarios: ROE and ME. In the ME scenario, the WIs are generally detailed and rigid, which roughly stipulates the operation processes, operation targets, and even operation methods during production business processes. In contrast, the requirements of WIs in ROE scenarios are relatively loose, and users can obtain a higher degree of freedom in their screen-touch operation.

Based on the above experimental results and analysis, one can see that the spatial-temporal information of consecutive screen-touch trajectory sequences can be utilized to effectively characterize the identity of the user. Therefore, the goal of the manufacturing cloud platform is to decide whether the user currently using MTs is a legitimate user or not, by exploiting the fine-grained features from the series of operation actions on the touchscreen of the MTs.

4.3.2 Threat Model

In the concerned network model, an attacker has access to physical MTs and can capture corresponding passcodes (e.g., smart cards, patterns, and fingerprints) to unlock MTs. Meanwhile, the attacker may be familiar with the business process and behavioral habits of the legitimate users, and attempt to access the IIoT system

through MTs by using identities (e.g., passcodes or other proofs) of legitimate users. As a result, sensitive information involving commercial confidentiality is typically exposed to the attacker through these MTs.

Table 4.1: Common styles of touch trajectories

Touch actions	**Description**
$T_{su,sl}^{\star}$, $T_{su,no}$, $T_{su,fa}$	sliding up slowly, sliding up normally, sliding up fast
$T_{sd,sl}$, $T_{sd,no}$, $T_{sd,fa}$	sliding down slowly, sliding down normally, sliding down fast
$T_{sl,sl}$, $T_{sl,no}$, $T_{sl,fa}$	sliding left slowly, sliding left normally, sliding left fast
$T_{sr,sl}$, $T_{sr,no}$, $T_{sr,fa}$	sliding right slowly, sliding right normally, sliding right fast

* : According to the touch speed of user sliding of the touchscreen, we employ the k-means clustering to divide the user's touch trajectories in a certain direction into three styles, i.e., sliding slowly, sliding normally and sliding fast.

4.4 User Identity Characterization Based on Consecutive Touch Trajectories

The focus of this section is mainly on designing a new authentication method utilizing the spatial-temporal characteristics of consecutive screen-touch trajectory sequences. We first characterize user identities based on time-varying characteristics and STTI characteristics of consecutive screen-touch trajectory sequences, and then apply the two characteristics to develop the corresponding classifiers. Finally, by jointly using the two designed classifiers, we present a weighted continuous user authentication method for the considered IIoT system.

4.4.1 User Identity Characterization Based on Time-varying Touch Trajectory Sequence

We divide touch trajectories involving a user into 12 types (as shown in Table 4.1) according to the sliding direction and speed of the touch trajectories. For a given WI, we then adopt the HMM to model the time-varying features of the successive touch behavior trajectory sequences, and determine user identities using the extracted behavior features.

We use $T_{Dir,Spe}$ to denote the type of a touch trajectory with Dir and Spe representing the set of direction and speed of the touch trajectory, respectively. For simplification, let Dir =$\{su,\ sd,\ sl,\ sr\}$ with su, sd, sl, and sr being sliding up, sliding down, sliding left, and sliding right, respectively. For the trajectories in a certain direction (e.g., su), we further divide the user's touch trajectories into three types by employing the k-means clustering model proposed in [83], i.e., Spe =$\{sl,\ no,\ fa\}$ with sl, no, and fa being slow sliding, normal sliding, and fast sliding, respectively. Hence, as shown in Table 4.1 we can obtain 12 types of trajectories. We use $\Phi = \{\phi_1, \cdots, \phi_M\}$ to denote the set of M styles of touch trajectories listed in Table 4.1 with $\phi_m \in \Phi$ being a certain touch trajectory style, for $m = 1, \cdots, M$, and use $\Psi = \{\psi_1, \ldots, \psi_N\}$ to denote the set of N work contents with $\psi_n \in \Psi$ being a work content specified in a WI (e.g., opening the motor control page, starting the motor self-test, and viewing the self-test messages), for $n = 1, \cdots, N$.

4.4.1.1 Time-Varying Screen-Touch Action Events

Given a specific WI, the resulting touch trajectory sequences from different users can reflect different temporal variation levels due to the individual differences such as muscle movement ability, operation habits, reaction speed, and knowledge level. As shown in Fig. 4.3, we adopt HMM to model the time-varying nature of user sequential touch trajectories.

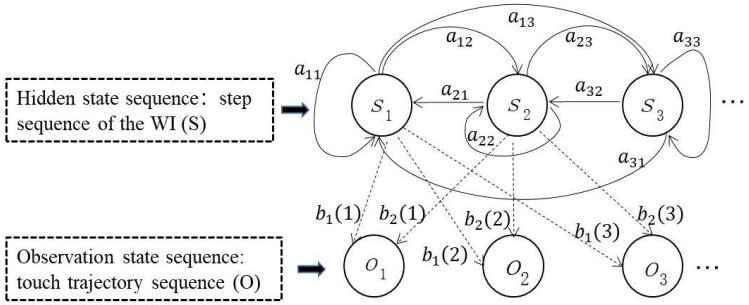


Figure 4.3: Network model for IIoT scenarios.

In particular, we use $O = (o_1, \cdots, o_L)$ to denote a touch trajectory sequence according to the WI, where L is the length of the sequence and $o_i \in \Phi$ denotes the i-th touch trajectory style in O, for $i = 1, \cdots, L$, and use $S = (s_1, \cdots, s_L)$ to denote work content sequences of the WI with $s_j \in \Psi$ denoting the j-th business content in the WI, for $j = 1, \cdots, L$. s_i can be viewed as a hidden state, which is not directly observable. Touch trajectory o_i that can be observed directly through the touchscreen of an MT is generated by the state s_i through a probabilistic function of the HMM. Therefore, O and S are referred as observation state sequence and hidden state sequence in the HMM, respectively.

4.4.1.2 Training of HMM

Similar to that in [26], we also adopt the standard HMM to model the time-varying nature of consecutive touch trajectory sequences, that is, $\lambda = (\pi, A, B)$, where π is the initial state matrix, A is the state transition matrix, and B is the observation matrix. We now present the expressions for π, A, and B. Concretely, we use π_i to denote the probability that the Markov chain will start in state ψ_i, use a_{ij} to denote the probability of transition from state ψ_i $(i = 1, \cdots, N)$ at time t to state ψ_j $(j = 1, \cdots, N)$ at time $t+1$, and use $b_j(k)$ to denote the probability of an observation ϕ_k $(k = 1, \cdots, M)$ being generated from a state ψ_j. Then we have

$$\pi = (\pi_i) = (P(s_i = \psi_i)), i = 1, \cdots, N, \tag{4.1}$$

$$A = [a_{ij}]_{N \times N} = [P(s_{t+1} = \psi_j | s_t = \psi_i)]_{N \times N}, \\ i = 1, \cdots, N, j = 1, \cdots, N, \tag{4.2}$$

$$B = [b_j(k)]_{N \times M} = [P(o_t = \phi_k | s_t = \psi_j)]_{N \times M}, \\ j = 1, \cdots, N, k = 1, \cdots, M. \tag{4.3}$$

Let $\alpha_t(i)$ denote the forward probability when the observation sequence is $\{o_1, o_2, \cdots, o_t\}$ and the state is ψ_i at time t, and then we have

$$\alpha_t(i) = P(o_1, o_2, \ldots, o_t, s_t = \psi_i | \lambda). \tag{4.4}$$

We use $\beta_t(i)$ to denote the backward probability [84] when the observation sequence is $\{o_{t+1}, o_{t+2}, \cdots, o_T\}$ from time $t+1$ to T and the state is ψ_i at time t, and then we have

$$\beta_t(i) = P(o_{t+1}, o_{t+2}, \ldots, o_T | s_t = \psi_i, \lambda). \tag{4.5}$$

Applying the Baum-Welch algorithm [85], the HMM training process is summarized in Algorithm 1.

Algorithm 1 HMM training algorithm

Input: $O = (o_1, \cdots, o_L)$

Output: $\lambda = (A, B, \pi)$

Initialisation :

For n=0, we set $\lambda^{(0)} = (A^{(0)}, B^{(0)}, \pi^{(0)})$.

Loop Process

for $n = 1$ to L **do**

calculate:

$$A^{(n+1)} = a_{ij}^{(n+1)} = \frac{\sum_{t=1}^{L-1} \xi_t(i,j)}{\sum_{t=1}^{L-1} \gamma_t(i)},$$

$$B^{(n+1)} = b_j(k)^{(n+1)} = \frac{\sum_{t=1, o_t=\phi_k} \gamma_t(j)}{\sum_{t=1}^{L} \gamma_t(j)},$$

$$\pi^{(n+1)} = \gamma_1(i),$$

where

$$\gamma_t(i) = \frac{\alpha_t(i)\beta_t(i)}{P(\lambda|O)} = \frac{\alpha_t(i)\beta_t(i)}{\sum_{j=1}^{N} \alpha_t(j)\beta_t(j)},$$

$$\begin{aligned} \xi(i,j) &= \frac{P(s_t = \psi_i, s_{t+1} = \psi_j, O|\lambda)}{P(O_L|\lambda)} \\ &= \frac{P(s_t = \psi_i, s_{t+1} = \psi_j, O|\lambda)}{\sum_{i=1}^{L}\sum_{j=1}^{L} P(s_t = \psi_i, s_{t+1} = \psi_j, O|\lambda)}. \end{aligned}$$

end for

return $\lambda^{(n+1)} = (A^{(n+1)}, B^{(n+1)}, \pi^{(n+1)})$

4.4.1.3 User Identity Verification Based on HMM

Based on the training process of HMM in Section 4.4.1.2, we can obtain the HMM for each user involved in the IIoT system, and authenticate the user identity through probability calculation for a given trajectory sequence. In particular, we use $\aleph = (\chi^{(1)}, \ldots, \chi^{(N_1)})$ to denote the set of screen-touch trajectory sequences from N_1 users in the IIoT system, where $\chi^{(i)}$ is the screen-touch trajectory sequences dataset of the i-th user, for $i = 1, \ldots, N_1$. Based on $\aleph$, we utilize Algorithm 2 to train the HMM for each user, and obtain the set of HMMs $\boldsymbol{\lambda}$ which consist of HMMs of N_1 users. Thus, $\boldsymbol{\lambda}$ is given by $\boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_{N_1})$, where λ_i is the HMM of the i-th user estimated through $\chi^{(i)}$.

Given an unknown touch trajectory sequence O^u, we calculate a probability vector

$$\nu = (P(O^u|\lambda_1), \cdots, P(O^u|\lambda_{N_1})), \tag{4.6}$$

where $P(O^u|\lambda_i)$ is the probability that the touch trajectory sequence O^u is most likely to be generated by λ_i according to forward algorithm [75]. In this work, ν is used to verify user identity based on a threshold predefined in the IIoT system.

4.4.2 User Identity Characterization Based on STTI

We explore the STTI features formed by consecutive touch actions of a user to characterize his identity for continuous user authentication across various IIoT application scenarios. Given a specific WI in IIoT production process, the consecutive touch trajectories from different users during routine work process could generate different images in the touchscreen, which can characterize unique behavioral characteristics of the individuals. We extract user behavior biometric features from STTIs of user touch actions, and conduct a systematic exploration of the stability and discriminability of these features. We then develop an XGBoost-based decision procedure

using one vs. all multi-class classification techniques to perform user authentication.

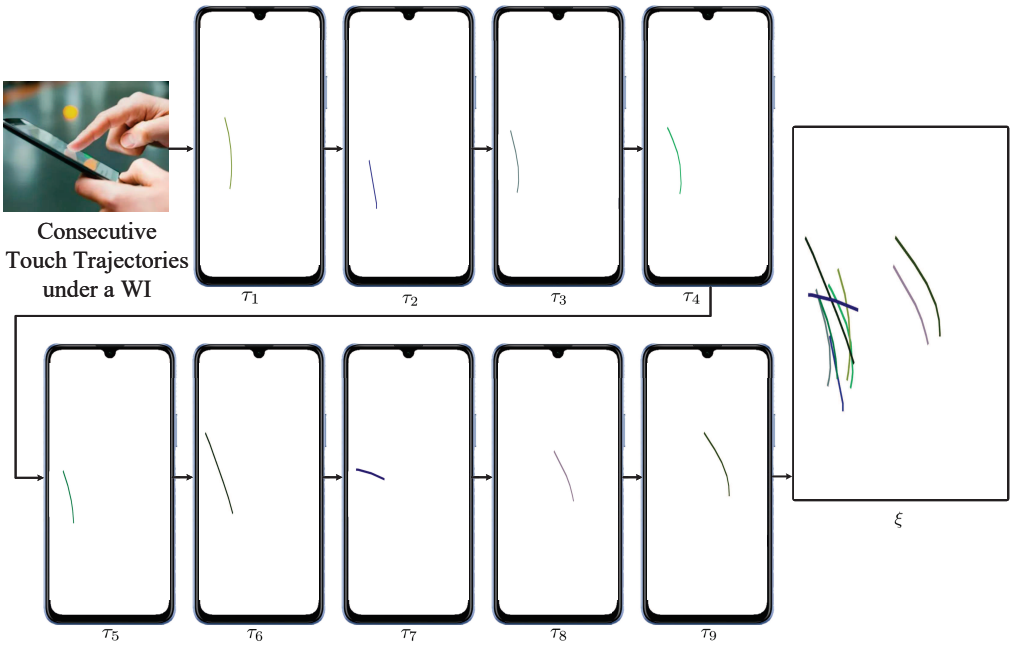


Figure 4.4: STTI consists of 9 trajectories (i.e., $L = 9$).

4.4.2.1 Construction of STTI

To depict the touch trajectory characteristics of length, shape, position, pressure, curvature, orientation, and superimposed effect of multi-touch trajectories formed by user routine touch actions, as shown in Fig. 4.4 we construct an STTI from the screen-touch trajectory sequence $O = (o_1, \cdots, o_L)$ and use $\mathcal{T} = (\tau_1, \cdots, \tau_L)$ to denote the STTI, where τ_i is the i-th trajectory in the STTI, $i = 1, \cdots, L$. We can see that $\mathcal{T}$ is an image which contains all the touch trajectories in O. In particular, the i-th trajectory τ_i in the STTI $\mathcal{T}$ is encoded as a structure $\tau_i =< \{(X_i, Y_i)\}, t_i, C_{\text{RGB}}(R_i, G_i, B_i) >$, where $\{(X_i, Y_i)\}$ is the set of coordinate points swiped by the touch action. $t_i = t_{i,1}, \cdots, t_{i,N_2}$ is the set of time stamp for elements of $\{(X_i, Y_i)\}$ with N_2 representing the sample times of the touchscreen sensor. R_i, G_i and B_i denote the RGB color values of τ_i.

To construct touch trajectories from users in the authentication server of the

IIoT system, we first use coordinate points $\{(X_i, Y_i)\}$ to fit the equation of touch trajectories for reproducing the length, shape, and position of the trajectory τ_i by executing Algorithm 1. Then we proceed to calculate the average pressure, average curvature, and average deviation degree when the user touches the screen to denote the red, green, and blue intensity (i.e., $C_{\mathrm{RGB}}(R_i, G_i, B_i)$) of the trajectory τ_i.

R_i **calculation**: The average pressure R_i of the trajectory τ_i is given by

$$R_i = \frac{1}{N_2} \sum_{j=1}^{N_2} \mathrm{Norm}(\mathcal{F}_{i,t_{i,j}}), \tag{4.7}$$

where N_2 is the pressure sample times. $\mathcal{F}_{i,t_{i,j}}$ is the j-th pressure value collected from the pressure sensor of the MT and $\mathrm{Norm}(\cdot)$ is normalization operation.

G_i **calculation**: We use average curvature of $f_{M_2}(x_i, a)$ to denote the green element value G_i of RGB color for the trajectory τ_i, and G_i is given by

$$G_i = \frac{1}{N_2} \sum_{j=1}^{N_2} \frac{|f''_{M_2}(X_{i,j}, a)|}{(1 + f'^2_{M_2}(X_{i,j}, a))^{\frac{3}{2}}}. \tag{4.8}$$

B_i **calculation**: We use the changes of orientation sensor values of the MT to denote the deviation degree during the formation of τ_i. An orientation sensor has three components Azimuth, Pitch, and Roll. We use Υ_{Az}, Υ_{Pi}, and Υ_{Ro} to denote the orientation sensor component values of Azimuth, Pitch, and Roll, respectively. The values of Υ_{Az}, Υ_{Pi}, and Υ_{Ro} can be obtained from the mobile device rotation angle around z-axis, x-axis, and y-axis, respectively. Let $t_i = \{t_{i,1}, \cdots, t_{i,N_2}\}$ denote the time stamp set of τ_i with $t_{i,j}$ denoting the time when $(X_{i,j}, Y_{i,j})$ is generated, $j = 1, \cdots, N_2$. We use $\boldsymbol{\beta}_{t_{i,j}}$ to represent the orientation sensor vector consisting of the corresponding component values during user routine touch actions. Thus, we have $\boldsymbol{\beta}_{t_{i,j}} = (\Upsilon_{Az,t_{i,j}}, \Upsilon_{Pi,t_{i,j}}, \Upsilon_{Ro,t_{i,j}})$. Let $D(t_{i,j}, t_{i,j+1})$ denote the deviation degree of the MT between time $t_{i,j}$ and $t_{i,j+1}$. $D(t_{i,j}, t_{i,j+1})$ is given by (4.9). Then, we use

$$D(t_{i,j}, t_{i,j+1}) = \sqrt{(\Upsilon_{Az,t_{i,j+1}} - \Upsilon_{Az,t_{i,j}})^2 + (\Upsilon_{Pi,t_{i,j+1}} - \Upsilon_{Pi,t_{i,j}})^2 + (\Upsilon_{Ro,t_{i,j+1}} - \Upsilon_{Ro,t_{i,j}})^2}. \tag{4.9}$$

average deviation degree to characterize the green element value of RGB color in the trajectory, and we have

$$B_i = \frac{1}{N_2 - 1} \sum_{j=1}^{N_2-1} D(t_{i,j}, t_{i,j+1}). \tag{4.10}$$

Algorithm 2 STTI construction

Input: Discrete touchscreen sensor values $\{(\mathcal{C}_X, \mathcal{C}_Y)\}$
Output: $f_{M_2}(x_i, a)$

1: **for** $j = 1$ to N_2 **do**
2: $X_{i,t_{i,j}} = \text{onTouchEvent.Get}(\mathcal{C}_X)$.
3: $Y_{i,t_{i,j}} = \text{onTouchEvent.Get}(\mathcal{C}_Y)$.
4: **end for**
5: Obtain discrete touchscreen sensor values of τ_i, i.e., $\{(X_i, Y_i)\} = \{(X_{i,t_{i,1}}, Y_{i,t_{i,1}}), \cdots, (X_{i,t_{i,N_2}}, Y_{i,t_{i,N_2}})\}$.
6: Adopt Least Squares Polynomial Fit to obtain the equation $f_{M_2}(x_i, a)$ of the touch trajectory τ_i in the screen-touch coordinate system:
$f_{M_2}(x_i, a) = \sum_{j=0}^{M_2} a_j x_i^j$,
7: where M_2 represents the degree of the polynomial and $a = (a_0, \cdots, a_{M_2})$.
8: **return** $f_{M_2}(x_i, a)$

Finally, we construct the STTI from the screen-touch trajectory sequence $O = (o_1, \cdots, o_L)$ by reproducing each element in $\mathcal{T} = (\tau_1, \cdots, \tau_L)$, and the procedure of STTI construction scheme is summarized in Algorithm 3.

4.4.2.2 SURF-based Feature Extraction

It is noticed that among several image feature descriptor and extraction techniques, SURF algorithm approximates or even outperforms previously proposed schemes (such as scale invariant feature transform (SIFT), Binary Robust Independent Elementary Features (BRIEF) and Features from Accelerated Segment Test (FAST))

Algorithm 3 STTI construction

Input: $\mathcal{T} = (\tau_1, \cdots, \tau_L)$
Output: STTI

1: **for** $i = 1$ to L **do**
2: Obtain the equation $f_{M_2}(x_i, a)$ of the touch trajectory τ_i using Algorithm 2.
3: Calculate $C_{\text{RGB}}(R_i, G_i, B_i)$ of the touch trajectory τ_i according to (4.7), (4.8) and (4.10).
4: Draw the image of the trajectory τ_i: Plot$(f_{M_2}(x_i, a), color = C_{\text{RGB}}(R_i, G_i, B_i))$.
5: Hold on
6: **end for**
7: Save the final image contains all trajectories in $\mathcal{T}$, i.e., STTI.
8: **return** STTI

with respect to repeatability, distinctiveness, and robustness [86], yet is adopted to extract STTI features for user identity classification in this work. Based on SURF, the feature extraction procedure for STTI can be summarized as follows.

Step 1: Given a point $\mathbf{X} = (x, y)$ and a scale σ for the STTI $\mathcal{T}$, we use $\mathcal{H}(\mathbf{X}, \sigma)$ to denote the Hessian matrix of point $\mathbf{X} = (x, y)$ at scale σ, and $\mathcal{H}(\mathbf{X}, \sigma)$ is given by

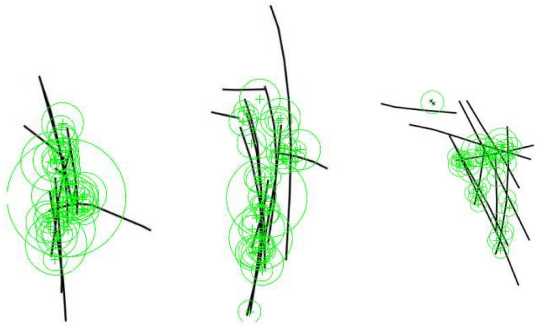
$$\mathcal{H}(\mathbf{X}, \sigma) = \begin{bmatrix} L_{xx}(\mathbf{X}, \sigma) & L_{xy}(\mathbf{X}, \sigma) \\ L_{xy}(\mathbf{X}, \sigma) & L_{yy}(\mathbf{X}, \sigma) \end{bmatrix}, \tag{4.11}$$

where $L_{xx}(\mathbf{X}, \sigma)$, $L_{xy}(\mathbf{X}, \sigma)$ and $L_{yy}(\mathbf{X}, \sigma)$ are the convolution of the Gaussian second order derivative $\frac{\partial^2 g(\sigma)}{\partial x^2}$, $\frac{\partial^2 g(\sigma)}{\partial xy}$ and $\frac{\partial^2 g(\sigma)}{\partial y^2}$ with the STTI $\mathcal{T}$ in point $\mathbf{X}$, respectively. Notice that values of $L_{xx}(\mathbf{X}, \sigma)$, $L_{xy}(\mathbf{X}, \sigma)$ and $L_{yy}(\mathbf{X}, \sigma)$ can be approximated through box filters to reduce the computation time. The determinant of the Hessian matrix for each pixel can be given by $det(\mathcal{H}) = D_{xx}D_{yy} - (\omega D_{xy})^2$, where D_{xx}, D_{xy} and D_{yy} are the approximation for the second order Gaussian partial derivative in x-direction, xy-direction and y-direction, respectively, and ω is an empirical constant.

Step 2: We use a non-maximum suppression in a $3 \times 3 \times 3$ neighbourhood to localize interest points of the STTI $\mathcal{T}$ [87]. In particular, we compare each pixel processed by the Hessian matrix (as demonstrated in Step 1) with all adjacent points

in the image domain and the scale domain (adjacent scale space). If the determinant of Hessian matrix for one pixel point is greater than (or less than) that of all the other adjacent points, the point is a candidate interest point. We provide in Fig. 4.5 the detected interest points for three STTIs. We can see from Fig. 4.5 that different users exhibit distinctive corners, blobs, and T-junctions in their STTIs.

Step 3: We select a reproducible orientation based on information from a circular region around the interest point, and construct a square region aligned to the selected orientation for extracting the SURF descriptor. Then, the square region is split up regularly into smaller 4×4 square sub-regions. Each sub-region has a four-dimensional descriptor vector $\mathbf{v}$ for its underlying intensity structure, and we have $\mathbf{v} = (\sum d_x, \sum d_y, \sum \|d_x\|, \sum \|d_y\|)$, where d_x and d_y are Haar wavelet response in the horizontal direction and vertical direction, respectively. Finally, we can obtain a descriptor vector for all 4×4 sub-regions of length 64 (i.e., SURF-64).



(a) (b) (c)

Figure 4.5: STTI and corresponding distinctive locations of 'interest points' from user interaction with the mobile terminal screen for the same WI. (a) STTI and interest points from User 1. (b) STTI and interest points from User 2. (c) STTI and interest points from User 3.

4.4.2.3 User Identity Verification Based on XGBoost

XGBoost is an ensemble learning method, which uses decision trees as base learners and combines many base learners to make a strong learner. By using the output

of many base classification models in the final prediction, XGBoost is an ideal blend of software and hardware optimization techniques to yield prevalent outcomes by using fewer computing resources in the shortest amount of time. Hence, we adopt XGBoost algorithm to realize the classification of STTI features by setting objective='multi:softprob', 'booster': 'gbtree', and 'max_depth': 20. In the training phase of the XGBoost model, we use $\mathbb{S}_{\mathcal{N}\times\mathcal{M}}$ to denote the training sample space with $\mathcal{N}$ samples and $\mathcal{M}$ labels (i.e., users, each label corresponding to a user). Then, we have $\mathbb{S}_{\mathcal{N}\times\mathcal{M}} = \{(V^{(1)}, L^{(1)}), \cdots, (V^{(\mathcal{N})}, L^{(\mathcal{N})})\}$, where $V^{(i)}$ and $L^{(i)}$ is the feature vector set of the STTIs from the i-th user and the label (user ID) of the user, respectively, $V^{(i)} = (v_1^{(i)}, v_2^{(i)}, \ldots, v_{N_3}^{(i)})$, $i = 1, \cdots, \mathcal{N}$. $\mathbb{S}$ is employed to train the XGBoost model, and the trained model is stored in the authentication server of the IIoT system. In the classification phase, for an STTI feature vector $V^{(u)}$ the XGBoost model outputs a probability vector ϑ, and ϑ is written as

$$\vartheta = (P(L_1|V = V^{(u)}), \cdots, P(L_{\mathcal{M}}|V = V^{(u)})), \tag{4.12}$$

where $P(L_i|V = V^{(u)})$ is the probability that $V^{(u)}$ belongs to label L_i, $i = 1, \cdots, \mathcal{M}$. Finally, the ϑ is leveraged to determine user identity under a preset threshold in the IIoT system.

4.4.3 User Authentication Utilizing Both Time-Varying and STTI Features of Consecutive Touch Trajectories

For a claimed identity X, let us consider his observation screen-touch trajectory sequence $O_X = (o_{X,1}, \cdots, o_{X,L})$ and its corresponding STTI feature V^X. According to (4.6) and (4.12), our authentication method determines if $(X, [O_X, V^X])$ belongs

to class ϖ_1 or ϖ_2 by

$$(X, [O_X, V^X]) \in \begin{cases} \varpi_1, & \omega_2 P(O_X|\lambda_X) + \omega_1 P(L_X|V = \\ & V^X) \geq \varphi, \\ \varpi_2, & \text{otherwise}, \end{cases} \tag{4.13}$$

where λ_X represents the HMM of user X and L_X is the label of X; ω_1 and ω_2 are two weights satisfying $\omega_1 + \omega_2 = 1$; φ is a preset threshold in the IIoT system; ϖ_1 indicates that the claim is true (a legitimate user) and ϖ_2 indicates that the claim is false (an impostor).

4.5 Experiment and Analysis

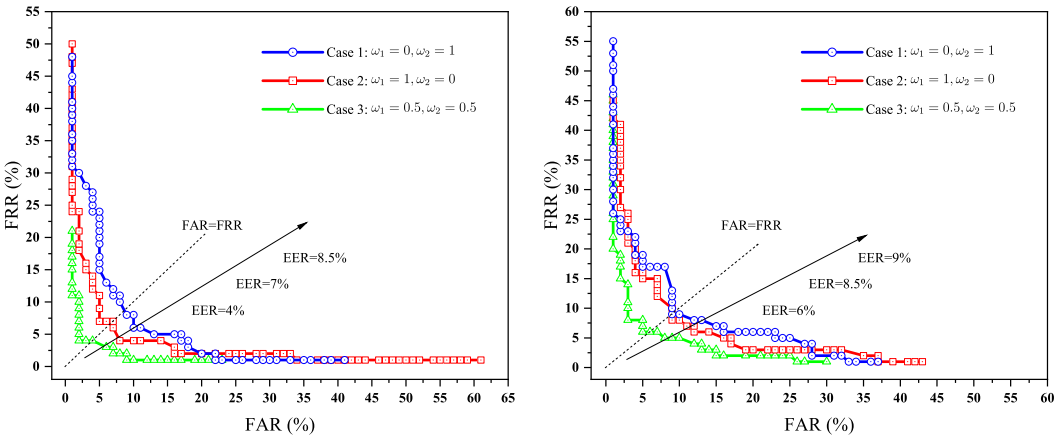
4.5.1 Data Acquisition and Performance Metric

To investigate the performance of the proposed continuous authentication method in practical IIoT systems, we collect 18000 screen-touch trajectory sequences with sequence lengths from 6 to 17 involving 50 users for both ME and ROE scenarios in Anhui Youkaipu Electronics Co., Ltd, which is an IIoT-based company that merges the IoT and cloud computing technologies for intelligent manufacturing. Specifically, in the ME environment users are required to execute an IIoT production process according to a specified WI, i.e., open the manufacturing APP, jump to the production process control page, call the parameter setting page, check the environmental parameter page, enter the number of qualified products, confirm the number of defective products, enter reasons for defective products, submit the form, and jump to the next process. In the ROE scenario, users are required to perform an IIoT R&D process according to a specified WI, i.e., open R&D process control APP, new product entry, product parameter entry, experiment process recording, product quality optimization, product quality evaluation, data analysis, and experiment report generation. In our

experiments, we collect these touch-screen trajectories during these WIs to construct spatial-temporal features of users for determining their identities.

To evaluate the performance of the proposed continuous authentication method, we first calculate three typical metrics, namely the FAR, FRR and EER [26]. Specifically, FAR is the ratio between the number of falsely accepted unauthorized users and the total number of imposters, and FRR is defined as the ratio between the number of falsely denied legitimate users and the total number of legitimate users. We then use FAR and FRR together to generate ROC curve to show the tradeoff between FAR and FRR under preset threshold values, and EER is calculated as the sensitivity of the classifier where FAR = FRR.

4.5.2 Authentication Performance Analysis



(a) ME scenario (b) ROE scenario

Figure 4.6: ROC curves of the proposed continuous user authentication framework for two IIoT scenarios (i.e., ROE and ME) under three cases (case 1: $\omega_1 = 0$, $\omega_2 = 1$, i.e., only using time-varying features based on HMM; case 2: $\omega_1 = 1$, $\omega_2 = 0$, only using STTI features based on XGBoost; case 3: $\omega_1 = 0.5$, $\omega_2 = 0.5$, i.e., jointly utilizing spatial-temporal touch-screen trajectory features based on HMM and XGBoost. (a) ROE scenario of IIoT. (b) ME scenario of IIoT.

To illustrate the impact of dimensions (i.e., time-varying nature dimension and spatial variation dimension of sequential screen-touch trajectories) of screen-touch

trajectory characteristics on the authentication performance, we consider three different combinations of weights (case 1: $\omega_1 = 0$, $\omega_2 = 1$, i.e., only using time-varying features based HMM; case 2: $\omega_1 = 1$, $\omega_2 = 0$, i.e., only using STTI features based on XGBoost; case 3: $\omega_1 = 0.5$, $\omega_2 = 0.5$, i.e., jointly utilizing time-varying features and STTI features based on HMM and XGBoost, respectively), and plot in Fig. 4.6 the corresponding ROC curves based on FAR and FRR. For the sake of clarity, we consider here the ROE and ME scenarios with touch trajectory sequence length 17 involving 50 users.

It is observed from Fig. 4.6 that under the proposed continuous authentication framework, authenticating users jointly utilizing spatial-temporal touch-screen trajectory features (case 3) outperforms the others in terms of ROC curves while authenticating users only utilizing the time-varying features (case 1) obtains the worst authentication performance in two IIoT scenarios. It indicates that utilizing spatial-temporal screen-touch trajectory features usually leads to a more accurate characterization of user identities. Specifically, in the ROE scenario, authenticating users integrating spatial-temporal identities has EER of 4%. Also, in the ROE scenario where users have high screen-touch operation freedom and interference during routine IIoT work process, the proposed method can still achieve EER of 6%. This indicates that the continuous user authentication using spatial-temporal screen-touch trajectory features from consecutive touch trajectories is promising to adapt various complicated IIoT application scenarios.

4.5.3 Sensitivity to Weights of ω_1 and ω_2

To explore how weights of two classifier weights ω_1 and ω_2 ($\omega_1 = 1 - \omega_2$) in (4.13) would affect the performance of the proposed continuous user authentication approach, we adopt the screen-touch trajectory length 17 and the number of users 50 in ME and ROE scenarios, and present in Fig. 4.7 the impact of ω_1 on EER by

varying ω_1 from 0 to 1 across the two IIoT scenarios. As shown in Fig. 4.7, the EER when $\omega_1 = 0$ (i.e., using only the time-varying features of scree-touch trajectory sequences) and the EER when $\omega_1 = 1$ (i.e., using only the STTI features of the touch trajectories) are always larger than that when $0 < \omega_1 < 1$ in both ROE and ME scenarios, and we can obtain the optimal EER values of 3.85% and 5.77% for ME and ROE scenarios with the settings of $\omega_1 = 0.4$ and $\omega_1 = 0.55$, respectively. Therefore, by reasonably adjusting the weights of the two weights (ω_1 and ω_2) the authentication performance of the proposed approach can be flexibly controlled to adapt to various IIoT scenarios. Another observation from Fig 4.7 is that due to more standardized operation action restrictions on work contents and work flows according to WIs in ME and thus a high discriminability among users there in terms of the spatial-temporal information, the EER in the ME scenario is always better than that in the ROE scenario.

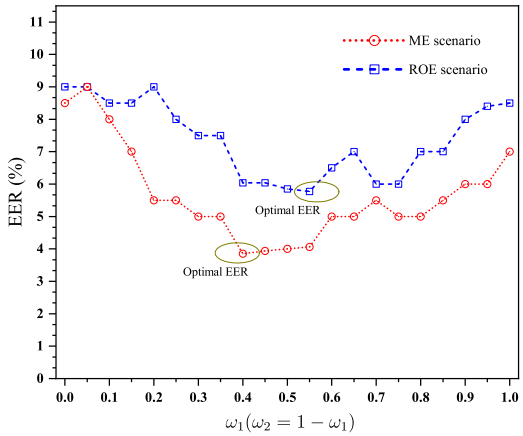


Figure 4.7: Authentication performance in terms of the usability to weights of two classifiers (i.e., HMM based classifier and XGBoost classifier).

4.5.4 Usability to Operation Length

We show in Fig. 4.8 the impact of the number of trajectories (i.e., the length of the successive screen-touch trajectories L used for user authentication) on the authentication performance under the settings of L from 7 to 17. We can see from

Fig. 4.8 that the EER of the proposed authentication framework monotonously decreases as L increases from 7 to 17, but such trend becomes less significant if we further increase the number of trajectories L. It indicates that when the length of the trajectory sequence is relatively small, we can get a significant improvement in the authentication performance in terms of EER by increasing the number of trajectories L, but a too large number of trajectories might not be cost efficient since using more screen-touch trajectories in the continuous authentication framework will lead to a long authentication time without yielding a significant authentication performance enhancement. Therefore, it is wise to select a suitable number of trajectories for various IIoT applications with different authentication performance requirements.

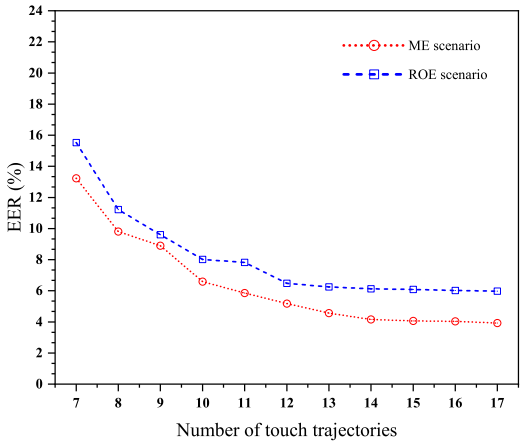


Figure 4.8: The impact of the length of the successive screen-touch trajectories L used for user authentication on EER.

4.5.5 Scalability to User Space

To evaluate the scalability to user space for the touch based continuous authentication approach, we present in Fig. 4.9 how EER varies with the user space size by considering randomly selected $10 \backsim 50$ users. The results show that in general EER increases as the size of the user space increases. Specifically, when the user space size increases from 10 to 50 the EER increases from 2.06% to 5.98% and 1.46% to 3.95% in ROE and ME scenarios under $L = 17$, respectively.

It shows that the value of EER increase as the size of the user space becomes larger, especially for small user space. Specifically, there is a significant increase in the authentication error rate in the interval between 10 and 30 users. This is as expected, since a larger number of legitimate users usually means a higher probability that two legitimate users have similar profiles. We also observe that when the user size is larger than about 30 users, the EERs become relatively stable, and only small fluctuations with the error range are apparent. These results indicate that the user size in our analysis should be (at least) larger than 30, in which case the influence of user space may be minimal. These results also indicate that our subject size is located in a range where the influence could be negligible.

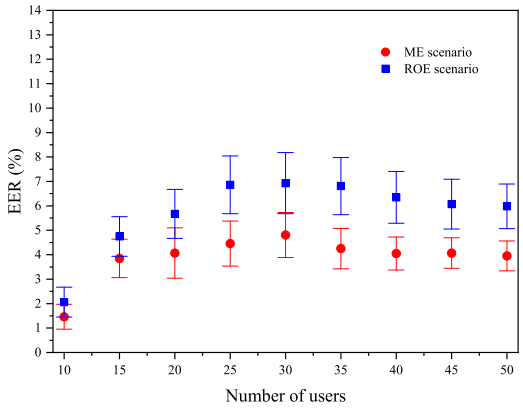


Figure 4.9: The scalability to user space for the touch-based continuous authentication framework.

4.6 Discussion

For IIoT authentication scenarios with high real-time requirement, we propose a new user authentication framework based on the spatial-temporal features of screen-touch trajectories for continuous user authentication in practical IIoT scenarios, in which every time a user touches the screen, the IIoT authentication system can verify the user's identity by analyzing the time-varying features of touch trajectory sequences and cumulative screen-touch trajectory image characteristics, thus ensuring

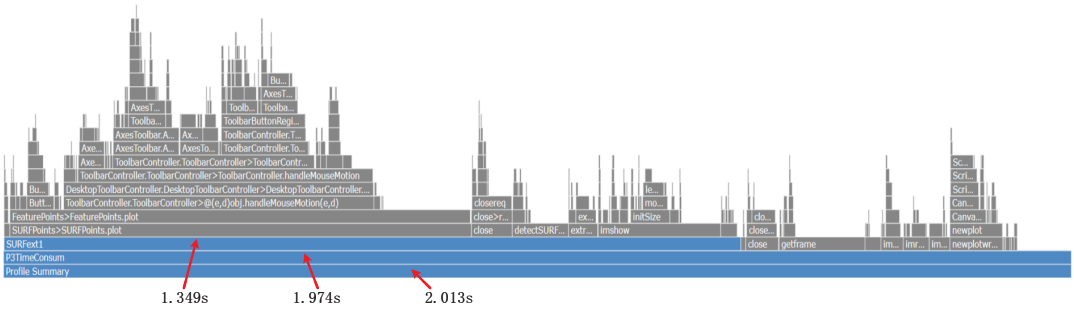


Figure 4.10: Flame graph of algorithm time consumption ($L = 9$).

the real-time user authentication. In this work, we draw the STTIs of multiple touch action trajectories of users at one time, and use SURF algorithm to quickly extract the identity features of STTIs. Every time a user touches the screen of MTs, we use the trajectories generated by the current screen-touch action to generate an STTI together with the previous multiple screen-touch trajectories, and use these STTI features and the time-varying features of touch trajectory sequences to quickly determine the legitimacy of the user's identity. As shown in Fig. 4.10, we present the flame graph of algorithm time consumption for STTI construction and feature extraction. We can see from Fig. 4.10 that our proposed authentication protocol can complete STTI construction and feature extraction in 1.974 seconds when the number of touch trajectories equals 9. Note that the classifiers for user identity classification are offline trained, and the authentication model can give the classification output within 1 second for a given test sample. This means that our algorithm can give the user's identity decision in about 3 seconds, and conduct the identity authentication almost every 3 seconds.

However, we also note that this algorithm that only relies on touch actions to perform user identity authentication lacks authentication basis from multiple perspectives (multiple modes or actions), so its security is slightly inadequate compared with the algorithm that uses multiple operations. In the future research work, we will try to add new features from different modes and operation actions to construct

the user identity, and improve the robustness and reliability of the algorithm while ensuring the real-time performance guarantee.

In the process of user screen-touch spatial-temporal feature enrollment, when a new user joins the authentication system, we require the user to complete specific screen-touch actions according to concerned certain WIs, thereby completing enrollment and storage of user screen-touch spatial-temporal features. In future work, we will further research and develop the dynamic registration and update method of user screen-touch spatial-temporal features to improve the usability of the authentication framework.

4.7 Summary

For IIoT authentication requirements of high real-time performance in the MC layer, this paper proposed a novel continuous authentication framework based on spatial-temporal screen-touch trajectory features. We demonstrated that the new framework enables a flexible and efficient authentication performance control to be achieved by adjusting the weights for classifiers, the screen-touch trajectory length, and the number of user space. Thus, the proposed framework is promising for satisfying different authentication performance requirements across various IIoT scenarios. Moreover, it is expected that the new authentication framework with the two proposed classifiers can serve as a good enhancement and complementary to the traditional authentication solutions for IIoT systems.

CHAPTER V

Authentication Utilizing Two-Dimensional Features for the Decision and Optimization (DO) Layer

5.1 Background and Related Work

The basic functions of the DO layer are decision making, optimization, description, diagnosis, business operations, and operation management. In the DO layer, the IIoT system encapsulates key technologies, algorithms, operational strategies, important customer information, financial and business data, manufacturing optimization strategies, and operational management technologies. With the integration of IIoT, cloud computing and big data, artificial intelligence platforms and intelligent decision-making systems, etc. are usually integrated in this layer. We can see that in the DO layer a large amount of confidential information and sensitive data (such as finance, core technology, core algorithms, operation and sales strategies, crucial customer information, and key management technical services) are generated, stored and exchanged. Therefore, the user authentication at this layer usually requires the authentication protocol to have high security performance.

By now, some research efforts have been devoted to the study of passive user

authentication with high security performance. In [26], the authors utilize kinematic information sequences of multi-motion sensor behavior for passive user authentication when the user interacts with his smartphone, and also propose a decision procedure based on Hidden Markov Model (HMM) to characterize the behavioral biometric feature space such that the continuous user identity verification can be implemented across various operational scenarios. The authors in [52] demonstrate the discriminability and robustness of features related to screen-touching behaviors, and then apply these features to develop a passive authentication solution for smartphone users. The authors in [53] show that it is possible to distinguish profiles of users by exploring CSI information even when they possess similar CSI fingerprints. They also design a practical user authentication approach based on the fine-grained CSI features to accurately determine the user identities in both lab and apartment environments. The literature [65] exploits the CSI of WiFi signals to extract the gesture features (like push, swing, and wave) and some identity-related imperceptible features, and then applies the HMM and Fresnel Model to develop a robust and efficient user authentication approach to determine user identities in IoT environments.

5.2 Motivation

It is notable that when applying existing one-dimensional feature-based user passive authentication approaches in modern IIoT systems, it is usually difficult to accurately depict the user identities and thus to achieve an acceptable user authentication performance based on only one-dimensional characteristics. First, users in IIoT systems usually just follow the requirements of industrial production businesses to conduct some basic operations over their MTs in a standardized manner, making it difficult to accurately characterize the user identities with only the time-varying behavioral biometric features extracted from their operation actions. Second, the IIoT system shares a relatively uniform electromagnetic and space environment, so users

there show a strong location correlation and thus a low discriminability in terms of the CSI spatial variation characteristics [53, 54]. However, our results in this work indicate that by jointly exploiting the two-dimensional features of the time-varying characteristics of user sequential operation actions and spatial variation characteristics of CSI caused by these actions, we can not only provide a full spatial-temporal characterization of user identities but also significantly improve the performance of passive user authentication.

In this chapter, we develop a novel two-dimensional passive authentication framework by jointly utilizing both the time-varying characteristics of the user sequential operation actions and spatial variation characteristics of CSI caused by these actions. In particular, by constructing time-varying operation action sequences from the routine work process of a user and adopting the HMM to model these sequences, we develop a new method to characterize the behavioral biometric characteristics of users in IIoT scenarios. We then propose a new approach to depict the spatial-temporal variations of CSI related to a user, in which the WiFi CSI data related to the user is first sliced to reduce the noise and interference from the random actions of the user, then the multi-domain features from the CSI data are extracted and the XGBoost model is applied to characterize these features. We further design two classifiers corresponding to the above two characteristics. By combining these two classifiers and assigning each classifier an appropriate weight, we thus develop a novel two-dimensional user authentication framework for passive, continuous and non-intrusive user authentication in IIoT scenarios. Finally, we conduct extensive experiments to evaluate the performance of the proposed authentication framework in terms of false acceptance rate, false rejection rate and equal error rate, and also examine the related authentication efficiency issues such as the sensitivity to the weights for classifiers, the sensitivity to authentication time and the capability of resisting against impersonation attacks.

5.3 Problem Formulation

5.3.1 Network Model

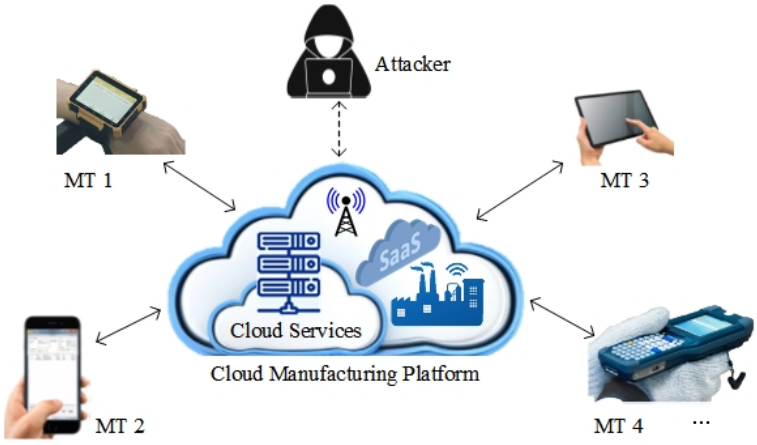
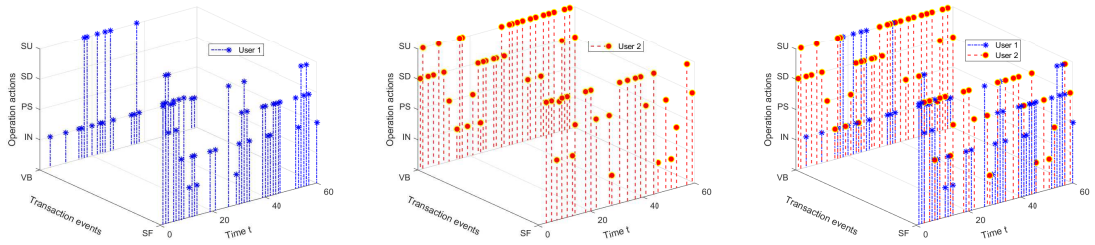


Figure 5.1: Network model for IIoT scenarios.

Consider an IIoT system consisting of multiple legitimate users with various MTs, a cloud manufacturing platform and a potential attacker, as shown in Fig. 5.1. In the IIoT system, users holding MTs always interact with the cloud manufacturing platform by performing some common operation actions (e.g., scanning, inputting, and sliding screen) on the MTs during industrial production processes. Legitimate users generally send/receive sensitive information (e.g., control instructions, business confidential data, and core technologies) to/from the cloud platform through operation actions on the MTs. The potential attacker may impersonate as legitimate users to launch a spoof attack by implementing a series of operation actions on the MTs, and thus hopes to acquire sensitive information (e.g., confidential information) from the IIoT system.

For a user, its operation action sequence (OAS) is a sequence of successive operation actions collected from his routine work process. It is noticed that time-varying OASs from each user can reflect a unique behavioral biometric characteristic of the user [26]. We provide in Fig. 5.2 the differences of OASs' time-varying properties

between User 1 and User 2, containing 60 sequential time points, 2 transaction events and 4 operation actions (the definition of transaction events and operation actions is listed in Table 5.1). We can see from Fig. 5.2 that the OAS from a user possesses its unique time-varying nature, which can be used to describe the user's identity in the IIoT scenarios. More importantly, the variations of WiFi CSI caused by operation



(a) (b) (c)

Figure 5.2: Differences of OASs' time-varying properties from different users. (a) The operation actions and the transaction events from User 1 change over time (t). (b) The operation actions and the transaction events from User 2 change over time (t). (c) The comparison of operation time-varying properties between User 1 and User 2.

actions from the user present a unique spatial-temporal characteristic due to the path loss and multi-path effects of the wireless channels [88]. Thus, we attempt to combine the characteristics of operation actions as well as channel CSI to design a passive user authentication framework for IIoT scenarios.

5.3.2 Threat Model

In the concerned network model, an attacker has access to physical MTs and can capture corresponding passcodes (e.g., smart cards, patterns, and fingerprints) to unlock MTs. Meanwhile, the attacker may be familiar with the business processes and behavioral habits of the legitimate users, and attempts to access the IIoT system through MTs by using identities (e.g., passcodes or other proofs) of legitimate users. As a result, sensitive information involving commercial confidentiality is typically exposed to the attacker through these MTs. Hence, the goal of our work is to design

a passive authentication framework for the IIoT system, which discriminates user identities continuously and non-intrusively through the tiny difference of the user's operation actions on MTs.

5.4 Proposed Passive Authentication Framework

In this section, a flexible and cost-effective passive user authentication framework is developed to determine user identities for the IIoT system, which exploits the behavioral biometric characteristics and the channel CSI patterns of users' operation actions from their routine work processes. As illustrated in Fig. 5.3, the proposed authentication framework consists of three processes: 1) User identity characterization based on behavioral biometric features; 2) User identity characterization based on CSI features; 3) User authentication jointly utilizing two-dimensional features.

5.4.1 User Identity Characterization Based on Behavioral Biometric Features

5.4.1.1 Time-varying operation action events

During the routine work process of a user, the user sequential operation actions with MTs are actually activated by a series of transaction events (such as submitting form, confirming operation, and browsing) encapsulated in the background of industrial APPs and/or cloud services to complete specific business functions in IIoT systems. Some common operation actions and transaction events are listed in Table 5.1.

To construct sequential OASs and transaction event sequences during user routine work processes, we use $\Phi = \{\mathrm{SC}, \mathrm{SR}, \mathrm{SL}, \mathrm{SU}, \mathrm{SD}, \mathrm{IN}, \mathrm{PS}\}$ to denote the set of operation actions and use $\Psi = \{\mathrm{SF}, \mathrm{CO}, \mathrm{VB}, \mathrm{PR}, \mathrm{DO}, \mathrm{UP}\}$ to denote the set of transaction events, where the definition of the elements in Φ and Ψ is described in Table 5.1. We

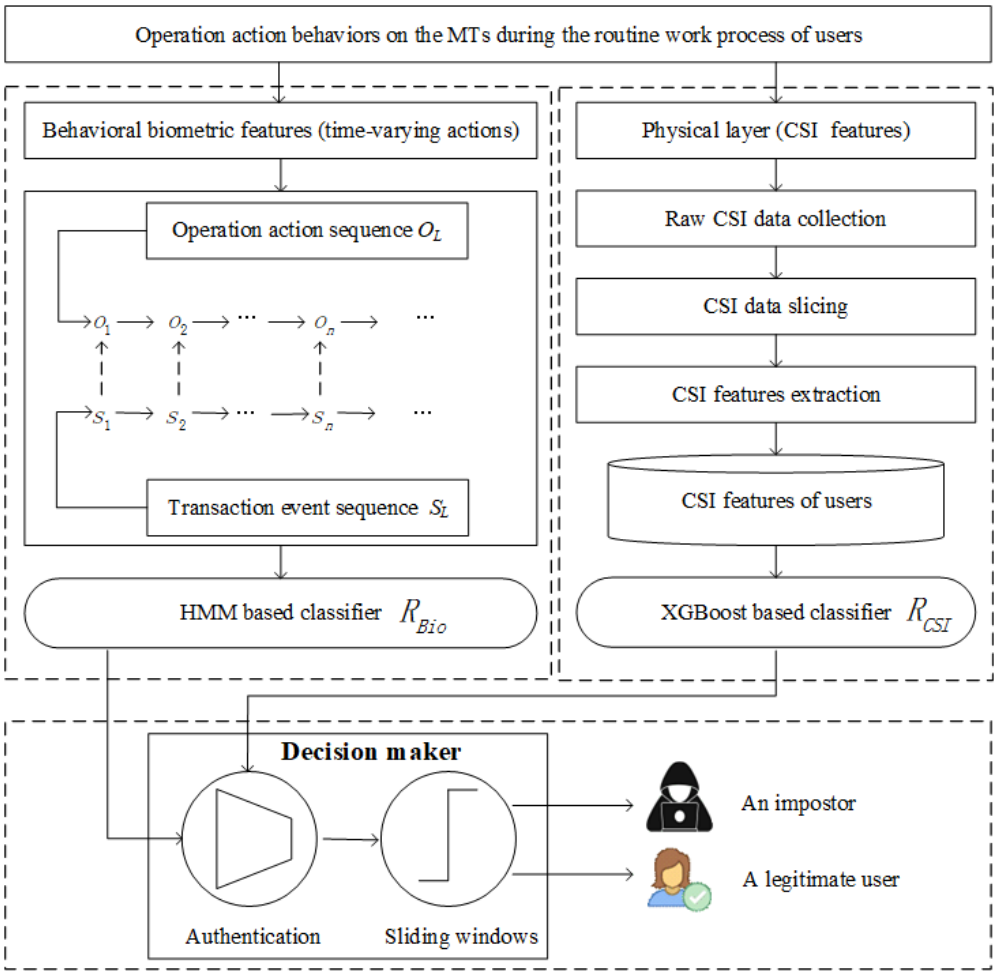


Figure 5.3: The processes of the proposed two-dimensional passive authentication framework for IIoT scenarios.

Table 5.1: Common operation actions and transaction events

Transaction events	**Description**	**Operation actions**	**Description**
SF	submitting form	SC	scanning Q/R
CO	confirming operation	SR	sliding right
VB	viewing and browsing	SL	sliding left
PR	printing	SU	sliding up
DO	downloading	SD	sliding down
UP	uploading	IN	inputting
		PS	pressing screen

use $O_L = (o_{L,1}, \cdots, o_{L,L})$ to denote an OSA of length L, and use $S_L = (s_{L,1}, \cdots, s_{L,L})$ to denote the corresponding transaction event sequence of O_L, where $o_{L,i} \in \Phi$ denotes the i-th operation action in the O_L and $s_{L,i} \in \Psi$ is the transaction event that activates the $o_{L,i}$, $i = 1, \cdots, L$. We further use t_i to denote the sampling time of $o_{L,i}$.

5.4.1.2 User identities modeling based on Hidden Markov Model

As shown in Fig. 5.4, we apply the HMM to characterize behavioral biometric characteristics of a user during his routine work processes. The state transition between operation actions in O_L can be observed directly, and a probabilistic function of the actual (hidden) states in S_L activates the observed states. Therefore, the transaction event sequence O_L is regarded as the observation state sequence, and the transaction event sequence S_L represents the hidden state sequence. Here, we use

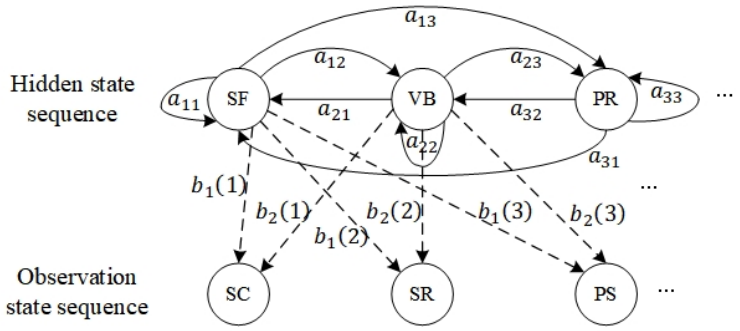


Figure 5.4: Observation states and hidden states of the HMM model.

$\lambda = (A, B, \pi)$ to denote the discrete HMM model in our framework with the parameters of the state transition probability matrix A, the observation probability matrix (obfuscation matrix) B and the vector of initial state probability π, as shown in Fig. 5.4. Applying λ, the process of operation action verification can be regarded as a probability evaluation of generating the observation sequence O under λ, based on three probability parameters of λ described as follows.

a) The state transition probability matrix A is written as

$$A = [a_{ij}]_{N \times N}, \tag{5.1}$$

where a_{ij} is the adjacency occurrence from transaction event (hidden state) $S_{L,i}$ at time t to the transaction event (hidden state) $S_{L,j}$ at time $t+1$, and a_{ij} is written as

$$a_{ij} = P(S_{L,j} = \emptyset_j | S_{L,i} = \emptyset_i), i = 1, \ldots, N; j = 1, \ldots, N, \tag{5.2}$$

where N is the number of elements in Ψ, $\emptyset_i$ and $\emptyset_j$ are the i-th and the j-th element in Ψ respectively.

b) The observation probability matrix B is denoted by

$$B = [b_j(k)]_{N \times M}, \tag{5.3}$$

where $b_j(k)$ is the probability of generating the operation action (observation state) ζ_k when the transaction event (hidden state) is $\emptyset_j$, and $b_j(k)$ is given by

$$b_j(k) = P(O_{L,k} = \zeta_k | S_{L,j} = \emptyset_j), k = 1, \ldots, M; j = 1, \ldots, N, \tag{5.4}$$

where M is the number of elements in Φ, ζ_k is the k-th operation action in Φ.

c) The vector of initial state probability denoted by π is given by $\pi = (\pi_i)$, where π_i is the probability of being in state $\emptyset_i$ at time $t = 0$ (initial time), and π_i is written as

$$\pi_i = P(i_1 = q_i), i = 1, 2 \ldots, N. \tag{5.5}$$

5.4.1.3 The multi-class classifier design based on HMM

According to [89–91], the authentication process for a user is regarded as a binary classification problem. Specifically, based on (5.1) $\backsim$ (5.5), the HMM model of the i-th user to be authenticated is given by $\lambda_i = (A_i, B_i, \pi_i)$. Given the observed OAS O_L from the user, the probability $P(O_L|\lambda_i)$ is calculated to authenticate the user's identity under a given threshold. Suppose the number of users in the IIoT system is N, we first develop N binary classifiers based on HMM for each legitimate user. Then, based on the N binary classifiers, we obtain a one-versus-all multi-class classifier denoted by R_{Bio} to determine user identities by using behavioral biometric features of users.

5.4.2 User Identity Characterization Based on CSI Features

According to [53, 54], the operation actions performed by a user present a unique spatial-temporal characteristic due to the path loss and multi-path effects of the wireless channels [88]. Hence, we characterize identities of users by exploiting the statistic-, frequency- and wavelet-domain features of channel CSI signals, and perform user authentication utilizing a well-known XGBoost algorithm based on the fine-grained CSI characteristics [92].

5.4.2.1 Raw CSI data collection

For various IIoT application scenarios, we leverage the CSI data corresponding to user operation actions to extract the unique spatial variation characteristics of CSI caused by these operation actions. In this work, the CSI data is collected from the Anhui Youkaipu Electronics Co., Ltd, an IIoT-based company that merges the IoT and cloud computing technologies for intelligent manufacturing. We provide the space layout of environments used for collecting raw CSI data in Fig. 5.5, where the location layout of rooms, equipments, walls, and interference sources are from

parts of real manufacturing environments (ROE and ME) of the company. In the IIoT application scenario, we utilize a public CSI tool [93] to record the WiFi data packets transmitted from a dual antenna commercial access point (namely TP-LINK-TL-CPE300D access point) to a desktop computer equipped with 3-antenna wireless network card (Intel 5300 NIC).

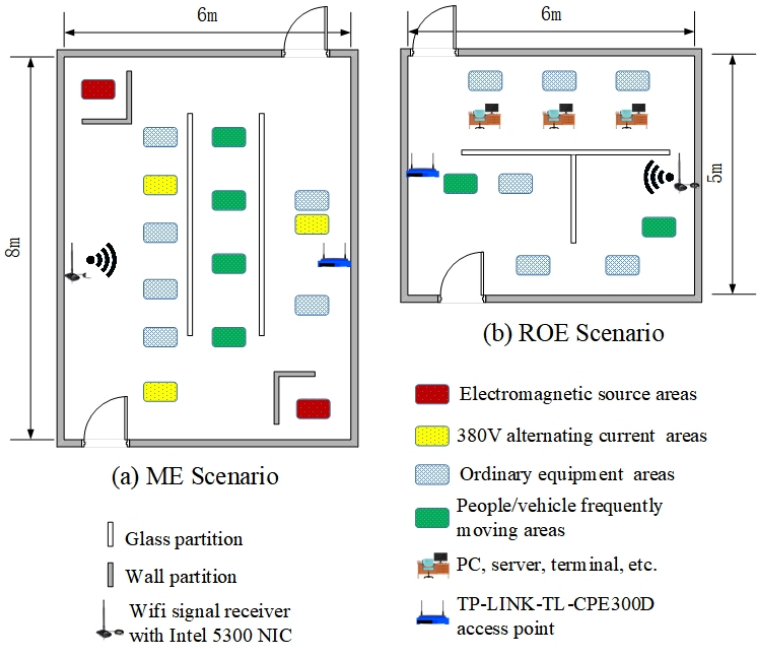
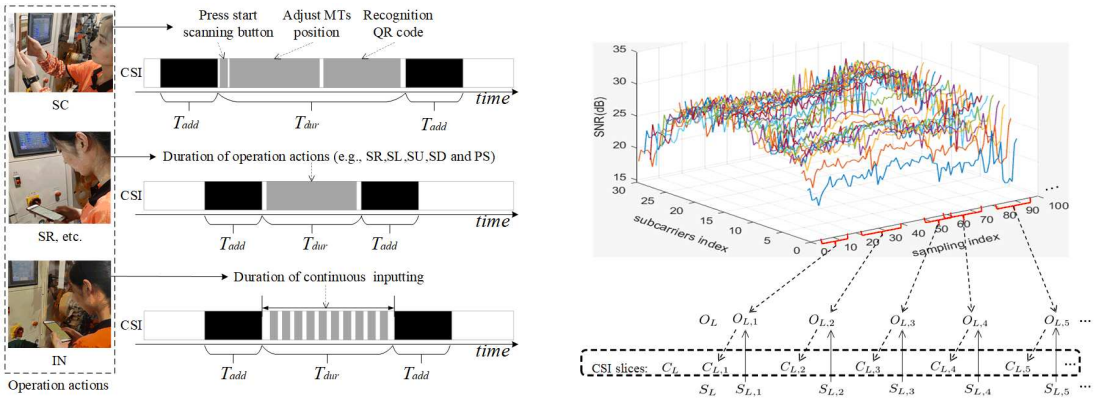


Figure 5.5: The layout of the rooms used to collect data in the IIoT system. (a) ME scenario. (b) ROE scenario.

5.4.2.2 CSI signal slicing

In this work, we use the CSI data corresponding to user routine operation actions to extract the unique spatial variation characteristics of user action behavior. In principle, we should use all CSI data generated during a user's work processes in the extraction of CSI features to obtain a better authentication performance. However, the CSI data outside of the operation action periods is commonly generated by random actions from the user, so it does not contain useful information of operation action features. Also, including such CSI data in the characterization of user identity

yields high computational complexity and random action interference without much real benefit in terms of authentication performance. To improve the authentication effectiveness and practicability of the proposed authentication framework, we perform CSI signal slicing shown in Fig. 5.6 and use a time offset before and after the time duration of an operation action to ensure that the CSI signal of our slice comes from the full cycle of user action.



(a) (b)

Figure 5.6: CSI signal slicing. (a) Intercepting the CSI signals of each operation action during the time $T_{dur}+2\times T_{add}$. (b) Slicing the CSI signals according to the time when the operation actions occur.

As shown in Fig. 5.6(a) that for the operation action $o_{L,i}$ in O_L, we use T_{dur_i} to denote its time duration and use T_{add_i} to denote a time offset before and after the time duration T_{dur_i}. As shown in Fig. 5.6(b) that we use $C_{L,i}$ to denote the CSI data slice during the time $T_{\mathrm{dur}_i} + 2 \times T_{\mathrm{add}_i}$, so the CSI slice sequence C_L of OAS O_L is determined as $C_L = (C_{L,1}, \cdots, C_{L,L})$. Finally, we use C_L to extract the spatial variation characteristics of CSI caused by user OAS O_L.

5.4.2.3 CSI feature extraction

To obtain fine-grained characteristics of CSI for each subcarrier to accurately depict user identities utilizing the channel CSI data, we extract three-domain features from CSI data C_L including 39 attributes. They are time domain features (ID: 1 ∽

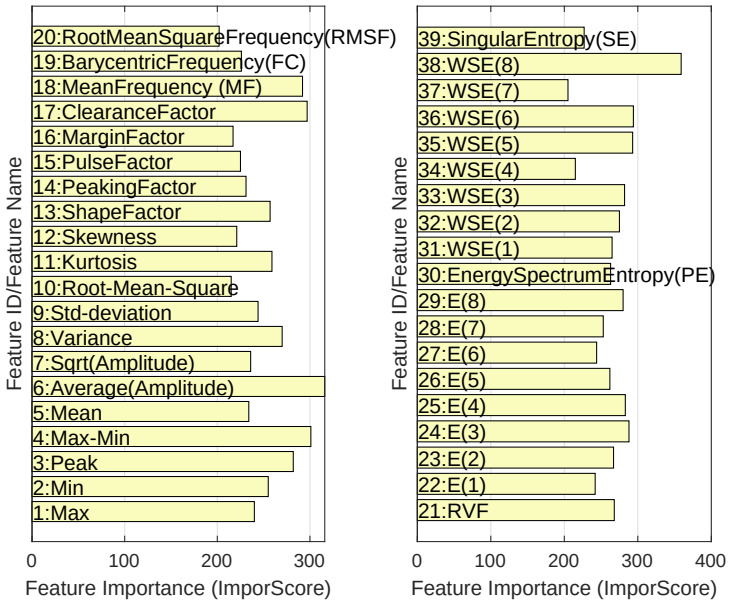


Figure 5.7: Three-domain CSI features extracted from C_L and corresponding feature importance.

17), frequency domain features (ID: 18 $\backsim$ 21), and wavelet domain features (ID: 22 $\backsim$ 39), as shown in Fig. 5.7. Specifically, the time domain features (ID: 1 $\backsim$ 17) are self-explanatory. For frequency domain features (ID: 18 $\backsim$ 21), we use F_{MF} to represent the mean of frequency f through FFT transform of CSI signal C_L. Let F_{FC}, F_{RMSF} and F_{RVF} denote the barycenter frequency, root mean square frequency and variance frequency, respectively. For a given frequency f, its frequency amplitude $s(f)$ can be obtained through FFT transform of CSI signal C_L, and its F_{FC}, F_{RMSF} and F_{RVF} can be written as

$$F_{\mathrm{FC}} = \frac{\sum fs(f)}{\sum s(f)}, \tag{5.6}$$

$$F_{\mathrm{RMSF}} = \sqrt{\frac{\sum f^2 s(f)}{\sum s(f)}}, \tag{5.7}$$

$$F_{\mathrm{RVF}} = \sqrt{\frac{(f - F_{\mathrm{FC}})^2 s(f)}{\sum s(f)}}. \tag{5.8}$$

To obtain the wavelet domain characteristics of CSI signals, we first perform three-layer wavelet packet decomposition on C_L to gain the wavelet packet tree denoted by τ. Then, we reconstruct the wavelet packets including 8 nodes in the third layer of τ, and obtain the wavelet packet coefficients of the 8 nodes represented by $\{\aleph_1, \ldots, \aleph_8\}$. We use Υ_i to denote the energy of $\aleph_i$, $i = 1, \cdots, 8$, and we have $\Upsilon_i = Norm(\aleph_i)$, where $Norm(\cdot)$ represents 2 norm calculation. The total energy of the 8 nodes denoted by E_{total} and the proportion of energy of the i-th node denoted by $E(i)$ are written as

$$E_{total} = \sum_{i=1}^{8} \Upsilon_i, \tag{5.9}$$

$$E(i) = \Upsilon_i / E_{\text{total}}, i = 1, 2, \ldots, 8, \tag{5.10}$$

respectively. We can see that $\{E(i)\}$ $(i = 1, 2, \ldots, 8)$ are the features corresponding to ID 22 $\backsim$ 29 in Fig. 5.7. The wavelet energy spectrum entropy for each node in the third layer of τ demoted by H_{PE} is calculated as

$$H_{\text{PE}} = -\sum_{i=1}^{8} E(i) \ln(E(i)), i = 1, 2, \ldots, 8. \tag{5.11}$$

We use H_{WSE} to denote wavelet Shannon entropy of 8 nodes in the third layer of τ, and H_{WSE} is given by

$$H_{\text{WSE}}(i) = -\sum_{i=1}^{N} \aleph_i^2 \ln(\aleph_i^2), \tag{5.12}$$

where i is the node in the third layer of τ, $i = 1, 2, \ldots, 8$ and N is the length of $\aleph_i$. We can see that $\{H_{\text{WSE}}(i)\}$ $(i = 1, 2, \ldots, 8)$ are the features corresponding to ID 31 $\backsim$ 38 in Fig. 5.7. Let H_{SE} denote wavelet singular entropy, and H_{SE} can be calculated as

$$H_{\text{SE}} = -\sum_{i=1}^{m} \frac{r(i)}{\sum_{i=1}^{m} r(i)} \ln \left(\frac{r(i)}{\sum_{i=1}^{m} r(i)} \right), \tag{5.13}$$

where $r(i)$ is the i-th singular value generated by the singular value decomposition [94] of τ and m is the number of singular values in $\{r(i)\}$.

5.4.2.4 The multi-class classifier based on XGBoost model

To solve multi-class classification problem for user authentication, we first construct feature datasets for each legitimate user by using the fine-grained three-domain CSI features. Supposing the number of users in the IIoT system is N, we then develop N binary classifiers based on XGBoost for each legitimate user. Finally, based on the N binary classifiers, we obtain a one-versus-all multi-class classifier denoted by R_{CSI} to determine user identities by using CSI features of users.

5.4.3 User Authentication Jointly Utilizing Two-Dimensional Features

The process of user authentication for the IIoT system can be regarded as a user identity verification process based on the two designed classifiers: R_{Bio} and R_{CSI}. To construct the sample space for training R_{Bio} and R_{CSI}, we use $\mathbb{S}_{N\times M}$ to denote the training sample space with N samples and M labels (users), use $\mathbb{S}^{+}_{\text{Lab}_m}$ to denote all the samples with label Lab_m ($m \in [1, M]$) in $\mathbb{S}_{N\times M}$, and use $\mathbb{S}^{-}_{\text{Lab}_m}$ to denote samples with labels except Lab_m. Then we can obtain M sub-sample sets denoted by $\mathbb{S}_{\text{sub}}$,

$$\mathbb{S}_{\text{sub}} = \left\{ \begin{bmatrix} \mathbb{S}^{+}_{\text{Lab}_1} \\ \mathbb{S}^{-}_{\text{Lab}_1} \end{bmatrix}, \begin{bmatrix} \mathbb{S}^{+}_{\text{Lab}_2} \\ \mathbb{S}^{-}_{\text{Lab}_2} \end{bmatrix}, \ldots, \begin{bmatrix} \mathbb{S}^{+}_{\text{Lab}_m} \\ \mathbb{S}^{-}_{\text{Lab}_m} \end{bmatrix} \right\}, \tag{5.14}$$

where $m \in [1, M]$, "+" and "-" denote the positive sample label and the negative sample label, respectively.

5.4.3.1 Training of R_{Bio}

For the m-th user with sub-sample set $\mathbb{S}_{(\text{sub},\text{lab}_m)}$ defined in (5.14) and $m = 1, 2, \ldots, M$, the HMM model for the user is given by $\lambda = (A, B, \pi)$. Given an OAS

O_L from the user, we attempt to train the HMM model for tuning the parameters of λ (namely the state transition matrix A, the observation matrix B, and the initial state distribution π), so that the model is maximally like the observed sequences O_L.

Under λ and O_L, we use $\gamma_t(i)$ to denote the probability that the state is $\emptyset_i$ at time t, and $\gamma_t(i)$ is given by

$$\gamma_t(i) = P(S_{L,t} = \emptyset_i | O_L, \lambda) = \frac{P(S_{L,t} = \emptyset_i, O_L|\lambda)}{P(O_L|\lambda)}. \tag{5.15}$$

Let $\alpha_t(i)$ and $\beta_t(i)$ denote the forward probability and backward probability [84] when the state is $\emptyset_i$ at time t, and then we have

$$\alpha_t(i) = P(O_{L,1}, O_{L,2}, \ldots, O_{L,t}, S_{L,t} = \emptyset_i|\lambda), \tag{5.16}$$

$$\beta_t(i) = P(O_{L,t+1}, O_{L,t+2}, \ldots, O_{L,L}|S_{L,t} = \emptyset_i, \lambda). \tag{5.17}$$

By combining (5.16), (5.17) and (5.15), $\gamma_t(i)$ can be re-written as

$$\gamma_t(i) = \frac{\alpha_t(i)\beta_t(i)}{P(\lambda|O)} = \frac{\alpha_t(i)\beta_t(i)}{\sum_{i=1}^{L} \alpha_t(j)\beta_t(j)}. \tag{5.18}$$

We user $\xi(i,j)$ to denote the probability that the state is $\emptyset_i$ at time t and the next state is $\emptyset_j$ at time $t+1$ under λ and O_L, and $\xi(i,j)$ is given by

$$\begin{aligned} \xi(i,j) &= \frac{P(S_{L,t} = \emptyset_i, S_{L,t+1} = \emptyset_j, O_L|\lambda)}{P(O_L|\lambda)} \\ &= \frac{P(S_{L,t} = \emptyset_i, S_{L,t+1} = \emptyset_j, O_L|\lambda)}{\sum_{i=1}^{L}\sum_{i=1}^{L} P(S_{L,t} = \emptyset_i, S_{L,t+1} = \emptyset_j, O_L|\lambda)}. \end{aligned} \tag{5.19}$$

According to Baum-Welch algorithm [95, 96], the training process of λ under

observation sequence O_L can be described as follow.

a) For n=0, the initial parameter values of the model λ are $A^{(0)} = a_{ij}^{(0)}$, $B^{(0)} = b_j(k)^{(0)}$ and $\pi_{(0)} = \pi_i^{(0)}$. Thus, the initial model denoted by $\lambda^{(0)}$ is given by

$$\lambda^{(0)} = (A^{(0)}, B^{(0)}, \pi^{(0)}). \tag{5.20}$$

b) For $n = 1, 2, \ldots, L$, we perform iterations

$$A^{(n+1)} = a_{ij}^{(n+1)} = \frac{\sum_{t=1}^{L-1} \xi_t(i,j)}{\sum_{t=1}^{L-1} \gamma_t(i)}, \tag{5.21}$$

$$B^{(n+1)} = b_j(k)^{(n+1)} = \frac{\sum_{t=1, O_{L,t}=\zeta_k} \gamma_t(j)}{\sum_{t=1}^{L} \gamma_t(j)}, \tag{5.22}$$

$$\pi^{(n+1)} = \gamma_1(i), \tag{5.23}$$

where $\gamma_t(i)$ and $\xi_t(i,j)$ can be calculated according to (5.18) and (5.19).

c) Finally, we obtain parameters A, B, and π of λ after all iterations are completed, and the training process is terminated.

By repeating the above training processes of HMM model for each user, we obtain M trained sub-classifiers in our R_{Bio} classifier. Then we can determine user identities based on the output of the classifier R_{Bio}.

5.4.3.2 Training of R_{CSI}

For a given sub-sample set $\mathbb{S}_{(\text{sub},\text{lab}_m)}$ defined in (5.14) with N samples, the training samples for the XGBoost model can be denoted by $\{(x_i, y_i)\}$, $i = 1, 2, \ldots, N$, $(x_i, y_i) \in \mathbb{S}_{(\text{sub},\text{lab}_m)}$. We can see that x_i and y_i represent the vector of CSI features and corresponding label of the i-th sample in $\mathbb{S}_{(\text{sub},\text{lab}_m)}$, respectively. Let $\hat{y}_i$ denote a tree ensemble model which uses K additive functions to predict the output corre-

sponding to x_i, and then we have

$$\hat{y}_i = \phi(x_i) = \sum_{i=1}^{K} f_k(x_i), f_k \in \digamma, \tag{5.24}$$

where $\digamma = \{f(x) = \upsilon_{\vartheta(x)}\}$ denotes the space of classification and regression tree (CART) [92, 97], ϑ is the structure of each regression tree that maps a training sample to the corresponding leaf index, f_k represents the input-output functional relationship of the k-th regression tree, each f_k corresponds to an independent tree structure ϑ and leaf weights υ. We use υ_i to represent score on the i-th leaf of the regression tree in ϑ.

To obtain f_k through the training of XGBoost model, we first define an objective function with regularization term $\mathcal{L}(\phi)$ as

$$\mathcal{L}(\phi) = \sum l(y_i, \hat{y}_i) + \sum \Omega(f_k), \tag{5.25}$$

where

$$\Omega(f) = \gamma T + \frac{1}{2}\chi\|\upsilon\|^2, \tag{5.26}$$

l is a differentiable convex function which represents the difference between the predicted value $\hat{y}_i$ and the actual value y_i, T is the number of leaves in the tree, $\Omega(f_k)$ is the penalty term of the complexity for each f_k to prevent the over fitting and also control the total number of leaf nodes.

According to boosting algorithm, we iterate Equation (5.24) as

$$\text{Iteration } 0 : \hat{y}_i^{(0)} = 0, \tag{5.27a}$$

$$\text{Iteration } 1 : \hat{y}_i^{(1)} = f_1(x_i), \tag{5.27b}$$

$$\text{Iteration } 2 : \hat{y}_i^{(2)} = f_1(x_i) + f_2(x_i) = \hat{y}_i^{(1)} + f_2(x_i), \tag{5.27c}$$

$$\vdots$$

$$\text{Iteration } t : \hat{y}_i^{(t)} = \sum_{i=1}^{t} f_k(x_i) = \hat{y}_i^{(t-1)} + f_t(x_i). \tag{5.27d}$$

By combining (5.27d) and (5.25), the objective function in (5.25) at the t-th iteration can be written as

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t). \tag{5.28}$$

Here, the second-order approximation is employed to quickly optimize the objective function by greedily adding the f_t according to (5.25) [98], and $\mathcal{L}^{(t)}$ is approximately calculated as

$$\mathcal{L}^{(t)} \approx \sum_{i=1}^{n} [l(y_i, \hat{y}_i^{(t-1)}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t), \tag{5.29}$$

where

$$g_i = \partial_{\hat{y}(t-1)} l(y_i, \hat{y}^{(t-1)}), \tag{5.30a}$$

$$h_i = \partial^2_{\hat{y}(t-1)} l(y_i, \hat{y}^{(t-1)}). \tag{5.30b}$$

Since the value of $\hat{y}_i^{(t-1)}$ has been calculated at iteration $t-1$, $l(y_i^{(t)}, \hat{y}_i^{(t-1)})$ is a known constant term. After removing the constant term, the objective function of the t-th iteration is given by

$$
\begin{aligned}
\mathcal{L}^{(t)} &\approx \sum_{i=1}^{n}[g_i f_t(x_i) + \frac{1}{2}h_i f_t^2(x_i)] + \gamma T + \frac{1}{2}\chi\|\upsilon\|^2 \\
&= \sum_{i=1}^{n}[g_i f_t(x_i) + \frac{1}{2}h_i f_t^2(x_i)] + \gamma T + \frac{1}{2}\chi\sum_{i=1}^{T}\upsilon_j^2 \\
&= \sum_{i=1}^{T}\left[\frac{1}{2}(\sum_{i\in I_j}h_i + \chi)\left(\upsilon_j + \frac{\sum\limits_{i\in I_j}g_i}{\sum\limits_{i\in I_j}h_i + \chi}\right)^2\right. \\
&\quad \left. -\frac{(\sum\limits_{i\in I_j}g_i)^2}{2(\sum\limits_{i\in I_j}h_i + \chi)}\right] + \gamma T, \qquad (5.31)
\end{aligned}
$$

where $I_j = i|\vartheta(x_i = j)$ represents the sample data sets assigned to the leaf node j, $\upsilon_i = f(x_i)$. For a fixed structure $\vartheta(x)$, we can calculate the optimal weight υ_j^* of leaf j as

$$
\upsilon_j^* = -\frac{\sum\limits_{i\in I_j}g_i}{\sum\limits_{i\in I_j}h_i + \chi}. \qquad (5.32)
$$

Let $\tilde{\mathcal{L}}^{(t)}(\vartheta)$ denote the optimal value of objective function $\mathcal{L}^{(t)}$ under $\upsilon_j = \upsilon_j^*$, and $\tilde{\mathcal{L}}^{(t)}(\vartheta)$ is given by

$$
\tilde{L}^{(t)}(\vartheta) = -\frac{1}{2}\sum_{i=1}^{T}\frac{(\sum\limits_{i\in I_j}g_i)^2}{\sum\limits_{i\in I_j}h_i + \chi} + \gamma T. \qquad (5.33)
$$

Equation (5.33) can be regarded as a scoring function to evaluate the quality of a tree structure ϑ for obtaining possible optimal tree structures.

In our XGBoost model, Equation (5.32) gives the optimal weight υ_j^* of arbitrary leaf j during the iteration processes. When all the iterations are completed, the

structure of decision tree is determined and the training process for the XGBoost is finished. By summing the output value v_j of all decision trees corresponding to a sample according to Equation (5.24), we can obtain the final label (one label indicates a class in our classifier) for the sample. To evaluate the contribution of each feature to the XGBoost model, we also provide in Fig. 5.7 the total number of times that a feature is used to split the training samples across all trees to show the feature importance of 39 CSI features in the XGBoost training process.

5.4.3.3 User authentication jointly utilizing two-dimensional identities

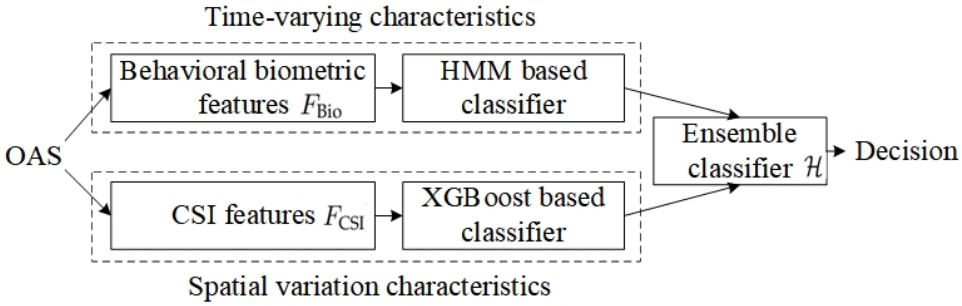


Figure 5.8: Passive user authentication utilizing two-dimensional features.

As shown in Fig. 5.8, the decision process for user authentication is as following. Given the two-dimensional features $[F_{\text{Bio}}, F_{\text{CSI}}]$ from a user (i.e., the behavioral biometric features F_{Bio} and CSI features F_{CSI} from the same OAS of the user) and a claimed identity $\mathbb{I}$, our authentication framework applies the ensemble classifier $\mathcal{H}$ designed based on the two classifiers R_{CSI} and R_{Bio} to determine if $(\mathbb{I}, [F_{\text{Bio}}, F_{\text{CSI}}])$ belongs to class ϖ_1 or ϖ_2 by using

$$\mathcal{H}(\mathbb{I}, [F_{\text{Bio}}, F_{\text{CSI}}]) \in \begin{cases} \varpi_1, & w_1 P_C(R_{\text{CSI}}(F_{\text{CSI}})) \\ & +w_2 P_B(R_{\text{Bio}}(F_{\text{Bio}})) \geq \theta, \\ \varpi_2, & \text{otherwise}, \end{cases} \tag{5.34}$$

where θ is a predefined threshold in the IIoT system; $P_C(\cdot)$ and $P_B(\cdot)$ are output

results of classifiers R_{CSI} and R_{Bio}, respectively; ϖ_1 indicates that the claim is true (a legitimate user) and ϖ_2 indicates that the claim is false (an impostor); w_1 and w_2 are two weights satisfying $w_1 + w_2 = 1$.

5.4.4 Security Analysis

Notice that the user authentication jointly utilizing two-dimensional features in (5.34) can be regarded as an ensemble learning strategy based on the voting technology and multiple base classifiers [99–101], where the vote of each base classifier has its own assigned weight (i.e., w_1 and w_2). Suppose we adopt an ensemble classifier $\mathcal{H}$ with total n base classifiers $h_1, \cdots, h_n$. Let $h_i(F)$ be the output label of base classifier h_i under features F (e.g., behavioral biometric features) and let Lab_{F} be the real label of F, where $i = 1, 2, \cdots, n$, $h_i(F), \text{Lab}_{\text{F}} \in \{\varpi_1, \varpi_2\}$. If the classification error rate $P(h_i(F) \neq \text{Lab}_{\text{F}})$ of base classifier h_i is ϵ, we have

$$P(h_i(F) \neq \text{Lab}_{\text{F}}) = \epsilon. \tag{5.35}$$

Since the ensemble learning strategy combines all base classifiers by voting, the ensemble learning will make a correct classification if more than half of the base classifiers are correct. We use $\mathcal{H}(F)$ to denote the output label of the ensemble classifier $\mathcal{H}$ under feature F, then $\mathcal{H}(F)$ is given by

$$\mathcal{H}(F) \begin{cases} = \text{Lab}_{\text{F}}, & \text{more than half of the base} \\ & \text{classifiers are correct,} \\ \neq \text{Lab}_{\text{F}}, & \text{otherwise.} \end{cases} \tag{5.36}$$

Notice that the classification accuracy of each base classifier should not do worse than that of random guessing, we have $\epsilon < 0.5$ (i.e., the classification error rate of each base classifier is less than 0.5). Thus, according to the Hoeffding inequality [102–104]

the error rate of the ensemble classifier $\mathcal{H}$ is determined as

$$\begin{aligned} P(\mathcal{H}(F) \neq \mathrm{Lab_F}) &= \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{k} (1-\epsilon)^k \epsilon^{n-k} \\ &\leqslant \exp\left(-\frac{1}{2}\, n(1-2\epsilon)^2\right). \end{aligned} \tag{5.37}$$

We can see from (5.37) that as the number of base classifiers n increases, the error rate decreases exponentially. Therefore, by integrating two classifiers (e.g., R_{CSI} and R_{Bio}) corresponding two independent characteristics (e.g., time-varying behavioral biometric characteristics and spatial variation characteristics of CSI) and assigning each classifier an appropriate weight, the proposed user authentication approach can lead to a significantly improved authentication performance than that of using only single classifier. In other words, the proposed two-dimensional user authentication approach can be used to obtain a more superior security performance for satisfying different authentication requirements across various IIoT scenarios.

5.5 Experiment and Analysis

Table 5.2: Main experiment datasets

Dataset Name	**IIoT Scenarios**	**Subjects**	**Data Acquisition**	**Operational Method**
Dataset ♯1	ME and ROE	45 users	The 45 users work in ME and ROE scenarios respectively. For each scenario of ME and ROE: Collecting 400 OASs and corresponding CSI data for each user. The length of each OAS is 9.	Hand-hold operation
Dataset ♯2	ME	18 users • (Group A: 9; Group B: 9;)	Sub-dataset $\wp_1$: Collecting 180 OASs and corresponding CSI data with the OAS length equal to 6∼14 for each user in the group A. Sub-dataset $\wp_2$: Collecting 180 OASs and corresponding CSI data with the OAS length equal to 6∼14 for each user in the group B. Sub-dataset $\wp_3{}^*$: Collecting 180 OASs and corresponding CSI data with the OAS length equal to 6∼14 for each user in the group A who impersonates the operation actions of his corresponding pair in the B group. Sub-dataset $\wp_4{}^\star$: Collecting 180 OASs and corresponding CSI data with the OAS length equal to 6∼14 for each user in the group B who impersonates the operation actions of his corresponding pair in the A group.	Hand-hold operation

• : The 18 users are not selected from the 45 users in dataset ♯1, and they are volunteers recruited to verify the performance of the proposed approach against impersonation attacks.

* : We conduct the one-to-one randomly pairing between users in the A group and B group, and each user in the A group will impersonate the operation actions of his corresponding pair in the B group.

⋆ : We conduct the one-to-one randomly pairing between users in the A group and B group, and each user in the B group will impersonate the operation actions of his corresponding pair in the A group.

5.5.1 Experiment Settings

We perform our experiments in Anhui Youkaipu Electronics Co., Ltd, an IIoT-based company that merges the IoT and cloud computing technologies for intelligent manufacturing. In the proposed passive user authentication, an OpenStack private cloud platform is adopted to provide storage, computing and authentication service encapsulation for feature extraction, classifier construction, training and user identity decision.

We consider two practical manufacturing environments (i.e., ROE and ME) with the space layouts shown in Fig. 5.5. In each environment, we collect user operation actions (i.e., the operation actions listed in Table 5.1) and raw CSI data caused by these actions during user routine work processes, and transfer these data to the OpenStack cloud platform through the API provided by OpenStack. The two datasets obtained are shown in Table 5.2. Specifically, based on the HBuilderX development environment running in the background of Android mobile terminals, we develop an APP to obtain operation actions of users and utilize the public CSI tool in [93] to record the CSI data caused by these actions. As shown in Fig. 5.5, the TP-LINK access point and the Intel 5300 NIC wireless network card are used for recording the CSI data. The operation actions and CSI data are transferred in real time to a virtual machine, which is running the Windows Server 2008 R2 operation system of the OpenStack private cloud platform. In the virtual machine, we adopt the Microsoft SQL Server 2008 R2 database to store the data, and employ the Matlab R2019a and Microsoft Visual C++ to implement the proposed passive user authentication.

5.5.2 Data Acquisition and Performance Metrics

1) Data acquisition

To investigate the performance of the proposed authentication approach in practical IIoT scenarios, we provide in Table 5.2 the main experiment data collected from

user routine operation actions. We can see from the dataset ♯1 in Table 5.2 that for each scenario (ROE or ME) and the action sequence length of 9, we collect 400 sequences of operation actions and corresponding physical CSI from 45 users (volunteers).

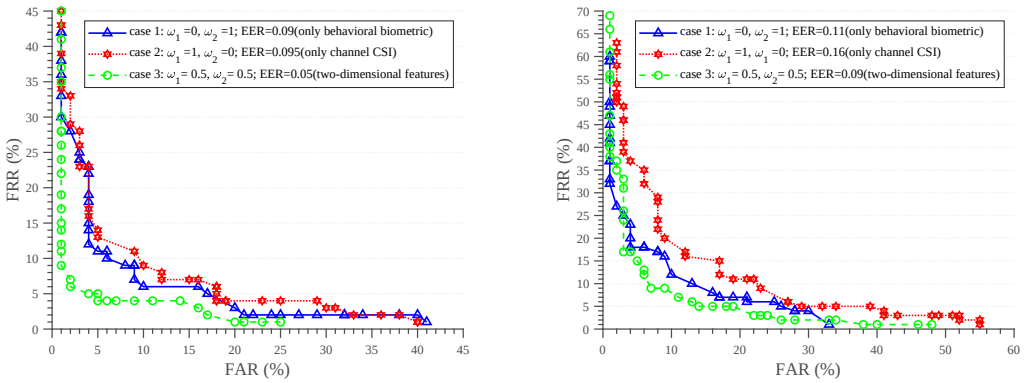
To explore the performance of the new authentication approach in resisting the impersonation attacks [105], in dataset ♯2 the 18 users under the ME scenario are evenly divided into two groups A and B to construct 4 sub-datasets denoted by $\wp_1$, $\wp_2$, $\wp_3$ and $\wp_4$. The sub-datasets $\wp_1$ and $\wp_2$ are constructed by collecting the operation actions and corresponding raw CSI data during routine work processes of users in group A and group B, respectively. To construct sub-datasets $\wp_3$ and $\wp_4$, we first conduct the one-to-one randomly pairing between users in group A and group B, then construct the sub-dataset $\wp_3$ (resp. $\wp_4$) by collecting the operation actions and corresponding raw CSI data generated by each user in group A (resp. group B) who impersonates the operation actions of his corresponding pair in the group B (resp. group A). Under each action sequence length in the range of [6,14], we collect 20 sequences of operation actions and corresponding physical CSI data for each user of a sub-dataset, so we finally obtain one normal dataset $\{\wp_1, \wp_2\}$ (without impersonation attack) and two impersonation attack datasets $\{\wp_3, \wp_2\}$ and $\{\wp_4, \wp_1\}$.

2) Performance metrics

To evaluate the performance of the proposed passive authentication framework, we first calculate three typical metrics, namely the FAR, FRR and EER [26]. Specifically, FAR is the ratio between the number of falsely accepted unauthorized users and the total number of imposters, and FRR is defined as the ratio between the number of falsely denied legitimate users and the total number of legitimate users. We then use FAR and FRR together to generate ROC curve to show the tradeoff between FAR and FRR under predefined threshold values, and EER is calculated as the sensitivity of the classifier where FAR = FRR. We also adopt the authenti-

cation accuracy to evaluate the performance for resisting the impersonation attacks of the proposed framework, here the authentication accuracy is defined as the probability that the system successfully distinguishes between the legitimate users and impersonation attacks [45, 65].

5.5.3 Authentication Performance Analysis



(a) ROE scenario (b) ME scenario

Figure 5.9: ROC curves of the proposed authentication approach for two IIoT scenarios (i.e., ROE and ME) under three cases (case 1: $\omega_1 = 0$, $\omega_2 = 1$, i.e., only using behavioral biometric features based on R_{Bio}; case 2: $\omega_1 = 1$, $\omega_2 = 0$, only using channel CSI features based on R_{CSI}; case 3: $\omega_1 = 0.5$, $\omega_2 = 0.5$, i.e., jointly utilizing behavioral biometric and channel CSI features based on R_{Bio} and R_{CSI}), respectively. (a) ROE scenario. (b) ME scenario.

To demonstrate the performance of the proposed authentication approach in both ROE and ME scenarios, we adopt the data of sequence length 9 in dataset ♯1 (Table 5.2) and show in Fig. 5.9 the corresponding ROC curves under three cases (each representing an authentication approach): (case 1: $\omega_1 = 0$, $\omega_2 = 1$, i.e., using only the behavioral biometric features based on R_{Bio}; case 2: $\omega_1 = 1$, $\omega_2 = 0$, i.e., using only the channel CSI features based on R_{CSI}; case 3: $\omega_1 = 0.5$, $\omega_2 = 0.5$, i.e., jointly utilizing both the behavioral biometric features and channel features based on classifiers R_{Bio} and R_{CSI}).

It is observed from Fig. 5.9 that in both ROE and ME scenarios the proposed authentication approach (case 3) outperforms the others in terms of ROC curves while the approach utilizing only the physical CSI features (case 2) obtains the worst authentication performance. Thus, the proposed combination approach jointly utilizing the two-dimensional features can lead to a more accurate characterization of user identities.

Another observation from Fig. 5.9 is that with the setting of sequence length 9, the corresponding EER values of three cases are all under 9.5% and 16% in ROE and ME scenarios, respectively. This demonstrates that when a large sequence length of operation action is adopted, the authentication with even one-dimensional feature might be enough to effectively discriminate user identities in IIoT scenarios.

Remark 1 *We can see from Fig. 5.9 that by assigning weight* $\omega_1 = 1$ *(resp.* $\omega_2 = 1$*), the proposed passive authentication framework reduces to the authentication with one-dimensional feature* F_{CSI} *(resp.* F_{Bio}*).*

5.5.4 Sensitivity to Weights of $R_{\mathbf{CSI}}$ and $R_{\mathbf{Bio}}$

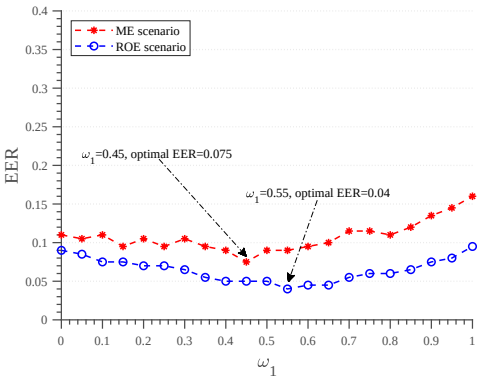
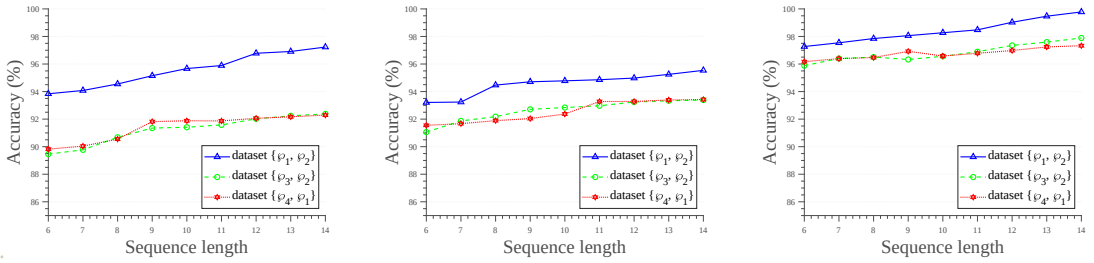


Figure 5.10: EER vs. weight ω_1 of classifier R_{CSI} (and thus weight ω_2=1-ω_1 of classifier R_{Bio}).

To explore how the weights ω_1 and ω_2 of the two classifiers R_{CSI} and R_{Bio} would affect the performance of the proposed authentication approach, we adopt the dataset

♯1 in Table 5.2 and present in Fig. 5.10 the impact of ω_1 (i.e., the weight of the classifier R_{CSI}) on EER by varying ω_1 from 0 to 1 across the two IIoT scenarios. As shown in Fig. 5.10, the EER when $\omega_1 = 0$ (i.e., using only the behavioral biometric features) and the EER when $\omega_1 = 1$ (i.e., using only the channel CSI features) are always larger than that when $0 < \omega_1 < 1$ in both ROE and ME scenarios, and we can obtain the optimal EER values of 4% and 7.5% for ROE and ME scenarios with the settings of $\omega_1 = 0.55$ and $\omega_1 = 0.45$, respectively. Therefore, by reasonably adjusting the weights of the two classifiers (R_{CSI} and R_{Bio}) the authentication performance of the proposed approach can be flexibly controlled to adapt to various IIoT scenarios. Another observation from Fig. 5.10 is that due to the more uniform electromagnetic and space environment in ME and thus a low discriminability among users there in terms of the CSI spatial variation, the EER in the ROE scenario is always better than that in the ME scenario.

5.5.5 Performance of Resisting Impersonation Attacks



(a) Case 1: $\omega_1 = 0$, $\omega_2 = 1$ (b) Case 2: $\omega_1 = 1$, $\omega_2 = 0$ (c) Case 3: $\omega_1 = 0.5$, $\omega_2 = 0.5$

Figure 5.11: Performance of resisting impersonation attacks for the ME scenarios under three cases. (a) Case 1: $\omega_1 = 0$, $\omega_2 = 1$. (b) Case 2: $\omega_1 = 1$, $\omega_2 = 0$. (c) Case 3: $\omega_1 = 0.5$, $\omega_2 = 0.5$.

To show the performance of resisting impersonation attacks for the proposed authentication approach, we present in Fig. 5.11 the impacts of weights of the two classifiers R_{CSI} and R_{Bio} on authentication accuracy based on 3 datasets (i.e., $\{\wp_1, \wp_2\}$, $\{\wp_3, \wp_2\}$ and $\{\wp_4, \wp_1\}$) of dataset ♯2 in Table 5.2 by considering three different

combinations of weights (case 1: $\omega_1 = 0$, $\omega_2 = 1$, i.e., only using behavioral biometric features based on the classifier R_{Bio}; case 2: $\omega_1 = 1$, $\omega_2 = 0$, i.e., only using channel CSI features based on the classifier R_{CSI}; case 3: $\omega_1 = 0.5$, $\omega_2 = 0.5$, i.e., jointly utilizing behavioral biometric features and channel features based on classifiers R_{CSI} and R_{Bio}, respectively). We can see from Fig. 5.11 that the performance of resisting impersonation attacks for case 3 significantly outperforms that of other cases in both the impersonation attack datasets (i.e, $\{\wp_3, \wp_2\}$ and $\{\wp_4, \wp_1\}$) and the common dataset (i.e, $\{\wp_1, \wp_2\}$). This indicates that user authentication jointly utilizing two-dimensional features can effectively resist impersonation attacks in the ME scenario, and the proposed passive authentication approach is promising to adapt various complicated IIoT application environments.

Another observation from the Fig. 5.11 is that for the IIoT scenario concerned, the authentication accuracy of all the 3 cases based on 3 datasets monotonously increases as the length of OASs increases from 6 to 14, but such trend becomes less significant if we increase the length of OASs further. It indicates that increasing the length of OASs can effectively improve the accuracy of the proposed user authentication approach, so as to better resist impersonation attacks in the IIoT scenario. Meanwhile, we can see from the Fig. 5.11 that when the length of OASs is relatively small, we can obtain a significant improvement in the authentication performance in terms of accuracy by increasing the length of OASs, but a too large sequence length might not be cost efficient since using more operation actions for user authentication will lead to a long authentication time without yielding a significant authentication performance enhancement. Therefore, it is wise to select a suitable OAS length for various IIoT applications with different authentication performance requirements.

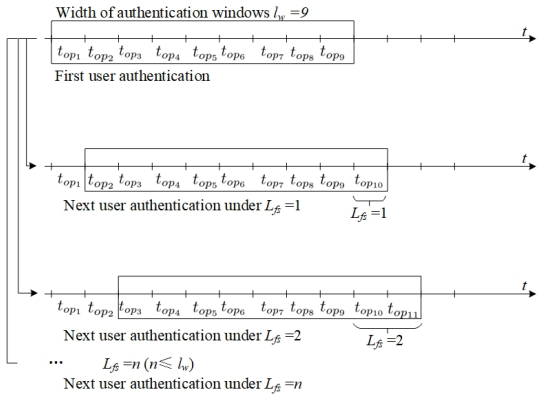


Figure 5.12: Authentication window sliding and operation action selection.

5.5.6 Sensitivity to Authentication Time

The authentication time of our authentication framework is defined as the time required for the overall authentication process, including operation action collection, operation action processing (i.e., operation action construction and CSI feature extraction), and classifier decision process. We use l_{op} to denote the length of the sequential operation actions and use $\bar{t}_{\text{op}}$ to denote the average time cost for collecting an operation action. Since we collect the user's operation actions one by one through the interface of MTs, the time consumption t_{col} used for the operation action collecting process is given by $t_{\text{col}} = l_{\text{op}} \times \bar{t}_{\text{op}}$. Regarding the time consumption t_{pro} in the operation action construction and CSI feature extraction, we can see from Section 5.4 that the operation action collection and operation action processing can be carried out parallelly, so we use t_{ove} to denote the overlap time between t_{col} and t_{pro}. Since the classifier training can be implemented offline, we ignore the time cost of classifier training here and use t_{der} to represent the classifier decision time. Thus, the authentication time t_{total} of our proposed authentication framework is determined as $t_{\text{total}} = t_{\text{col}} + t_{\text{pro}} + t_{\text{der}} + t_{\text{net}} - t_{\text{ove}}$, where t_{net} denotes the time consumption of network transmission.

Our experiment results based on dataset ♯1 indicate that the average authenti-

cation time of the proposed approach t_{total} is no more than 12 seconds when the length of the sequential operation actions l_{op} is no more than 9, so we can authenticate the identity of a user every 12s in the background of the IIoT system during user's routine work process when $l_{\text{op}} \leq 9$. Note that the proposed passive user authentication approach performs user authentication in a continuous and non-intrusive manner, so confirming the user's identity at the frequency of every 12 seconds should be acceptable to secure for the secure operation of IIoT systems.

Notice that the time consumption used for the operation action collecting process (i.e., t_{col}) usually exceeds 60% of the total time consumption, to meet the real-time user authentication requirements for some critical IIoT applications, we design here an authentication window sliding mechanism to further reduce the time for operation action collection and thus the overall authentication time of the proposed approach. As shown in Fig. 5.12 we first define an authentication window with width l_{w} equaling to the length of operation actions used for user authentication. Let L_{fs} be the steps (i.e., the number of operation actions) that the window slides backward after completing current authentication, we then use the operation actions in the windows to perform the next user authentication. Extensive experiment results based on dataset ♯1 show that the average authentication time using the authentication window sliding mechanism is no more than 5.5 seconds under l_{w} =9 and L_{fs} =2, which can satisfy the requirements of the real-time user authentication in some critical IIoT applications.

5.5.7 Comparison of Existing and Our Proposed Approaches

Table 5.3: Qualitative comparison with extant works for passive authentication

Source Study	**Features**	**Classifiers (methods)**	**Performance**		**Users**	**Resisting impersonation**	**Passive / For IIoT**
			Results	**Time (s)**			
Our work	Behavioral biometric and Physical CSI.	R_{Bio} and R_{CSI}	EER(ROE): 4% EER(ME): 7.5%	$\leqslant$12 ($l_{op} \leqslant$9); $\leqslant$5.5 (L_{fs} =2)	63 (Dataset: ♯1:45; ♯2:18)	Yes	Yes/Yes
C. Shen *et al* [26]	Behavioral biometric	HMM-based 1-class classifier	EER:4.51%	8	102	NA *	Yes/NA
M. Frank *et al* [52]	Behavioral biometric	SVM and kNN	EER:2%$\sim$6%	NA	NA	Yes	Yes/NA
H. Liu *et al* [53]	Physical CSI	SVM 2-class classifier	Accuracy:$\geq$98% (under profile samples $\geq$ 10)	NA	NA	NA	Yes/NA
M. Muaaz *et al* [105]	Behavioral biometric	DTW distance	EER:13%	NA	35	Yes	Yes/NA
Q. Zou *et al* [45]	Behavioral biometric	SVM	EER: $\leq$ 5%	NA	50	NA	Yes/NA
Y. Zhao *et al* [65]	Physical CSI	HMM and Fresnel model	Accuracy: 93%	2.08 ($\pm$0.4)	10	Yes	Yes/NA
Z. Wang *et al* [106]	Physical CSI	Deep neural networks	Accuracy: 98%	NA	8	NA	Yes/NA

* : The corresponding indexes, data or scenarios are not explicitly reported.

The comparison between existing and our proposed approaches is shown in Table 5.3. We can see from the Table 5.3 that compared with the existing methods in [26, 45, 52, 53, 65, 105, 106], the EER of the two IIoT scenarios (ROE and ME) for our approach is below 8% and the authentication time is within 12s under $l_{op} \leq 9$. This indicates that the proposed approach is promising for various complicated IIoT application scenarios. Particularly, in the ROE scenario the authentication performance (i.e., EER) of the proposed approach outperforms that of existing methods even in industrial production environments, while in the ME scenarios the authentication performance of our approach is slightly close to that of existing methods. Moreover, we also can see from Table 5.3 that compared with the existing methods, the performance of resisting impersonation attacks in terms of accuracy of the proposed approach is above 96% (under l_{op}=9) in the ME scenario. This indicates that the proposed approach is promising in countering impersonation attacks for various complicated IIoT application scenarios.

This is due to the following reasons. First, by modeling sequential operation actions from the routine work process of a user as a Markov process and applying the HMM model to characterize behavioral biometric features of the user, the proposed passive authentication framework can nicely depict the time-varying nature and dynamic properties of sequential operation actions from the user in challenging IIoT scenarios. Second, by slicing the CSI signals to reduce the noise and interference generated by the user's random actions and modeling the CSI profile of operation actions from the user based on XGBoost with superior performance, the proposed passive authentication framework achieves an accurate characterization of the user's channel CSI identity. Third, the proposed authentication framework jointly utilizes behavioral biometric characteristics and channel CSI characteristics to depict user identities, where the weight of classifiers R_{CSI} and R_{Bio} can be adjusted adaptively, and thus can determine user identities stably and accurately in complicated IIoT ap-

plication environments. As a result, the proposed passive user authentication framework can be applied to most IIoT scenarios with a promising performance guarantee.

Remark 2 *The results in Table 5.3 indicate that in comparison with the authentication based on one-dimensional characteristics, our proposed authentication approaches based on two-dimensional characteristics will not bring significantly higher computational complexity and are applicable to various IIoT scenarios with a promising authentication performance guarantee.*

5.6 Discussion

It is noticed that time-varying OASs from each user can reflect a unique behavioral biometric characteristic of the user [26]. We provide in Fig. 5.2 that the differences of OASs' time-varying properties between User 1 and User 2. More importantly, the variations of WiFi CSI caused by OAS from the user present a unique spatial-temporal characteristic due to the path loss and multi-path effects of the wireless channels [88]. Thus, by jointly exploiting the two-dimensional features of user sequential operation actions, we can not only provide a full spatial-temporal characterization of user identities but also significantly improve the security of the proposed passive user.

Note that using both the time-varying and CSI features of OAS can accurately characterize user identities. More importantly, it is very difficult for an attacker to imitate both the time-varying and CSI features of user actions in IIoT systems. Therefore, this work uses two-dimensional features to describe user identities, which can meet the IIoT application requirement for high security performance. However, the construction, preprocessing, extraction, and joint authentication of two-dimensional operation action features will bring a certain storage burden and more time consumption. Therefore, in the follow-up research, we should design a better multi-dimensional feature fusion algorithm, which can reduce the storage and time consumption of the

algorithm while obtaining high security performance.

In the process of user two-dimensional feature enrollment, when a new user joins the authentication system, we require the user to complete specific operation actions according to concerned work business process, thereby completing enrollment and storage of user two-dimensional features. In future work, we will further research and develop the dynamic registration and update method of user two-dimensional features to improve the usability of the authentication framework.

5.7 Summary

For IIoT authentication requirements of high security performance in the DO layer, this work proposed a novel two-dimensional passive authentication framework by exploiting both the time-varying characteristics of the user sequential operation actions and spatial-temporal variation characteristics of WiFi CSI caused by operation actions. We demonstrated that the new framework enables a flexible and efficient authentication performance control to be achieved by adjusting the system parameters like the weights of the two designed classifiers. Thus, the proposed framework is promising for satisfying different authentication performance requirements across various IIoT scenarios. Moreover, it is expected that the passive authentication solution developed in this work can be used as a promising supplement or alternative to the traditional pin-based and pattern-based active authentication methods to achieve security enhancement in the IIoT systems.

CHAPTER VI

Conclusion

To satisfy three general requirements for user authentication in IIoT systems, we developed corresponding three schemes to ensure the secure operation of IIoT systems. First, for user authentication of the ME layer, this dissertation explored the common behavioral biometrics from user sequential operation actions in IIoT systems to propose a passive authentication framework, which provided continuous/non-intrusive user authentication and posed good anti-interference capability in the interference-intensive environment of the ME layer. Second, for user authentication of the MC layer, we explored the user consecutive screen-touch actions during routine work processes and proposed a passive authentication method based on both the time-varying characteristics and spatial image characteristics of the user touch trajectory sequences, which provided implicit/non-intrusive user identity verification and can meet the real-time authentication requirement of the MC layer. Finally, for user authentication of the DO layer, we developed a novel two-dimensional passive authentication framework by jointly utilizing both the time-varying characteristics of the user sequential operation actions and spatial variation characteristics of CSI caused by these actions, which applied to the authentication of the DO layer with high security requirement.

We studied in Chapter III an IIoT application scenario with more electromagnetic

interference, where legitimate users interacted with IIoT systems through mobile devices (industrial-level terminals) in the presence of a potential attacker. We used sensor motion characteristics of multiple operation actions to characterize user identities, and leveraged the Kalman filtering and Wavelet techniques for interference elimination and the singular value decomposition method for the dimensionality reduction of characteristics. We then developed a multiple characteristics-based passive authentication framework for continuous and non-intrusive user identity verification for IIoT systems. Our results showed that the new passive authentication framework enables a flexible authentication performance control to be achieved by adjusting the system parameters like the length of operation sequence, number of features, size of user space, and the proportion of a certain action feature. Thus, the proposed framework is promising for satisfying different performance requirements across various IIoT scenarios.

For authentication solution exploiting touch-based features of time-varying screen-touch trajectory sequences and cumulative consecutive screen-touch trajectory images from user touch actions during routine work processes, we investigated in Chapter IV an IIoT system with the high real-time requirement. We explored touch-based features of time-varying screen-touch trajectory sequences and cumulative consecutive screen-touch trajectory images from user touch actions during routine work processes in IIoT systems and developed a touch-based passive authentication framework for continuous user identity verification. In the authentication solution, every time a user touched the screen, the IIoT authentication system can verify the user's identity by analyzing the time-varying features of touch trajectory sequences and STTI characteristics. We further demonstrated that the performance of the proposed authentication framework in terms of false acceptance rate, false rejection rate and equal error rate, and also examined the related authentication efficiency issues such as the sensitivity to the weights for classifiers, the sensitivity to authentication time and the capability

of resisting against impersonation attacks.

In Chapter V, we addressed passive user authentication issue with high security requirement. We first developed a new method to characterize the behavioral biometric characteristics of users in IIoT scenarios. We then proposed a new approach to depict the spatial-temporal variations of CSI related to a user, in which the WiFi CSI data related to the user was first sliced to reduce the noise and interference from the random actions of the user. We further designed two classifiers corresponding to the above two characteristics. By combining these two classifiers and assigning each classifier an appropriate weight, we thus developed a novel two-dimensional user authentication framework for passive, continuous and non-intrusive user authentication in IIoT scenarios. We conducted extensive experiments to evaluate the performance of the proposed authentication framework in terms of false acceptance rate, false rejection rate and equal error rate, and also examined the related authentication efficiency issues such as the sensitivity to the weights for classifiers, the sensitivity to authentication time and the capability of resisting against impersonation attacks.

It is notable that, this thesis explores common behavioral biometrics from sequential user operation actions, spatial-temporal touch-based features, and two-dimensional features involving user routine operation actions to develop novel user authentication frameworks for passive, continuous and non-intrusive user authentication in IIoT scenarios. We demonstrated that the new framework enables a flexible and efficient authentication performance control to be achieved by adjusting the system parameters like the weights of the two designed classifiers. Thus, the proposed framework is promising for satisfying different authentication performance requirements across various IIoT scenarios. Moreover, it is expected that the passive authentication solution developed in this thesis can be used as a promising supplement or alternative to the traditional pin-based and pattern-based active authentication methods to achieve security enhancement in the IIoT systems.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] A. Gilchrist, *Industry 4.0: the Industrial Internet of Things*. Springer, 2016.

[2] T. Qiu, J. Chi, X. Zhou, Z. Ning, M. Atiquzzaman, and D. O. Wu, "Edge computing in Industrial Internet of Things: Architecture, advances and challenges," *IEEE Communications Surveys Tutorials*, vol. 22, no. 4, pp. 2462–2488, Jul. 2020.

[3] S. Mumtaz, A. Bo, A. Al-Dulaimi, and K. Tsang, "Guest editorial 5G and beyond mobile technologies and applications for Industrial IoT (IIoT)," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 6, pp. 2588–2591, Jun. 2018.

[4] IIC. Industrial internet security framework. IIC:PUB:G4:V1.0:PB:20160926. [Online]. Available: https://www.iiconsortium.org/IISF/

[5] I. I. I. Alliance. Industrial internet architecture (version 2.0). Accessed on Apr. 23, 2020. [Online]. Available: https://www.aii-alliance.org/index/c315/n45.html

[6] P. Zhang, Y. Wu, and H. Zhu, "Open ecosystem for future industrial Internet of things (IIoT): Architecture and application," *CSEE Journal of Power and Energy Systems*, vol. 6, no. 1, pp. 1–11, Mar. 2020.

[7] H. M. O. Canilang, A. C. Caliwag, and W. Lim, "Design, implementation, and deployment of modular battery management system for IIoT-Based applications," *IEEE Access*, vol. 10, pp. 109 008–109 028, Oct. 2022.

[8] P. Liu, Z. Wang, S. Wei, Y. Bo, and S. Pu, "Recent developments of modulation and control for high-power current-source-converters fed electric machine systems," *CES Transactions on Electrical Machines and Systems*, vol. 4, no. 3, pp. 215–226, Sep. 2020.

[9] V. Sharma, G. Choudhary, Y. Ko, and I. You, "Behavior and vulnerability assessment of drones-enabled Industrial Internet of Things (IIoT)," *IEEE Access*, vol. 6, pp. 43 368–43 383, Jul. 2018.

[10] S. Misra, C. Roy, T. Sauter, A. Mukherjee, and J. Maiti, "Industrial Internet of Things for safety management applications: A survey," *IEEE Access*, vol. 10, pp. 83 415–83 439, Jul. 2022.

[11] M. H. Syed, E. B. Fernandez, and M. Ilyas, "A pattern for fog computing," in *Proceedings of the 10th Travelling Conference on Pattern Languages of Programs*, 2016, pp. 1–10.

[12] N. Tyagi, "A reference architecture for IoT," *International Journal of Computer Engineering and Applications*, vol. X, Jan. 2016.

[13] R. Roman, J. Zhou, and J. Lopez, "On the features and challenges of security and privacy in distributed Internet of Things," *Computer Networks*, vol. 57, no. 10, pp. 2266–2279, Jul. 2013.

[14] L. H. Nunes, J. C. Estrella, C. Perera, S. Reiff-Marganiec, and A. C. Botazzo Delbem, "Multi-criteria IoT resource discovery: a comparative analysis," *Software: Practice and Experience*, vol. 47, no. 10, pp. 1325–1341, Oct. 2017.

[15] A. P. Fournaris, C. Alexakos, C. Anagnostopoulos, C. Koulamas, and A. Kalogeras, "Introducing hardware-based intelligence and reconfigurability on industrial IoT edge nodes," *IEEE Design and Test*, vol. 36, no. 4, pp. 15–23, Aug. 2019.

[16] D. E. Boubiche, A.-S. K. Pathan, J. Lloret, H. Zhou, S. Hong, S. O. Amin, and M. A. Feki, "Advanced industrial wireless sensor networks and intelligent IoT," *IEEE Communications Magazine*, vol. 56, no. 2, pp. 14–15, Feb. 2018.

[17] B. R. Haverkort and A. Zimmermann, "Smart industry: How ict will change the game!" *IEEE Internet Computing*, vol. 21, no. 1, pp. 8–10, Jan.-Feb. 2017.

[18] M. Abdallah, O. A. Dobre, P.-H. Ho, S. Jabbar, M. J. Khabbaz, and J. J. P. C. Rodrigues, "Blockchain-enabled Industrial Internet of Things: Advances, applications, and challenges," *IEEE Internet of Things Magazine*, vol. 3, no. 2, pp. 16–18, Jun. 2020.

[19] G. Premsankar, M. Di Francesco, and T. Taleb, "Edge computing for the Internet of Things: A case study," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 1275–1284, Apr. 2018.

[20] A. Molina Zarca, J. Bernal Bernabe, I. Farris, Y. Khettab, T. Taleb, and A. Skarmeta, "Enhancing IoT security through network softwarization and virtual security appliances," *International Journal of Network Management*, vol. 28, no. 5, p. e2038, Jul. 2018.

[21] H. Sedjelmaci, S. M. Senouci, and T. Taleb, "An accurate security game for low-resource IoT devices," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 10, pp. 9381–9393, May. 2017.

[22] L. D. Xu, W. He, and S. Li, "Internet of things in industries: A survey," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 4, pp. 2233–2243, Nov. 2014.

[23] Y. Chen, Y. Sun, B. Yang, and T. Taleb, "Deep reinforcement learning-based joint caching and computing edge service placement for sensing-data-driven IIoT applications," in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 4287–4292.

[24] J. Kim, S. Kim, T. Taleb, and S. Choi, "RAPID: Contention resolution based random access using context ID for IoT," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 7, pp. 7121–7135, Jul. 2019.

[25] I. Farris, T. Taleb, Y. Khettab, and J. Song, "A survey on emerging SDN and NFV security mechanisms for IoT systems," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 812–837, Aug. 2019.

[26] C. Shen, Y. Li, Y. Chen, X. Guan, and R. A. Maxion, "Performance analysis of multi-motion sensor behavior for active smartphone authentication," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 1, pp. 48–62, Jan. 2018.

[27] T. Zhu, Z. Qu, H. Xu, J. Zhang, Z. Shao, Y. Chen, S. Prabhakar, and J. Yang, "Riskcog: Unobtrusive real-time user authentication on mobile devices in the wild," *IEEE Transactions on Mobile Computing*, vol. 19, no. 2, pp. 466–483, Feb. 2020.

[28] A. Roy, N. Memon, and A. Ross, "Masterprint: Exploring the vulnerability of partial fingerprint-based authentication systems," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 9, pp. 2013–2025, Sep. 2017.

[29] S. Li and A. C. Kot, "Fingerprint combination for privacy protection," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 2, pp. 350–360, Feb. 2013.

[30] D. F. Smith, A. Wiliem, and B. C. Lovell, "Face recognition on consumer devices: Reflections on replay attacks," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 4, pp. 736–745, Apr. 2015.

[31] S. Thavalengal, I. Andorko, A. Drimbarean, P. Bigioi, and P. Corcoran, "Proof-of-concept and evaluation of a dual function visible/nir camera for iris authentication in smartphones," *IEEE Transactions on Consumer Electronics*, vol. 61, no. 2, pp. 137–143, May. 2015.

[32] C. Shen, Y. Chen, X. Guan, and R. A. Maxion, "Pattern-growth based mining mouse-interaction behavior for an active user authentication system," *IEEE Transactions on Dependable and Secure Computing*, vol. 17, no. 2, pp. 335–349, Mar. 2020.

[33] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 4–20, Jan. 2004.

[34] G. Zhao, P. Zhang, Y. Shen, and X. Jiang, "Passive user authentication utilizing behavioral biometrics for IIoT systems," *IEEE Internet of Things Journal*, vol. 9, no. 14, pp. 12 783–12 798, Jul. 2022.

[35] S. Keykhaie and S. Pierre, "Mobile match on card active authentication using touchscreen biometric," *IEEE Transactions on Consumer Electronics*, vol. 66, no. 4, pp. 376–385, Oct. 2020.

[36] V. M. Patel, R. Chellappa, D. Chandra, and B. Barbello, "Continuous user authentication on mobile devices: Recent progress and remaining challenges," *IEEE Signal Processing Magazine*, vol. 33, no. 4, pp. 49–61, Jul. 2016.

[37] M. Shahzad and M. P. Singh, "Continuous authentication and authorization for the Internet of Things," *IEEE Internet Computing*, vol. 21, no. 2, pp. 86–90, Mar. 2017.

[38] M. Abuhamad, T. Abuhmed, D. Mohaisen, and D. Nyang, "Autosen: Deep-learning-based implicit continuous authentication using smartphone sensors," *IEEE Internet of Things Journal*, vol. 7, no. 6, pp. 5008–5020, Jun. 2020.

[39] Z. Qin, L. Hu, N. Zhang, D. Chen, K. Zhang, Z. Qin, and K. R. Choo, "Learning-aided user identification using smartphone sensors for smart homes," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 7760–7772, Oct. 2019.

[40] Z. Sitová, J. Šeděnka, Q. Yang, G. Peng, G. Zhou, P. Gasti, and K. S. Balagani, "HMOG: New behavioral biometric features for continuous authentication of smartphone users," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 5, pp. 877–892, May. 2016.

[41] Y. Li, H. Hu, and G. Zhou, "Using data augmentation in continuous authentication on smartphones," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 628–640, Feb. 2019.

[42] A. Acien, A. Morales, J. V. Monaco, R. Vera-Rodriguez, and J. Fierrez, "Typenet: Deep learning keystroke biometrics," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 4, no. 1, pp. 57–70, Sep. 2022.

[43] W. Jia, Y. Qi, A. Huang, F. Zhou, and S. Gao, "High security user authentication based on piezoelectric keystroke dynamics applying to multiple emotional responses," *IEEE Sensors Journal*, vol. 22, no. 3, pp. 2814–2822, Dec. 2022.

[44] B. Ayotte, M. Banavar, D. Hou, and S. Schuckers, "Fast free-text authentication via instance-based keystroke dynamics," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 2, no. 4, pp. 377–387, Jun. 2020.

[45] Q. Zou, L. Ni, Q. Wang, Q. Li, and S. Wang, "Robust gait recognition by integrating inertial and RGBD sensors," *IEEE Transactions on Cybernetics*, vol. 48, no. 4, pp. 1136–1150, Apr. 2018.

[46] J. Zhang, J. Pu, C. Chen, and R. Fleischer, "Low-resolution gait recognition," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 40, no. 4, pp. 986–996, Aug. 2010.

[47] L. Wu, J. Yang, M. Zhou, Y. Chen, and Q. Wang, "Lvid: A multimodal biometrics authentication system on smartphones," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 1572–1585, Sep. 2020.

[48] L. Lu, J. Yu, Y. Chen, H. Liu, Y. Zhu, L. Kong, and M. Li, "Lip reading-based user authentication through acoustic sensing on smartphones," *IEEE/ACM Transactions on Networking*, vol. 27, no. 1, pp. 447–460, Feb. 2019.

[49] G. Peng, G. Zhou, D. T. Nguyen, X. Qi, Q. Yang, and S. Wang, "Continuous authentication with touch behavioral biometrics and voice on wearable glasses," *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 3, pp. 404–416, Nov. 2017.

[50] S. M. Kolly, R. Wattenhofer, and S. Welten, "A personal touch: Recognizing users based on touch screen behavior," *in Proc. Association for Computing Machinery*, pp. 1–5, 2012.

[51] R. C. Prati, G. E. A. P. A. Batista, and M. C. Monard, "Evaluating classifiers using ROC curves," *IEEE Latin America Transactions*, vol. 6, no. 2, pp. 215–222, Jun. 2008.

[52] M. Frank, R. Biedert, E. Ma, I. Martinovic, and D. Song, "Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 1, pp. 136–148, Jan. 2013.

[53] H. Liu, Y. Wang, J. Liu, J. Yang, Y. Chen, and H. V. Poor, "Authenticating users through fine-grained channel information," *IEEE Transactions on Mobile Computing*, vol. 17, no. 2, pp. 251–264, Feb. 2018.

[54] R. Alazrai, A. Awad, B. Alsaify, M. Hababeh, and M. I. Daoud, "A dataset for Wi-Fi-based human-to-human interaction recognition," *Data in Brief*, vol. 31, p. 105668, Aug. 2020.

[55] S. Hardy. Mobile devices and the Industrial Internet of Things (IIoT). Accessed on Feb. 26, 2016. [Online]. Available: https://www.csoonline.com/article/3249265/mobile-devices-and-the-industrial-internet-of-things-iiot.html

[56] H. Mouratidis and V. Diamantopoulou, "A security analysis method for Industrial Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 9, pp. 4093–4100, Sep. 2018.

[57] R. P. Guidorizzi, "Security: Active authentication," *IT Professional*, vol. 15, no. 04, pp. 4–7, Jul-Aug. 2013.

[58] A. Perala. Unifyid says smartphone-based gait authentication system rivals finger, face biometrics. Accessed on Aug. 26, 2020. [Online]. Available: https://findbiometrics.com/unifyid-says-smartphone-based-gait-authentication-system-rivals-finger-face-biometrics-082604/

[59] V. P. R, A. A., and U. Gopalakrishnan, *A Report on Behavior-Based Implicit Continuous Biometric Authentication for Smart Phone.* Advances in Intelligent Systems and Computing, Jan. 2020, vol. 1155, pp. 169–184.

[60] Y. Yang, J. Sun, and L. Guo, "Personaia: A lightweight implicit authentication system based on customized user behavior selection," *IEEE Transactions on Dependable and Secure Computing*, vol. 16, no. 1, pp. 113–126, Jul. 2019.

[61] K. Yoshio, M. Kazunari, and I. Keiichi, "Identity continuance in single sign-on with authentication server failure," in *2011 Fifth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, 2011, pp. 597–602.

[62] G. Zhao, D. Zheng, and K. Chen, "Design of single sign-on," in *IEEE International Conference on E-Commerce Technology for Dynamic E-Business*, 2004, pp. 253–256.

[63] F. Zhu and H. Diao, "Single sign-on assistant: An authentication broker for web applications," in *2010 Third International Conference on Knowledge Discovery and Data Mining*, 2010, pp. 146–149.

[64] P. Harding, L. Johansson, and N. Klingenstein, "Dynamic security assertion markup language: Simplifying single sign-on," *IEEE Security Privacy*, vol. 6, no. 2, pp. 83–85, Apr. 2008.

[65] Y. Zhao, R. Gao, S. Liu, L. Xie, J. Wu, H. Tu, and B. Chen, "Device-free secure interaction with hand gestures in wifi-enabled IoT environment," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5619–5631, Apr. 2021.

[66] M. Vermucht. The keys to securing Industrial IoT (IIoT) environments. Accessed on Jan. 29, 2020. [Online]. Available: https://www.vdoo.com/blog/industrial-iot-security

[67] X. Z. L. Li and G. Xue, "Unobservable re-authentication for smartphones," *in Proc. 20th Network and Distributed Syst. Security Symp.*, pp. 1–16, Apr. 2013.

[68] A. Roy, T. Halevi, and N. Memon, "An hmm-based behavior modeling approach for continuous mobile authentication," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 3789–3793.

[69] G. Welch, G. Bishop *et al.*, "An introduction to the Kalman filter," 1995.

[70] F. Xie, H. Wen, Y. Li, S. Chen, L. Hu, Y. Chen, and H. Song, "Optimized coherent integration-based radio frequency fingerprinting in Internet of Things," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3967–3977, Oct. 2018.

[71] C. Shen, Y. Zhang, X. Guan, and R. A. Maxion, "Performance analysis of touch-interaction behavior for active smartphone authentication," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 3, pp. 498–513, Mar. 2016.

[72] Y. Shu, Y. J. Gu, and J. Chen, "Dynamic authentication with sensory information for the access control systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 2, pp. 427–436, May. 2014.

[73] S. Arunthavanathan, S. Kandeepan, and R. J. Evans, "A Markov decision process-based opportunistic spectral access," *IEEE Wireless Communications Letters*, vol. 5, no. 5, pp. 544–547, Oct. 2016.

[74] L. E. Baum, T. Petrie, and S. N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains," *Annals of Mathematical Stats*, vol. 41, no. 1, pp. 164–171, Apr. 1969.

[75] L. Rabiner and B. Juang, "An introduction to hidden Markov models," *IEEE ASSP Magazine*, vol. 3, no. 1, pp. 4–16, Jan. 1986.

[76] A. Gretton, K. M. Borgwardt, M. J. Rasch, B. Schölkopf, and A. Smola, "A kernel two-sample test," *J. Mach. Learn. Res.*, vol. 13, no. 1, pp. 723–773, Mar. 2012.

[77] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011, software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm.

[78] N. B. Karayiannis and Y. Xiong, "Training reformulated radial basis function neural networks capable of identifying uncertainty in data classification," *IEEE Transactions on Neural Networks*, vol. 17, no. 5, pp. 1222–1234, Sep. 2006.

[79] C. Holz and M. Knaust, "Biometric touch sensing: Seamlessly augmenting each touch with continuous authentication." Association for Computing Machinery, 2015, p. 303–312.

[80] T. Dee, I. Richardson, and A. Tyagi, "Continuous transparent mobile device touchscreen soft keyboard biometric authentication," in *2019 32nd International Conference on VLSI Design and 2019 18th International Conference on Embedded Systems (VLSID)*, 2019, pp. 539–540.

[81] I. The MathWorks. Linear or rank correlation. Accessed on Oct. 23, 2022. [Online]. Available: https://www.mathworks.com/help/stats/corr.html

[82] D. Wang, J. Yang, W. Cui, L. Xie, and S. Sun, "Caution: A robust wifi-based human authentication system via few-shot open-set recognition," *IEEE Internet of Things Journal*, vol. 9, no. 18, pp. 17 323–17 333, Sep. 2022.

[83] J. A. Hartigan and M. A. Wong, "A k-means clustering algorithm," *Journal of the royal statistical society. series c (applied statistics)*, vol. 28, no. 1, pp. 100–108, 1979.

[84] V. Digalakis, J. Rohlicek, and M. Ostendorf, "Ml estimation of a stochastic linear system with the em algorithm and its application to speech recognition," *IEEE Transactions on Speech and Audio Processing*, vol. 1, no. 4, pp. 431–442, Oct. 1993.

[85] P. Baggenstoss, "A modified baum-welch algorithm for hidden markov models with multiple observation spaces," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 4, pp. 411–416, May. 2001.

[86] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Speeded-up robust features (SURF)," *Computer Vision & Image Understanding*, vol. 110, no. 3, pp. 346–359, Dec. 2008.

[87] H. Bay, T. Tuytelaars, and L. V. Gool, "Surf: Speeded up robust features," in *European conference on computer vision*. Springer, 2006, pp. 404–417.

[88] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Predictable 802.11 packet delivery from wireless channel measurements," *ACM SIGCOMM Computer Communication Review*, vol. 40, no. 4, pp. 159–170, Aug. 2010.

[89] M. Liu, D. Zhang, S. Chen, and H. Xue, "Joint binary classifier learning for ecoc-based multi-class classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 11, pp. 2335–2341, Nov. 2016.

[90] A. X. Liu, C. R. Meiners, and E. Torng, "Packet classification using binary content addressable memory," *IEEE/ACM Transactions on Networking*, vol. 24, no. 3, pp. 1295–1307, Apr. 2016.

[91] N. García-Pedrajas and D. Ortiz-Boyer, "An empirical study of binary classifier fusion methods for multiclass classification," *Information Fusion*, vol. 12, no. 2, pp. 111–130, Apr. 2011.

[92] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, Aug. 2016. [Online]. Available: http://dx.doi.org/10.1145/2939672.2939785

[93] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: Gathering 802.11n traces with channel state information," *SIGCOMM Comput. Commun. Rev.*, vol. 41, no. 1, pp. 53–53, Jan. 2011. [Online]. Available: https://doi.org/10.1145/1925861.1925870

[94] O. Edfors, M. Sandell, J.-J. van de Beek, S. Wilson, and P. Borjesson, "OFDM channel estimation by singular value decomposition," *IEEE Transactions on Communications*, vol. 46, no. 7, pp. 931–939, Jul. 1998.

[95] S. E. Levinson, L. R. Rabiner, and M. M. Sondhi, "An introduction to the application of the theory of probabilistic functions of a markov process to automatic speech recognition," *The Bell System Technical Journal*, vol. 62, no. 4, pp. 1035–1074, Apr. 1983.

[96] K. W. Choi and E. Hossain, "Estimation of primary user parameters in cognitive radio systems via hidden markov model," *IEEE Transactions on Signal Processing*, vol. 61, no. 3, pp. 782–795, Feb. 2013.

[97] A. Suarez and J. Lutsko, "Globally optimal fuzzy decision trees for classification and regression," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 12, pp. 1297–1311, Dec. 1999.

[98] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: a statistical view of boosting," *The annals of statistics*, vol. 28, no. 2, pp. 337–407, Apr. 2000.

[99] G. Webb and Z. Zheng, "Multistrategy ensemble learning: reducing error by combining ensemble learning techniques," *IEEE Transactions on Knowledge and Data Engineering*, vol. 16, no. 8, pp. 980–991, Aug. 2004.

[100] Q. Sun and Z. Ge, "Deep learning for industrial kpi prediction: When ensemble learning meets semi-supervised data," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 1, pp. 260–269, Jan. 2021.

[101] Y. Dang, C. Benzaïd, B. Yang, T. Taleb, and Y. Shen, "Deep ensemble learning based GPS spoofing detection for cellular-connected UAVs," *IEEE Internet of Things Journal*, vol. 9, no. 24, pp. 25 068–25 085, Dec. 2022.

[102] R. J. Serfling, "Probability Inequalities for the Sum in Sampling without Replacement," *The Annals of Statistics*, vol. 2, no. 1, pp. 39–48, Jan. 1974.

[103] M. A. Wiering and H. van Hasselt, "Ensemble algorithms in reinforcement learning," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 38, no. 4, pp. 930–936, Aug. 2008.

[104] Y. Gao, G. Ji, Z. Yang, and J. Pan, "A dynamic adaboost algorithm with adaptive changes of loss function," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 6, pp. 1828–1841, Nov. 2012.

[105] M. Muaaz and R. Mayrhofer, "Smartphone-based gait recognition: From authentication to imitation," *IEEE Transactions on Mobile Computing*, vol. 16, no. 11, pp. 3209–3221, Nov. 2017.

[106] Z. Wang, W. Dou, Z. Huang, C. Zhang, and D. Chen, "Wau: A user authentication system based on channel state information and deep neural networks," *Journal of Physics: Conference Series*, vol. 1748, no. 3, p. 032019, Jan. 2021.

Publications

Jounal Articles

[1] Zhao Guozhu, Zhang Pinchang, Shen Yulong and Jiang Xiaohong. " Passive User Authentication Utilizing Behavioral Biometrics for IIoT Systems," *IEEE Internet of Things Journal,* vol. 9, no. 14, pp. 12783-12798, Jul. 2022.

[2] Zhao Guozhu, Zhang Pinchang, Shen Yulong , Peng Limei and Jiang, Xiaohong. " Passive User Authentication Utilizing Two-Dimensional Features for IIoT Systems," *IEEE Transactions on Cloud Computing*, (Accept, Early Access Article doi=10.1109/TCC.2022.3227171).

Conference Papers

[3] Guozhu Zhao, Pinchang Zhang and Xiaohong Jiang. " On the Applicability of Users' Operation-action Characteristics for the Continuous Authentication in IIoT Scenarios". 2020 International Conference on Networking and Network Applications (NaNA), Haikou City, China, pp. 124-129, Oct. 2020.

[4] Guozhu Zhao, Pinchang Zhang and Lisheng Ma. " On the Applicability of Multi-Characteristics for the Continuous Authentication in IIoT Scenarios". 2021 International Conference on Networking and Network Applications (NaNA), Lijiang City, China, pp. 200-205, Oct. 2021.

[5] Guozhu Zhao, Pinchang Zhang, Yulong Shen and Xiaohong Jiang. " Passive User Authentication Utilizing Consecutive Touch Action Features for IIoT Systems". The 4th International Conference on Science of Cyber Security (SciSec 2022), Matsue city, Shimane, Japan, Aug. 2022.