DOI: 10.32347/2412-9933.2025.64.171-176
UDC 004.056.52:004.93

Paliy Sergiy
PhD (Engineering), Associate Professor of the Information Systems and Technologies Department,
https://orcid.org/0000-0001-9742-1116
Taras Shevchenko National University of Kyiv, Kyiv
Boiko Dmytro
PhD Student of the Information Systems and Technologies Depar3tment,
https://orcid.org/0009-0006-1195-1065
Taras Shevchenko National University of Kyiv, Kyiv
Shabala Yevheniia
PhD (Engineering), Associate Professor of the Cybersecurity and Computer Engineering Department,
https://orcid.org/0000-0002-0428-9273
Kyiv National University of Construction and Architecture, Kyiv

USING THE VIDEO CHANNEL OF ATM DEVICES TO ENHANCE THE LEVEL OF CYBERSECURITY DURING CUSTOMER INTERACTIONS

Abstract. *The rapid growth of automated banking and cardless financial transactions has increased the need for advanced user authentication mechanisms that can withstand fraud, social engineering, and biometric spoofing. Traditional authentication methods based solely on cards and PINs are increasingly vulnerable to theft, duress, and replay attacks. In this context, biometric technologies, including facial recognition combined with real-time behavioral analysis, represent a promising direction for enhancing transaction security. This study proposes an integrated method to improve user authentication in automated financial terminals by utilizing the device's video feed. The approach focuses on continuous identity verification and action monitoring during user interaction, supported by a formally modeled interaction protocol between the terminal and the decision-making system. The proposed method integrates video-based biometric identification, activity detection, and behavioral analysis into a single authentication workflow. A protocol model was developed that defines secure data exchange, time constraints, and decision logic between the user, the terminal, and the server component. The model supports adaptive decision outcomes, including approval, re-verification, or rejection, based on trust levels derived from biometric and behavioral features. The use of edge computing principles reduces latency and network load by transmitting compact feature vectors instead of raw video streams. The approach also includes degradation modes for network failures, anomaly detection mechanisms, and compliance-oriented data minimization strategies. A comparative analysis with existing global practices demonstrates that the proposed solution is in line with current trends in multifactor authentication, extending them through protocol formalization and real-time video analytics. The results show that integrated video-based authentication can significantly improve the security and resilience of financial terminals without compromising usability or accessibility. The proposed interoperability protocol enables scalable deployment, supports regulatory compliance, and provides the flexibility to adapt authentication levels to risk conditions. This method is applicable not only to ATMs but also to self-service kiosks, remote banking terminals, and mobile financial infrastructures, offering a robust foundation for next-generation secure financial transaction systems.*

Keywords: Face Recognition; biometric validation; ATM

Introduction

In modern conditions, ensuring the security of financial transactions through automated terminals such as ATMs requires the application of biometric technologies, particularly facial identification and real-time user behavior analysis. Dynamic authentication methods that consider factors such as head movement or proximity to the camera have proven to be highly effective: for instance, systems using three-dimensional object features (liveness detection) achieve accuracy levels above 97% in verifying vitality and recognizing identity [1].

Within the category of multi-factor biometrics, combinations of fingerprints, iris recognition, and facial recognition are applied, significantly increasing the

reliability of authentication in cardless ATM systems [2]. Moreover, there are solutions that integrate the visual channel in such a way that an active prompt or pattern is displayed, requiring the user to respond in a specific manner–for example, keeping their face within a designated frame area or repeating a movement–which helps detect fraudulent activities involving video replay, fake images, or deepfake models [3].

This work proposes a method for enhancing user authentication security during the processing of financial information through digital terminals. A key component is the use of the device's video channel to determine the identity and actions of the user during interaction with the terminal. An interaction protocol between the user and the decision-making system has been modeled, enabling effective integration of recognition algorithms into the framework of secure financial transactions.

Models available on Ukraine and Global Practices

In Ukraine, various models aimed at improving ATM security and usability are being actively introduced. A well-known example is a system that enables access to ATMs via physical or virtual payment cards with support for Apple Wallet and Google Wallet, ensuring 24/7 availability without requiring the physical presence of bank staff [4].

In Japan, a major company has implemented ATMs equipped with palm vein recognition technology [5]. Users can authenticate themselves without a card or PIN code by simply placing their hand on a dedicated scanner. This approach offers strong resistance to fraud, as the unique vein pattern cannot be copied or forged.

An inclusive solution was introduced in Ukraine through the deployment of voice assistants for ATMs and terminals. By 2024, more than 600 such devices were installed across the country, providing services for visually impaired clients. This innovation represents a unique functionality that aligns with the official methodological guidelines of the National Bank on financial inclusion.

In the United States, companies have launched next-generation ATMs supporting cardless transactions via smartphones [6]. Customers initiate withdrawals through a mobile application and confirm their identity at the terminal using facial recognition or a QR code. This significantly improves security by eliminating the risks associated with compromised physical cards.

There are also solutions based on video analytics, such as distributed video control systems that integrate cameras with ATM transaction systems. They provide continuous monitoring, recording, data transmission and storage, as well as instant alerts in case of suspicious activity.

China has seen wide adoption of real-time face recognition ATMs integrated with the national citizen database. For instance, at ICBC (Industrial and Commercial Bank of China), a transaction can only be completed after identity verification through the ATM camera.

In India, ATMs equipped with multi-factor biometrics are being tested, combining fingerprint scanning, facial recognition, and PIN code entry. These systems are designed for regions where social engineering attacks and card data theft attempts are common.

In Europe, particularly in Germany, intelligent fraud-detection ATMs are being developed, where user behavior video analysis identifies suspicious activity, such as the presence of bystanders or unusual movements. This technology functions as a supplement to classical biometric authentication methods.

Beyond video and contactless technologies, the market also offers hardware access control solutions. For example, systems that automatically switch on lighting when a client enters an ATM cabin and turn it off when empty, provide voice prompts, and are compatible with all types of payment cards.

Additionally, some models are designed to support a wide range of peripheral devices. One such solution includes embedded computers with serial expansion, allowing reliable control of additional security devices, cameras, biometric sensors, and card readers.

As a result, the ATM market now offers comprehensive security ecosystems–ranging from contactless authentication and voice assistance to video monitoring and hardware access control–together forming a robust and advanced infrastructure of financial transaction security.

Technical and software solutions currently available

Today, technical solutions for face recognition encompass both classical algorithmic approaches and modern neural network models, the latter having become the standard in practical applications.

The earliest implementations were based on geometric methods, where key facial points (such as the distance between the eyes, the shape of the nose, or the contours of the mouth) were represented as feature vectors. Examples include the Eigenfaces approach (Principal Component Analysis, PCA) and Fisherfaces (Linear Discriminant Analysis, LDA). These methods were computationally efficient but sensitive to lighting, head rotation, and facial expression changes.

The next advancement came with local descriptors such as Local Binary Patterns (LBP) and Histogram of Oriented Gradients (HOG), which captured texture-based features of the face. These improved robustness under varying external conditions but still fell short in accuracy compared to more modern techniques.

A true breakthrough arrived with the introduction of deep convolutional neural networks (CNNs). Systems based on models like DeepFace (Facebook, 2014), FaceNet (Google, 2015), and DeepID (Chinese University of Hong Kong) achieved accuracy exceeding 99% on benchmark datasets such as LFW (Labeled Faces in the Wild). These models map each face into a high-dimensional feature vector, enabling highly reliable identification despite differences in pose, lighting, or age.

In current practice, open-source libraries and frameworks are widely used. Dlib is among the most popular, providing CNN-based face detection and comparison. Another example is OpenFace, an open-source project inspired by FaceNet. For commercial use, cloud-based services such as Microsoft Azure Face API, Amazon Rekognition, and Face++ offer ready-to-use modules for real-time face detection and identification.

A key trend is the adoption of edge-computing solutions, where recognition algorithms run directly on devices (smartphones, cameras, or even ATMs) without transmitting data to the cloud. Technologies such as Apple Face ID and Huawei 3D Face Recognition employ specialized sensors to construct 3D depth maps of faces, significantly enhancing protection against spoofing via photos or videos.

Problem definition

In modern conditions of automated customer service at banking terminals and ATMs, a few pressing challenges arise that significantly affect the reliability, security, and user experience of these devices. One common issue is vandalism and deliberate obstruction of terminal operation by inserting foreign objects into the intake slots. Often, malicious actors insert strings, wires, glue, foreign cards, or other items that interfere with the functioning of card readers or cash acceptors (Figure 1).

Such attacks not only disrupt transactions but also create a risk of mechanical damage and necessitate costly maintenance, especially when it comes to recycling ATMs. The second challenge is analyzing the emotional state of the customer as a security element. In situations where another person is nearby, or the customer has just received a call from someone attempting to coerce cash withdrawal, the client is under intense emotional pressure. Modern video-stream analysis algorithms should be able to detect signs of stress or fear-such as trembling hands, rapid breathing, or restless facial expressions-in order to take prompt action: display warning signals on the screen, trigger a “silent” alert to security, temporarily block the terminal, or switch it to an additional authentication mode.

The second challenge is analyzing the emotional state of the customer as a security element. In situations where another person is nearby, or the customer has just received a call from someone attempting to coerce cash withdrawal, the client is under intense emotional pressure.

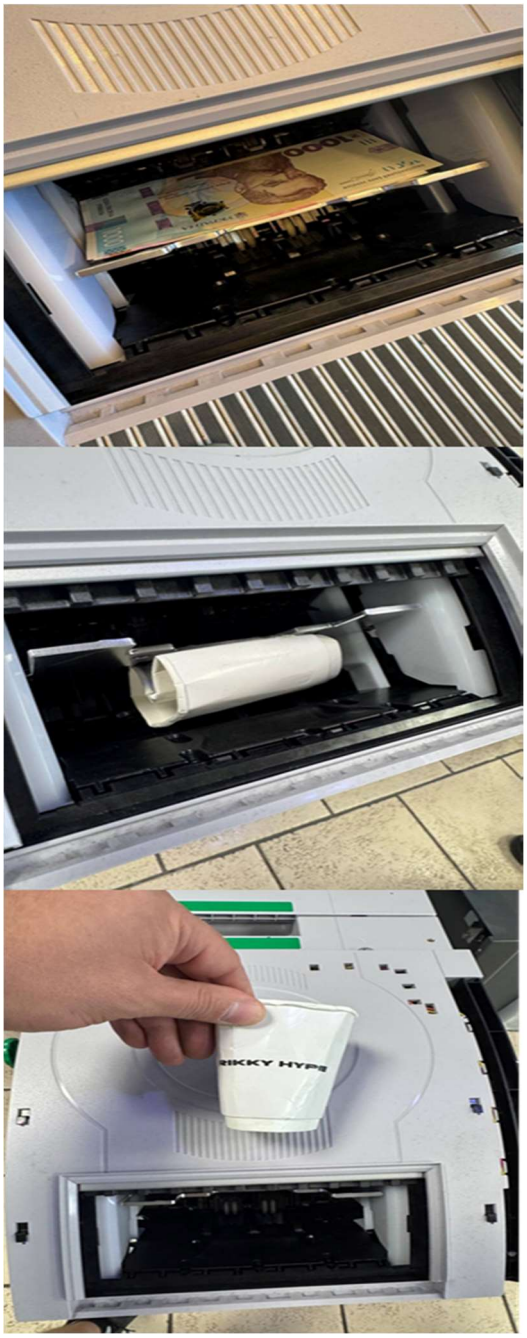

Figure 1 – Examples incorrect use of cash acceptor

Modern video-stream analysis algorithms should be able to detect signs of stress or fear – such as trembling hands, rapid breathing, or restless facial expressions–in order to take prompt action: display warning signals on the screen, trigger a “silent” alert to security, temporarily block the terminal, or switch it to an additional authentication mode.

A third, equally important issue is ensuring accessibility for people with visual impairments. Increasingly, modern service systems incorporate voice commands and audio guidance that instruct the user where to insert the card, which PIN code to enter, and which buttons to press.Additionally, the combination of voice prompts with an interactive voice assistant significantly enhances the accessibility of such terminals.

However, developing such voice interfaces must consider sensitivity to background noise, language localization, clarity of instructions, and confidentiality (e.g., the option for private headphone interaction). While this adds complexity to the software component of the system, it substantially improves inclusivity.

Another contemporary challenge is integrating marketing offers into the customer service process while maintaining a balance between personalization and security. For example, after successful authentication, adaptive offers–such as new financial products, credit lines, insurance, or cashback programs–may appear on the ATM screen. However, it is important that such advertising messages are unobtrusive, relevant to the customer's profile (considering their previous service history), and do not slow down critical operations. It is also necessary to ensure that marketing elements do not create additional risks–for instance, distracting the customer while entering the PIN code or encouraging manipulative psychological pressure (e.g., excessive advertising stimulation prompting quick reactions).

Modeling the interaction protocol

Modeling the interaction protocol involves formalizing the exchange between three entities–the user, the terminal, and the decision-making system–as a controlled state machine with clearly defined messages, time constraints, and security policies (Figure 2).

The session is initiated by starting the terminal session (T0), during which a one-time session token and server certificate are obtained. The terminal then launches the video channel and local liveness-check procedures (blink/pose/texture), constructs a feature vector zt, and sends it over a secure channel (TLS 1.3 + mutual authentication) as a compact packet with a timestamp and nonce.

After initial validation (T1), the server performs dual-channel authentication: biometric (matching zt against templates or via one-shot/Siamese comparison) and behavioral (analyzing gaze trajectories, keypress latencies, gesture geometry) with specified trust thresholds, while also checking for signs of spoofing (replay, print/video attacks) using scene signatures and optical indicators. If sufficient confidence is achieved, a decision

$$at \in \{APPROVE, CHECK{-}AGAIN, REJECT\}.$$

is generated and returned to the terminal along with a policy for further actions (e.g., requesting a repeated gesture or OTP via a linked channel), with the total latency for the "frame → decision" cycle not exceeding the defined SLA (e.g., 150-200 ms with local precomputation).

The protocol includes a degradation mode: in case of network loss, critical liveness and detection checks are performed locally on the terminal, while the server component switches to asynchronous transaction verification with delayed confirmation. To reduce bandwidth usage, video is transmitted in the form of features and metadata (RoI vectors, keypoints, scene hashes), while the raw stream is stored locally in a cyclic buffer and transmitted only in case of an incident or auditor request. Confidentiality is ensured through a data minimization policy (on-device templates, differential privacy for telemetry, pseudonymization), and integrity is maintained via message signatures and order verification (sequence numbers, nonce-based freshness).

Decision-making logic is modeled as an MDP with rewards for speed and security, where actions include selecting the verification level, re-requesting a frame, escalating to an operator, or blocking; policy training can be performed offline on simulated attack scenarios and online under safe constraints. Anomaly detection in the protocol is implemented via a separate loop that analyzes event sequences (e.g., atypical cancellation rates, unusual gesture trajectories, timestamp inconsistencies) and adjusts adaptive thresholds. Session termination is accompanied by a final signed audit record containing hashes of key artifacts, storage policies, and user consent markers, ensuring reproducibility and regulatory compliance.

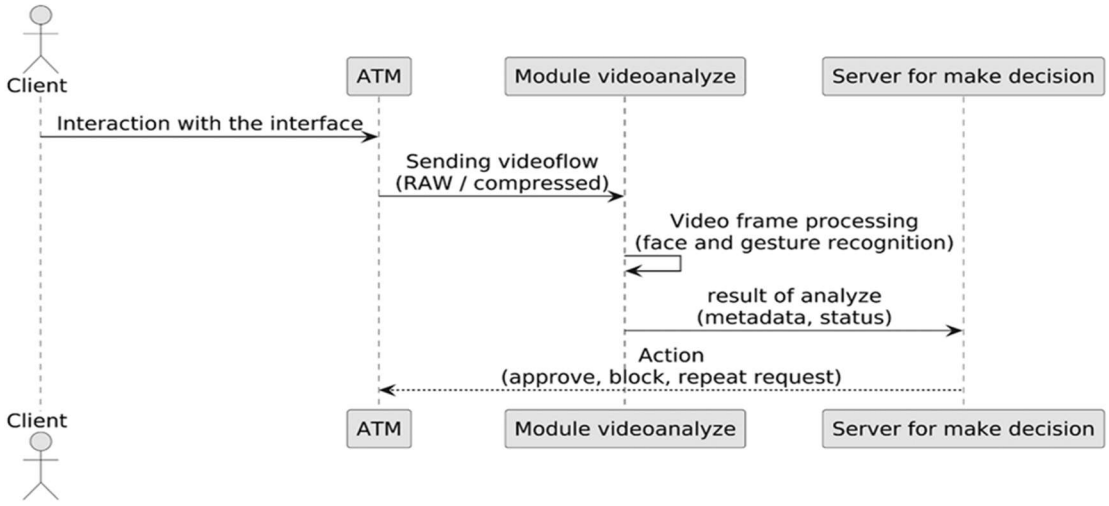

Figure – Schema of model of the interaction protocol

This model enables the integration of video analytics, multi-factor authentication, and cybersecurity into a single protocol with formally defined correctness, latency, and resilience properties against common attacks.

Risk analysis of biometric data processing and legal compliance considerations

The deployment of facial recognition and behavioral biometrics in ATM entails specific risks to privacy, data protection, and regulatory compliance. Below is a structured analysis aligned with common principles of personal data protection laws (including GDPR-like frameworks and national legislation on personal data protection).

The main risks consist in re-identification of pseudonymized templates, cross-context linkage with external datasets, over-retention of raw streams, template theft enabling cross-system attacks, third-party processor misuse.

To prevent the reproduction of the listed threats, it is proposed to use both hardware and software solutions, including:

– On-screen notice or sticker about video surveillance (most ATMs are in public places, which provides a legal basis for video monitoring);

– Use network-level solutions to encrypt transmitted data (HSM, TLS, least-privilege access, digital signatures);

– Incident management (72-hour window), forensics with minimal exposure;

– Clear separation between security data and marketing.

Conclusion

The method of integrated video authentication of users in financial terminals has a range of practical application scenarios. One basic example is the enhancement of a conventional ATM, where the standard card and PIN verification is complemented by automatic biometric identification via facial recognition and behavioral features, reducing the risk of PIN and card compromise. A similar approach can be applied in self-service payment kiosks, for instance, when paying utility bills or topping up accounts, where authentication through a video channel lowers the likelihood of using stolen or forged credentials.

Another example is the integration of the method into remote banking office systems, where the client interacts with the terminal without an operator present: real-time video analysis allows identity verification during transactions or confirmation of large transfers. A promising application is in mobile ATMs deployed in crowded areas or field conditions, where conventional security mechanisms are limited. The method can also serve as an additional layer of protection for corporate financial terminals in large organizations, where video authentication of employees helps prevent access by unauthorized individuals.

Overall, these examples demonstrate that this method is suitable wherever there is a risk of payment fraud and where rapid user identity verification is critical without the need for staff intervention.

References

1. Castelblanco, A., Rivera, E., Solano, J., Tengana, L., López, C., & Ochoa, M. (2022). Dynamic face authentication systems: Deep learning verification for camera close-up and head rotation paradigms. *Computers & Security*, 115, 102629. https://doi.org/10.1016/j.cose.2022.102629.

2. Aljuaid, S. M., & Ansari, A. S. (2022). Automated Teller Machine Authentication Using Biometric. *Computer Systems Science and Engineering*, 41 (3), 1009–1025. https://doi.org/10.32604/csse.2022.020785.

3. Ironvest, Inc. (2024). *User authentication and transaction verification via a shared video stream* (World Intellectual Property Organization Patent No. WO 2024194747A1). URL: https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2024194747.

4. Viatec. (2024). *U-PROX – convenient and secure access to ATMs using cards*. Technical Report.

5. Yamada, S., Aoki, T., Shimoyama, T., & Mori, S. (2019). Biometric authentication technology facilitating protection and management of biometric data. *Fujitsu Scientific & Technical Journal*, 55 (5), 53–58.

6. Diebold Nixdorf & Samsung SDS America. (2017). *Diebold Nixdorf and Samsung SDS demonstrate next step in cardless ATM transactions with mobile-based biometric authentication* [Press release].

7. Biloshchytskyi, A. O., Dikhtiarenko, O. V., & Paliy, S. V. (2015). Searching for partial duplicate images in scientific works. *Management of Development of Complex Systems*, 21, 149–155.

8. Tsiutsiura, S. V., Tereikovskyi, I. A., & Paliy, S. V. (2013). Application of neural networks for recognizing "the ideal interlocutor" among social network users. *Control, Navigation and Communication Systems*, 4 (28), 123–126.

9. Tsiutsiura, S. V., Tereikovskyi, I. A., & Paliy, S. V. (2014). Modification of a classical neural network of probabilistic type for recognizing an "ideal interlocutor" among social network users. *Management of Development of Complex Systems*, 19, 118–123.

10. Hozak, Ya., & Paliy, S. (2024). Modern small networks for image classification. Feature analysis. *Management of Development of Complex Systems*, 60, 221–229. DOI: 10.32347/2412-9933.2024.60.221-229.

The article has been sent to the editorial board 30.11.2025

Палій Сергій Володимирович
Кандидат технічних наук, доцент кафедри інформаційних систем та технологій,
https://orcid.org/0000-0001-9742-1116
Київський національний університет імені Тараса Шевченка, Київ
Бойко Дмитро Михайлович
Аспірант кафедри інформаційних систем та технологій,
https://orcid.org/0009-0006-1195-1065
Київський національний університет імені Тараса Шевченка, Київ
Шабала Євгенія Євгеніївна
Кандидат технічних наук, доцент кафедри кібербезпеки та комп'ютерної інженерії,
https://orcid.org/0000-0002-0428-9273
Київський національний університет будівництва і архітектури, Київ

ВИКОРИСТАННЯ ВІДЕОКАНАЛУ БАНКОМАТІВ ДЛЯ ПІДВИЩЕННЯ РІВНЯ КІБЕРБЕЗПЕКИ ПІД ЧАС ВЗАЄМОДІЇ З КЛІЄНТАМИ

Анотація. *Стрімке зростання автоматизованих банківських послуг та безкаркових фінансових операцій посилило потребу в удосконалених механізмах автентифікації користувачів, здатних протидіяти шахрайству, соціальній інженерії та біометричному спуфінгу. Традиційні методи автентифікації, що базуються виключно на картках та PIN-кодах, стають дедалі вразливішими до крадіжок, примусу та атак повторного відтворення. У цьому контексті біометричні технології, зокрема розпізнавання обличчя у поєднанні з поведінковим аналізом у реальному часі, є перспективним напрямом підвищення безпеки транзакцій. Це дослідження пропонує інтегрований метод покращення автентифікації користувачів в автоматизованих фінансових терміналах шляхом використання відеопотоку пристрою. Підхід зосереджений на безперервній верифікації особи та моніторингу дій під час взаємодії з користувачем, що підтримується формально змодельованим протоколом взаємодії між терміналом та системою прийняття рішень. Запропонований метод інтегрує відео-біометричну ідентифікацію, виявлення активності та поведінковий аналіз у єдиний робочий процес автентифікації. Було розроблено модель протоколу, яка визначає безпечний обмін даними, часові обмеження та логіку прийняття рішень між користувачем, терміналом та серверним компонентом. Модель підтримує адаптивні результати рішень, включаючи схвалення, повторну верифікацію або відхилення, на основі рівнів довіри, отриманих з біометричних та поведінкових ознак. Використання принципів периферійних обчислень (edge computing) зменшує затримку та навантаження на мережу шляхом передачі компактних векторів ознак замість необроблених відеопотоків. Підхід також включає режими деградації на випадок мережевих збоїв, механізми виявлення аномалій та стратегії мінімізації даних, орієнтовані на дотримання нормативних вимог. Порівняльний аналіз із існуючими світовими практиками демонструє, що запропоноване рішення відповідає сучасним тенденціям мультифакторної автентифікації, розширюючи їх через формалізацію протоколу та відеоаналітику в реальному часі. Результати показують, що інтегрована автентифікація на основі відео може значно підвищити безпеку та стійкість фінансових терміналів без шкоди для зручності або доступності. Запропонований протокол взаємодії забезпечує масштабованість впровадження, підтримує відповідність нормативним вимогам та надає гнучкість для адаптації рівнів автентифікації до умов ризику. Цей метод застосовний не лише до банкоматів, а й до кіосків самообслуговування, терміналів дистанційного банківського обслуговування та мобільних фінансових інфраструктур, пропонуючи надійну основу для захищених систем фінансових транзакцій наступного покоління.*

Ключові слова: розпізнавання обличчя; біометрична валідація; банкомат

Link to publication

APA Paliy, S., Boiko, D., & Shabala, Ye. (2025). Using the video channel of ATM devices to enhance the level of cybersecurity during customer interactions. *Management of Development of Complex Systems*, 64, 171–176, dx.doi.org\10.32347/2412-9933.2025.64.171-176.

ДСТУ Палій С. В., Бойко Д. М., Шабала Є. Є. Використання відеоканалу банкоматів для підвищення рівня кібербезпеки під час взаємодії з клієнтами. *Управління розвитком складних систем.* Київ, 2025. № 64. С. 171 – 176, dx.doi.org\10.32347/2412-9933.2025.64.171-176.