Don't Forget to Include that Camera in the Threat Model: Vulnerability of ATM Systems due to Surveillance Cameras

Piyumi Seneviratne#1, Dilanka Perera[2], Harinda Samarasekara[3], Chamath Keppitiyagama[4], Kasun De Soyza[5], Kenneth Thilakarathna[6], Primal Wijesekera[7]

Abstract— **Video Surveillance Systems (VSS) that are used to provide physical protection to assets and personnel of organizations open up new information channels, but they are often not considered an integral part of the organization's information system. Therefore, more often than not, VSS is not considered when designing and evaluating organizations' information security. Hence, a VSS may weaken the information security of an organization while strengthening physical security. We present such a threat that the VSS used in ATM kiosks of Sri Lankan banks can severely weaken the ATM PIN security due to the ad hoc placement of cameras. While we have observed that in some installations, the video camera directly captures the PIN-pad, we show that forearm movements' visibility is sufficient to infer PINs with a significant level of accuracy. We used a mock-up of an ATM kiosk for our analysis, and we show that a human observer can guess a PIN with 22.5% accuracy within 3 attempts without the PIN Pad's visuals. A computer can infer the PIN using the same footages with an accuracy of 50% using a straightforward algorithm. Critical processes in the banks, such as authentication, are built around the assumption of the confidentiality of the PIN thus invest heavily in the PIN generation process. This well-protected PIN is exposed to the VSS when entering the PIN, thus violating a crucial assumption. However, this violation has hitherto gone unnoticed by the banks' security audits because VSS is not considered an inalienable component of the information system.**

Keywords **- Video Surveillance, Inferring Keyboard Inputs, Side-Channel Vulnerabilities, Computer Vision-Based Attacks, ATM PIN Security, Video Analysis, Shoulder Surfing, Threat Modelling, Information Security.**

I. INTRODUCTION

Security plays a crucial role in an information system. Implementation of Video Surveillance Systems (VSS) has come to a point where it is accepted as a common practice to have VSS indoors and outdoors as a security control [1][2].VSS is used to provide physical security to the assets of organizations. One such scenario is the implementation of VSS inside ATM kiosks.

While VSS provides a mechanism to increase security and thwart any physical security breaches in an ATM kiosk, there is a lack of literature on whether a VSS setup could possibly be leaking confidential and security sensitive data & information of an ATM user, in turn, jeopardizing the overall ATM system's security assumptions.

The Central Bank of Sri Lanka (CBSL), the local governing body for banks, suggests and encourages banks to have VSS [3] in ATM environments. However, it does not provide any specifications or guidelines on the VSS installation. Further, Payment Card Industry PIN Transaction Security Point of Interaction (PCI PTS POI) Security Requirements Guidelines [4] states, "the location for camera installation at ATMs should be carefully chosen to ensure that images of keypad entry are not captured and the camera should support the detection of the attachment of alien devices to the ATM front view and possess the ability to generate an alarm for remote monitoring if the camera is blocked or otherwise disabled". This was the only relevant guideline available regarding the specific instructions on installing VSS in an ATM.

We believe that this specific PCI guideline is there to address the possible issue of leakage of user PIN, from a camera that is directly focused on the pin-pad of an ATM. To this end, we examined whether banks are adhering to the previously mentioned PCI standards and guidelines when installing and managing VSS in ATM environments and the adherence to these guidelines is sufficient to prevent possible PIN leakage through VSS footage of ATM kiosks.

We had interviews with 9 employees from all 3 local banks whose duties are related to the ATM system and its functions and a technical expert in the ATM systems domain. After the interviews, we found out that banks do not follow any formal guidelines when installing the VSS in ATM kiosks. This has resulted in ad hoc placement of security cameras. In our 3 onsite visits from each of the 3 banks, we observed that there are many incidents where the keypad and the keypress events are visible through the VSS footage due to this ad hoc installation practice of surveillance cameras in ATM kiosks.

To further investigate the possibility of predicting the passcode through VSS footage, we extended our study to investigate whether it is possible to infer the typed PIN from a VSS footage. For this we have developed an experimental design setup mimicking an ATM kiosk, having standard prototypes of an ATM pin pad and a camera. We used this to gather video data of users entering different PINs. We were able to show that it is possible to infer the ATM PIN during

Correspondence: P. Seneviratne#1 (E-mail: piyumis@scorelab.org)
Received: 15-02-2021 Revised:24-03-2021 Accepted: 07-04-2021

P. Seneviratne#1, D. Perera[2], H. Samarasekara[3], C. Keppitiyagama[4], K. De Soyza[5], K. Thilakarathna[6] are form University of Colombo School of Computing. (piyumis@scorelab.org, dilankap@scorelab.org, harindas@scorelab.org, chamath@ucsc.cmb.ac.lk, kasun@ucsc.cmb.ac.lk, kmt@ucsc.cmb.ac.lk)
P. Wijesekera[7] is from University of California, Berkeley & International Computer Science Institute, Berkeley, USA (primal@cs.berkeley.edu)

This paper is an extended version of the paper "*Impact of video surveillance systems on ATM PIN security*" presented at the ICTer Conference (2020)
DOI: http://doi.org/10.4038/icter.v14i3.7229

situations even when the keypad and fingertips of the user are not visible to the attacker. We were able to infer the passcode with a 22.5% accuracy which indicated that there is indeed a possibility of inferring the PIN of a VSS footage even if the pin-pad is not visible.

Further, we explored the possibility of intensifying the accuracy of successful PIN inference by having automated the footage analysis rather than naked-eye observation. For this, we used a programme that tracks the forearm movement of the person who is entering the PIN and then feeding the x, y coordinates of the forearm in adjacent keypress events, to our heuristic algorithm to try and infer the PIN. Using this approach, we were able to increase the accuracy level to 50%, relatively higher than the human-observer approach.

Furthermore, this solidifies the risk of revealing security sensitive information by having ad-hoc implementations of CCTV cameras in ATM environments. To the best of our knowledge, this study is the first work that studied the real-world practice of installing VSS inside ATM kiosks and addressing possible threats and implications. And also, previous studies performed by Balzarotti et al. [5] and Xu et al. [6] have shown the possibility of reconstructing typed input with such video recordings. They have tracked the fingertip movements when there is a direct or indirect line of sight to the fingertips during the typing process to reconstruct the inputs. In our study, we show that the visibility of forearm movements is sufficient to infer PINs with a significant level of accuracy.

In this paper, section II describes the background study which was focused on the current practice of installing VSS at ATM kiosks, section III gives the details of the previous work related to reconstructing typed inputs taking different approaches. Section IV and Section V explain our approach to show the existing threat to the ATMs in terms of qualitative analysis and quantitative analysis respectively. Results of our work and the discussion of the results are presented in Section VI and Section VII respectively. Section VIII presents the conclusion and future work in Section IX.

II. Background

We performed a detailed background analysis on the rules and regulations of installing at ATM kiosks. Policies, guidelines, rules, and regulations by both local and global regulatory bodies concerning the ATM system were reviewed to identify if there are any regulatory framework that should follow when surveillance cameras in ATM kiosks.

Policies, guidelines, rules, and regulations by both local and global regulatory bodies concerning the ATM system were reviewed to identify if there are any regulatory framework that should follow when surveillance cameras in ATM kiosks.

CBSL is the governing authority for the financial sector which regulates and supervises banks and selected non-bank financial institutions in Sri Lanka. According to the CBSL, the number of ATM terminals at the end of Quarter 3, 2018 is recorded as 4,296 [7]. The analysis of regulations imposed by local regulatory bodies revealed that there is only one circular issued by CBSL in 2006 concerning the surveillance cameras placed in ATM cubicles [3]. This circular was published addressing all licensed banks of Sri Lanka to encourage the installation of CCTV cameras to enhance the surveillance at ATM kiosks. However, this circular has not referred to any recommendations on the positioning of the surveillance cameras in ATM kiosks.

When analyzing the global regulations relating to the surveillance camera installation at ATM kiosks, PCI ATM Security Guidelines found a greater emphasis on addressing the information security in an ATM kiosk [4]. PCI ATM Security Guidelines encourage the installation of surveillance cameras at ATM kiosks in situations where possible and allowed by law. However, the Guidelines and Best Practices of PCI ATM Security Guidelines state that the location of the surveillance camera installation needs to be carefully identified to ensure that the images of keypad entry are not captured. The surveillance camera needs to be installed to assist in identifying the attachment of alien devices to the ATM front and should possess the capability to produce an alert for remote monitoring if the camera is blocked or disabled.

III. Related Work

We have considered previous studies related to both surveillance systems and inference of keyboard inputs using camera footage.

Andrei Costin [9], has conducted a systematic review of existing and novel threats and vulnerabilities in video surveillance, CCTV and IP-camera systems based on publicly available data to identify security and privacy risks associated with the development and deployment of these systems. The paper also presents a set of recommendations and mitigations to enhance the security and privacy aspects of video surveillance systems. Michael et al. [10] address vulnerabilities associated with Wi-Fi IP based CCTV systems. The authors have considered the relevant vulnerabilities and significance based on confidentiality, integrity, availability and present a framework that can be utilized to minimize the security risks associated.

Mowery et al. [11] in their research, discuss the usage of thermal camera footage of a keypad after a user's typing session, to derive the possible keys pressed. ATM PIN is one of the main focused areas of their study. First, the PIN recovery results from human analysis have been presented. Secondly, the researchers incorporated computer vision techniques to automatically extract the code from the created heat map of the thermal camera data. Even though the automated analysis only slightly outperformed the manual analysis it has demonstrated the potential to scale such an attack scenario in practice.

Balzarotti et al. [5] present a tool named Clearshot which automatically constructs the most probable text from video footage of a user's keyboard typing process. The video is captured using an over the keyboard video camera which has a full view of the hand movements and the keyboard. In constructing the process to recover the text being typed, firstly, the video recording of the typing session of the user is analyzed to identify the hand or finger movements and possible keypresses. The researchers then follow an occlusion-based analysis to find out keypresses. Then, a text analysis process constructs the most probable text that has been typed. Clearshot was able to extract a substantial proportion of typed text by the user and to suggest around 80% words correctly within the first 50 choices of correct words proposed.

The study by Maggi et al. [12] on the dynamic conditions present that key magnifying feedback provided by iPhone, Android and Blackberry mobile devices while typing a PIN are vulnerable to shoulder-surfing attacks. Hence, there is no specific positioning of the attacking camera and the target device. The discussed attack was facilitated by computer vision and image processing techniques to identify possible key magnifying events. This study proposed a fast method to infer keystrokes from either online or offline videos. It concludes that the key magnification feedback is not suitable for applications that require high security. However, both these attacks by Balzarotti et al. [5] and Maggi et al. [12] require the camera to record the typing process that captures a complete view of the hand and finger movement on the desktop keyboard where our study focuses on a situation where only the forearm movements are visible.

Qinggang Yue et al. [13] have used Google glass-based spy camera attack on touch screens to decode the typed input, where the input is not visible to the naked eye. Yet, it needs to have a direct sightline to the fingertips of the user. The basic idea of this approach is to track the movement of the fingertip and use its relative position on the touch screen to detect the pressed keys. By applying the optical flow, deformable part-based model, k-means, and clustering and other computer vision techniques to automatically track and analyze fingertip movements. They were able to decode more than 90% of the typed passcodes.

The studies by Maggi et al. [12], Balzarotti et al. [5] and Qinggang Yue et al. [13] show that the requirement of an undisturbed view of the keypad or typing fingers to infer keypresses and to construct a probable text.

Jin et al. [14] also present a novel vision-based attack towards keyboard inputs. In this study, the researchers have created a tool called ViviSnoop which uses video recording of a typing session to construct the typed phrase. They have analyzed the video with image processing to identify the subtle vibrations of the desk where the keyboard is placed, which occurs with each keypress. It emphasizes the fact that even though they have used a mobile phone camera, a webcam or ordinary surveillance cameras can be used to infer typed inputs despite having a direct sightline on the keyboard.

A recent study conducted by Chen et al. [15] prototypes and evaluates that gaze-based attacks exploit with a video are possible within a short distance and a small angle between the camera and the victim. In this work, it proposes a novel keystroke inference method exploited by recording the eye gaze. As per the study, it was able to infer PINs, unlock patterns and text input to mobile devices. Both of these studies by Jin et al. [14] and Chen et al. [15] states that there is a threat to keyboard inputs from video recordings, even if there is an indirect view of the keyboard or keypad or the user's input scenario. However, they have considered eye gaze and the subtler vibrations without focusing on the keypad to create attack vectors while we use forearm movements to guess the PIN of an ATM transaction present the information security threat.

Shukla et al. [16] show that typing inputs can be reconstructed even if the keyboard is invisible. It analyzes around 200 videos of the typing process which was captured using an HTC phone with a camera focusing on the rear side of the target device. It selects the Spatio-temporal dynamics of the hand of the user typing on the smartphone to reconstruct the typed text. This paper emphasizes the scenario in which a user holds the smartphone in one hand and typing using the other. The scenario does not depend on having a direct or complete view of the keypad of the target device. Shukla et al. were able to infer an average of over 50% PINs in the first attempt and up to an average of over 85% PINs in the tenth attempt. The study by Shukla et al. [16] can be considered similar to our work where both studies do not require the visibility of either the fingertips or keypad to infer the PIN. As we consider the ATM as our threat scenario, the visibility area of the placed surveillance camera inside the ATM kiosk might not be able to either grasp the subtle vibrations of the ATM [14] or record the eye gaze of the subject [15] to possibly infer the user's PIN input. In contrast, our study focuses on how surveillance cameras can become a threat to information security. By taking a scenario where both PIN pad and the fingertips of the ATM user not visible through the surveillance camera and only using the forearm movement of the user during the typing process which intuitively thought to be as an unfavourable situation for an adversary.

IV. Qualitative Analysis

As the first step, we conducted a preliminary study to qualitatively analyse the current state of installing VSS inside ATM kiosks. We have conducted interviews with employees from 03 banks and one technical expert to further study the context and validate our hypothesis. Commencement of interviews was done only after receiving the consent from relevant authorities of the banks. And also the authorities of banks provide us with their consent to publish the analysed results while maintaining the anonymity of the bank. After this qualitative analysis, we have built an experimental design to quantitatively analyse the threat.

A. Interviews with Three Banks

Besides the circular issued by CBSL in 2006, no other specific regulation was found that addresses the conservation of information security which is at stake through surveillance cameras. Hence, we conducted interviews with three leading banks of Sri Lanka (two state-owned banks and one private bank) to explore more information on the purposes of installing surveillance cameras at ATM kiosks, procedures followed by the bank when installing surveillance cameras at ATM kiosks, and overall management of ATM surveillance camera footages. The banks were selected based on convenience and acceptance of our request.

TABLE I
Interview Questions

Question No	Question
Q1	How many ATMs are under operation by the bank?
Q2	Is there a separate vendor(s) to install surveillance cameras at ATMs?
Q3	Are you aware of the PCI ATM Security Guidelines and CBSL regulations concerning the placement of CCTV in ATM kiosks?
Q4	How the monitoring procedure of ATM surveillance camera footage is being managed?
Q5	Who in the bank has access to the ATM kiosk's surveillance camera streams?

Interviews were carried out with employees whose duties are related to the ATM system and its functions. As interviewees, a CIO, a security divisional head, an IT manager, 4 employees in the ATM division at the bank's head office and 2 employees from the centralized surveillance monitoring room were participated in total. Employees were selected based on convenience and face-to-face semi-structured interviews were carried out with a set of predetermined questions (Table I). All the interviews were transcribed and analyzed to identify processes and operations concerning the surveillance cameras at ATM kiosks and their installation procedure.

Based on the responses of the interviews, we identified that there are mainly two requirements for the installation of surveillance cameras at ATM kiosks; 1.) to monitor and capture the face of the person who enters the ATM kiosk 2.) to focus on the cash dispensing area of the ATM as a precautionary against any dispute.

However, it was identified that not all the ATM kiosks in Sri Lanka are equipped with surveillance cameras and there are some ATM kiosks set up with multiple surveillance cameras.

Based on the interviews we had with 03 banks, it was found that Bank A and Bank C install surveillance cameras inside ATM kiosks to meet the above two requirements and have their own documented guidelines for the installation of surveillance cameras, while Bank B does not have any documented guidelines (Table II). Bank A and Bank C have live monitoring of the surveillance video of the ATM kiosks both at branch level and the centralized in head offices. Security division personnel and branch managers have access to those videos, while in Bank B, only branch managers have access to surveillance video footage.

Also, it was stated that all three banks are aware of the PCI ATM Security Guidelines and CBSL guidelines concerning the installation of surveillance cameras at ATM kiosks.

During the interviews, real video streams of ATM surveillance cameras were observed on screens at surveillance control rooms. During this inspection, it was identified that there are incidents where the PIN Pad is visible fully or partially when a customer is entering the PIN due to the ad-hoc installations of surveillance cameras at ATM kiosks. Although the real video streams were not observed in Bank B, it was stated the PIN Pad may be captured through the surveillance cameras at ATM kiosks.

In particular, interviewees from Bank C stated that the bank uses high-quality HD cameras wherein it is even possible to determine the colour of the note that is dispatched from the ATM.

However, with the analysis of information gathered from the three banks, it reveals that the current practices of installing surveillance cameras do not comply with the PCI ATM Security Guidelines and are contradicting with ATM PIN security controls. The banks have not considered the security threat that might affect the ATM PIN security through surveillance cameras.

B. *Interview with the Technical Expert*

A face-to-face unstructured interview was conducted with one of the technical experts of the ATM System Information Technology and services industry in Sri Lanka to explore the knowledge on functions of the ATM system and ATM PIN management. Based on the interview, we gathered detailed insights concerning the operations of the ATM system and its use of HSM (Hardware Security Module).

Banks employ HSM for PIN generation, management, and validation. HSM is a part of the ATM system which is employed for the ATM PIN management validation process. In an ATM system, all the PINs are stored in the HSM. In general, there are two ways for a user to create a PIN. One is the random PIN generated by the HSM through the PIN Mailer and the other way is that the user creates his/her own PIN via ATM. PIN Mailer used to securely print the PIN without revealing it to anyone except the user who owns it. Either way, the PIN Offset is securely stored in this HSM. 3PIN offset is the reference key to the PIN Block which is stored in the client information database whereas PIN Block is the encrypted PIN that is stored in the HSM.

In the user authentication process of the ATM system, when the user enters the PIN, the system validates the entered PIN by matching the PIN Block together with the relevant PIN Offset and HSM. With this mechanism, the user PINs

TABLE II

SUMMARY OF THE INTERVIEW FINDINGS

Description	**Bank A**	**Bank B**	**Bank C**
Date(s) of the interview	24/10/2019	04/11/2019 and 06/11/2019	05/11/2019
National Ratings by Fitch Ratings (Lanka) Ltd. [8]	AA+ (lka)	AA+ (lka)	AA (lka)
Type of bank	Licensed Specialised Banks	Licensed Commercial Bank	Licensed Commercial Bank
Number of ATM terminals owned	310	1141	853
Installer	External party	External party	Own staff
Follows any policies or guidelines when installing cameras	Yes (to meet the requirements mentioned in section IV A)	No (minor feasibility study is conducted with a new ATM installation)	Yes (to meet the above mentioned two requirements; consider the distance and height, the location and the size of the ATM kiosks)
Awareness of PCI DSS guidelines and CBSL guidelines on surveillance camera installation at ATMs	Yes	Yes	Yes
Mechanism to monitor live ATM video stream	Both central and at branch level	Only branch level	Both central and at branch level.
Who has the access to ATM video stream	Security division employees in the control room and branch manager.	Branch manager and higher authorities of the bank.	Branch manager and higher authorities of the bank.

are protected and stored securely to ensure confidentiality, integrity and availability. Hence, the ATM system is mainly operating under the trust assumption that ATM PIN is secured and kept confidential from both the banking system and employees of the bank.

C. Threat Model for the existing ATM System

Based on the information and insights gathered from all four interviews, we developed a threat model for the existing ATM system considering the ATM PIN as the asset.

The existing ATM system operates under the assumption that the ATM PIN is secured, and the system provides confidentiality for the PIN. It also assumes that the PIN is securely stored and is not revealed even to the internal employees of the banking system who are considered to be in the trusted boundary. As shown in Fig. 1, the threat model of the existing ATM system, the banking system functions within a Trust Boundary. ATM kiosks are located in an untrusted environment as their operation needs a publicly open interface. Banks control the communication between trusted and untrusted environments using channel encryption as a security control. Banks have employed HSMs to securely manage PINs, and also, they use physical and technical access controls to ensure the security of the PIN. Hence, banks are under the assumption that the protection of ATM PIN is ensured.

Hence, it is assumed that the ATM PIN is secured within the threat model illustrated in Fig. 1 and further confidentiality of ATM PINs are guaranteed.

However, banks have not considered the VSS in the ATM kiosk in the threat model (Fig. 1). The VSS is only regarded as a physical control to ensure operational security at ATM kiosks by banks and the consequences resulting from the physical security control which might cause threats to ATM PIN security are not considered. Hence, with the observations and finding we gathered during the interviews, we state that there is a threat to the PIN, as well as to the banking information system from the installation of VSS.

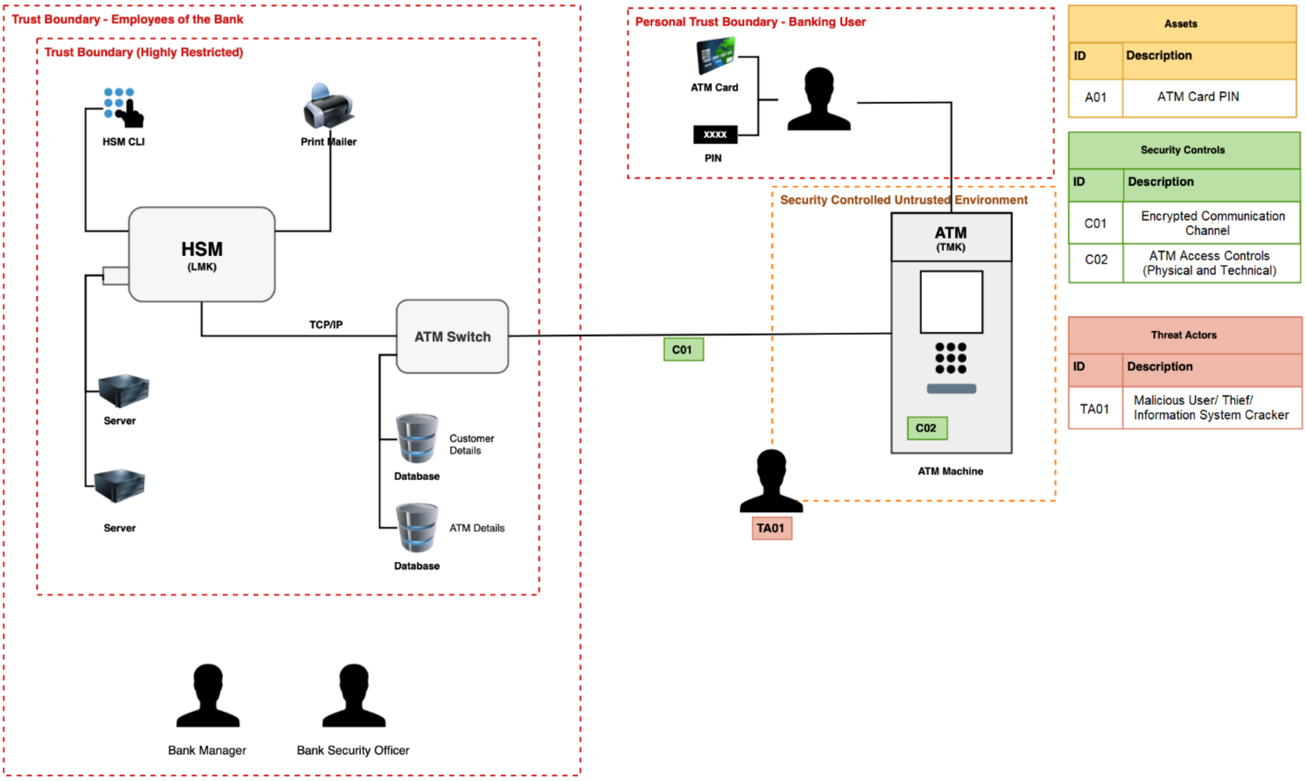

Trust Boundary - Employees of the Bank
Trust Boundary (Highly Restricted)
HSM CLI
Print Mailer
HSM
(LMK)
TCP/IP
ATM Switch
Server
Server
Customer Details
Database
ATM Details
Database
Bank Manager
Bank Security Officer
Personal Trust Boundary - Banking User
ATM Card
XXXX
PIN
Security Controlled Untrusted Environment
C01
ATM
(TMK)
C02
ATM Machine
TA01
Assets
ID
Description
A01
ATM Card PIN
Security Controls
ID
Description
C01
Encrypted Communication Channel
C02
ATM Access Controls (Physical and Technical)
Threat Actors
ID
Description
TA01
Malicious User/ Thief/ Information System Cracker

Fig. 1 Existing Threat Model

V. Quantitative Analysis

As our second step of the research, we demonstrate that the current practice of installing surveillance cameras in ATM kiosks possess a threat to ATM PIN security, even in the scenarios where the PIN Pad is not visible through the surveillance camera. Incorporating the information gathered during the preliminary study, we created an experimental setup to simulate the PIN entering scenario at an ATM kiosk due to legal constraints to collect or obtain actual surveillance camera footage from banks. Fig. 2 shows the implementation flow of our approach and the experimental setup. The actual implementation should consider the VSS camera footage but due to the constraints, we use footage recorded with a mobile phone camera. In our PIN inferencing process, we used the video clips that were obtained by employing the experimental setup and demonstrated the possibility of inferring the PIN from captured surveillance footage.

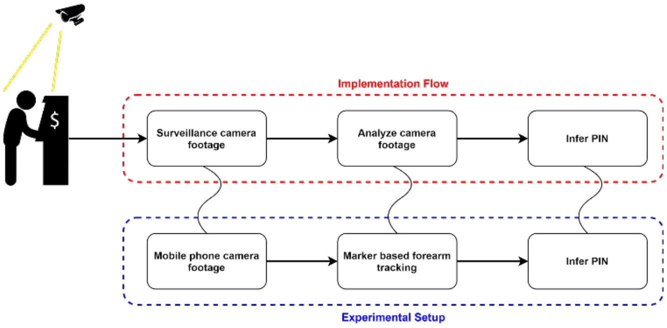

Implementation Flow
Surveillance camera footage
Analyze camera footage
Infer PIN
Mobile phone camera footage
Marker based forearm tracking
Infer PIN
Experimental Setup
$

Fig. 2 Implementation Flow

A. Experimental Setup

1) PIN Pad

We used a PIN Pad prototype (Fig. 3) built using the dimensions of JUSTE6021 ATM PIN Pad [17] (Fig. 4) which is commonly used for ATMs in Sri Lanka as stated by officials, working at a leading commercial bank in SL in an interview on November, 2019. We have created this PIN Pad prototype as an electronic keypad using 12 push buttons and a dot board. The size of the keys and distance between keys were kept identical to the JUSTE6021 PIN Pad (Fig. 4). We set up an LED to the PIN Pad prototype as an indicator to identify the actual keypress events.

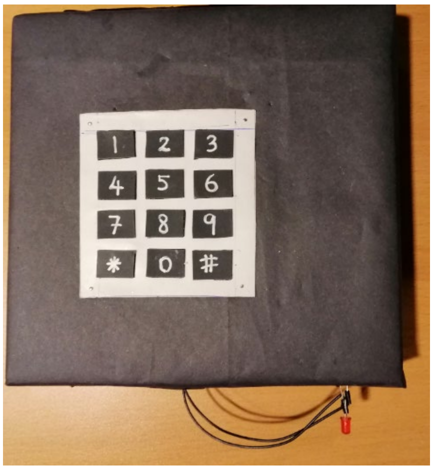


Fig. 3 Prototype of the PIN Pad

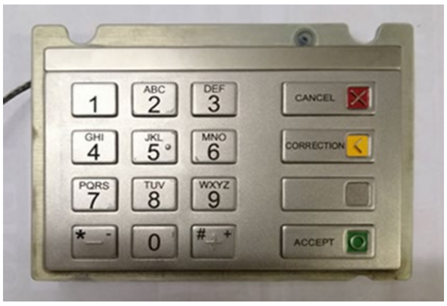


Fig. 4 JUSTE6021 PIN Pad

2) PINs

20 randomly generated PINs were used, and each PIN was containing 04 unique digits.

3) Camera Settings

We captured the PIN entering process using a mobile phone camera which has an f/22 and 26mm focal length. We recorded the videos without any camera effects, zooming, filters, or flashers. In the real-world scenario, some banks have full HD surveillance cameras with 30fps. Therefore, the frame rate was set to 30fps while the resolution was set to 1080p. However, the sensor sizes of the surveillance cameras used in banks are bigger than the mobile phone camera sensor. Hence, actual footage from cameras in ATM kiosks produce better quality video with high resolution and more detail. Unlike the footage we capture, the quality of actual VSS streams is good enough to observe the number of dispense notes, their colour and sometimes even the serial number on the note.

4) Camera Position

During the background study, we were able to observe live footage from VSS control rooms. Two banks of the study state that they focus their camera to capture the cash dispense area. In such cases, forearm movements are clearly visible through the surveillance footage. Considering the observations and the facts gathered, we selected an arbitrary position on the left side of an ATM user who is facing towards the prototype of the PIN Pad to place the camera. As shown in Fig. 5, from that position, forearm movements were visible during the PIN entering process but nor the PIN Pad, neither the fingertips of the ATM user. Hence, it creates a comparatively unfavourable situation for the attacker.

Two users took part in the simulation of the PIN entering process. Both of them only used the index finger of their right hand to enter the PIN. The two users (referred to as User A, User B later) were given all 20 PINs and using this experimental setup we have collected 20 video footages for each user entering given PINs. All those footages were labelled with the respective PIN (referred to as ground truth) entered.

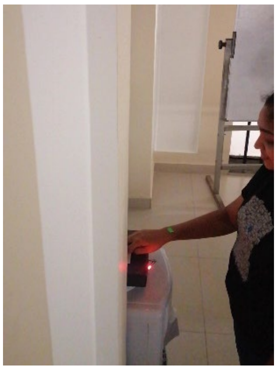


Fig. 5 PIN entry process

B. PIN Guessing

We used two methods to find out the possibility of revealing a PIN by using the video footage of the PIN entering process that we captured using the experimental setup where the PIN Pad or the fingertips are not visible through the video footage.

5) PIN Guessing Method 01 – Human Observer

In this phase, a lab study was conducted with twenty participants of a convenient sample of undergraduates from the University of Colombo School of Computing. Here we have used randomly selected 10 video footages out of 40 recorded. In the first attempt given to the participant, all ten videos without the PIN label were put on a laptop computer display, one at a time. The participant was instructed to observe the forearm movement of the user who is entering the PIN and guess the PIN at the end of each video. Without providing any feedback on whether they have guessed the correct digit(s) or not, three such attempts were given to each participant to watch each video consecutively and to guess the digit(s) of the

PIN. A structured document was given for the participant to record each digit(s)/ PIN they guessed at each attempt.

TABLE IIII
PIN GUESSING PHASES

PIN Guessing Phase	Description
Method 01 – human observer	PIN pad prototype with actual dimensions of JUST E6021 ATM PIN Pad. Lab study on ATM PIN guessing with human observation.
Method 02– computer analysis	PIN pad prototype with actual dimensions of JUST E6021 ATM PIN Pad. An algorithmic approach for automated PIN guessing.

6) *PIN Guessing Method 02 – Computer Analysis*

In the automated analysis, we automated the PIN guessing process by incorporating an algorithmic approach. This phase subsumed the exact video footage from the PIN Guessing Method 01 and the rest of the video footage we collected using an experimental setup. The Computer Analysis approach for PIN guessing contains 04 basic steps as shown in Fig. 6.

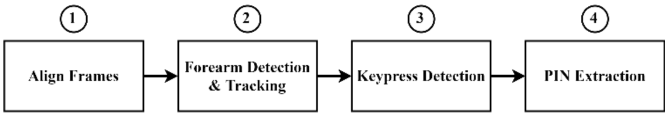


Fig. 6 Steps of Computer Analysis

Step 1: Align Frames: We rotated each frame in the video in a way that rows of the PIN pad align in the X direction and columns of the PIN pad align in the Y direction of the OpenCV coordinate system. Therefore, increment in X direction and increment in Y direction in the forearm movement corresponds to increment in the number values on the PIN pad on the grounds. Ultimately it is straightforward to map the direction of the forearm movements to the PIN patterns.

Step 2: Marker-Based Forearm Detection and Tracking: The main objective of the research was to test whether we can identify PIN patterns entered by a person using forearm movements. Therefore, tracking forearm movements was done by placing a marker that has a contrasting colour from the surroundings on the user's forearm, during the PIN entering process. An OpenCV python program was written to extract the center coordinates of the minimum enclosing circle of the placed marker. The x, y coordinates of the center of the minimum enclosing circle were recorded concerning the frame number to trace the forearm movement during the entire video footage. Alongside, the actual keypress events or the ground truth was also detected using the attached LED to the prototype of the PIN pad. We use an additional OpenCV programme to identify the blinks of the LED more accurately and mark corresponding video frame numbers as "Actual Keypress Events". A researcher may extend this work to fine-tune the code to infer PIN without using a marker to track forearm movements. Fig. 9 show the x, y data point of the forearm in each frame. Actual keypress events identified using the attached LED also marked in the same graph.

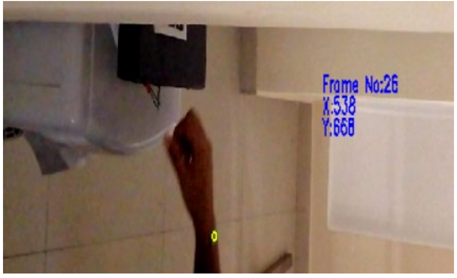


Fig. 7 Orientation of the Video Frame After Alignment

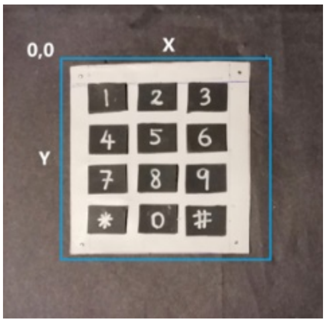


Fig. 8 Aligned PIN pad with OpenCV Coordinate System

Step 3: Keypress Detection: In this paper, we propose a novel method to detect Keypress events considering the gradient of the forearm movements during the PIN entry process. Here, we have applied a heuristic, that the movement of the forearm (in the aligned frames) in the X direction and the Y direction is very little when compared to the movement that happens during the travelling between keys or no movement happens at all when the ATM user is pressing a key. Therefore, the gradient of the forearm movement is close to zero during a keypress event. The net gradient of the forearm was calculated by taking the gradient of the forearm movement in the X direction (dx) and gradient of the forearm movement in the Y direction (dy) using the following equation (1).

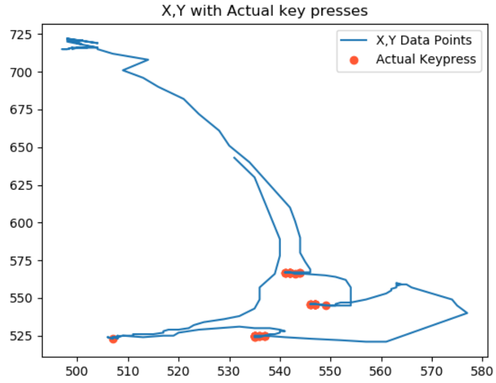


Fig. 9 X, Y coordinates of forearm with actual Key press events

$$Net\ Gradient = \sqrt{dx^2 + dy^2} \qquad (1)$$

According to the heuristic we applied, the keypress events fall on the local valleys of the gradient graph where the gradient is near to zero. Keypress events were identified by extracting these local valleys in the gradient graph. For particular footage, the related frame numbers were backtracked for the extracted local valleys. These frame numbers were clustered into 04 because there are 04 digits in the PIN. Corresponding x, y coordinates of cluster centroids taken as the position of the forearm during a keypress event. Fig 10 depicts forearm movement for PIN 0631 with actual keypress events in dark blue dots. Fig. 11 shows the Gradient graph correspond to the graph in Fig. 10. In Fig. 11, actual keypress events were marked in orange dots. Hence, it proves the heuristic is applicable to find the keypress events using gradient.

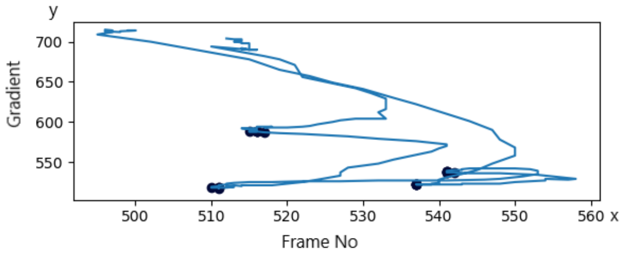


Fig. 10 Traced forearm movement for PIN - 6031

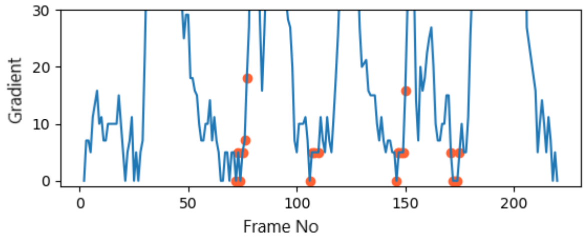


Fig. 11 Gradient Graph of PIN - 6031

C. PIN Guessing Algorithm

As our main focus in this study is on identifying and exhibiting the existence of threats exploiting the VSS and due to the time-consuming data collection process involved in the machine learning-based approach, we opted not to follow a machine learning-based approach for the study. Despite this, we came up with an algorithmic approach to guess the PIN using the x, y coordinates obtained from Step 3. Those coordinates were used to calculate the input parameters to the PIN extraction algorithm. First, the change matrix or the matrix containing the unit difference of rows & columns between adjacent keypresses was calculated for User B.

1) Calculating Change Matrix:

Taking a scenario where a user enters the PIN as 1789 into consideration and following X, Y coordinates were extracted for each key pressed.

- 1 → (517, 513)
- 7 → (523, 560)
- 8 → (532, 567)
- 9 → (545, 567)

And the changes were as follows.

- 1 →7; 0 column change, +2 row down
- 7 →8; +1 column to right, 0 row changes
- 8 →9; +1 column to right, 0 row changes

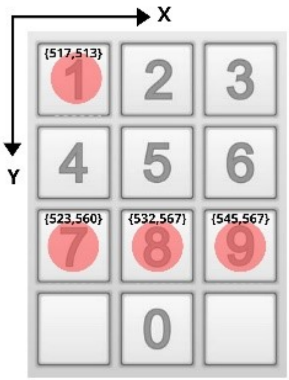


Fig. 12 Coordinates of PIN 1789

These changes were put up as {0, +1, +1} for columns and {2,0,0} for rows. This change matrix is then mapped to possible matching PINs.

To determine the column & row changes between two presses, the coordinates of adjacent presses were subtracted, and the difference was calculated. As shown in Table III, 0 change of columns or rows could still record a movement or difference of X & Y coordinates because of camera tilt and relative hand movement difference even for the same button pressed. In the 1st change, it has recorded a -6 change for X-axis while the ground truth for the change of columns remains 0. Further, it has recorded +47 of a difference of the Y-axis in the 1st change. This leads to two main challenges as follows.

- Identify actual movement in columns or rows. (Did delta of the coordinates for two adjacent keypresses actually depicts a change in rows & columns?)
- Differentiate between a single unit of change and two or more units of change in columns & rows.

2) Calculating Threshold

To address the aforementioned challenges, we established an approach to defining two separate threshold values for columns (tx) and rows (ty). Thus, if only the difference between two adjacent keypresses exceeds the tx or ty threshold value, it was considered as a possible change of columns or rows. An algorithm was developed to determine the most suitable tx and ty using the keypress coordinates of User A.

Each PIN was labelled with its actual value and with keypress coordinates inferred from Step 3: Keypress Detection. Then each PIN was iterated to find the coordinate difference of adjacent two keypresses. Therefore, the resulting row or/and column change was compared with the ground truth. If it results in the same value, the true counter for that specific threshold was incremented. The threshold values which had the highest number of the true count was then determined as the best matching values for tx and ty.

TABLE IV
Column & Row Changes Of Adjacent Keypresses Derived From Coordinates Of User

With the data set of User A which is used to determine the best matching threshold values,

Change	T x+1-Tx	Change of x	Actual change in columns	Ty+1 - Ty	Change of y	Actual change of rows
1st change	517-523	-6	0	560-513	47	2
2nd change	532-517	+15	+1	567-560	7	0
3rd change	545-532	+13	+1	567-567	0	0

- For columns (change of X-axis) →tx = 7 || 8
- For rows (change of Y-axis) →ty = 11

Using the threshold values (tx and ty) derived from the data set of User A, the change matrix for column & row changes between the keypresses were calculated for the dataset of User B. That change matrix was passed to the algorithm and the most probable PIN was matched accordingly. This process was automated for the data set of User B for a total number of 20 PINs.

D. Limitations

We have used a marker-based forearm tracking method mainly due to two reasons.

1) The implementation of more accurate forearm tracking methods was not considered.

2) Due to the video quality and the camera tilt, using bound box methods focused on the forearm was not successful.

The gradient-based key-press detection method is only applicable when the hand is not resting on the PIN pad while the keypresses and the hand is freely moving during the keypresses.

VI. Results

A. PIN Guessing Method 01- Human Observer-based

The responses provided by ten participants during the Lab Study were analyzed to determine the accuracy against the actual PIN in each sample video of the PIN entering process. The responses for all three attempts were considered to discover the number of one, two, three and four (exact PIN) digits correctly identified.

Following equation (2) was applied for the calculation of the accuracy of the Lab Study responses.

$$A(d,n,i) = \frac{\sum_{i=1}^{n}(C_{(d,i)})}{\text{SxV}} \qquad (2)$$

Where *A(d,n,i)* indicates the accuracy of correctly guessing at least d number of digits within n number of attempt(s), $C_{(d,i)}$ is the number of correctly guessed instances with at least d number of digits in i^{th} attempt, *S* and *V* represent the number of subjects and number of sample videos respectively.

As shown in Fig. 13, the analysis of the responses of Lab Study indicates that there is 6% accuracy in guessing all four digits of the PIN by all ten volunteers within the first attempt while the accuracy has been increased to 16% and 22.5% within the second attempt and third attempt respectively. Also, the accuracy of guessing at least one digit of the PIN within the first attempt is 76%. This has been increased to 83% and 85.5% in the second and third attempts. This emphasizes that the accuracy of guessing four digits (PIN), three digits, two digits and one digit of the actual PIN has been improved with the increase in the number of attempts.

B. PIN Guessing Method 02- Computer Analysis

Out of the 20 PINs, the automated PIN guessing derives the correct PINs in 10 instances. This result (refer to Fig. 14) shows that 50% accuracy has been obtained in guessing the entire PIN while the accuracy has been increased to 80% for at least one digit.

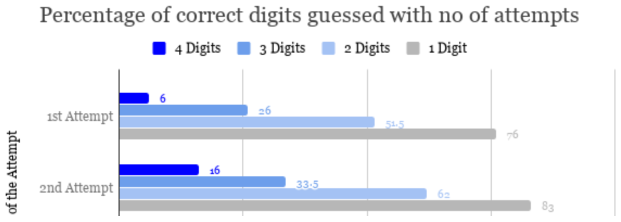


Fig. 13 Results of the Human observer-based Approach

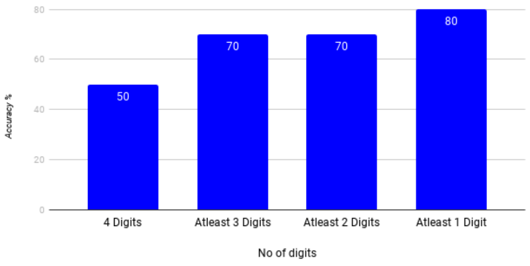
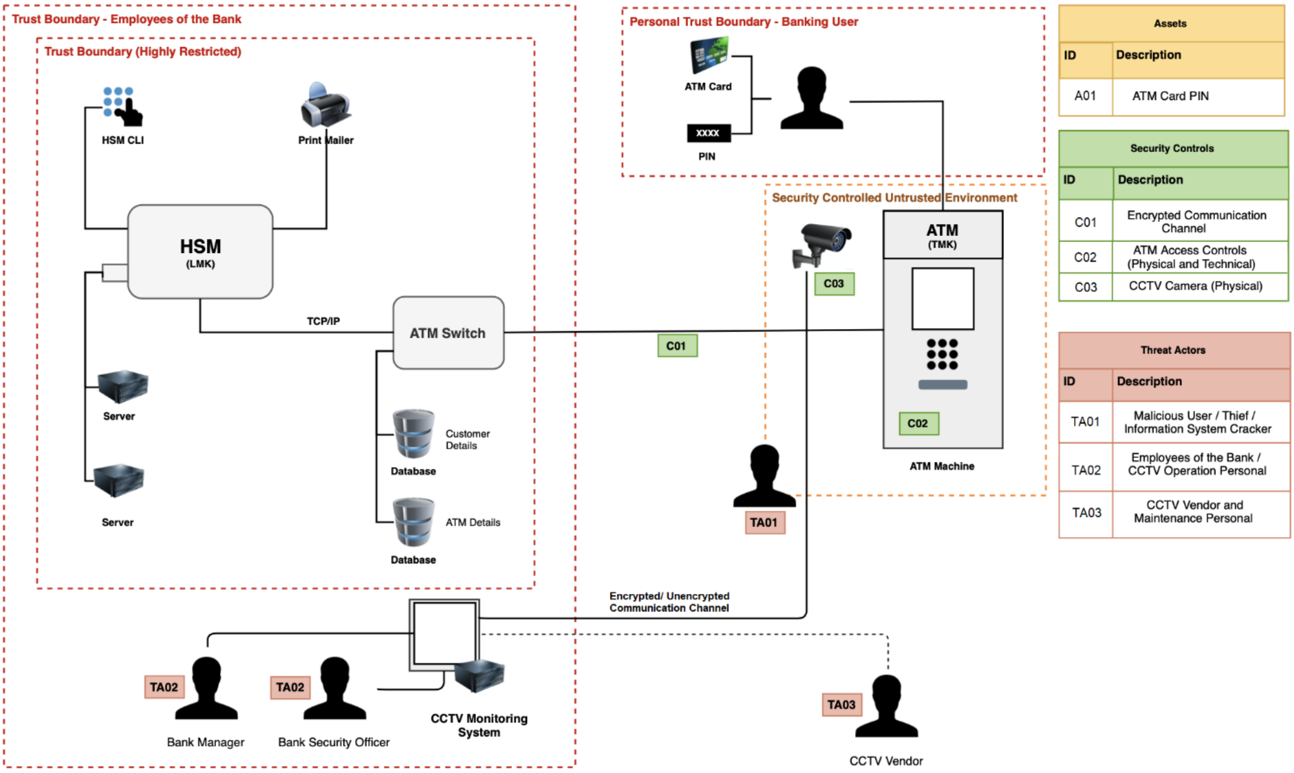


Fig. 14 Results of the Human observer-based Approach



Assets	
ID	Description
A01	ATM Card PIN

Security Controls	
ID	Description
C01	Encrypted Communication Channel
C02	ATM Access Controls (Physical and Technical)
C03	CCTV Camera (Physical)

Threat Actors	
ID	Description
TA01	Malicious User / Thief / Information System Cracker
TA02	Employees of the Bank / CCTV Operation Personal
TA03	CCTV Vendor and Maintenance Personal

Fig. 15 Proposed Threat Model

The Human Observer-based approach indicates that the sight of the forearm during the PIN entry process higher the chances of guessing the PIN correctly compared to a random guess. A random guess has an accuracy of 0.001% (1/10000=0.001%) while the Human Observer-based approach produces 6% accuracy in one attempt. This accuracy improved with the increment of the number of attempts given. This has been further increased with the computer analysis. The human observers considered in this PIN Guessing Method 01 are not experts. Hence, a well-trained human attacker might succeed in guessing the PIN with much higher accuracy.

VII. Proposed Threat Model

In this study, we identified that the ad-hoc installation of surveillance cameras at ATM kiosks results in revealing the ATM PIN itself. It was shown that PIN can be inferred even in situations where both the PIN pad and finger-tips are not visible in the surveillance camera footage during the PIN entering process.

However, the threat model of the existing ATM system (Section IV C. Threat Model for the existing ATM System) does not consider the implications of possessing surveillance cameras at ATM kiosks and the potential threat for ATM PIN security.

Threat modelling requires a holistic approach to accurately identify threats and vulnerabilities. Therefore, the side-channel vulnerability of revealing the PIN through surveillance video footage can only be identified by including the surveillance camera as an actor in the threat model. Hence the potential threats for ATM PIN Security can be identified.

Even though surveillance cameras are installed as a physical control at ATM kiosks to enhance security, it is important to rigorously consider these consequences and to have a holistic approach when developing the threat model. Therefore, it is needed to consider surveillance cameras when defining the threat boundary of the ATM systems’ threat model. Fig. 15 shows an improved version of the threat model that considers this. A holistic approach would assist the banking systems to comply with the basic principles of information as well as physical security when deploying ATM systems.

VIII. Conclusion

PIN is one factor of the two-factor authentication system used in ATM transactions. Banks invest heavily to ensure that a PIN is generated inside an HSM and revealed only to the customer. This indicates that banks operate under the assumption that the PIN is known only to the customer. However, surveillance cameras installed inside ATM kiosks to improve physical security open up a side-channel that can potentially reveal the PIN to third parties. We demonstrated that it is sufficient to capture the forearm movement to infer the PIN with a significant level of accuracy and it is also possible to automate this inference process.

If we consider an ATM transaction such as cardless cash, which allows the user to withdraw a limited sum of money without producing the debit card, the pre-assigned PIN is the sole authentication factor. In such situations, the impact is very high on the security of the banking information system if their VSS possesses side-channel vulnerabilities towards ATM PIN security.

We have further analysed the PINs guessed by human observers, even though those guessed PINs did not match with the actual PIN. When analysing the responses, we have identified that the participants of the lab study have been guessed PINs that have similar trajectory movement to the actual PIN of the related sample video. We have conscientiously considered the angle of the movement from

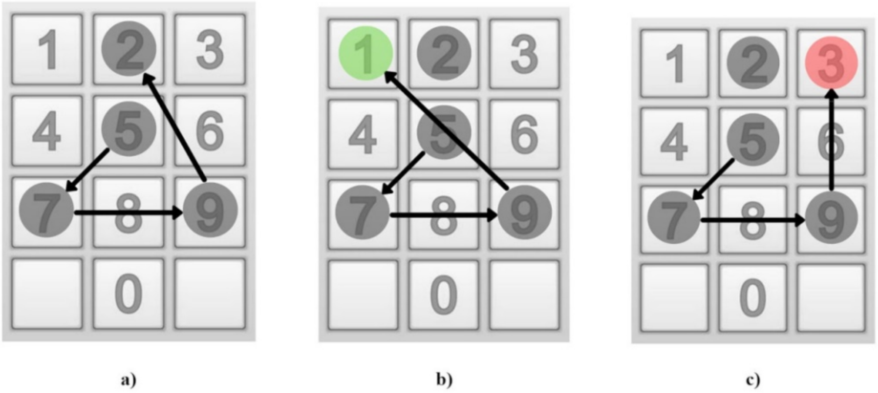


Fig. 16 Determining Trajectories

one digit to another digit when determining the trajectory of the actual PIN to the guessed PIN.

For example, when considering the actual PIN like 5792, the shift from digit 9 to digit 2 has a left-angled movement as shown in Fig. 16 (a). Therefore, when considering a trajectory of a guessed PIN, these facts are analysed. Hence, a guessed PIN like 5731 was regarded as a similar trajectory (Fig. 16 (b)) and a guessed PIN was like 5793 (Fig. 16 (c)) as dissimilar as it has a straight angle movement from 9 to 3.

Accordingly, for each PIN guessed by the participants in the survey for all three attempts, the corresponding trajectory was plotted to identify PIN patterns. The outcome of this further analysis shows that human observers can identify the trajectory of the PIN with 44.5% accuracy. This implies that even the sight of forearm movements can make favourable situation to guess the PIN rather than a random guess.

In this paper, we have taken the PIN entry process at ATM kiosks as a case study to exhibit the possible threats towards information security due to the implementation of VSS. Our algorithmic approach is not the optimal way to infer the PIN using forearm movements, but we present the possibility of PIN guessing simply by tracking the forearm movements.

As a solution to this unforeseen threat toward information systems security, we suggest that proper guidelines and standards on the placement of VSS can mitigate this vulnerability to a certain extent. However, to date, there are no such guidelines, standards, or regulations to govern the surveillance camera placement in Sri Lanka. As conscientious researchers, we took necessary measures to convey the problems disclosed through the research to banks and banking authorities prior to publishing them.

IX. Future Work

Our work was carried out to expose the vulnerability of having VSS in the environments where security-sensitive data is being used. This case study is an example of one such scenario. However, there are other VSS in banking systems that we did not consider for this study. We have observed that banks also use VSS inside their work environment where employees log in to their user accounts in the core banking system. The passwords and usernames might be visible through VSS or it might be possible to infer those data by analysing the VSS footage as well. We plan to extend this study to identify the impact of having VSS inside the bank and as well as other work environments where security-sensitive data and/or the data entry process is revealed through the VSS footage.

Our algorithm brings to light that there is at least one possible method to infer the PIN, even though the PIN pad and fingertips are not visible through VSS footage. Yet, this is not the ideal algorithm to guess the PIN by tracking forearm movements. We intend to improve this work with the use of alternative methods for forearm detection and tracking along with a better algorithm to improve the PIN guessing accuracy.

We followed an experimental study with a limited amount of data. Therefore, we did not consider applying machine learning techniques for computer analysis. We plan to collect a dataset with a large number of PINs collected for different users to take a machine learning approach to guess the PIN. The purpose of this proposed extension to work is to develop a PIN-guessing algorithm that auditors can use to evaluate the security of a VSS.

References

[1] Sikandar, T., Ghazali, K. and Rabbi, M., 2018. ATM crime detection using image processing integrated video surveillance: a systematic review. Multimedia Systems, 25(3), pp.229-251

[2] R. Mandal and N. Choudhury, "Automatic video surveillance for theft detection in ATM machines: An enhanced approach," 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, 2016, pp. 2821-2826.

[3] Supervision Department Central Bank of Sri Lanka, "Directions, Determinations, and Circulars issued to Licensed Commercial Banks", Sri Lanka , 2021 [Online]. Available: https://www.cbsl.gov.lk/sites/default/files/cbslweb_documents/laws/cdg/bsd_LCB_Up_to_30_Nov_2013_compressed_0.pdf. [Accessed: Jan-27- 2021].

[4] Standard: "PCI PIN Transaction Security Point of Interaction Security Requirements (PCI PTS POI)," PCI Security Standards Council, January 2013.

[5] D. Balzarotti, M. Cova, and G. Vigna, "Clearshot: Eavesdropping on keyboard input from video," in 2008 IEEE Symposium on Security and Privacy (sp 2008). IEEE, 2008, pp. 170–183.

[6] Y. Xu, J. Heinly, A. M. White, F. Monrose, and J.-M. Frahm, "Seeing double: Reconstructing obscured typed input from repeated compromising reflections," in Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security, pp. 1063-1074, ACM, 2013.

[7] Payments and Settlements Department Central Bank of Sri Lanka, "Payment Bulletin Third Quarter 2019", Sri Lanka , 2021. [Online]. Available: https://www.cbsl.gov.lk/sites/default/files/Payments_Bulletin_3Q2019.pdf. [Accessed: Jan-27- 2021].

[8] Lmd.lk. 2020. National Ratings by Fitch Ratings (Lanka) Ltd. As at 3 February 2020. [online] Available at: <https://lmd.lk/wp-content/uploads/2020/02/Fitch-monthly_-Feb-2020.pdf>.

[9] A. Costin, "Security of cctv and video surveillance systems: Threats, vulnerabilities, attacks, and mitigations," in Proceedings of the 6th

international workshop on trustworthy embedded devices, pp. 45-54, ACM, 2016.

[10] M. Coole, A. Woodward, and C. Valli, “Understanding the vulnerabilities in wi-fi and the impact on its use in cctv systems" 5th Australian Security and Intelligence Conference, Western Australia, December 2012.

[11] K. Mowery, S. Meiklejohn, and S. Savage, “Heat of the moment: Characterizing the efficacy of thermal camera-based attacks," in Proceedings of the 5th USENIX conference on Offensive technologies, pp. 6-6, USENIX Association, 2011.

[12] F. Maggi, A. Volpatto, S. Gasparini, G. Boracchi, and S. Zanero, “A fast eavesdropping attack against touchscreens,” in 2011 7th International Conference on Information Assurance and Security (IAS). IEEE, 2011, pp. 320–325.

[13] Q. Yue, Z. Ling, X. Fu, B. Liu, W. Yu, and W. Zhao, “My google glass sees your passwords,” Proceedings of the Black Hat USA, 2014.

[14] K. Jin, S. Fang, C. Peng, Z. Teng, X. Mao, L. Zhang, and X. Li, “Vivisnoop: Someone is snooping your typing without seeing it!,” in2017 IEEE Conference on Communications and Network Security (CNS), pp. 1–9, IEEE, 2017.

[15] A. T.-Y. Chen, M. Biglari-Abhari, I. Kevin, and K. Wang, “Context is king: Privacy perceptions of camera-based surveillance,” in2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), pp. 1–6, IEEE, 2018.

[16] D. Shukla, R. Kumar, A. Serwadda, and V. V. Phoha, “Beware, your hands reveal your secrets!,” in Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, pp. 904–917, ACM, 2014.

[17] "Wincor ATM RKL Encrypting Pin Pad E6021 (J6, J6.1) PCI5.0 Approved", *Justtide*, 2021. [Online]. Available: https://www.justtide.com/product/wincor-atm-rkl-encrypting-pin-pad-e6021-j6-j61-pci50-approved.html. [Accessed: 24- Mar- 2021].