Mehrnezhad M, Hao F, Shahandashti SF. Tap-Tap and Pay (TTP): Preventing the Mafia Attack in NFC Payment. *In: 2nd International Conference on Research in Security Standardisation (SSR'15)*. 2015, Tokyo, Japan: Springer.

Tap-Tap and Pay (TTP): Preventing the Mafia Attack in NFC Payment

Maryam Mehrnezhad, Feng Hao, and Siamak F. Shahandashti

School of Computing Science, Newcastle University,
Newcastle upon Tyne, United Kingdom
{m.mehrnezhad,feng.hao,siamak.shahandashti}@newcastle.ac.uk

Abstract. Mobile NFC payment is an emerging industry, estimated to reach $670 billion by 2015. The Mafia attack presents a realistic threat to payment systems including mobile NFC payment. In this attack, a user consciously initiates an NFC payment against a legitimate-looking NFC reader (controlled by the Mafia), not knowing that the reader actually relays the data to a remote legitimate NFC reader to pay for something more expensive. In this paper, we present "Tap-Tap and Pay" (TTP), to effectively prevent the Mafia attack in mobile NFC payment. In TTP, a user initiates an NFC payment by physically tapping her mobile phone against the reader twice in succession. The physical tapping causes transient vibrations at both devices, which can be measured by the embedded accelerometers. Our experiments indicate that the two measurements are closely correlated if they are from the same tapping, and are different if obtained from different tapping events. By comparing the similarity between the two measurements, we can effectively tell apart the Mafia fraud from a legitimate NFC transaction. To evaluate the practical feasibility of this solution, we present a prototype of the TTP system based on a pair of NFC-enabled mobile phones and also conduct a user study. The results suggest that our solution is reliable, fast, easy-to-use and has good potential for practical deployment.

Keywords: Near Field Communication, Mobile NFC payment, Mafia attack, MITM attack, Mobile sensor, Accelerometer, Security, Usability

1 Introduction

NFC payment: Near Field Communication (NFC) payment is an upcoming technology that uses Radio Frequency Identification (RFID) to perform contactless payments. An RFID system has two parts: the RFID tag (card) that can be attached to any physical object to be identified; and the RFID reader that can interrogate a tag within physical proximity, via radio frequency communication. An NFC-enabled payment card has an embedded RFID tag. To make an NFC payment, the user just needs to hold the card in front of an NFC reader for a short while and wait for confirmation. NFC payments are usually limited to rather small-value purchases[1].

[1] For instance, the contactless limit increased from £20 to £30 in 2015 in the UK.

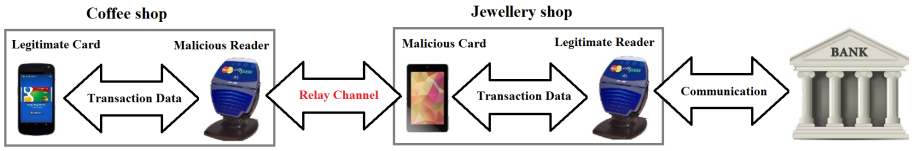

Fig. 1. The Mafia attack: a malicious reader colludes with a malicious card and fools the honest card to pay for something more expensive to a legitimate reader

A mobile phone can also be used as an NFC payment card. HSBC Hong Kong Mobile Payment[2], Google Wallet[3], Apple Pay[4], and Android Pay[5] are examples of NFC payment mobile apps. Using a mobile phone for NFC payment is considered convenient since people can save all of their cards in their phones. It is estimated that mobile payments using NFC will total 670 billion US dollars by 2015 [9]. To support this trend, new generations of smart phones have commonly been equipped with NFC sensors. In this paper, we focus on mobile payment using NFC. Hence unless stated otherwise, by "NFC card", we refer to an NFC-enabled mobile phone functioning as a payment card. By "NFC reader", we refer to a payment terminal that communicates with the card via NFC. A *legitimate* NFC reader is one that is authorised by the banking network and is connected to the back-end banking network for payment processing.

It is known that NFC payment is vulnerable to different types of Man-In-The-Middle (MITM) attacks [21], also known in the literature as relay, or wormhole attacks [19]. In a simple form of a relay attack known as ghost-and-leech attack [22], the attacker places an NFC reader so as to secretly interrogate the user's NFC card without the user's awareness, and relays the card response to a remote NFC reader to obtain a payment from the victim's account. Such an attack is demonstrated in [20] and [21].

Relay attacks can be countered in a number of ways. A simple solution is to put the NFC card within an NFC protective shield such as Id Stronghold[6]. Equivalently, one can add an activation button so that the NFC function on the phone is only turned on with an explicit user action. More advanced countermeasures are proposed in the literature, including *Secret Handshakes* [18], *UWave* [32], *Still and Silent* [37], and *Tap-Wave-Rub* [30]. However, none of these solutions can prevent a more severe type of attack as we explain below.

Mafia attack: Another type of the MITM attack is called the Mafia attack, which is also known as Mafia fraud [19] or the reader-and-ghost attack [38, 22]. In this more severe attack, the user consciously initiates an NFC payment with a legitimate-looking reader controlled by the Mafia; but the reader actually relays the card response to a remote legitimate NFC reader – via a malicious card – to

[2] `www.hsbc.com.hk`
[3] `wallet.google.com`
[4] `www.apple.com/iphone-6/apple-pay`
[5] `www.android.com/intl/en_us/pay`
[6] `www.idstronghold.com`

pay for something more expensive. Figure 1 shows an example of such an attack. This attack has been shown to be feasible in [19].

Unlike simple relay attacks, the Mafia attack cannot be prevented by using a protective shield or an activation button since the user consciously initiates the payment. For the same reason, various user-movement-based unlocking mechanisms [18, 32, 37, 30] cannot stop the attack either. We will explain the current countermeasures to this attack by first reviewing the NFC payment standards and specifications.

NFC payment standards and specifications: EMV is the primary protocol standard for smart card payments in Europe. The EMV standards are managed by EMVCo[7], a consortium of multinational companies such as Visa, Mastercard, and American Express. These standards use smart-cards including contact and contactless cards and are based on ISO/IEC 7816 [4] and ISO/IEC 14443. Mobile NFC payment technologies, such as Android Host-based Card Emulation (HCE)[8], are also based on ISO/IEC 14443, which is an international standard in four parts, defining the technology-specific requirements for proximity cards used for identification [2, 3, 7, 8].

The extensive EMV specifications—presented in 10 books: A [10], B [11], C1–C7 (e.g. [12, 13]), and D [14]—provide the details of EMV-compliant payment system design. Furthermore, EMVCo provides a book on security and key management [1] as a part of EMV 4.3 specifications as well as additional security guidelines for acquirers [5] and issuers [6] of EMV payment cards.

The risk of MITM attacks in payment systems has been generally neglected in the above standards and specifications (except in a recent 2015 EMV Contactless payment specifications Book C-2 [12], as we will explain). As explained by Drimer et al. in [19], such attacks are commonly perceived to be too expensive to work. However, in the same paper, Drimer et al. show this is a misperception by demonstrating practical MITM attacks in a set of live experiments against the UK's EMV system. Given the practicality of deploying such attacks [19] and the projected rapid growth in the size of the contactless payment industry [9], we believe that it is important for the payment industry to seriously consider the security concerns posed by such attacks and the countermeasures that are needed.

Distance bounding protocols: Distance bounding protocols have been considered a potential solution to this problem. In the latest MasterCard EMV specifications (Book C-2 [12] released in March 2015), a distance bounding protocol (called the Relay Resistance Protocol in the specifications) is defined. This protocol starts with the reader sending the card a random challenge and the card replying with a digitally signed response. The reader verifies the digital signature and also checks the response time is within a specified range. This protocol requires an additional private key and a public key certificate installed on the card. Furthermore, the card needs to perform expensive public key operations, which may incur a notable processing delay. To minimize the processing delay on

[7] `www.emvco.com`

[8] `developer.android.com/guide/topics/connectivity/nfc/hce.html`

the card, most distance bounding protocols defined in the literature [16, 19] resort to using only symmetric key operations, such as hash and symmetric-cipher encryptions. However, applying those solutions to NFC payment would require the card and the reader to have a pre-shared symmetric key. In the current practice, the card only has a pre-shared key with the issuing bank. By contrast, our solution does not require any additional cryptographic keys. In fact, it is orthogonal to distance bounding protocols and can be used in conjuction with any one of them.

Other countermeasures: Other countermeasures to the MITM attack have been actively explored by a number of researchers. One straightforward solution is to require user vigilance at the time of making the NFC payment. However, it has been generally agreed that user vigilance alone is not sufficient [22, 33, 38]. It is desirable to design a countermeasure that can effectively prevent Mafia attacks without having to rely on user vigilance. Current solutions generally involve using ambient sensors to measure the characteristics of the surrounding environment, such as light [22], sound [22], location via GPS [33] and a combination of temperature, humidity, precision gas, and altitude [38]. The underlying assumption is that the malicious and legitimate readers will be in two different locations with distinct ambient environments. However, the validity of this assumption may be challenged in some situations where the two readers are in similar environments (e.g., nearby stalls in the same mall).

Overview of our idea: Our idea is based on the following observation: as a result of the physical tapping between a pair of devices (a card and a reader quipped with accelerometers), the tapping creates transient vibrations, which can be measured using embedded accelerometer sensors. By comparing the similarity of the two measurements, we are able to determine if the two devices were involved in the same tapping event. This effectively distinguishes the Mafia attack from a normal NFC transaction.

In contrast to the mentioned solutions, we do not assume that the attacker's reader is in an environment different from that of the legitimate reader. Thus our threat model considers a more severe attack.

Contributions: Our main contributions are summarised below:

1. We propose "Tap-Tap and Pay" (TTP) as a new countermeasure to prevent Mafia attacks. Our solution is the first that works even if the malicious and legitimate readers are in similar environments.
2. We present a proof-of-concept implementation of TTP by using a pair of NFC-enabled smartphones. Experiments confirm that vibrations induced from the same tapping event are closely correlated between the card and the reader, while they are not if originating from different tapping events.
3. We conduct user studies to evaluate the usability of our TTP prototype. Based on the feedback, users generally find the suggested solution fast and easy to use.

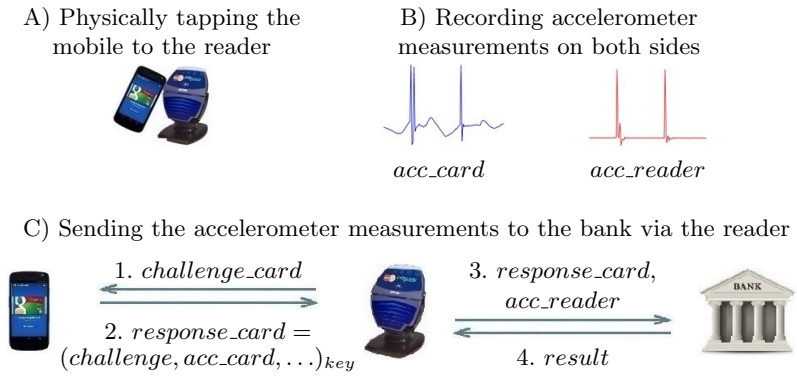

Fig. 2. Overview of the proposed solution: Tap-Tap and Pay

2 Our Solution: Tap-Tap and Pay (TTP)

2.1 Threat model

We assume a user consciously initiates an NFC payment against a legitimate-looking NFC reader without realizing that it is a malicious one controlled by the Mafia. The difference between the malicious reader and the legitimate reader is that the former is not connected to the back-end banking network while the latter is. We assume the Mafia does not want to directly connect to the banking network, as that will run the risk of being caught by the bank. The malicious reader relays the victim's card to a remote legitimate reader to pay for something more expensive, through the help of an accomplice who holds a legitimate-looking NFC card (see Figure 1). From the perspective of the legitimate merchant, there is nothing suspicious – a customer uses a mobile phone to make an NFC payment. The amount of the payment may be near the upper end of the limit, but that is perfectly acceptable (see [19] for a demonstration of successful Mafia attacks on the UK's EMV payment system using contact chip-and-PIN cards; the attacks on the contactless payment work in the same way).

Furthermore, we assume the attacker is able to put the NFC reader in an ambient environment that is very similar to the legitimate reader. In one scenario, the attacker may set up a mobile temporary stall near a shopping mall. He may pretend to sell cheap items such as coffee, tea or confectionery, and show the buyer a small amount on the reader's screen. While accepting the buyer's NFC payment, the attacker relays it to one of his accomplices in nearby shops to buy something more expensive. The attacker and his accomplices can avoid detection by constantly changing the location. Once they make enough profit in a day, they will disappear and repeat the same attack at a different place. Under the above threat model, previous ambient-sensor-based solutions may fail completely. However, despite the assumption of a stronger attacker, we will present a solution that can effectively prevent Mafia attacks under the same condition.

The practical feasibility of such Mafia attacks [19], compounded by the fact that they are undetectable by banks in the backend, can prove problematic. This can have serious implications on the liability if the security of the system only depends on user vigilance. In practice, if any dispute arises regarding the discrepancy of the amount charged for an NFC payment, users will be to blame by default since they are required to be "vigilant". We believe this is not fair to users. Our solution addresses this problem by providing banks more evidence so they can tell apart a legitimate NFC payment from a Mafia fraud. This is done at the minimum inconvenience to users, as we explain in the next section.

2.2 Overview of the solution

An overview of our solution is shown in Figure 2. First, the user physically taps the mobile phone against the reader twice to make an NFC payment. The tapping causes transient vibrations at both devices, which are measured by the embedded accelerometer sensors. The user then holds the card close to the reader. At this point, the reader detects the presence of an NFC card within physical proximity and starts a standard challenge-and-response process for the NFC payment. At a high level, this involves the reader sending a challenge to the card, and the card replying with a response authenticated by a pre-shared key via MAC with the issuing bank. Our solution does not alter this existing data flow; but within the card response, we propose to add an additional item *acc_card* to the items being sent by the card. This new item represents the measurement of the vibration by the card accelerometer. When the reader forwards the card's response to the issuing bank through a secure back-end network, it appends *acc_reader*, which is the measurement of the vibration by the reader accelerometer. The bank compares the two measurements and approves the transaction only if it finds the two sufficiently similar. Recall that in Figure 1, the user's NFC card and the legitimate NFC reader are honest devices and can perform trustworthy measurements.

TTP suggests two taps because we found it to be the minimum number of taps needed to obtain both sufficiently correlated measurements of the same tapping, and at the same time sufficiently uncorrelated measurements of different tappings. Of course with more than two taps, more features can be extracted, but at the expense of user convenience. Hence, we chose double-tap as the default setting for our solution.

2.3 Sensor data preprocessing

To enable data collection, we developed two Android apps: *Card app* and *Reader app* and installed them on two NFC-enabled smartphones, two Nexus 5 devices[9], which are equipped with a range of different sensors.

[9] Prototyping of our TTP protocol requires the facility of bidirectional NFC using Host-based Card Emulation (HCE). At the time of experiments, Nexus 5 was the only device allowing that facility.

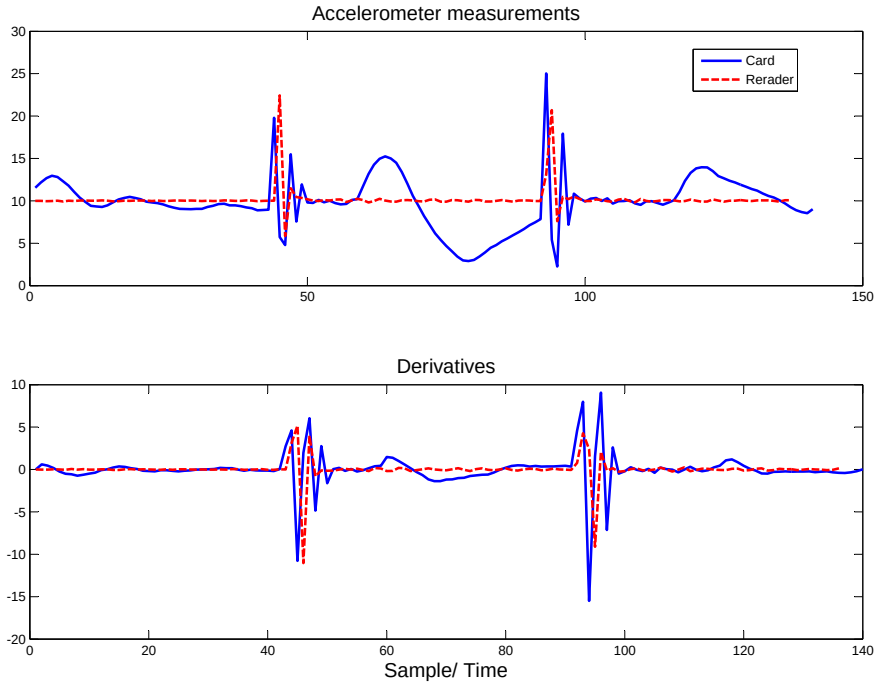

Fig. 3. Final sequences obtained from Equation 1 (top), and their derivatives from Equation 2 (bottom) of a sample of double tapping

Accelerometer data: We use the embedded accelerometer sensor on the mobile phone to capture vibration changes during physical tapping. The accelerometer sensor returns acceleration data in three dimensions, obtained by measuring forces (including the force of gravity) applied to the sensor along the local x, y and z axes. The coordinate system is defined with reference to the phone screen in its portrait orientation; x is horizontal in the plane of the screen from left of the screen towards right, y vertical from the bottom of the screen towards up, and z perpendicular to the plane of the screen from inside the screen towards outside. We consider the sequence representing the length of the three-dimensional vector obtained through accelerometer measurements calculated from Equation 1 where the components represent the i-th measurement in the three dimensions $(acc_{xi}, acc_{yi}, acc_{zi})$:

$$acc_i = \sqrt{acc_{xi}^2 + acc_{yi}^2 + acc_{zi}^2} \tag{1}$$

Fig. 3 (top) shows the above vector length sequences acc_i for a typical double-tapping as measured on a card and a reader. From now on, we refer to this vector length sequence acc_i simply as accelerometer measurement.

Derivatives: As shown in Fig. 3 (top), the accelerometer measurement made by the card is more vibrant than that by the reader, since the card is moving in the hand of the user. They are also different in scale, depending on the start status of accelerometers. In order to smooth out irrelevant movements specially on the card side, we apply the following equation (based on [26]) to approximate the first derivatives of the sequences. The results are displayed in Fig. 3 (bottom).

$$D_i = \frac{(acc_i - acc_{i-1}) + ((acc_{i+1} - acc_{i-1})/2)}{2} \tag{2}$$

Sequence alignment: After obtaining the derivatives, we align the two sequences by identifying the peaks. This can be simply achieved by searching for the extreme values (max or min) with a minimum gap between them. The two sequences are then aligned based on the first peak (with a few linear shifts to get the best matching by trial-and-error). Based on our evaluation of the collected data, we found that this simple alignment algorithm is accurate and fast.

After the alignment of the two sequences, we cut a segment of each sequence, starting from 0.2 seconds before the first peak until 0.2 seconds after the second peak. This covers the whole significant variation of the accelerometer data. Our analysis shows that with this setting, the whole recording time is in the range of 0.6 and 1.5 seconds.

2.4 Similarity comparison

Suggested sensor data comparison methods include correlation coefficients, covariance, cross covariance (e.g. [15]) and cross correlation (e.g. [18] and [22]) in the time domain, and coherence (e.g.[36]) in the frequency domain. We found the correlation coefficients in the time domain and the coherence in the frequency domain to be the two most effective ones on our collected data. Here we use them along with the energy of the series as well as the distance between the two peaks as the inputs of our suggest TTP decision maker.

Correlation coefficient (Time domain). The correlation coefficient is commonly used to compare the similarity of the shapes of two signals. The intuition is that if the two measurements originate from the same double-tap, their signal shapes, especially their tap shapes, would be highly correlated, and otherwise they would not be correlated. Given two sequences X and Y and $\mathrm{Cov}(X, Y)$ denoting covariance between X and Y, the correlation coefficient is computed as below, where $\mathrm{Cov}(X, X) = \sigma_X^2$ and $\mathrm{Cov}(Y, Y) = \sigma_Y^2$:

$$R_{XY} = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Cov}(X, X) \cdot \mathrm{Cov}(Y, Y)}} \tag{3}$$

Coherence (Frequency domain): To obtain a similarity measure in the frequency domain, we apply the coherence method which indicates the level of matching of features in the frequency domain between two time series. Given two sequences X and Y, we compute the magnitude squared coherence based on the following equation, where $P_{XX}(f)$ and $P_{YY}(f)$ are power spectral densities of X and Y, and $P_{XY}(f)$ the cross power spectral density between X and Y:

$$C_{XY}(f) = \frac{|P_{XY}(f)|^2}{P_{XX}(f) \cdot P_{YY}(f)} \tag{4}$$

We define the similarity rate between the two signals based on magnitude squared coherence as the sum of the squares of the magnitudes of coherence values at all frequencies as follows:

$$F_{XY} = \sum_f C_{XY}(f) \tag{5}$$

Energy Difference: Our analysis shows that different users tap devices with different strengths; some taps are very gentle, some are of medium strength, and some are very strong. We found that the total energy levels of the card and reader sequences of the same tap are strongly correlated, while they are distinctive if obtained from different taps. Hence, we use the following measure to capture the distance of two signals X and Y in term of the total signal energy levels:

$$D_{XY} = \left| \sum_t X(t)^2 - \sum_t Y(t)^2 \right| \tag{6}$$

Peak Gap Difference: Last but not least, the distance between the two peaks in each measured sequence is an important factor in deciding if two measurements come from the same double tapping or not. We define G_{XY} in Equation 7 where Gap_X is the distance between the two extremums of sequence X and Gap_Y is the distance defined similarly for sequence Y:

$$G_{XY} = |Gap_X - Gap_Y| \tag{7}$$

TTP Decision Engine: Our TTP decision engine has two steps. First, we have an initial check according to the peak gap defined in Equation 7 and then we use a combined method to include the other three similarity measures. We suggest a simple linear fusion method by using the weighted sum of the three measures: correlation coefficient, coherence, and the energy similarity. Therefore, the ultimate decision is made based on comparing the peak gap against a threshold and if successful comparing the weighted sum of the combined method against another threshold. Hence according to the output of the decision engine, the bank decides to authorize or decline the transaction.

We use a simple linear normalisation that maps the three values to the interval $[0, 1]$. Let us denote these normalised versions by $\bar{R}_{XY}$, $\bar{F}_{XY}$, and $\bar{D}_{XY}$,

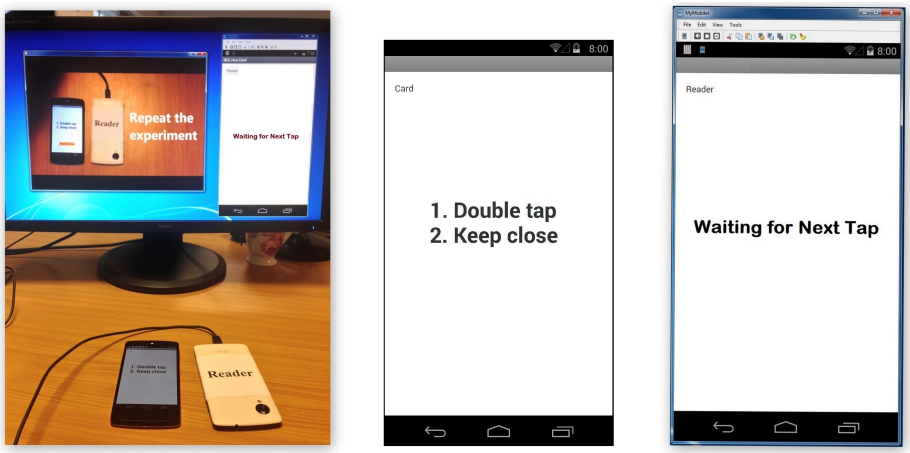

Fig. 4. Data collection environment (left), Card app (centre), and Reader app (right)

respectively. Since unlike the other two measures, $\bar{D}_{XY}$ decreases with similarity, we define $\bar{E}_{XY}$ as below. Note that $\bar{E}_{XY}$ is also a normalised value belonging to the interval $[0, 1]$.

$$\bar{E}_{XY} = 1 - \bar{D}_{XY} \tag{8}$$

Given $\bar{R}_{XY}$, $\bar{F}_{XY}$ and $\bar{E}_{XY}$, T_{XY} calculates the total similarity rate of two signals X and Y as below, where a, b and c are the weights of each method:

$$T_{XY} = a \cdot \bar{R}_{XY} + b \cdot \bar{F}_{XY} + c \cdot \bar{E}_{XY} \tag{9}$$

The weight parameters are determined through experiments based on the collected user data by testing all possible weights up to two decimal places for a, b, and c – under the condition that the sum of them is equal to 1 – and observing the equal error rate. The values which gave us the best error rate have been fixed as $a = 0.45$, $b = 0.21$, and $c = 0.33$.

3 System evaluation

3.1 Experiment setup and Data collection

We implemented a proof-of-concept prototype for the TTP system by developing two Android apps (card and reader). When the user taps the reader, the two apps independently record the accelerometer data. Once the NFC card is detected by the reader in close proximity, the two devices start a two way NFC communication and simulate an NFC payment.

In order to evaluate the system performance based on real user data, we recruited 23 volunteers (university students and staff, 10 males and 13 females) to participate in the data collection, each performing five double tapping actions.

Method	Equal error rate
Correlation coefficients	19.15%
Coherence	27.91%
Total energy	23.48%
Peak gap	14.09%
TTP decision algorithm	***9.99%***

Table 1. Equal error rates for different suggested methods

We made a short self-explanatory training video to demonstrate how to do the double-tap and showed it to the users before the experiment. Users generally found the video guide useful in helping them quickly grasp the instruction of "Tap-Tap and Pay".

We fixed the reader phone to the table using double-sided tape, as shown in Fig. 4 (left). The front of the phone faced downwards and the back was labelled "Reader". We used MyMobiler[10] to operate the reader through a USB connection. The GUIs of the reader and card apps are shown in Fig. 4 , right and centre, respectively. After launching the card app, the user just double tapped the phone to the reader and kept it close to complete an NFC payment. Once she was notified of a successful completion, she could repeat the experiment. The recorded sensor data were saved into a file for further analysis in Matlab.

3.2 Results

We use the False Negative Rate (FNR) and the False Positive Rate (FPR) to evaluate the performance. The FNR is the rate that two measurements from the same tap event are determined as not matching. The FPR is the rate that two measurements from two different tap events are determined as matching. FNR and FPR vary according to a threshold. The Equal Error Rate (EER) is the rate where the FNR and the FPR curves intersect. The EER is commonly used as a measure to evaluate the overall performance of a system. We computed the EERs based on the similarity comparison methods as described in Section 2.4. The results for EER are presented in Table 1. Overall, the Equal Error Rate of our prototype system is 9.99% using the combined method (Table 1). Therefore with this setting, we have FNR= FPR= 9.99%. Hence, a legitimate NFC transaction may be falsely rejected with a probability of 9.99%. Then the user would need to try again. On average, it takes $1/(1 - 0.099) = 1.1$ attempts for a legitimate user to complete an NFC payment transaction. On the other hand, if the Mafia attack takes place during the NFC payment, the transaction is more likely to be denied by the bank due to inconsistent data measurements. The Mafia may trick the user to try again, but it would require on average $1/0.099 = 10$ attempts to get a fraudulent transaction to come through. However, consecutively failed verifications for a single NFC transaction will likely trigger an alert at the back-end banking network, prompting an investigation. Furthermore, when the user

[10] `www.mymobiler.com`

gets repeated denials from the NFC payment (say three times), she might not try further and may choose to query her bank instead. All this can significantly increase the chance of having the Mafia attack exposed.

3.3 Online and offline modes

So far, the description of our TTP solution assumes that the NFC transaction is *online* i.e., the reader is connected to the banking network, so that the backend system is able to evaluate the received measurements and authorize the payment in real-time. The same assumption is made in other researchers' solutions [22, 33, 38] (which we will detail in Section 5).

However in practice, an NFC transaction may be performed *offline*. According to the EMV specifications, an EMV transaction flow includes several steps including *offline* data authentication and *online* transaction authorisation. Depending on the result of the negotiation between the card and the reader, the card may decide to go with offline authorisation. This decision is based on different factors including the transaction value, the type, and the card's record of recent offline transactions. Our solution will be less effective in the offline mode, however, we believe it still provides important added value in preserving critical evidence when a dispute regarding Mafia attacks occurs and a retrospective fraud investigation is needed.

4 Usability study

4.1 Experiment setup and Data collection

We performed a second experiment to evaluate usability aspects of the system. We asked 22 different users (partially overlapped the previous user set, university students and staff, 15 males and 7 females) to perform two NFC payments; first by using the contactless method, and second by using TTP. We developed two Android apps (card and reader) to simulate the two tasks. Before the experiment, we presented users a study description, including a short introduction of mobile contactless payment using NFC, followed by a general description of mobile payment using TTP. In the first task, the user was asked to hold the phone near the reader and wait for the confirmation message. In the second task, the user was asked to double-tap the reader, keep the phone near the reader and wait for the confirmation. Figure 5 shows the GUIs of the two tasks in this experiment.

4.2 Findings

After completing the two tasks, the users were asked to fill in a questionnaire and rate the level of convenience, speed, and feeling of the security of each payment method in a Likert scale from level 5 to 1 (corresponding to "strongly agree", "agree", "neutral", "disagree", and "strongly disagree"). They were also asked to write free comments about their experience in this experiment. Figure 6 shows the average user rating of using the contactless payment and the TTP method.

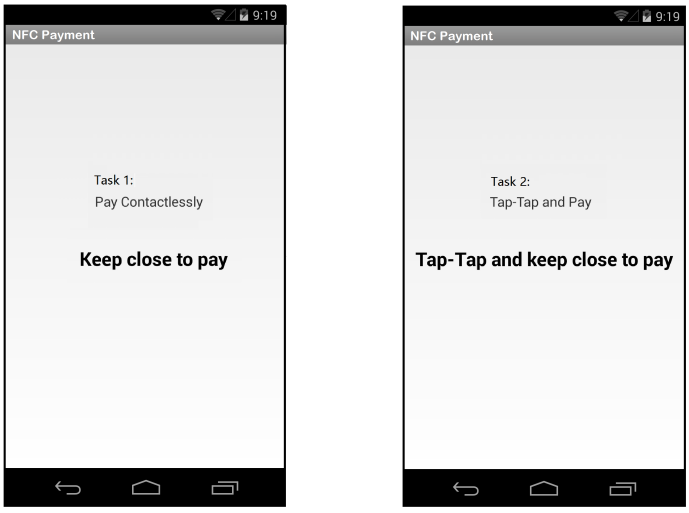

Fig. 5. User study Card app; Task 1: Contactless payment (left), Task 2: TTP (right)

As shown in Fig. 6, users generally found contactless payment more convenient than TTP. Including a physical action makes it less convenient for some users. As one user commented: "... the fact that I need to keep the device close to the reader after tapping made the experience less convenient".

However, in contrast to convenience, many users considered TTP faster than the contactless method, since they were able to precisely sense the start of the action by tapping, while it took them some time to find the proper distance in contactless payment. The uncertainty about when contactless payment would start made some people feel that the process took longer than how long it actually took. As one user commented: "Even [though] I had to tap twice, but the process felt faster comparing to the first one. I feel after tapping I automatically bring the phone close enough to the reader, but in first task, my phone was not close for a while and it took longer".

Moreover, users felt TTP is more secure than contactless payment. By performing a physical tapping action, users felt in control of the transaction and would worry less about accidental payments. As one of the users commented: "As before [i.e. task 1] payment is very easy. I like the action of tapping the reader as this made me feel more in control of when the transaction took place. I felt this method [TTP] was more secure due to the action of tapping to start the transaction. This meant I know when the transaction took place". A similar view was expressed by another user in the comment: "The payment [in task 1] is very easy, but I don't know when the connection between wallet and reader is made; range or time, so I would keep my payment device away from the reader to be sure until I want to pay."

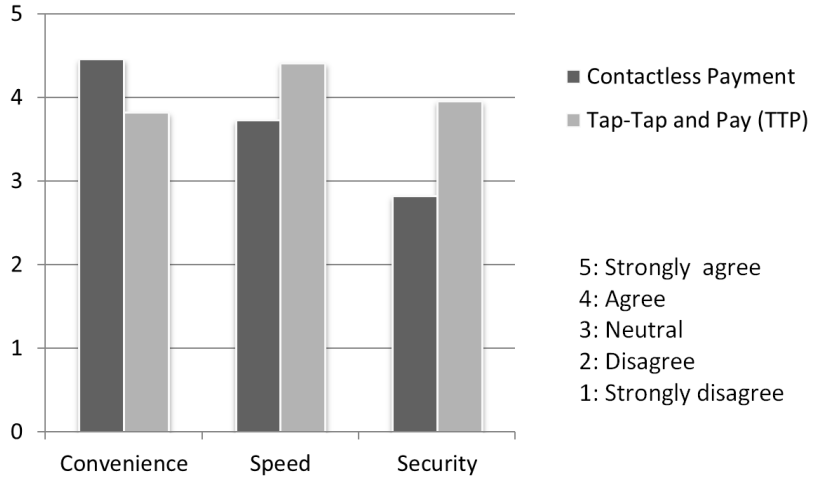

Fig. 6. User study: average user rating of contactless payment and TTP

5 Comparison with previous works

Table 2 briefly compares TTP with previous ambience sensing based solutions. In terms of security, TTP is the first solution able to prevent the Mafia attack even when both readers share the same ambient environment. Ambience sensing solutions are inherently incapable of detecting the attack in this condition.

We now review the error rates reported in the previous works based on measuring the ambient environment. Halevi et al. [22] (sensors: audio and light) report false positive and false negative rates of 0% for audio sensor, and around 5% for light sensor for distinguishing different business types (such as library, concert hall, restaurant, etc.). Ma et al. [33] (sensor: GPS) report a 0% false negative rate under the assumption that the attacker is located 20 meters or farther, 67.5% when the distance is more that 5 meters, and 100% when the distance is less than one meter. False positive rates are not reported in their work. Shrestha et al. [38] (sensors: multiple sensors) report false negative rates approximately in the range of 10%–25% and false positive rates approximately in the range of 15%–30% for individual sensors. By combining the sensor readings, they achieve a false negative rate of about 3% and a false positive rate of about 6%.

The equal error rate of 9.99% in our result is comparable to those reported in the previous works. However, when the two readers are in nearby locations and share the same or similar ambient environments, the reported error rates in [22] [33] [38] are no longer meaningful and all previous ambient-sensor based solutions may fail completely. By contrast, our TTP solution works regardless of whether or not the two readers share similar ambient environments.

In terms of usability, our protocol needs a sensor recording of only 0.6 to 1.5 seconds which is sufficiently fast for contactless payment. Schemes based

Sensor/ Solution	Prevents attacker at same environment	Recording duration (sec)	Embedded mobile sensors	Based on ambience or device
Audio [22]	✗	1	✓	Ambience
Light [22]	✗	2	✓	Ambience
GPS [33]	✗	10	✓	Ambience
Temperature (T) [38]	✗	instant	✗	Ambience
Precision Gas (G) [38]	✗	instant	✗	Ambience
Humidity (H) [38]	✗	instant	✗	Ambience
Altitude (A) [38]	✗	instant	✗	Ambience
THGA [38]	✗	instant	✗	Ambience
Accelerometer (TTP)	✓	0.6–1.5	✓	Device

Table 2. Comparing TTP with ambient sensors based solutions

on audio and light sensors [22] achieve similar timings. However, the GPS-based protocol [33] requires 10 seconds of sensor recording which makes the system not suitable for contactless payment. Our scheme is based on accelerometer which is readily available on most mobile devices, as are microphones (audio), light sensors, and GPS. However, meteorological sensors [38] are only available on specialised devices which is a barrier in adopting such protocols in practice.

In summary, our solution presents a new approach in tackling the Mafia attack with promising initial results in terms of security, efficiency and usability. Being orthogonal ways to solve the same problem, TTP and ambient-sensor-based solutions could potentially be combined to achieve even better results. We leave this as a subject for further investigation in future.

6 Further related works

In this section, we present some other related works that either use *Tap* gesture, or accelerometer sensor data for other security purposes, and explain how TTP differs from them.

Bump. Using the tap gesture to establish device to device communication has been suggested before. Bump[11] is probably the most well-known example in this category. Two users bumps their mobile phones together to exchange contacts, photos and files. Each phone sends a set of data to a remote server, including the device's location (via GPS), the IP address, the timestamp of bumping and the accelerometer measurement. The server matches the devices based on the received data and transfers the data between the two matched devices. Bump and TTP are clearly distinct as they solve different problems and they assume different threat models. Our threat model assumes a malicious reader, whereas in the Bump model, the two devices bumped to each other are assumed to be both legitimate. Consequently, our main goal is to protect against MITM adversaries whereas Bump's main goal is to identify devices being bumped together. In fact,

[11] `www.bu.mp`

it has been shown that Bump is vulnerable to MITM attacks [39] due to timing issues. It is worth mentioning that privacy concerns that arise from environment sensing also apply to Bump since at least the locations and IP addresses of all users in the system are communicated with the Bump server each time the app is used. Since January 2014, Bump has been discontinued with all apps removed from App Store and Google Play [31].

Tap identification proposals. Performing a tap gesture in order to synchronise multiple devices has been proposed in *Synchronous Gestures* [24]. Tap identification using mobile accelerometer is another problem which could also be applied for security purposes. For example *Tap-Wave-Rub* [30] suggests a system for malware prevention for smarphones. Although similar sensors are used in these proposals, they are in general orthogonal to our solution since they are designed to solve an identification problem for legitimate devices, whereas our solution is designed to resist Mafia attacks in an environment where one of the devices behaves maliciously. Consequently, these solutions can be used alongside our proposal to provide a system in which tapping is used to both unlock the device and secure the payment.

Shake to pair. The idea of shaking two devices for device pairing has been suggested by multiple works [35, 36, 34, 15, 27, 28]. While both TTP and the mentioned works use accelerometer, the amount of entropy produced by shaking, the eventual application, the threat model, and the problem solved by these works are all different from ours. In these works, the user needs to shake the two devices together for a while until both devices generate and agree on a shared key, whereas in our work we do not aim to generate shared keys and we only need the user to tap her device to the reader twice. Device pairing, and more generally key exchange cannot prevent Mafia attacks due to the involvement of the malicious reader. Device pairing and securing NFC payments are distinct security problems. While the former has been explored by researchers for a long time [17, 25, 29], the latter is less explored. However, with the impending global deployment of NFC payments, we believe the security of NFC payments deserves more attention by the security community.

7 Conclusion

In this paper, we have proposed a simple and effective solution, called "Tap-Tap and Pay" (TTP), to prevent the Mafia attack in NFC payment by sing mobile sensors. Our solution leverages the characteristics of vibration when an NFC card is physically tapped on an NFC reader. We observed that the accelerometer measurements produced by both devices were closely correlated within the same tapping, while they were different if obtained from different tapping events. The experimental results and the user feedback suggest the practical feasibility of the proposed solution. As compared with previous ambient-sensor based solutions, ours has the advantage that it works even when the attacker's reader and the legitimate reader are in nearby locations or share similar ambient environments.

The TTP solution can be easily integrated into existing EMV standards and requires minimal infrastructural change to the EMV system. The structure of the payment protocol remains the same; only an extra string of accelerometer measurement is added in the transmitted message. In terms of hardware, deploying TTP requires the integration of accelerometer sensors in contactless readers. This can be done progressively by equipping the next generation of the readers with accelerometer sensors which are quite inexpensive (e.g., iPhone 4 accelerometers are estimated to cost 65 cents each [23]). Furthermore, TTP can be rolled out gradually since the protocols remain backward compatible.

In future work, we plan to investigate how to further improve system performance by e.g., combining different sensor measurements and using more precise sensors on newer mobile phones. Moreover, it will also be interesting to explore if it is feasible to apply TTP to other NFC-based payment solutions such as NFC-enabled credit/debit cards, and Barclays bPay band[12] to defend against the Mafia attack by retrofitting accelerometers to such devices.

8 Acknowledgements

We thank all the participants who contributed to our experiments. We also thank the anonymous reviewers of this paper. The second and the third authors are supported by ERC Starting Grant No. 306994.

References

1. Book 2 - Security and Key Management. 2011. Available at `www.emvco.com /specifications.aspx?id=223`.
2. International Organization for Standardization, BS ISO/IEC 14443-1:2008+A1:2012 Identification cards. Contactless integrated circuit cards. Proximity cards. Physical characteristics. 2012. Available at `www.bsol. bsigroup.com`.
3. International Organization for Standardization, BS ISO/IEC 14443-2:2010+A2:2012 Identification cards. Contactless integrated circuit cards. Proximity cards. Radio frequency power and signal interface. 2012. Available at `www.bsol.bsigroup.com`.
4. International Organization for Standardization, BS ISO/IEC 7816-4:2013, Identification cards. Integrated circuit cards. Organization, security and commands for interchange. 2013. Available at`www.bsol.bsigroup.com`.
5. EMV Acquirer and Terminal Security Guidelines. 2014. Available at `www.emvco.com/specifications.aspx?id=71`.
6. EMV Issuer and Application Security Guidelines. 2014. Available at `www.emvco.com/specifications.aspx?id=71`.
7. International Organization for Standardization, BS ISO/IEC 14443-3:2011+A6:2014 Identification cards. Contactless integrated circuit cards. Proximity cards. Initialization and anticollision. 2014. Available at `www.bsol.bsigroup.com`.

[12] `www.bpayband.co.uk`

8. International Organization for Standardization, BS ISO/IEC 14443-4:2008+A4:2014 Identification cards. Contactless integrated circuit cards. Proximity cards. Transmission protocol . 2014. Available at www.bsol.bsigroup.com.
9. Mobile payment strategies: Remote, contactless & money transfer 2014–2018. Market leading report by Juniper Research, July 2014. Available online at http://www.juniperresearch.com/reports.php?id=726.
10. EMV Contactless Specifications for Payment Systems, Book A: Architecture and General Requirements. 2015. Available at www.emvco.com /specifications.aspx?id=21.
11. EMV Contactless Specifications for Payment Systems, Book B: Entry Point. 2015. Available at www.emvco.com/specifications.aspx?id=21.
12. EMV Contactless Specifications for Payment Systems, Book C2: Kernel 2 Specification. 2015. Available at www.emvco.com/specifications.aspx?id=21.
13. EMV Contactless Specifications for Payment Systems, Book C3: Kernel 3 Specification. 2015. Available at www.emvco.com/specifications.aspx?id=21.
14. EMV Contactless Specifications for Payment Systems, Book D: Contactless Communication Protocol. 2015. Available at www.emvco.com /specifications.aspx?id=21.
15. D. Bichler, G. Stromberg, M. Huemer, and M. Löw. Key generation based on acceleration data of shaking processes. In J. Krumm, G. Abowd, A. Seneviratne, and T. Strang, editors, *UbiComp 2007: Ubiquitous Computing*, volume 4717 of *Lecture Notes in Computer Science*, pages 304–317. Springer Berlin Heidelberg, 2007.
16. S. Brands and D. Chaum. Distance-bounding protocols. In T. Helleseth, editor, *Advances in Cryptology — EUROCRYPT 93*, volume 765 of *Lecture Notes in Computer Science*, pages 344–359. Springer Berlin Heidelberg, 1994.
17. M. K. Chong and H. Gellersen. How users associate wireless devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1909–1918, New York, NY, USA, 2011. ACM.
18. A. Czeskis, K. Koscher, J. R. Smith, and T. Kohno. RFIDs and secret handshakes: Defending against ghost-and-leech attacks and unauthorized reads with context-aware communications. In *Proceedings of the 15th ACM conference on Computer and communications security*, pages 479–490. ACM, 2008.
19. S. Drimer and S. J. Murdoch. Keep your enemies close: Distance bounding against smartcard relay attacks. In *Proceedings of 16th USENIX Security Symposium on USENIX Security Symposium*, SS'07, pages 7:1–7:16, Berkeley, CA, USA, 2007. USENIX Association.
20. M. Emms and A. van Moorsel. Practical attack on contactless payment cards. In *HCI2011 Workshop-Heath, Wealth and Identity Theft*, 2011.
21. L. Francis, G. P. Hancke, K. Mayes, and K. Markantonakis. Practical relay attack on contactless transactions by using nfc mobile phones. *IACR Cryptology ePrint Archive*, 2011:618, 2011.
22. T. Halevi, D. Ma, N. Saxena, and T. Xiang. Secure proximity detection for NFC devices based on ambient sensor data. In *Computer Security–ESORICS 2012*, pages 379–396. Springer, 2012.
23. A. Hesseldahl. Apple iPhone 4 parts cost about $188. *Bloomberg Business*, June 2010. Available at www.bloomberg.com/bw/technology/content/jun2010 /tc20100627_763714.htm.
24. K. Hinckley. Synchronous gestures for multiple persons and computers. In *Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology*, UIST '03, pages 149–158, New York, NY, USA, 2003. ACM.

25. I. Ion, M. Langheinrich, P. Kumaraguru, and S. Čapkun. Influence of user perception, security needs, and social factors on device pairing method choices. In *Proceedings of the Sixth Symposium on Usable Privacy and Security*, SOUPS '10, pages 6:1–6:13, New York, NY, USA, 2010. ACM.
26. E. J. Keogh and M. J. Pazzani. Derivative dynamic time warping. In *the 1st SIAM Int. Conf. on Data Mining (SDM-2001), Chicago, IL, USA*. SIAM, 2001.
27. D. Kirovski, M. Sinclair, and D. Wilson. The martini synch. Technical Report MSR-TR-2007-123, Microsoft Research, September 2007.
28. D. Kirovski, M. Sinclair, and D. Wilson. The martini synch: Device pairing via joint quantization. In *Information Theory, 2007. ISIT 2007. IEEE International Symposium on*, pages 466–470, June 2007.
29. A. Kobsa, R. Sonawalla, G. Tsudik, E. Uzun, and Y. Wang. Serial hook-ups: A comparative usability study of secure device pairing methods. In *Proceedings of the 5th Symposium on Usable Privacy and Security*, SOUPS '09, pages 10:1–10:12, New York, NY, USA, 2009. ACM.
30. H. Li, D. Ma, N. Saxena, B. Shrestha, and Y. Zhu. Tap-Wave-Rub: Lightweight malware prevention for smartphones using intuitive human gestures. In *Proceedings of the Sixth ACM Conference on Security and Privacy in Wireless and Mobile Networks*, WiSec '13, pages 25–30, New York, NY, USA, 2013. ACM.
31. D. Lieb. All good things... 2014. Available at `blog.bu.mp/post/71781606704/all-good-things`.
32. J. Liu, L. Zhong, J. Wickramasuriya, and V. Vasudevan. uWave: Accelerometer-based personalized gesture recognition and its applications. *Pervasive and Mobile Computing*, 5(6):657–675, 2009.
33. D. Ma, N. Saxena, T. Xiang, and Y. Zhu. Location-aware and safer cards: Enhancing RFID security and privacy via location sensing. *Dependable and Secure Computing, IEEE Transactions on*, 10(2):57–69, 2013.
34. R. Mayrhofer. The candidate key protocol for generating secret shared keys from similar sensor data streams. In *Security and Privacy in Ad-hoc and Sensor Networks*, volume 4572 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2007.
35. R. Mayrhofer and H. Gellersen. Shake well before use: Authentication based on accelerometer data. In *Pervasive Computing*, Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2007.
36. R. Mayrhofer and H. Gellersen. Shake well before use: Intuitive and secure pairing of mobile devices. *Mobile Computing, IEEE Transactions on*, 8(6):792–806, 2009.
37. N. Saxena and J. Voris. Still and silent: motion detection for enhanced RFID security and privacy without changing the usage model. In *Radio Frequency Identification: Security and Privacy Issues*, pages 2–21. Springer, 2010.
38. B. Shrestha, N. Saxena, H. T. T. Truong, and N. Asokan. Drone to the rescue: Relay-resilient authentication using ambient multi-sensing. In *Proc. Eighteenth International Conference on Financial Cryptography and Data Security*, 2014.
39. A. Studer, T. Passaro, and L. Bauer. Don't bump, shake on it: The exploitation of a popular accelerometer-based smart phone exchange and its secure replacement. In *Proceedings of the 27th Annual Computer Security Applications Conference*, ACSAC '11, pages 333–342, New York, NY, USA, 2011. ACM.