The Cards Aren’t Alright: Detecting Counterfeit Gift Cards Using Encoding Jitter

Nolen Scaife, Christian Peeters, Camilo Velez, Hanqing Zhao, Patrick Traynor, David Arnold
University of Florida
{scaife, cpeeters, camilovelez, hanqingzhao, darnold}@ufl.edu, traynor@cise.ufl.edu

Abstract—Gift cards are an increasingly popular payment platform. Much like credit cards, gift cards rely on a magnetic stripe to encode account information. Unlike credit cards, however, the EMV standard is entirely infeasible for gift cards due to compatibility and cost. As such, much of the fraud that has plagued credit cards has started to move towards gift cards, resulting in billions of dollars of loss annually. In this paper, we present a system for detecting counterfeit magnetic stripe gift cards that does not require the original card to be measured at the time of manufacture. Our system relies on a phenomenon known as jitter, which is present on all ISO/IEC-standard magnetic stripe cards. Variances in bit length are induced by the card encoding hardware and are difficult and expensive to reduce. We verify this hypothesis with a high-resolution magneto-optical microscope, then build our detector using inexpensive, commodity card readers. We then partnered with Walmart to evaluate their gift cards and distinguished legitimate gift cards from our clones with up to 99.3% accuracy. Our results show that measurement and detection of jitter increases the difficulty for adversaries to produce undetectable counterfeiters, thereby creating significant opportunity to reduce gift card fraud.

I. INTRODUCTION

Gift and prepaid cards\(^1\) are an increasingly popular payment mechanism. In the United States alone, such cards enable over $130 billion in sales annually, with predictions of more than $200 billion in transactions annually by 2020 [1]. Retailers like such cards because they often drive consumers to their store, and encourage those customers to make additional purchases beyond the value of the card. Consumers like gift cards because they reduce the guilt of purchasing indulgences or that they allow them to better control their finances [2]. Finally, many security experts recommend using gift cards over traditional credit/debit cards to retain anonymity and reduce the risks associated with retail data breaches.

Unfortunately, gift cards now represent the fastest growing source of fraud for retailers, with suspected losses in the billions of dollars [3], [4]. Much of this rise can be attributed to the rollout of the Europay, Mastercard and Visa (EMV) standard for credit/debit cards in the United States. Like pre-EMV credit/debit cards, gift cards rely on a simple magnetic stripe to encode a static account identifier which can easily be recorded, cloned and replayed in a later transaction. However, the generally short lifetime of gift cards (i.e., many are intended for a single use) makes the use of EMV extremely unlikely - whereas magnetic stripe cards cost less than $0.08 to manufacture, EMV cards cost approximately $2.00 each [5].\(^2\) EMV also cannot be added to gift cards that have already been sold, which represent billions of dollars in unspent balances that must legally be honored by the card issuers [6]. Finally, unlike industry-mandated EMV for credit and debit cards, gift card transaction processing is developed and handled by the merchant ad-hoc or by one of many gift card processors who are free to develop any standard (or none) they wish. As such, EMV can not necessarily be easily dropped into many of these systems.

In this paper, we develop a mechanism to detect cloned gift cards by identifying artifacts of the analog encoding process. Our hypothesis is that cards written in quality-controlled, automated facilities will exhibit more consistent bit lengths on their magnetic stripes. Cloned gift cards, which the majority of investigative reports argue are created by hand through inexpensive encoders [7], [8], [9], provide valid encodings as a strong indicator that a gift card has been encoded in a legitimate facility. Cards exhibiting high jitter are almost uniformly the result of cloning by hand. We show how this observation can be used to detect counterfeiters.

We make the following contributions:

- **Identify Analog Phenomenon Associated with Card Cloning:** We identify low variance in bit length, or jitter, as a strong indicator that a gift card has been encoded in a legitimate facility. Cards exhibiting high jitter are almost uniformly the result of cloning by hand. We show how this observation can be used to detect counterfeiters.

- **Develop Commodity Detector:** We use a range of tools to observe and then test our hypothesis, including imaging via a high resolution magneto-optical microscope capable of quantitative measurement of magnetic microstructures and an array of our own analog measurement tools. We then develop a detector using inexpensive, commodity hardware. Developing this system required expertise in computer science, signal processing and materials science.

- **Experimental Analysis:** We partnered with Walmart to verify the utility and effectiveness of our system. Using 650 gift and stored value cards, our system exhibits accu-

\(^{1}\)Known within the industry as “closed” and “open” loop cards, we use these terms interchangeably unless necessary to describe a specific card in our experiments.

\(^{2}\)A company selling a $10 gift card can easily absorb the cost of a $0.08 (<1% of the total value); at $2.00 for a one-time use EMV card, the company is much less likely to accept a 20% loss.
racy up to 99.9% (99.3% TPR/0.6% FPR) distinguishing copies from high-quality originals. We also perform a confirmation experiment to show that this system also works on legacy magnetic stripe credit/debit cards and demonstrate statistical significance for our results and sample sizes.

The remainder of this paper is organized as follows: Section II provides background information on magnetic stripe cards; Section III more formally defines our hypothesis; Section IV provides an initial confirmation of our hypothesis; Section V develops more realistic tools for determining jitter in real time; Section VI provides our experimental results; Section VII provides recommendations and discussion; Section VIII discusses related work in payment systems security; and Section IX gives concluding remarks.

II. BACKGROUND

Magnetic Stripe Encoding: Magnetic stripe cards contain a band of magnetic material that allows them to store small amounts of data. Data is encoded as a sequence of in-plane magnetization states (bits) imprinted into the magnetic stripe. The data bits are sequentially written into the stripe using a magnetic write head that creates a localized magnetic field to magnetize the bits into the stripe along the length of the card. This data is then recoverable by a read head, which relies on the swiping of the card to induce voltages that can be amplified and converted into plaintext characters.

Card writers encode data on magnetic stripe cards via frequency/double frequency (F2F) encoding. Figure 1 shows F2F in detail. Flux transitions, or changes in polarity (i.e., a 0 becoming a 1, and vice versa), occur on every rising edge of the clock signal. A single transition per clock cycle is used to encode a 0, whereas two transitions in a clock cycle encode a 1. Note that this is true regardless of whether the polarity begins as a low or high signal. Encoded magnetic stripes begin and end with a repeated series of 0s so that the reader can determine the clock for the card. The series is then followed by a start symbol and the data itself.

Fraud Protection: Magnetic stripe cards offer no protection from duplication by themselves. All data contained on a card’s tracks are written as plaintext, and an adversary with access to the magnetic stripe (e.g., with a skimmer) can create a legitimate card. These cloned cards then work in exactly the same way as the originals. A number of techniques have been deployed to try to prevent cloning attacks. The most widely-used method for major-network (e.g., Visa) cards is Card Verification Value (CVV) codes, of which there are two types: CVV1 codes are encoded directly onto the magnetic stripe and transmitted with a card to prove the presence of a card in a card present transaction. This code prevents an adversary from cloning a card from a photo. The more familiar CVV2 codes are printed on the card itself and are used to prove possession of the card in online transactions. In both cases, an adversary capable of physically scanning a card can easily recover both CVV codes.

Fig. 1: F2F Encoding: A polarity transition per clock cycle encodes a 0, whereas two encode a 1.

Fig. 2: A gift card purchased at retail with an unmasked PIN hidden behind a paper sleeve. Such PINs can be easily copied by an adversary, who waits until the card is purchased to steal the card’s funds.

Magnetic stripe gift cards rely on a similar mechanism. Similar to CVV2 codes, gift cards optionally have a PIN embossed on them. To prevent access to the PIN prior to sale, these cards frequently have scratch-off material obscuring the PIN.

Gift Card Fraud: Physical access to the cards allows attackers to trivially bypass these controls. Gift card fraud works as follows:

- The attacker enters a store and gains physical access to the gift cards. Gift cards are often made available as “grab-and-go” items throughout a store.
- In some cases, the magnetic stripe is physically accessible on the rack as shown in Figure 2. If not, the attacker will
open the package to gain access to the magnetic stripe.
• The attacker will read the magnetic stripe and record the data.
• If a PIN is present for the card, the attacker will record the PIN. If it is obscured by scratch-off material, the attacker will scratch-off the material then record the PIN. Figure 2 shows a card where the PIN is not obscured.
• The attacker replaces the scratch-off label and reseals the package, if necessary.
• Finally, the attacker puts the card back on the shelf. The attacker checks the gift card balance to see if it has been activated.

In some cases, the attacker can obtain gift card data over the Internet (e.g., via data breaches or as payment for illicit goods or services). Others have discovered that some gift card numbers are easily guessable [4]. Once the attacker has valid, activated gift card data, the money can be laundered:
• The attacker encodes a new, counterfeit gift card using the obtained data.
• The counterfeit card is used in-store to purchase goods.
• The goods are sold (either online or in-person) for cash.

III. HYPOTHESIS

Legitimate, mass-produced magnetic stripe cards are encoded in facilities with a high degree of automation and process control [10]. Cloned cards are not encoded in such facilities and therefore will exhibit artifacts not seen in machine encoded cards. Furthermore, the quality of the physical cards influences the quality of their magnetic signal.

A. Jitter

After the magnetic stripe card has been physically manufactured, account data is recorded onto the magnetic material using F2F encoding as described in Section II. While we expect this clock to be evenly-distributed across the magnetic stripe, the process of writing bits on an analog medium creates variations in the physical distance (the bit length) between transitions. This clock variation is known as jitter.

Jitter is a normal and expected phenomenon for magnetic stripe cards. Cards are generally permitted to have ±10% variation (via ISO/IEC standards [11], [12]) in the placement of clocking flux transitions with respect to the expected clock rate. While minimal jitter is expected, jitter beyond this tolerance can prevent a card reader from reading a card. Manufacturers of magnetic stripe card encoders frequently include a specification to describe the jitter that is induced by the encoder.

B. Encoding

In manufacturing plants, plastic cards are loaded into a hopper and sent down an assembly line that includes a magnetic encoder. Machines move cards over the write head at a high, consistent speed. By contrast, cloned cards are often encoded by hand, using an inexpensive commodity writer [7], [8], [9]. Instead of moving the card through the track at a consistent rate, these rely on the user to pull the card across the write head by hand. These hand-swiping encoders have a small, rubberized wheel (a rotary encoder) in the card track to measure the instantaneous speed of the card. Without this function, the writer would not be able to write a consistent clock as it would not know the position of the card on the head.

This measurement system in hand-swiping encoders is not without error, however. Limitations in the resolution of the measurement, latency in internal processing, and slippage in the card track influences the encoder’s ability to respond to changes in speed. However, cards encoded in these devices adhere to the ISO standard and can still be read successfully. We expect that cards encoded at constant speeds in controlled manufacturing processes will have less jitter. The presence of this additional jitter is critical to our work, and has not previously been tested as a means of identifying counterfeit cards.

IV. CONFIRMATION OF PHENOMENON

We first seek to confirm that jitter is indeed present on magnetic stripe cards and quantify it. With these measurements, we can confirm if there exists any difference between original cards and copies. In order to do this, we examined the cards using a high-resolution magneto-optical microscope capable of measuring and quantifying the magnetic field produced by the recordings on the card’s magnetic stripe.

A. Magneto-Optical Microscope

The simplest way to read a magnetic stripe card is to swipe it in a reader. However, as we will discuss in Section V, this method is dependent on the reading/swiping speed introducing

---

3 Scratch-off labels are readily available on the Internet.
4 The cards are often given to another person for this step to reduce risk of being caught.
artifacts that may interfere with the accurate measurement of the magnetic card’s bits. Consequently, magneto-optical imaging (MOI) is used to directly observe/image stray magnetic fields written on various card stripes. MOI is a magnetic field measurement technique that enables measurement of magnetic fields over a two-dimensional image plane [13]. MOI is commercially used for qualitative measurement of magnetic field patterns for back-end inspection of magnets, imaging of magnetic inks, data recovery from magnetic media storage, and forensic analysis/recovery. It is also used for high-spatial-resolution, time-resolved measurements in complex scientific experiments [14].

Commercially-available MOI tools exist for imaging the entire magnetic stripe on a single card, but these lack the capability for quantifying the magnetic field and lack the spatial resolution to measure magnetic features with micrometer accuracy. To overcome these limitations, we gained access to a magneto-optical microscope [15] that adapts the MOI imaging technique onto a conventional metallurgical microscope for quantitative, microscopic magnetic metrology. Our system enables measurement of stray fields with 6 µm spatial resolution over a 2.7 mm × 2.1 mm field of view.

The microscope, shown in Figure 3, uses a magneto-optical indicator film (MOIF) that leverages the Faraday effect to measure the perpendicular (z-direction) magnetic field at the location of the film. The Faraday effect is an optical phenomenon wherein a rotation of the plane of polarization in a light wave, caused by the interaction between light and the MOIF, is proportional to the external magnetic field. Proper calibration and validation will yield a quantification mechanism of the z component of the magnetic flux density ($B_z$) in units of Teslas.

The MOIF is a bismuth substituted yttrium iron garnet growth over a gadolinium gallium garnet substrate (transparent and with no contribution to the Faraday rotation) and covered by an aluminum reflective layer and a sapphire protection layer. Two calibrated types of MOIF were used during this work with a 5x magnification microscope: 1) 45 mT magnetic field range with ±0.5 mT field resolution and ±6.2 µm spatial resolution, and 2) ±230 mT magnetic field range with ±1 mT field resolution and 20.1 µm spatial resolution. These two MOIFs are capable of imaging the weak stray fields produced by the magnetic stripe and at the same time provide spatially resolved, quantitative measurements.

Cards are loaded into the microscope stage and affixed in place to ensure flatness and prevent movement during imaging. Multiple images of 2660 µm × 2128 µm are stitched together to obtain an image over the entire card. A section of this image is displayed in Figure 4, where positive fields (perpendicular to the card plane) are represented by red and negative fields in blue. The positive and negative peaks correspond with the magnetic bit transitions, as shown in Figure 1. Because this technique directly images the magnetization (stored bits), it provides a reliable and accurate technique to measure the distance between those transitions (to within 6 µm accuracy). By comparing images of an original card with its copy, it is possible to measure dimensional variations that we hypothesize arise in the magnetic stripe writing mechanism of the card encoder.

B. Encoders

We purchased three manual card encoders from three manufacturers to make copies of cards: the Lanora LNR910®, Osayde MSR605®, and the Misiri MSR750®, shown in Figure 5. We chose these devices because they are inexpensive (~$80-100 USD) and can be purchased from a number of sellers without any type of verification. These encoders are visually similar to those seen in a number of card cloning instructional videos [7], [8], [9]. While purchasing the encoders, we deliberately attempted to choose devices that were dissimilar (e.g., had separate vendor web pages, different model numbers, feature sets, etc.). We also contacted other encoder manufacturers/resellers and discovered that higher-quality commercial and scientific encoders (including linear optical encoding equipment) cannot be purchased without an existing relationship with the vendor or identity verification because of the high risk of fraud. We study adversarial attempts to use available mechanical solutions (and the failure of these automated techniques to produce low-jitter cards) in Section VI-C.

Upon disassembling the devices, we discovered that they are functionally and visually identical with minor modifications to circuit design to accommodate different features (e.g., removing the requirement for an external power supply) or components. Surprisingly, the Lanora device’s circuit board contains the inscription “Misiri 1605A,” identifying Misiri as the manufacturer of the board. Finally, the software that shipped with each device is functionally identical and visually similar, as shown in Figure 6.

To read the speed of the card as it passes over the write head, each of the encoders has a small rotary encoder attached to a wheel in the track. As the card passes through the track, the wheel turns, causing the motor to output a continuous sinusoidal wave. Increased frequency in the waveform corresponds to faster speeds. We measured the resolution of the Misiri motor at 50 µm and confirmed that the Lanora and Osayde devices use the same component by measuring the motors’ output with a logic analyzer.

Once the source card is swiped, the data appears on-screen and creating the copy is as simple as swiping a blank card through the same track. The destination card is verified to ensure the data on the card is correct. The encoder does not make a perfect analog copy of the original analog encoding; it simply reads the binary data on the source card and writes a new analog track onto the destination card. As a result, the jitter present on the destination card is not related to any jitter present on the original. Instead, it is the result of the mechanical limitations of the rotary encoder. Attempts to perfectly copy the analog waveform would be unsuccessful with this encoder.

3 http://www.lndevice.com
4 http://www.osayde.com
7 http://www.msrdevice.com
Fig. 4: A subset of stitched images of a card from our magneto-optical microscope. Track 1 (210 bits per inch) is shown at the top of the image and track 2 (75 bits per inch) at the bottom. In total, 43 images were taken. Positive $B_z$ is represented by red and negative fields in blue. The repeated artifacts in the image are imperfections in the MOIF. Each card took over four hours of effort to image.

Anyway, precisely because its 50µm resolution simply cannot accurately measure the position of the card.

Using the software provided with each, we performed a confirmation experiment to verify that each encoder produces expected amounts of jitter during encoding. After creating clones of 10 cards, we inspected the clones to verify they were encoded with higher jitter variance than the originals. As expected, each encoder produces virtually identical results. Based on the hardware, software, and visual similarities, the Misiri-style encoders represent a substantial portion of readily-available card encoding equipment. Accordingly, we use the Misiri MSR705 to create clones in the remainder of the paper.

C. Experiment

We examined an original card and a copy of that card created with our encoder. We captured 43 images of each card using a ±230 mT MOIF, with each picture slightly overlapping to capture the entirety of the cards’ second tracks.

Figure 4 shows one of the images we captured. Each pixel in the image is exactly 2.08µm. These images show “ground truth” measurements.

Next, we recover the clocking flux transitions. The image processing code averages adjacent pixels in the image to reduce the ability for imperfections in the MOIF to influence the results. The output of this process can be described as a waveform, with the most intense regions in the image as peaks in the waveform.

The remainder of the process is identical to reading a card. We identify the locations of the flux transitions using the peaks of the waveform, measure the distance between them, and decide whether each transition is on the clock or the half-clock. After discarding any half-clock transitions, the remaining data contains only those transitions which represent the clock.

Figure 7 shows images from an original card and its corresponding clone. The images are of the same section of data and show the measured distances between each of the clock transitions. While some of the transitions may look the same, a difference of only 5 pixels is 10.4µm in the image. Ultimately, this figure shows the difference in jitter over a small section of both cards.

Figure 8 shows these same clock transition distances over the entirety of both cards. The copied card has highly-variable distances compared to the original, and so the copied card has a higher amount of jitter. Accordingly, this confirms our hypothesis that variances in jitter can be used to distinguish original cards and copies.

We recognize, however, that the process of using sophisticated scientific equipment to examine a magnetic stripe is too slow and expensive for practical use. In particular, scanning and processing a single card took more than four hours of effort. Having demonstrated that such jitter exists and is different between a candidate pair, the next section explores how magnetic read heads can be used to more economically and practically detect this phenomenon.

V. BUILDING A DETECTION SYSTEM

Having confirmed the phenomenon using the magneto-optical microscope, we now develop a detection system for
Fig. 6: Each of the encoders we examined was bundled with an 80 mm mini-CD containing the software pictured above. Although the software is vendor-branded, each is functionally identical.

Fig. 7: The measured clock distances between the same section of an original card and a copy of that same card. The copy has a higher $V_j$ than the original. Figure 8 shows this measurement over the length of these two cards. Each pixel in these images is exactly 2.08 $\mu$m.

jitter using only commodity hardware. To accomplish this, we seek to understand how jitter manifests in the analog waveforms recorded from widely deployed magnetic read heads. We develop a system for recording and analyzing these waveforms and show how jitter can be measured while reading a card.

A. Audio Recording

As the card’s magnetic flux transitions pass over the read head, a voltage is created in accordance with Faraday’s Law. When the read head is connected to an audio sink (e.g., a microphone input), this voltage creates a waveform. Processing this waveform recovers the underlying F2F encoded binary bits.

We attempted to use several varieties of Square Readers to obtain high-fidelity audio waveforms. Since first-generation Square Readers did not encrypt the analog waveform, they seemed to be a turn-key solution. However, we faced several problems with this hardware. First, these early-model readers are now out of production, so availability is limited to used devices in generally poor condition. Second, the track in this Square Reader is small compared to other readers ($\sim 2.5$ cm), making consistent read speeds difficult. Finally, the friction created by the read head in the reader varies greatly depending on the thickness and material of the card, making it more difficult to swipe some cards in our experiments.

To alleviate these issues, we built our own reader. We purchased an inexpensive magnetic stripe reader, removed the electronic components, and connected the read head and a resistor (to reduce noise) to a 3.5 mm audio jack. The circuit, shown in Figure 9, is functionally identical to smartphone magnetic stripe readers and the longer 17.5 cm track provides more consistent swipes.

The analog signal from the reader must then be converted to a digital signal. Most consumer-grade audio analog-to-digital (ADC) hardware (e.g., those found in laptops and
smartphones) supports audio capture at a maximum sampling rate at 44.1 kHz or 48 kHz. Our initial testing found that these rates are insufficient for accurately detecting jitter. Higher-resolution equipment is able to more accurately measure jitter on a wider variety of swipe speeds. We therefore used a higher-resolution audio capture device (i.e., Sound Blaster Audigy 2 NX) that supported a 96 kHz sampling rate. We connected our reader to the microphone input on this device, and the audio hardware was connected via USB to a laptop running Ubuntu Linux. Audio recording software then captured the microphone input while a card is swiped. Popular magnetic stripe ICs (e.g., Magtek 21006516) in deployed readers output clock distances, however developing our own PC-based solution allowed us to more rapidly prototype software and hardware changes.

The analog waveform is decoded from F2F to binary, and the binary is decoded to plaintext. Our system also verifies the card’s checksums and discards any swipes that could not be read correctly. The system then measures the number of samples between each clocking transition and outputs a vector of distance (in samples).

B. Speed Variance

The use of commodity reading equipment introduces several additional factors which might create error in the results: the average swipe speed, acceleration, magnetic field strength, and curvature of the card material.

Large inconsistencies in the speed of the card as it moves across the read head induces jitter in the audio waveform. This artifact is intuitive; as the flux transitions pass the read head slower or faster, the distance between them in the resulting waveform respectively increases or decreases. Therefore, we wish to capture swipes at a consistent speed.

To demonstrate how swipe speeds might affect the accurate recovery of jitter, we instrumented the rotary encoder in our card encoder and connected it to an oscilloscope. This motor generates a constant sinusoid wave which compresses the periods as the speed of the card increases. We swiped a single card at a reasonable, fast speed and at a deliberate, slow speed. Both times we attempted to maintain a constant speed in the track. We found that the fast swipe, though it continues accelerating in the track, provides a relatively consistent speed when compared to the slow swipe. For our experiments, we create all copied cards with fast swipes through the encoder to attempt to reduce jitter induced by speed inconsistency.

C. Measurement

The value we seek to measure is precisely the variance of the differences in distances between clocking flux transitions. Cards created in high-precision manufacturing lines are expected to have more consistent, physical distances between these transitions; this will result in a lower variance value. Cards with greater fluctuation in these distances will have a greater variance value. We construct this value as follows:

1) Once the absolute locations $p_0, \ldots, p_m$ of the clocking peaks have been detected, we store this in a vector $D$ containing the distance (in samples) between peaks:

$$D = \{p_0 - p_1, \ldots, p_m - p_{m-1}\}$$

2) We then take the distance of this vector, which is the measured jitter of the card. We store this in $J$, a vector of length $m - 1$:

$$J = \{|D_1 - D_0|, \ldots, |D_{m-1} - D_{m-2}|\}$$

3) Finally, we discard the first 10% of the values in $J$ to ignore the effect of initial rapid acceleration of the card in the track. This acceleration can overinfluence the results as the speed sharply increases before becoming steady. We compute the variance $V_J$ of $J$, where $\mu_J$ is the mean of $J$ and $|J|$ is the length of $J$ after discarding values:
\[ V_J = \frac{\sum_{k=0}^{\lfloor J \rfloor - 1} (J_k - \mu_J)^2}{\lfloor J \rfloor} \]

\( V_J \) is a single value that describes how much a card’s measured jitter differs from the mean. Greater values of \( V_J \) indicate more extreme fluctuation of jitter, and so we built our detector to measure this value and record it for each swipe of a card. The recorded values of \( V_J \) are then compared to characterize the jitter between original cards and copies.

D. Confirmation Experiment

To confirm the variance in clock-symbol placement is detectable with our commodity system, we swiped an original credit card and a copy of that card. We then compared the output from both. The original card generates a smoother curve, indicating that it has less jitter than the copy. The measured \( V_J \) for the original card was 0.531 and the copy was 0.709, also showing the expected difference for both cards. Therefore, \( V_J \) can be measured with commodity hardware.

VI. EXPERIMENTAL ANALYSIS

In this section, we measure the jitter on multiple real-world magnetic stripe cards. We demonstrate how our system can distinguish original gift, stored value, credit, and debit cards from counterfeits.

A. Gift and Stored Value Cards

We partnered with Walmart to test the effectiveness of our system. The company provided us with 5 types of open- (e.g., Visa) and closed-loop (same retailer only) cards consisting of 650 individual cards. We cloned each card and swiped/recorded each card at least 10 times, discarding any unreadable swipes. In total, we obtained 12,919 audio waveforms for analysis. Table I shows the breakdown of each type of card.

1) Physical Examination: The cards we obtained were manufactured in a wide range of qualities. The reloadable cards are made from a typical glossy card stock and stripe material, whereas the non-reloadable cards are matte and a much softer grade of plastic. We noticed after swiping the non-reloadable cards that the swiping process had slightly shaved down the plastic. The lower quality of these cards underscores that they are intended for a single use. These cards often have low-coercivity stripes, which are more sensitive to magnetization. As a result, these stripes are often noisy, easily damaged, and produce non-ideal waveforms (i.e., rounder peaks). This makes accurate measurement of \( V_J \) more difficult by reducing the amplitude of peaks and therefore making peak detection more difficult. The ISO/IEC standards [11], [12] state that in high-sensitivity systems such as ours, that the magnetic characteristics of high- and low-coercivity cards cause higher peak amplitude on high-coercivity cards than low. These cards have visible characteristics of being low-coercivity, and so we discuss the results for these cards below by their quality.

2) Results: In this section, we examine the effectiveness of our detector by card ID. This matches the model for how gift and stored value cards are typically used. As opposed to credit and debit cards, where all merchants accept any card in a payment network, many of these cards are intended for use at a specific merchant. Since the merchant is both issuing and accepting the cards, it may wish to set more specific detection thresholds or policies by card.

High Quality. Figure 11 shows the receiver operating characteristic (ROC) curve for each of the high-quality cards. Using only one swipe, the detector is able to distinguish cards with accuracy ranging from 96.9% to 99.9%. Figure 13 shows a kernel density estimate for all swipes for ID 4 using the ROC’s computed optimal threshold. Generally, the figure shows that our detector is able to distinguish these cards with extremely high accuracy, corroborating our results in previous sections.

Low Quality. As expected, the non-reloadable, lower-quality cards we obtained performed worse. Figure 12 shows the
TABLE I: Gift/Stored Value Card Experimental Results

<table>
<thead>
<tr>
<th>Quality</th>
<th>ID</th>
<th>Card Type</th>
<th># Cards</th>
<th># Good Swipes</th>
<th>Finish</th>
<th>Reloadable</th>
<th>Accuracy</th>
<th>TPR</th>
<th>FPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0</td>
<td>Open (Visa)</td>
<td>100</td>
<td>1,970</td>
<td>Matte</td>
<td>No</td>
<td>96.8%</td>
<td>93.4%</td>
<td>6.7%</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Open (Visa)</td>
<td>100</td>
<td>1,990</td>
<td>Matte</td>
<td>No</td>
<td>93.7%</td>
<td>85.8%</td>
<td>14.2%</td>
</tr>
<tr>
<td>High</td>
<td>2</td>
<td>Open (MasterCard)</td>
<td>100</td>
<td>2,000</td>
<td>Glossy</td>
<td>Yes</td>
<td>99.2%</td>
<td>98.6%</td>
<td>1.7%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Open (Visa)</td>
<td>100</td>
<td>1,990</td>
<td>Glossy</td>
<td>Yes</td>
<td>98.5%</td>
<td>96.9%</td>
<td>2.6%</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Closed</td>
<td>250</td>
<td>4,969</td>
<td>Glossy</td>
<td>Yes</td>
<td>99.9%</td>
<td>99.3%</td>
<td>0.6%</td>
</tr>
</tbody>
</table>

3) Statistical Confirmation of Sample Size: To demonstrate that the number of cards we used in these experiments provide a statistically significant result, we performed two-group (original, copy) independent means difference t-tests. This test is a null hypothesis test, where the null hypothesis is that the two means are equal. In our case, the null hypothesis states that there is no statistical difference between the jitter measured on original cards and copies. We performed this test twice: once for the sets of all copies and all originals (i.e., we seek to distinguish any original from any copy) and again for the set of copies and originals for card ID 4 (i.e., we seek to distinguish a copy of card ID 4 from an original).

**All cards.** The calculated Cohen $d$-value the sets of all original and all copies was 1.192 and a $r$-effect size of 0.51, indicating a very large effect size. Our p-value was $< 0.0001$ with power of 1.0, which indicates an extremely high likelihood that there is a statistical difference between original gift cards and copies. Accordingly, we were able to reject the null hypothesis and confirm that our results are statistically significant.

**ID 4.** The calculated Cohen $d$-value the sets of original and copies for ID 4 was 2.385 and a $r$-effect size of 0.77, indicating a huge effect size. Our p-value was $< 0.0001$ with power of 1.0, which indicates an extremely high likelihood that there is a statistical difference between the originals and copies. Accordingly, we were able to reject the null hypothesis and confirm that these results remain statistically significant.

B. Credit and Debit Cards

For completeness, we also performed an analysis on credit and debit cards. We solicited faculty and students at our university to allow us to swipe and measure their credit and debit cards, provided they have one of the four major payment network logos on them: Visa, MasterCard, American Express, and Discover. We contacted our IRB, who noted that because the subject of the experiment was magnetic encodings and not people themselves, *no further IRB review or approval was necessary*. In total, we were able to access 55 credit and debit cards from a variety of issuers.

We cloned each of the cards with our encoder, then swiped both cards using our detector 10 times each. For each swipe, our system recorded the duration of the swipe and the distance (both in number of samples) between each clocking flux transition. We manually inspected each card for physical defects and extreme wear (e.g., cracks). We then securely deleted all sensitive data and physically destroyed the copies (in the presence of the cards’ owners) to protect the security of the payment cards. The machine used to capture this information was not connected to a network during our work.
Rejected Original Card Swipes

Detected Counterfeit Card Swipes

Receiver operating characteristic

Credit/Debit (AUC = 0.978)
Credit/Debit [Majority] (AUC = 1.000)

Fig. 15: The ROC curve for credit and debit cards. The curve is plotted twice, once by comparing all swipes and another with a simple majority voting scheme. With the voting scheme, our TPR increases to 100% and the FPR decreases to 0%.

Figure 15 shows the receiver operating characteristic (ROC) curve for our set of both original and cloned credit and debit cards. Figure 16 shows a kernel density estimate for all swipes from the above experiment along with the computed optimal threshold using the ROC. Generally, the figure shows that our swipes demonstrate a significant difference between jitter in original and copied cards, which corroborates our results in earlier sections.

To reduce outliers caused by swipes with highly variable speeds, we re-analyzed each card using a n-majority voting system. In this system, we fix the optimal threshold determined in the generation of the ROC curve (where TPR−FPR is minimized). We then examine all combinations of n swipes and consider that the detector alerts when > \( \frac{n}{2} \) swipes are above the threshold. With \( n = 3 \), we achieve a TPR of 100% and a FPR of 0%. In dip-style readers, such as those typically found at gas pumps and some ATMs, the act of dipping the card produces two swipes, so in some cases no additional swipes may be needed.

1) Statistical Confirmation of Sample Size: To demonstrate that the number of cards we obtained in this experiment provides a statistically significant result, we performed a two-group (original, copy) independent means difference t-test. This test is a null hypothesis test, where the null hypothesis is that the two means are equal. In our case, the null hypothesis states that there is no statistical difference between the jitter measured on original cards and copies. The calculated Cohen d-value from our individual swipe sets was 2.287 and a r-effect size of 0.75, indicating a very large effect size. Our p-value was < 0.0001 with power of 1.0, which indicates an extremely high likelihood that there is a statistical difference between original credit/debit cards and copies. Accordingly, we were able to reject the null hypothesis and confirm that our results are statistically significant.

C. Mechanical Swiping

We recognize that like all security research, our detector is not a panacea and will spark an arms race. One obvious evasion tactic is to develop or acquire an automatic encoding machine to remove jitter induced by hand-swiping cards. Attaining the required micron-scale precision, however, is much more difficult than a simple, do-it-yourself motorized card track. The equipment must move the card at a precise, constant rate. While we are unable to prove a negative (i.e., that there exists no commercial encoder capable of producing low-jitter cards), below we examine two publicly-available motorized magnetic stripe devices.

ZCS Technology ZCS150 (Price: $250): We purchased this unit (shown in Figure 17) online. It is designed to be used in ATM-style terminals and pulls the card over the magnetic head with a motor as the card enters or exits the device. Although this inexpensive device is not capable of encoding magnetic stripe cards, we initially attempted to modify it to do so. The hardware contains security features to prevent modification, and our attempts to augment the capabilities of this device were unsuccessful. An attacker
trying to repurpose this device would face similar difficulty. However, we do not need to convert this device into an encoder to determine whether or not it could produce more consistent cards if the attacker could evade the security features. Instead, because the ZCS150 does not contain the control system to precisely detect card position and simply interacts with the card assuming consistent movement rate within the ISO specified range, we can measure the acceleration of a card in the ZCS150 and determine if it is substantially less variable than doing so by hand.

We attached an accelerometer to a card, then allowed the device to draw in the card 10 times. For comparison, we attached the same accelerometer to a wristband and hand-swiped a card 10 times. Cards read by the ZCS150 exhibited an average acceleration of $\sim 50 \text{ mm/s}^2$ whereas those done by hand were $\sim 30 \text{ mm/s}^2$. This means that contact with magnetic stripe cards was consistent (i.e., the speed varied little in each case), making hand-swiping nearly identical (and, in fact, slightly more consistent) to adding a simple motor to the encoding process. We believe that the additional variation in the ZCS150 can be attributed to sources of friction within the unit (e.g., a ledge that pushes that card against the read head). Accordingly, there is no obvious reduction in jitter simply by adding a small motor and an adversary would not be able to avoid our detection via this simple strategy.

**HID Fargo DTC5500LMX (Price: $8,000):** Our university uses this device (shown in Figure 18) to automatically print, laminate, and encode student and faculty ID cards. An attacker with such a device could produce realistic-looking gift card clones complete with a branded face. While we were unable to produce arbitrary cards with this device, we examined the $V_f$ values of two cards produced by this device. Here, we consider the cards produced by the Fargo device to be counterfeits by assuming the attacker has control of this device.

Figure 19 shows the distribution of 150 swipes each of two university ID cards as it compares with gift card ID 4. Despite this device being nearly two orders of magnitude more expensive than a Misiri-style encoder, it also produces cards with high jitter compared to the gift cards.

There are a wide range of technologies for encoding cards, however developing feedback control systems to overcome issues in jitter is both expensive and requires major engineering skills. Any attacker with the means and expertise to develop an encoding system comparable to large card manufacturers are likely to be more productive selling those services to make legitimate cards. As we have shown, however, our system increases the difficulty and expense of creating counterfeits, and as such can reduce card fraud.

**VII. Recommendations and Discussion**

Our results show clear changes in the detectability of counterfeit cards based on the quality of manufacturing. Accordingly, to increase the detection rate of gift cards, we make these recommendations:

**A. High-Quality Manufacturing**

We discussed in Section III that cards manufactured in assembly lines move at more consistent speeds than hand-swiped cards. As a result, machined cards are more likely to have consistent distances between clocking flux transitions. Our initial assumption was that all manufactured cards were high-quality, but our analysis shows that this is not strictly true.

We make two recommendations for manufacturers and card issuers about the quality of encoding:

1) Manufacturers should provide expected measures of jitter from their production lines. This problem is compounded by vague marketing language about the quality of cards, further obfuscating the actual quality of the produced cards.

2) Card issuers should insist on manufacturing processes that output low-jitter cards. Our results show that original
cards with higher values of jitter are more difficult to detect when compared with copies.

B. High-Coercivity Stripes

During our analysis in Section VI, we discovered that the coercivity of a card’s magnetic stripe makes distinguishing original and counterfeit cards more difficult. Whereas high jitter is an output solely of the encoding process, inexpensive, low-coercivity stripes are easily damaged, producing noise as seen in Figure 14. Furthermore, these low-quality stripes also produce smoother curves when swiped. These two phenomena interfere with the peak detection process and diminish detection accuracy. We therefore make an additional recommendation regarding the quality of stripe:

1) Card issuers should require high-coercivity cards in order to reduce noise and increase signal quality.

C. Deployability

Our detection system does not require measurement of the original card at the time of manufacture. This allows our system to be deployed and used to distinguish existing original cards and any copies. In some cases, point-of-sale terminals at merchants may only need software updates to support capturing and analyzing the raw analog waveform, as many deployed magnetic stripe ICs output clock data capable of measuring jitter.

We believe that the cost of deploying our system is significantly less than that of deploying EMV for gift cards. First, high-coercivity magnetic stripes cost substantially less than adding a chip to a card ($0.08 vs $2.00 [5]). Moreover, because gift cards (unlike credit/debit cards) are not governed by a single international standard, EMV simply would not be compatible with the vast majority of backend processing systems. Such a migration would force the widespread and expensive replacement of those systems. Second, EMV can not be simply added to cards that have already been sold; however, we believe that a small terminal software update will allow retailers to use the approach proposed in this paper on high-coercivity gift cards immediately. For these reasons, we believe that our techniques are practical and of interest to those being impacted by gift card fraud.

VIII. RELATED WORK

The most accessible and attacked function in the retail environment is the payment system. The development of secure electronic payment systems falls largely into two categories [16], [17]: token-based systems are cash-like; value is transferred directly between parties as part of the transaction. Token-based systems such as NetCash [18], [17], DigiCash/Ecash [19], [20], [17], [21], Millicent [22], Mondex [23], [24], and Chipper/Chipknip [25] do not rely heavily on intermediaries. As a result, revocation and counterfeit currency detection are difficult to perform, restricting consumer and retailer trust and limiting their adoption. Account-based electronic payment systems use accounts to store the value with an intermediate (e.g., banks) to process payments between the consumer and retailer. Technologies such as NetCheque [26], [27], [17], NetBill [17], First Virtual [17], Bitcoin [28], PayPal [29], and Square Cash [30] rely on an online system to verify and authorize transactions. In the United States, these systems have failed to achieve the level of success of the ubiquitous credit/debit card system, where transaction authorizations are backed by a bank or credit account. In developing economies, systems such as M-PESA, Oxigen Wallet, and Airtel Money, that are designed to provide service in developing economies, have been shown to have egregious flaws [31].

Electronic payment systems encompass a wide range of attacks, including transaction snooping [32], [33], fraudulent accounts [34], [35], counterfeit/tampered transactions, and double spending [21], [36]. With the continued use of the inexpensive magnetic stripe card for credit/debit/gift cards, counterfeit payment cards remain a major problem. The magnetic stripe does not offer any security features, and as a result its data is easy to copy [37]. Data stolen (via keyloggers or cameras [38]) or obtained via the Internet can be used to create a counterfeit card to use in a physical store [9], [39].

MagnePrint [40] attempts to resolve this problem by authenticating the physical magnetic material. The system calculates a fingerprint using the noise present between peaks in the analog waveform and matches it to a known value. Unlike our system, MagnePrint requires the card to be measured at the time of manufacture and it requires the merchant to transmit the calculated signature during the authorization process. This requires modification of the authorization network protocol to support additional authentication. Since magnetic stripes offer practically no security, most academic research in this area has focused on replacement technologies for the magnetic stripe such as EMV. Widely known as “Chip-and-PIN,” tamper-resistant EMV cards run code to perform authentication of an original card with the issuer. While the security features of the EMV-chipped portion of the card offer more protection than magnetic stripes, it has proven vulnerable to attacks [41], [42] including stripe-only cloning, relay attacks [43], [44], PIN bypass [45], and replay attacks [46].

Overall, EMV and MagnePrint do not solve the problem of securing previously-issued payment cards. In this paper, we design and build a system which can detect counterfeit cards with high accuracy. Our system does not require prior measurement of the card and detection can be performed directly on a point-of-sale terminal, allowing rapid deployment. Accordingly, we show that counterfeit magnetic stripe cards can be quickly detected with minimal effort on both the merchant and consumer side.

IX. CONCLUSION

In this paper, we develop a mechanism for detecting cloned gift cards. We explore the phenomenon of jitter on magnetic stripe cards, which is an artifact induced at the time of encoding. We argue that such techniques are critical given that gift cards rely on magnetic stripes and no suitable replacement exists. After measuring and verifying this phenomenon using a
high-resolution magneto-optical microscope, we built a practical system for detecting counterfeit cards using commodity hardware. Our detection measure, the variance of the jitter, describes the amount of fluctuation in jitter on a card. We then partnered with a Walmart and ran our detector on a variety of their cards and found that using a simple majority voting scheme, our system is able to detect counterfeit cards with 99.9% accuracy. Accordingly, we show that counterfeiting magnetic stripe gift cards with currently-available equipment can be effectively detected.

ACKNOWLEDGMENTS

The authors wish to thank Walmart for providing us with the uncirculated cards used in this work. The authors are also grateful to GMW Associates for granting us access to a Matesy minimo magnetic imaging device. This work was supported in part by the National Science Foundation under grant number CNS-1526718. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

