An Assessment of SMS Fraud in Pakistan

Fahad Pervaiz[1], Rai Shah Nawaz[2], Muhammad Umer Ramzan[2], Maryem Zafar Usmani[2], Shrirang Mare[1], Kurtis Heimerl[1], Faisal Kamiran[2], Richard Anderson[1], Lubna Razaq[2]

[1]University of Washington, USA
[2]Information Technology University, Pakistan

ABSTRACT

SMS fraud has become a growing concern for those working toward financial inclusion, however, it is often unclear how widespread such threats are in practice. This multi-method study investigates SMS fraud in Pakistan through identification and categorization of fraudulent messages as well as the impact on those who receive such messages. We collect fraudulent SMS messages by various means, including byway of a custom-built Android smartphone application. To complement this, we interview people exposed to SMS fraud and representatives of mobile network operators. Based on our analysis, lottery type fraud schemes dominate SMS fraud in Pakistan, and these schemes have the greatest impact on vulnerable low-income, rural populations. We offer a simple heuristic for fraud detection that has a high accuracy rate and is adaptable to evolving fraud schemes, and conclude with a recommendation for a fraud mitigation strategy to target fraudster call back numbers.

KEYWORDS

ICTD; mobile money; fraud; SMS; finance; mBanking; human factors

ACM Reference Format:
Fahad Pervaiz[1], Rai Shah Nawaz[2], Muhammad Umer Ramzan[2], Maryem Zafar Usmani[2], Shrirang Mare[1], Kurtis Heimerl[1], Faisal Kamiran[2], Richard Anderson[1], Lubna Razaq[2]. 2019. An Assessment of SMS Fraud in Pakistan. In *ACM SIGCAS Conference on Computing and Sustainable Societies (COMPASS) (COMPASS '19), July 3–5, 2019, Accra, Ghana.* ACM, New York, NY, USA, 11 pages. **https://doi.org/10.1145/3314344.3332500**

1 INTRODUCTION

Financial inclusion, the access to the formal economy through banking, loans, and credit, is recognized as an important development objective [1, 11, 33]. As the world economy becomes more digital through services such as electronic banking, credit cards, mobile money, digital payments, and other mechanisms, development organizations promote these digital financial services (DFS) as a primary means to achieve financial inclusion. With this ongoing promotion and adoption, research has shifted to assess the myriad barriers to DFS uptake. These include network and service outages, insufficient agent liquidity, complex user interfaces, poor customer recourse, inadequate data privacy, non-transparent fees, and customer-targeted fraud [16, 24]. Fraud, in particular, is problematic because low-income and marginalized populations are more impacted by financial loss [15].

Anecdotes regarding SMS-based fraud are prevalent in Pakistan and have a wide circulation among DFS researcher networks [9, 28, 35]. For those who have spent time in Pakistan, personal observations of incoming SMS messages that attempt to initiate fraud are common. Even though mobile money is not directly implemented through SMS, many components of mobile money systems such as transaction receipts rely on SMS [29]. Both SMS and mobile money are associated with basic mobile phones, and because of this those who promote DFS perceive SMS fraud as widespread with mobile money operators and customers as targets [26].

To investigate the scope and scale of the problem of SMS fraud in Pakistan, we developed and deployed a SMS data collection application (app) at a university. This app gathered users' SMS and sends them to be anonymized and analyzed at our in-country research server. Given the sampling bias inherent in such a method with majority university students, we extended this data set with an advertised SMS-forwarding service, widening our participant base outside of the university setting. We categorized the SMS into ten types and three classes.To complement our SMS corpus and to better understand the experience of fraud within vulnerable populations, we conducted interviews with low-income rural and urban Pakistanis. These combined datasets provided a wide base to better understand the practice and impact of SMS fraud in Pakistan. In particular, we wanted to address the following questions:

- What are the types of SMS fraud? What sorts of attacks are *common*?
- Are *DFS* a common component of SMS-based fraud?
- Can we *easily* and *accurately* detect and classify SMS fraud in Pakistan?
- What are the properties of SMS fraud in Pakistan?
- How is fraud experienced among *rural and low-income* Pakistanis?

Our results indicate a large ecosystem of fraudulent SMS in Pakistan. Contrary to initial assumptions, financial services are *not* a major component of this ecosystem, despite the recent push for DFS adoption in Pakistan. We found ten different fraud schemes of three different types. We demonstrate that a simple classifier can detect a large majority of fraud messages with a small false positive rate. We conclude that the SMS fraud ecosystem in Pakistan has a number of unique features, including language-based targeting, and find that rural users regardless of income level who have just come online may be primed as victims for future SMS-based fraud.

COMPASS '19, July 3–5, 2019, Accra, Ghana

ACM ISBN 978-1-4503-6714-1/19/07…$15.00
https://doi.org/10.1145/3314344.3332500

2 RELATED WORK

2.1 Spam

The detection, avoidance, and blocking of spam is one of the most studied problems in computer science. Spam can take many forms, including email [7, 13], web sites [14], and even voice [23]. Likewise, attackers target SMS as a communication medium. Several studies have explored the possibility of adapting email spam filtering techniques for SMS spam detection [10, 12]. As spam and SMS traffic evolve, sophisticated algorithms will be necessary for detection [19, 27, 31]. These systems are highly language-dependent, although there have been efforts for SMS spam filtering using non-context based features [37]. These filtering systems are often deployed within cellular networks, which block at the SMSC [18].

Our research departs from previous work in that we do not focus on solving the problem of SMS spam in Pakistan. Indeed, it is likely that an off-the-shelf filtering algorithm is already present (but disabled) in the telecom messaging center. Our discussions with telecom representatives indicated that the government considered enabling such filtering as a form of "censorship." As such, the problem is largely structural rather than technological. This provides a unique opportunity to explore an active fraud and spam SMS ecosystem.

2.2 Fraud

Digital fraud, much like spam, is a heavily studied area within computer science. Scammers use mediums like email [17], voice [25], and SMS to reach a wide audience through socially engineered attack messages [2]. To successfully scam people, scammers deploy various strategies to make their messages appear official and trustworthy [25, 34]. Several works have evaluated how versions of the Nigerian email scam or similar fake lottery scams have spread over email and the consequences when a victim contacts a scammer [5, 17]. This paper extends prior work by exploring attacks present in a previously unstudied region (Pakistan) and the unique properties of attacks found there.

2.3 Digital Financial Services

Digital financial services (DFS) are a set of services that provide access to formal banking solutions through mobile technologies [20]. These services, such as mobile wallets or credit solutions, rely on existing cellphone communication channels like USSD, SMS, or data. As a nascent technology in many markets, the technical aspects of DFS systems do not appear as the subject of many studies. Reaves et al. [32] examined and summarized vulnerabilities present in developing world DFS apps due to insecure connections or data leakage. Similarly, Castle et al. [6] expanded the threat model and pointed out SMS as a vulnerable communication channel in DFS due to the lack of number verification that can lead to SMS spoofing. Phipps et al. [30] explored the potential for ThinSIM-based attacks on mobile money systems. Our work provides a supporting view into ongoing attacks on DFS systems, finding that these types of attacks are largely absent in the current ecosystem.

2.4 Security in the Developing World

Numerous researchers have explored security in developing contexts [3, 36]. Common themes include social mismatches between technologies and cultures [21], providing security in light of infrastructure failure [8], and the application of developed-world best practices to other regions [21]. Our research advances this area of inquiry in that it analyzes users' experiences with attempted SMS-based attacks in a developing context (Pakistan).

3 SMS FRAUD BACKGROUND

We first define fraud in SMS and then provide examples of SMS fraud that motivated this study.

3.1 Definition

For this paper, we define fraud as an act where one person is attempting to deceive another person to get money or other items of value.

Items of value may include credentials such as account numbers, PINs, or personal identifying information that can be used to acquire other things of value. SMS *fraud* thus refers to fraud that is initiated or executed through SMS. We distinguish SMS *fraud* from other forms of unwanted SMS messages such as unsolicited advertisements or *spam*. We discuss this distinction in more detail in the Data Analysis section.

3.2 Examples of SMS Fraud

To familiarize the reader with the context for this work, we present and discuss several examples of SMS fraud gathered from online sources prior to embarking on our work in Pakistan.

One of the most common types of SMS fraud is what we refer to as the *lottery fraud*: A fraudster sends a message informing the recipient that he or she has won some money, and that the person must contact a certain number to receive it. Once the victim calls back, the fraudster convinces the victim that they must pay a fee to obtain the prize money. A payment is made by some mechanism, but the prize money is never delivered. This is also a staple among email fraudsters who announce winnings of implausible lotteries such as the BillGates lottery. Typical examples include:

> **Ghana (Twitter):** *Valued customer, ur number is one of our lucky winner of Gh12,000 on Airtel Wo Mner3 Nie promo! Call Mr Owusu to cash it out on 026263xxxx*[1]
> **Kenya (Twitter):** *CONGRATULATIONS from SAFARICOM MAISHA NI MPESA TU! Promotion, You have Won. Ksh 100,000.00 your secret code 555555 Call (078349xxxx) for more information. NB: DO NOT PAY Anything.*

A second common SMS fraud is *receipt fraud.* In this case, a fraudster sends a fake receipt for funds added to a targeted subscriber's account. He then calls the subscriber to ask for that money back, stating that it was accidentally sent to the subscriber's mobile wallet. The receipt format mimics the form of a legitimate mobile money receipt. Depending on the sophistication of the fraudster, there may be an attempt to spoof the sending address, although the majority of the examples we have seen appear to originate from a mobile number. Receipt fraud is a direct attack on individuals, as opposed to the lottery fraud which relies on sending a large number

[1] We obscure the final digits of phone numbers for privacy reasons, though the numbers themselves are likely no longer valid.

of messages to collect responses. Because receipt fraud is dependent on mobile money systems, it can be classified as DFS fraud, whereas lottery fraud is typically not. Examples of receipt fraud are as follows:

Kenya (Twitter): *MPESA LHR9VQ7DKE Confirmed.You have received Ksh9,730.00 from BEN ONYANGO 070671xxxx on 15/8/17 NEW M-PESA balance is Ksh*(Pending)*Pay Bills via M-PESA.*

Uganda (Twitter): *MTNMobMoney. Y'ello. You have received UGX973,000.00 FRrom: KTM LTD. Token ID: 79864532991. Remember to get secret code from sender to access your funds.*

4 DATA COLLECTION

Gathering a broad and representative sample of SMS fraud in Pakistan required a range of different data collection methodologies. We began with the development and deployment of an app-based solution to provide "ground truth" of SMS spam and fraud rates by recording and transmitting all SMS messages received by 246 study participants. Following this, to ensure that we discovered all types of fraudulent messages, we set up and advertised a fraud SMS forwarding phone number for users outside of the study, with a total count of 518 messages received. In order to broaden our study to the voices of non-smartphone users, we conducted eight interviews outside of Lahore with low-income, basic phones users.

IRB approval was obtained for all aspect of this study, with rigorous safeguards in place to ensure anonymity, security, and privacy for the participants. Additionally, when the smartphone app users opened the app for the first time, they had to go through several screens that explained the research objectives and the data to be uploaded. On the last screen, they were asked to consent to the data upload and the privacy agreement by checking a box on that screen.

4.1 Smartphone App-based Collection

We began our data collection by developing an app for Android phones, the most commmon smartphone platform in Pakistan. This platform records some received SMS, as decided by the user, and forwards them to research servers for collection and analysis. Users set the app as their default SMS handler, allowing it to investigate all messages that pass through the phone.

Our application, known as "Safe SMS," was built as an extension of the existing "QKSMS" open-source SMS handling app [4]. Our version was enhanced to provide the following features to:

(1) Upload messages and labels to a secure research server;
(2) Allow users to label a message as either *Fraud*, *Spam*, or *Ok*;
(3) Allow users to mark a conversation as private so it is never uploaded; and
(4) Offer an in-app tutorial explaining the research project, organizations, and the additional features.

The app kept a database (DB) of supplemental variables in parallel to the Android system-wide SMS DB. The local app DB tracked a user's marked label for a given thread, the upload status of messages, the SMS header of messages received after a user agreed to the app's consent, the location of the device when an SMS was received, and the list of threads that a user wanted to maintain as private.

Furthermore, all data was sent over a secured connection to ensure that it was encrypted in transit to our verified server. The data upload

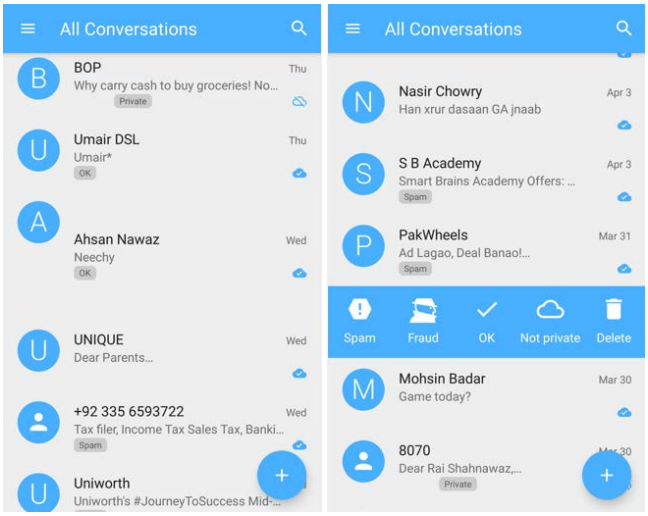

Figure 1: Screenshots from the Safe SMS app that illustrate how to label a conversation

happened in the background after a user visited the upload screen and manually pressed the upload button. The user received a notification once the upload was complete.

Labels: We introduced labels as a way to ascertain the user's perspective about the nature of the messages. Users could label a conversation with one of three labels: 1) *Fraud*, indicating that the message is an intentional effort to defraud the receiver, 2) *Spam*, indicating that the message was a broadly sent advertisement, and 3) *OK*, indicating a normal conversational SMS. The goal of this labelling was not to provide a "ground truth" of the content of the messages (instead we tagged them manually) but to explore if our definition matched the perception of our participants. The app also had a screen to sort and to view all of the unlabelled messages or messages with specific labels, to encourage consistent labelling and reduce satisficing [22].

Private Messages: Given the volume of personal information present in SMS communications, we sought to uphold privacy. We did this through an upfront option to mark a conversation as "private." Marking a conversation as "private" excluded any message in the conversation with that person (with the phone number as the identifier) from being uploaded to the research server. Similarly, if any message was uploaded in a previous sync and later marked as "private," the app would send this change to the server resulting in the data being deleted. Data analysis was done at the conclusion of the study, so no intermediate data (uploaded but then marked "private") was used in the analysis.

Tutorial: Given the extensive additions to the default "QKSMS" app, we built a tutorial in our app, "Safe SMS," that introduced the user to our team and research agenda, as well as how to mark a conversation as private, how to label conversations, and how to upload data. The tutorial opened when the app launched for the first time on a device. A user had the option to review the tutorial at anytime by pressing the question button, which is available in the menu on all screens.

The app was uploaded to the Play store, which is the official app store for Android smartphones. This made it easy for us to distribute the app with a single URL in which the app install page on the Play store opens. We setup a research server in Pakistan with firewall and security certificates in place to ensure all communication between the app and the server was encrypted and secure.

4.1.1 App Testing. We conducted a series of iterative design sprints to shake out any bugs or confusion about our design. This was done in two rounds: first internally with seven members of the research team and then externally with eight recruited outside users from our home university. The external testers were asked to install the app and upload some SMS conversations. After this we interviewed these users for feedback. The users were tested with improvements implemented from the previous phase's feedback.

Our findings pointed to the the privacy implications of the app. Users were uncomfortable with the app's repeated requests for permissions during the install. The app recorded the location of the users to be able to categorize the user as rural or urban. We removed this feature to require fewer permissions. Second, users were uneasy about the messages being uploaded in the background. To resolve this, we introduced a screen that lists all the unsynced messages, showing users that only selected messages were being uploaded. The user could search and filter the conversation based on the marked labels to make it easy to select all intended conversations. This was also the step where a user could further exclude sensitive or personal messages from being uploaded, even if they were not initially marked as private. This could introduce bias in our data as a user might select only the fraud related messages for upload, which relies on their perception of fraud. Several users reported confusion about label icons and how to track which conversations had been labeled. We removed all the label icons and only made them visible in the menu that had full text. Moreover, we changed the view that lists all the conversations. We added label names, sync status, and a privacy tag on each conversation so the user could scroll through the list quickly to look up the message status rather than having to select and view each individual conversation.

4.1.2 App Deployment. Once the system was generally accepted by our testers, we expanded to a wide-scale deployment outside of our research group. We advertised the app at a local university in Pakistan with the initial target being students, whom we considered among the most text-savvy. Also, we encouraged faculty and researchers at the university to advertise the research project in their research labs and class rooms as a means to drive adoption. Some faculty requested that we provide a five minute talk in their class about the purpose of the app as well as how to use it.

We became concerned that only asking university students to test the app could skew our results in unpredictable ways. To resolve this, we used personal connections in different rural and urban localities to generate a more diverse set of participants. In urban areas, we approached 125 individuals from different offices and incubation centers. We also forwarded the app to 40 people through our personal contacts, akin to a convenience sampling. In rural areas, we reached out to 20-25 individuals from various villages through similar personal connections. In the end we had 246 users who downloaded and installed our app.

4.2 SMS Forwarding Service

Although the app provided a robust base to explore fraudulent and spam SMS, our participant selection methodology introduced a sampling bias and caused us to potentially miss other classes of messages. To resolve this, we set up a Pakistan-based phone number where anyone, including those who did not participate though the installation of the app, could forward fraudulent messages through SMS or WhatsApp. This allowed us to gather fraudulent SMS from a broader range of people who had a variety of phone types (basic, feature, or smart), who may have had limited technical skills (forwarding an SMS is more broadly understood as compared to downloading and installing an app), and who resided outside the initial geographic range of our study. While the forwarding service did not completely eliminate sampling biases,this intervention did provide insight from a much broader sample.

The forwarding service was implemented on a smartphone with an active SIM and an enabled WhatsApp service. Our advertised number could be used to send SMS or WhatsApp messages to this device from anywhere in Pakistan. The phone received all the forwarded SMS messages as well as all the message screenshots or texts sent over WhatsApp. The SMS messages were uploaded to our "Safe SMS" app server under a specific user that represented this forwarding service. All the WhatsApp messages were human transcribed from images into a database on the same server.

In Punjab province, home to over half of Pakistan's population, the service was widely advertised through social and print media both in Urdu and English languages. The advertisement asked citizens to forward the fraudulent messages for a public good, and they were advised to append "Forwarded from: ..." at the beginning of the message or take a screenshot of the SMS to send through WhatsApp or SMS. Over the course of seven weeks, we received 746 fraud messages from 351 users.

One shortcoming of the forwarding technique was that it lost additional data around the forwarded SMS, such as the sender's number, the receipt time, and any messages previously received from that number. Although we requested that people send the number from which they received the fraudulent SMS number, some only sent the fraudulent SMS content.

4.3 Interviews

In order to investigate further how fraudsters successfully defraud people, we conducted informal interviews with eight participants. We focused on conducting interviews with low-income people with the assumption that fraud schemes may target them. Three of these participants were from a major city while five were from a rural village, both in Punjab. The interviewees were randomly approached in markets and public areas of the village, and may not have been aware of the SMS fraud research prior to us approaching them. All of the village participants were from the home village of one of the researchers. Due to privacy concerns, we took field notes but did not record conversations.

5 DATA ANALYSIS

Our "Safe SMS" smartphone app had 246 installs with 106 users uploading some data. Out of those 106 users, 100 self labeled part of their data with tags *Fraud*, *Spam* and *Ok*. Figure 2 shows the number

of conversations available on each user's phone and the number of messages uploaded. A conversation is a thread of messages sent and received between the unique pair of a user and an external phone number. Collectively, users uploaded 4057 conversations that consisted of 52,169 total messages.

5.1 Authoritative Researcher Labels

While we asked participants to label messages, we never intended to use their selections to evaluate the ecosystem as users could interpret the labels in different ways. Instead, we sought to create an authoritative, "ground-truth" set of *researcher labels*. To do this, the research team labeled all the conversations. The data was labeled by two reviewers, irrespective of whether a user labeled it or not. Those who labeled the messages are native speakers of the local languages (Urdu and Punjabi) in which the messages were composed. After each reviewer went through the labeling exercise individually, they merged their labels, discussing any mismatches and inconsistencies to reach consensus.

Through the labeling process we discovered that the original categories of *fraud*, *spam*, and *OK* were insufficient. As such, we expanded the categories to include the following:

- **Ok:** Conversation between two real people.
- **Fraud:** SMS that deceive a user to defraud them of their money.
- **Spam:** A generic advertisement for products and unsolicited public messages.
- **Status:** A system notification for a specific user about their account status, action taken on a service, or delivery update.
- **Spam with Status:** The same number sends spam with advertisements as well as sends account status updates. This is common for services where a user has an account, like a telco, a bank, or a ride sharing service.

A user might consider a *Status* message as *Spam* or as *Ok*, so we created the more specific label. Moreover, some spam could be considered fraud. A misleading advertisement about credit or insurance that over promises, and, hence, can be regarded as fraud, would still be labeled as spam. For example, this message is misleading because permanent residency processes take longer than one month yet we labeled it as spam:

> *Australia Permanent Residency Approval in 1 Month BA/MA Are Eligible 2018 Relax Policy 100% Success Embasy Fee Also Return If Rejection 03211818190 Natasha.*

In general, users have a reasonable understanding of fraud and spam messages. Table 1 lists the conversations that users labeled with the corresponding messages' count in brackets. Users labeled *Fraud*, *Ok*, and *Spam* with 75%, 58%, and 92% accuracy, respectively. We adjusted the numbers to remove the confusing messages with labels *Status* and *Spam with Status* that could be considered as *Spam* or *OK* by the user. We found that accuracy of user labeling increased to 75%, 71%, and 97% for *Fraud*, *Ok*, and *Spam*, respectively. This illustrates that we can rely on user labeling to a certain degree, but, more importantly, indicates that our definition of labels is comprehensible given that overall 93% of user-labeled conversations agreed with our labels. These users were invested in the goals of our research considering that they were motivated to contribute, comfortable with determining a label, and spent time marking data.

5.2 Forwarding Service

Our SMS forwarding service significantly expanded our fraud message corpus. We received 152 messages from the app, 228 from people forwarding the fraud message to a central number and 518 from text or screenshot of the text via WhatsApp. While these numbers were impressive, 39% of forwarded messages did not include the address of the sender. Hence, we did not obtain the sender's information for 289 out of 898 fraud messages.

5.3 Data Collection Challenges

We had challenges in data collection due to lack of trust as well as possibly from our strategy to incentivize using social and public good rather than offering monetary benefits. This may have hindered mass adoption of the app resulting in 246 users out of which only 106 uploaded any data. Moreover, the forwarding service received only limited data since people often delete fraud and spam messages quickly, especially those received on feature or basic phones due to memory limitations.

6 RESULTS

6.1 A Rough Taxonomy of SMS Fraud in Pakistan

In consolidating the fraud messages gathered from various methods, we found ten repeated fraud schemes as shown in Table 2. We classified these schemes into three categories. The most common type is the lottery type where the fraudster announces that the user has won or received something from a lucky draw or a scholarship or public program. The second type, we call "Damsel in Distress," in which a fraudster poses as a vulnerable young woman who is in need of help. The fraudster appeals to the user to send a few mobile credits that she will return later. Finally, we saw a few instances of a fraudster trying to steal credentials by stating that the victim's bank services, like SMS notification or ATM card, had been disabled. The fraudster invites the victim to call to reactivate the services, but provides a number through which the fraudster can intercept the call and obtain the victim's personal information.

Each scheme has a generic story that can be broadcast to anyone. Nonetheless, we observed fraudsters who change the formatting with new lines or adjust the wording or spacing slightly. Fraudsters employ several techniques to appear legitimate. Sometimes they will add a user's number to the message to make it specific to that user. We also found a message that claimed to be from Zong telecom service and provided the actual website address in the message, however, the call back number did not route to Zong's customer service number. A few UK award messages were sent from a number with a +44 country code (the UK's code), while the majority of messages with that scheme were sent from numbers that originated from Pakistan. Some fraudsters are willing and able to make their messages appear more authentic.

Most fraudsters request that their victims place a call. The call, in turn, initiates the fraud and makes it challenging to trace. Moreover, up to five fraud messages in our dataset, complete with unique call back numbers within the message, originated from a single number. This suggests that a huge cache of unique SIMs are being used in order to conduct fraud. We did not anticipate this finding because the Pakistani government regulates SIM registration and requires

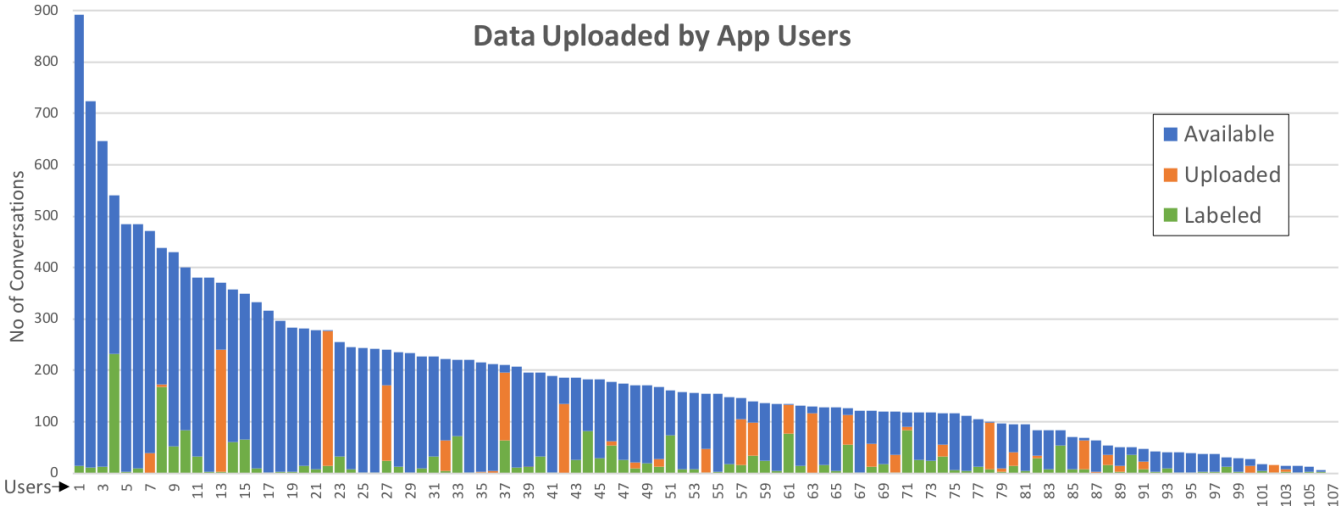

Figure 2: Number of conversations available on each user's phone, user-uploaded conversations, and user-labeled conversations

Table 1: Summary of User Labeled Data.

User Label	Total Uploaded Conversations (Messages)	Actual Labels for Conversations (Messages)					
		Fraud	Ok	Spam	Spam with Status	Status	Unknown
Unlabeled	1869 (33233)	20 (22)	816 (28198)	788 (2403)	21 (819)	188 (1481)	36 (310)
Fraud	148 (215)	**111 (113)**	3 (8)	33 (93)	0 (0)	1 (1)	0 (0)
Ok	264 (9661)	1 (1)	**152 (8525)**	62 (615)	9 (114)	40 (406)	0 (0)
Spam	1776 (9060)	12 (12)	29 (146)	**1641 (6228)**	22 (1471)	62 (1182)	10 (21)
Total	4057 (52169)	144 (148)	1000 (36877)	2524 (9339)	52 (2404)	291 (3070)	46 (331)

Table 2: Summary of Fraud Types.

Fraud Scheme	Fraud Type	Count	R. Urdu	Urdu	Eng.
ARY Jeeto	Lottery	557	553	4	0
BISP	Lottery	215	127	88	0
Waseela-e-haq	Lottery	46	46	0	0
UK Award	Lottery	40	0	0	40
Easyload	Damsel	31	27	4	0
Scholarship	Lottery	4	4	0	0
Bank Service	Steal Creds	2	2	0	0
ATM Card	Steal Creds	1	0	0	1
Pak Army	Lottery	1	1	0	0
Zong	Lottery	1	1	0	0

biometric verification linked to a person's national ID card for a person to receive a new SIM. In addition to this, an individual can only have up to eight SIMs (five for voice and three for data). Therefore, fraudsters with large caches of SIMS might indicate a black market for registered SIMs linked to the identities of people who are unaware that fraudsters had obtained SIMs in their names without their permission.

6.2 DFS-Specific Fraud

Surprisingly, fraud over SMS in Pakistan is not related to DFS. We found more general call based fraud since the majority of fraud schemes ask victims to call instead of directly send money. The easyload fraud, where the fraudster asks for money directly, requests that victims send airtime credits instead of money over a mobile wallet. We observed only one instance in that scheme where the message asked for "Jazz Cash" instead of easyload. Also, from our interviews, discussed in next section, we learned that a call to the lottery based schemes results in the victim being asked for a small fee in the form of airtime credits through prepaid scratch card numbers. A few instance revealed that sophisticated fraudsters had tried to obtain bank account details or credentials from victims, but those are very rare in our dataset.

6.3 Detecting Fraud

Specific features of messages allow us to distinguish fraud from regular and spam messages. Figure 3 shows the percentage of different labels that were positive for each feature. A fraud message is never sent from a short code and typically has a call back number in

Table 3: Examples of fraudulent messages that were collected. English translations are given for messages sent in Roman Urdu.

Fraud Scheme	Example	English Translation
ARY Jeeto	ARY JEETO PAKISTAN K show me apne is 03047227028 se 8038per SMS kiya tha Qrandzi me ap ki 1adad bike or 5lakh cash nikla he ap is no 0303769xxxx pr call karen	In ARY JEETO PAKISTAN show, you sent SMS from 03047227028 to 8038. From the lucky draw, you got 1 bike or 5 hundred thousand cash. You should call this no 0303769xxxx
BISP	BENAZIR INCOME support ki taraf se apko Rs.25200 mubarek ho.apka ye number0323759xxxx BISP mein Register tha.ap is number per 0306709xxxx rabita karen..	From BENAZIR INCOME support, congratulations on Rs.25200.Your this number0323759xxxx was register in BISP.You should contact this number 0306709xxxx..
UK Award	CONGRATS! YOUR MOBILE NUMER HAS WON 500,00 POUNDS IN THE 2018 PEPSI PROMO. TO CLAIM YOUR PRIZE. SEND UR NAME, AGE AND MOBILE NUMBER TO: ppeawd@hotmail.com	
Easyload	mery is number 0304946xxxx py 50 ka Mobilink load karwa do main bad me wapis kar don gi call nne.pleasa.saba	Send Mobilink load of 50 to my this number 0304946xxxx. I will return it later call nne.pleasa.saba
Bank service	Dear Customer! Ap Ki HBL Ki SMS ALEART Service Khtm Ho Rahi Hai Dubara Free SMS ALEART Active Krne K liye Is Pr Visit Krain . Visit www.smshbl.com	Dear Customer! Your HBL SMS ALEART Service is ending. To make Free SMS ALEART Active again, visit this . Visit www.smshbl.com
ATM Card	Dear Coustmores,your ATM card has been blocked Because you did not have an update yet. If you want your ATM card to work properly, then contact this	

Table 4: Summary of Heuristic Results

True Label	Total Messages	ID'd as Fraud	Percentage
Ok	36877	35	00.10%
Spam	9367	10	00.11%
Fraud	898	891	99.22%
Status	3070	16	00.52%
Spam/Status	2404	0	00.00%

the message. Relative to other messages, fraud messages are likely to have congratulatory words, phrasing related to receiving something, and terms about a lucky draw. Congratulatory words included various spellings of "congratulations" or "mubarak" (Urdu for congratulations) both in English and Urdu. Words related to receiving included "won," "awarded," "nikla" (Urdu for got), "mila hai" (Urdu for received), and "aye hain" (Urdu for came). Furthermore, we identified certain keywords related to the type of fraud, like "lucky draw," "qrandazi" (Urdu for lucky draw), and "load" as these types of fraud aim to get money through easyload or announce that the victim has won a lottery. Even with some features better correlated with fraud, there is no one feature that can distinctively indicate fraud.

We explore several heuristics based on the insight we gained from the data. The one that proves most effective for detecting fraud is as follows:

$$fx\ NxORMxANDCxORRxORLx$$

where N, M, C, R and L are functions that return as true if the phone number, currency, congratulatory words, receiving words, or lucky draw-related words are present. The results of the heuristic are shown in Table 4. It detects 99.22% of fraud messages as fraud and less than 1% of other labeled messages as fraud. This is promising in that a simple heuristic-based, unsupervised algorithm can detect fraud without knowing specific details about different fraud schemes. This makes it generic enough that if a new lottery scheme appears, the algorithm would likely be able to detect it. The heuristic has its limitations as fraud vocabulary evolves, and, eventually, it could be difficult to differentiate fraud from spam. This can be avoided by building heuristics specifically for spam detection, an area that is well researched. Moreover, we believe that the heuristic for fraud has to be able to evolve.

6.4 Fraud Specific to Pakistan

In Pakistan the majority of fraud messages are composed in Roman Urdu script, while only one scheme is consistently sent in both Arabic Urdu script and Roman Urdu script. This exception is the Benazir Income Support Program (BISP) scheme. BISP is an unconditional cash transfer program for low-income women, which is arguably why it is sent using both scripts. Two other schemes send a few messages in Arabic Urdu script, yet they more commonly use Roman Urdu script. The "UK Award" scheme and the "ATM Card" fraud are sent in English only, as they are intended for a different audience with ATM cards or those who have email savvy. This illustrates that fraudsters adjust the language and scripts of messages depending on their intended audience.

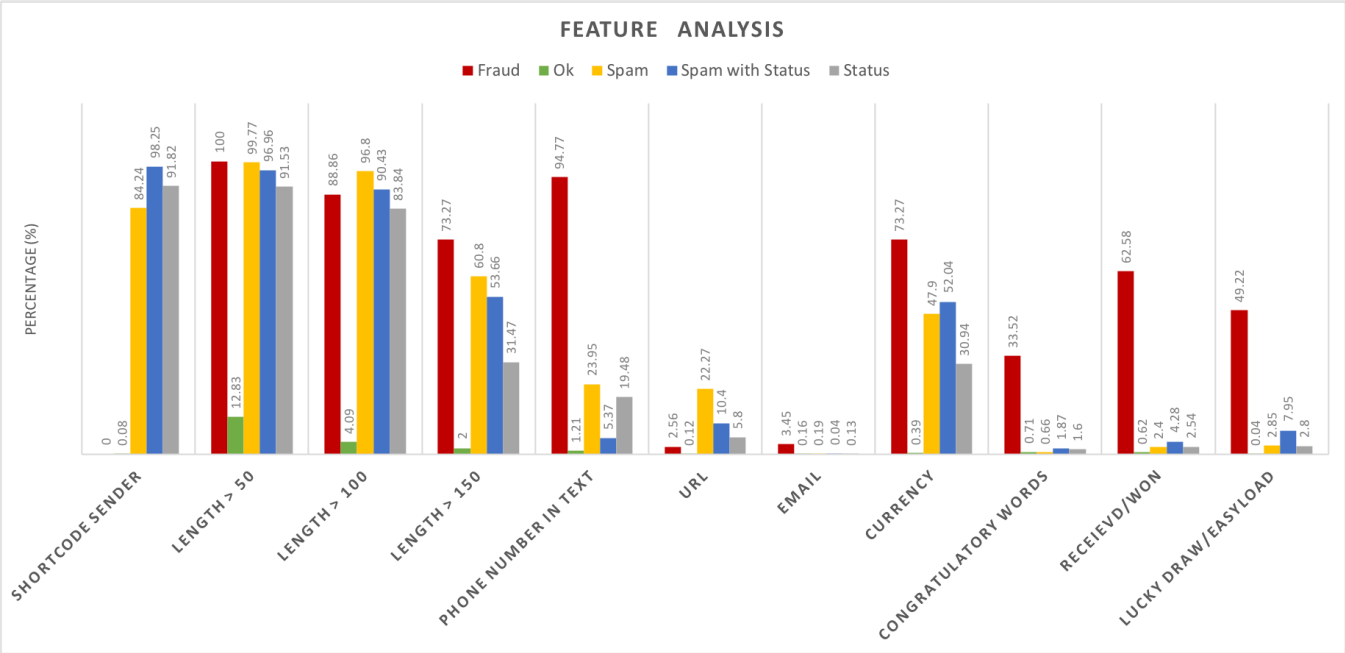

Figure 3: Presence of different features with important ones highlighted

Furthermore, our intuition is that some of these fraud messages are hand-typed on mobile phones because of the tremendous amount of variation in the transliterations of words. This generally arises due to multiple authors transliterating into roman script. For example, "Mariam" and "Mariyam" represent the same name in Urdu. "Chak," "check," "chek," and "cheak" are various spellings that refer to a bank check or cheque. We had difficulty running our queries for standard Urdu transliterations like "Mubarak" (congratulations) because of the huge variation in transliterations found in the data, which included, "mubarik," "mobarak," "mubarek," "mubarak," "mubark," "mobarik," and "mubraka."

7 QUALITATIVE ANALYSIS

Here we describe the findings based on our interviews in low-income areas of Punjab. These results, by in large, reinforce our quantitative analysis, however, the five interviews with people from rural areas were especially insightful for understanding who may be most at risk of being defrauded.

7.1 Urban Participants

The interviews with three participants from urban areas revealed that all SMS users received promotional spam messages as well as fraud messages. Each participant was aware of fraud SMS messages and none of them had ever responded or fallen for fraudulent schemes circulated over SMS. They were reluctant to share their fraud messages, explaining that they were not sure about our intentions for asking to view those messages. Overall, they were aware of fraud over SMS and took a suspicious approach to SMS activity received from unknown numbers.

7.2 Rural Participants

The five participants from the rural village were semi-literate with education ranging from 8th to 10th grade, with the exception of one who attended college. Nearly all of them reported that they are low-income or belong to low-income families. They could read and write messages in Urdu (Arabic script) and transliterated Urdu (Roman script). Their smartphones had internet capabilities.

Four out of the five rural participants reported that they had not received spam or promotional messages. The exception was the person who went to a college in a nearby city. Nonetheless, all of them reported that they had received multiple fraud messages. We hypothesize that the fraud messages are potentially being sent to a wider list of (perhaps randomly generated) numbers as opposed to spam messages that target phone numbers collected from various services that often urban dwellers intentionally or unintentionally subscribe to such as pizza delivery, weekly deals, or online shopping order notifications.

All of the rural participants had experienced fraud or had witnessed someone fall victim to fraud (or at least engage in making an initial response to fraud). Three participants reported successful fraud attempts, while two discussed how they responded at first but them realized the fraudulent nature of the communication. The following are reports from the five rural participants:

Rural Participant One reported that he received a call from a regular (non-shortcode) number and was told that he had money waiting for him from Benazir Income Support Program (BISP). He was asked to pay Rs. 500 as a registration fee. The money was requested in the form of teleco prepaid credits by sending an SMS with the number from the balance reload scratch card. The participant

sent the SMS but subsequently did not hear back from the fraudster. Later he tried to call the number, but it was disconnected.

Like the first participant, Rural Participant Two received an SMS about collecting money from BISP. After he called the number as requested in the message, he was told that he needed to pay a fee to initiate the disbursement process and that the fee should be paid in the form of easy credit reload. He discussed the scenario with an imam at his local mosque who warned him that this was fraud. The fraudster kept calling him even though he told the fraudster that he suspected fraud. Eventually, the participant gave the number to a trusted person to report to authorities.

Rural Participant Three received a BISP fraud message in Arabic Urdu script. The fraudster informed him that he had received Rs 30,000, but to receive the money he was required to pay Rs 2,000 using Telenor's Easypaisa, which is a telco mobile wallet. The fraudster sent him a number to dial for whomever has an Easypaisa account. The participant went to an Easypaisa agent for help to make the transfer from his account. The agent explained to him that dialing the given code would transfer all the money from his account. This made the participant realize that this was fraud.

Rural Participant Four informed us that he had learned about different frauds from his friends who had been victims. He recounted how three of his friends had received an urgent message from a women who claimed that she had run out of credits on her phone and desperately needed to increase her balance. She promised that she would return the credits. All his friends who responded to these pleas and sent money never received any money in return. When they called the number, it played a recorded message in a woman's voice that repeated the same plea as had been previously texted.

Similarly, Rural Participant Five did not experience fraud himself but instead narrated an account of another person he knew who had been a victim of fraud. This acquaintance was told that he had won 400,000 rupees. The fraudster asked for Rs 2,500 in the form of scratch card numbers for balance reload in order to initiate the payment. The victim went to Participant Five because he worked at a general store which sells scratch cards. They only had 12 cards worth Rs 100 but the participant went with him to nearby village to buy the additional 13 cards. The victim sent the card numbers to the fraudster and they responded by giving him an address in a major city and the name of the person he should meet. He went to that address and found nobody with the given name there.

7.3 Summary of Interview Findings

Our interviews revealed that urban dwellers have a more suspicious attitude toward fraud messages and are familiar with fraud schemes. On the other hand, people from rural areas are more vulnerable due to lack of awareness about fraud schemes. They generally do not receive spam messages but regularly experience fraud SMS. The fraudster typically collects from their victims in the form of scratch card numbers for prepaid airtime. The most common scheme appears to be the BISP fraud.

8 DISCUSSION

8.1 The Fraud Ecosystem

Departing from our initial assumptions, SMS fraud in Pakistan as represented in our data is overwhelmingly related to lottery fraud schemes that are relevant to local programs and products. In these schemes, fraudsters aim to collect money in the form of airtime credits, which are more universal than mobile money wallets. Another fraud that was prominent in our dataset is the "damsel in distress" scheme, in which a woman pleads for help in the form of easyload credits. This scheme uses a simple narrative and requires no sophistication to convince the victim. The third most frequent fraud scheme, of which we obtained a few examples, is credential stealing where the victim is told their services are disabled and that they must call to reactivate them. This scheme is targeted at more tech-savvy bank users who would enable SMS notification, use ATM cards for their account, and potentially have an online login. While we did not see this this third type of fraud frequently, it alludes to the potential for the expansion of SMS fraud in Pakistan.

Our analysis suggests that mobile users in urban areas are well aware of the existence of fraudulent schemes and are skeptical of such schemes. The SMS app users could label SMS messages with over 70% accuracy in each category, while the verbal requests to low-income urban dwellers to share their fraudulent messages did not succeed. People's suspicions aid them in staying vigilant against fraud attempts, but a fraudster can overcome this by using more sophisticated and developed narratives. We ascertain that those who are low-income and live in rural areas are most vulnerable to current attacks. All of our rural interview participants had fallen for fraud attacks or knew someone who had been the victim of fraud.

Fraudsters use simple, socially engineered messages to defraud and convince their victims to call them. The majority of SMS fraud becomes a call-based fraud which contributes to the regular fraud ecosystem. They do not appear to use automation, and they appear to hand type the messages because the messages are visually similar with slight variations and errors. The fraudsters have access to a surprisingly large cache of phone numbers, some of which are employed to send the fraud SMS and others to list as call back numbers for their schemes. We found that fraudsters take advantage of a loophole in Pakistan's Biometric Verification System (BVS) to obtain verified SIMs. These SIMs are linked to real people, mostly in rural villages, who have no idea that their information is being misused. Most of the misuse of the BVS system happens at franchises and small outlets, and because of this we would advise telcos to enforce stricter practices at these centers. This also indicates how those in rural areas are at risk for being victims directly of fraud and also by implication with their credentials being stolen.

We had several discussions with stakeholders, including officials from telecommunication organizations, Pakistan Telecommunication Authority (PTA), and other security experts. Our conversations revealed that telcos have no interest in processing and filtering messages or taking on additional equipment costs. They did not see how this would lead to a return on investment. Many countries have a regulatory system that targets spam traffic and have deployed spam filters in their GSM networks. PTA does not enforce any such regulation, stating that this would be a form of censorship. Moreover, the PTA officials mentioned that more fraud happens when a fraudster directly calls their victims rather than by sending an SMS and waiting for the victim to call back. They explained that fraud over phone calls can reach upwards to millions of rupee. They also claim that educated people are vulnerable to responding to these frauds.

Fraudsters operate in gangs, and once a victim responds to an attack, multiple attacks will follow.

8.2 Mitigation Strategies

Education is the most effective strategy to mitigate fraud attacks. From our data, we observed that the PTA educated the public about BISP fraud by sending universal notices. Several banks also sent messages warning their users to never give out their passwords. All these messages arrived in threads where banks or the PTA were spamming users with other non-educational messages, so the educational part may have been overlooked by the recipients. Moreover, these are very specific examples of fraud awareness, and we recommend a more broad approach that would address current and potential schemes. A regular schedule for informing the public about fraud schemes could also alleviate people being defrauded by new schemes.

Disabling fraudulent phone numbers quickly is key to stopping all call initiated fraud. This strategy would especially help the most vulnerable. To collect active numbers, a government authority could establish a service where people could forward a potential fraud message. It is easy to identify a message of a known fraud scheme with the heuristic we developed. Phone numbers identified in this way could be automatically reported while the other unknown fraud-type messages could be evaluated case by case. Our SMS forwarding service that we advertised widely was the major source for our fraud corpus. During interviews, our participants requested a way to report fraud messages but without having to forward the messages. This indicates that a mitigation strategy could rely on reaching out to public for supportive data.

Preventing fraud via a fraud detection app at the user end is a more robust and distributed approach. The smartphone app with our defined heuristic could be effective in warning users as fraud schemes evolve. The heuristic is a simple algorithm that does not rely on heavy machine learning and would work locally on a user's phone instead of requiring heavy phone processing or cloud computation. The challenge from our experience is that either users do not trust lesser known apps or they lack the expertise to download and install a new app. At the same time, users who lack trust for a new app or use apps like Truecaller, a caller identification app that warns them about malicious phone numbers, typically recognize fraud SMS. We recommend explaining to users how to install a fraud detection app as well as the app's purpose.

8.3 Fraud amongst New Users and in Rural Areas

People who are relatively new users of mobile services are most vulnerable to SMS based fraud. They lack awareness about how messages could be spammed to random numbers and how fraudsters use this to their advantage to spread their fraud schemes to a large audience. Hence, these users trust these messages. As demonstrated in our interviews, they will engage in the fraud until someone in their community alerts them. The number of vulnerable newcomers is expanding as more people are coming online in this digital and mobile age. The majority of these newcomers are part of marginalized and rural communities. If new users have greater interaction with experienced users, then they are more likely to learn quickly about how to avoid SMS scams.

Those who are at risk include urban dwellers who might be educated but less tech-savvy (e.g., the elderly, women, those new to DFS etc.). They can recognize a simple lottery scheme but are not aware of or suspicious enough to pick up sophisticated attacks that are trying to steal banking details or credentials. For instance, they might not recognize that instead of "https://www.hbl.com", the given URL is "www.smshbl.com". This could lead to them signing onto a fake website and hence giving away their bank credentials. Moreover, they might not be aware of or know how to install a caller identification app like Truecaller, which is effective at identifying malicious numbers.

9 CONCLUSION

This paper examines the SMS fraud ecosystem in Pakistan using data collected through a smartphone SMS app, a SMS forwarding service, and interviews. It identifies ten fraud schemes that represent the following three categories of fraud: lottery, "damsel in distress," and stealing credentials. It concludes that new users and people in rural regions are the most vulnerable to these frauds. The majority of these frauds require the victim to call the fraudster, and most ask for money via airtime credits instead of through DFS. The paper also presents a heuristic that can detect fraud with a very high accuracy and is generic enough to be adapted to new fraud schemes as they evolve. Finally, it proposes a mitigation strategy drawn from the data analysis, experiences of defrauded individuals, and discussions with various stakeholders.

ACKNOWLEDGMENTS

We thank Jake Kendall, Samia Ibtasam, Sam Castle, and Kushal Shah for their feedback during the early stages of the project. We also thank Siyu Pan, Hung-Lin Yeh, and Sean Jaffe for their help in building the Android app, and Saksham Aggarwal for his help in labeling the data. We are grateful to Jennifer Webster for thoroughly reviewing and editing this paper. We would like to acknowledge all the participants who contributed their messages for our analysis and shared their fraud experiences. The work was supported in part by grants from the Financial Services for the Poor program at the Bill and Melinda Gates Foundation and from Karandaaz Pakistan.

REFERENCES

[1] Frank W. Agbola, Angelito Acupan, and Amir Mahmood. 2017. Does microfinance reduce poverty? New evidence from Northeastern Mindanao, the Philippines. *Journal of Rural Studies* 50 (Feb. 2017), 159–171. DOI:http://dx.doi.org/10.1016/j.jrurstud.2016.11.005

[2] David S Anderson, Chris Fleizach, Stefan Savage, and Geoffrey M Voelker. 2007. *Spamscatter: Characterizing internet scam hosting infrastructure*. Ph.D. Dissertation. University of California, San Diego.

[3] Yahel Ben-David, Shaddi Hasan, Joyojeet Pal, Matthias Vallentin, Saurabh Panjwani, Philipp Gutheim, Jay Chen, and Eric A Brewer. 2011. Computing security in the developing world: A case for multidisciplinary research. In *Proceedings of the 5th ACM workshop on Networked systems for developing regions*. ACM, 39–44.

[4] Moez Bhatti. 2019. QKSMS. https://github.com/moezbhatti/qksms. (2019). Accessed March 2019.

[5] Jenna Burrell. 2008. Problematic Empowerment: West African Internet Scams as Strategic Misrepresentation. *Information Technologies and International Development* 4, 4 (2008), 15–30.

[6] Sam Castle, Fahad Pervaiz, Galen Weld, Franziska Roesner, and Richard Anderson. 2016. Let's Talk Money: Evaluating the Security Challenges of Mobile Money in the Developing World. In *Proceedings of the 7th Annual Symposium on Computing for Development (ACM DEV '16)*. ACM, New York, NY, USA, 4:1–4:10. DOI: http://dx.doi.org/10.1145/3001913.3001919

[7] Gordon V Cormack and others. 2008. Email spam filtering: A systematic review. *Foundations and Trends® in Information Retrieval* 1, 4 (2008), 335–455.

[8] Henry Corrigan-Gibbs and Jay Chen. 2014. Flashpatch: spreading software updates over flash drives in under-connected regions. In *Proceedings of the Fifth ACM Symposium on Computing for Development*. ACM, 1–10.

[9] DAWN. 2018. FIA finds the educated "equally vulnerable" to online bank fraud. (2018). https://www.dawn.com/news/1444243

[10] Sarah Jane Delany, Mark Buckley, and Derek Greene. 2012. SMS spam filtering: methods and data. *Expert Systems with Applications* 39, 10 (2012), 9899–9908.

[11] Ficawoyi Donou-Adonsou and Kevin Sylwester. 2016. Financial development and poverty reduction in developing countries: New evidence from banks and microfinance institutions. *Review of Development Finance* 6, 1 (June 2016), 82–90. DOI:http://dx.doi.org/10.1016/j.rdf.2016.06.002

[12] José María Gómez Hidalgo, Guillermo Cajigas Bringas, Enrique Puertas Sánz, and Francisco Carrero García. 2006. Content based SMS spam filtering. In *Proceedings of the 2006 ACM symposium on Document engineering*. ACM, 107–114.

[13] Joshua Goodman, Gordon V Cormack, and David Heckerman. 2007. Spam and the ongoing battle for the inbox. *Commun. ACM* 50, 2 (2007), 24–33.

[14] Zoltan Gyongyi and Hector Garcia-Molina. 2005. Web spam taxonomy. In *First international workshop on adversarial information retrieval on the web (AIRWeb 2005)*.

[15] Rasmus Heltberg, Naomi Hossain, and An Reva. 2012. *Living through crises: How the food, fuel, and financial shocks affect the poor*. The World Bank.

[16] Samia Ibtasam, Hamid Mehmood, Lubna Razaq, Jennifer Webster, Sarah Yu, and Richard Anderson. 2017. An Exploration of Smartphone Based Mobile Money Applications in Pakistan. In *Proceedings of the Ninth International Conference on Information and Communication Technologies and Development (ICTD '17)*. ACM, New York, NY, USA, Article 1, 11 pages. DOI:http://dx.doi.org/10.1145/3136560.3136571

[17] Jelena Isacenkova, Olivier Thonnard, Andrei Costin, Aurélien Francillon, and David Balzarotti. 2014. Inside the scam jungle: A closer look at 419 scam email operations. *EURASIP Journal on Information Security* 2014, 1 (2014), 4.

[18] Nan Jiang, Yu Jin, Ann Skudlark, and Zhi-Li Zhang. 2013. Greystar: Fast and Accurate Detection of {SMS} Spam Numbers in Large Cellular Networks Using Gray Phone Space. In *Presented as part of the 22nd {USENIX} Security Symposium ({USENIX} Security 13)*. 1–16.

[19] Amir Karami and Lina Zhou. 2014. Improving static SMS spam detection by using new content-based features. (2014).

[20] Dean Karlan, Jake Kendall, Rebecca Mann, Rohini Pande, Tavneet Suri, and Jonathan Zinman. 2016. *Research and impacts of digital financial services*. Technical Report. National Bureau of Economic Research.

[21] Amy K Karlson, AJ Brush, and Stuart Schechter. 2009. Can i borrow your phone?: understanding concerns when sharing mobile phones. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1647–1650.

[22] Jon A Krosnick, Sowmya Narayan, and Wendy R Smith. 1996. Satisficing in surveys: Initial evidence. *New directions for evaluation* 1996, 70 (1996), 29–44.

[23] Robert MacIntosh and Dmitri Vinokurov. 2005. Detection and mitigation of spam in IP telephony networks using signaling protocol analysis. In *IEEE/Sarnoff Symposium on Advances in Wired and Wireless Communication, 2005.* IEEE, 49–52.

[24] Kate McKee, Michelle Kaffenberger, and Jamie Zimmerman. 2015. Doing Digital Finance Right. (June 2015). http://www.cgap.org/publications/doing-digital-finance-right

[25] Najmeh Miramirkhani, Oleksii Starov, and Nick Nikiforakis. 2016. Dial one for scam: A large-scale analysis of technical support scams. *arXiv preprint arXiv:1607.06891* (2016).

[26] Joseck Luminzu Mudiri. 2013. Fraud in mobile financial services. *Rapport technique, MicroSave* (2013), 30.

[27] Akshay Narayan and Prateek Saxena. 2013. The curse of 140 characters: evaluating the efficacy of SMS spam detection on android. In *Proceedings of the Third ACM workshop on Security and privacy in smartphones & mobile devices*. ACM, 33–42.

[28] The Nation. 2018. Internet banking frauds on the rise. (2018). https://nation.com.pk/29-Nov-2018/internet-banking-frauds-on-the-rise

[29] Rowan Phipps, Shrirang Mare, Peter Ney, Jennifer Webster, and Kurtis Heimerl. 2018a. ThinSIM-based Attacks on Mobile Money Systems. In *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*. ACM, 23.

[30] Rowan Phipps, Shrirang Mare, Peter Ney, Jennifer Webster, and Kurtis Heimerl. 2018b. ThinSIM-based Attacks on Mobile Money Systems. In *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies (COMPASS '18)*. ACM, New York, NY, USA, Article 23, 11 pages. DOI:http://dx.doi.org/10.1145/3209811.3209817

[31] Bradley Reaves, Logan Blue, Dave Tian, Patrick Traynor, and Kevin RB Butler. 2016. Detecting SMS spam in the age of legitimate bulk messaging. In *Proceedings of the 9th ACM Conference on Security & Privacy in Wireless and Mobile Networks*. ACM, 165–170.

[32] Bradley Reaves, Nolen Scaife, Adam Bates, Patrick Traynor, and Kevin R. B. Butler. 2015. Mo(bile) Money, Mo(bile) Problems: Analysis of Branchless Banking Applications in the Developing World. In *USENIX Security*. 17–32.

[33] Johan Rewilak. 2017. The role of financial development in poverty reduction. *Review of Development Finance* 7, 2 (Dec. 2017), 169–176. DOI:http://dx.doi.org/10.1016/j.rdf.2017.10.001

[34] Frank Stajano and Paul Wilson. 2009. *Understanding scam victims: seven principles for systems security*. Technical Report. University of Cambridge, Computer Laboratory.

[35] Pakistan Today. 2018. Fraudsters continue scamming unsuspecting citizens in BISP's name. (2018). https://bit.ly/2F8kelG

[36] Aditya Vashistha, Richard Anderson, and Shrirang Mare. 2018. Examining Security and Privacy Research in Developing Regions. In *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies (COMPASS '18)*. ACM, New York, NY, USA, Article 25, 14 pages. DOI:http://dx.doi.org/10.1145/3209811.3209818

[37] Qian Xu, Evan Wei Xiang, Qiang Yang, Jiachun Du, and Jieping Zhong. 2012. Sms spam detection using noncontent features. *IEEE Intelligent Systems* 27, 6 (2012), 44–51.