A Study of Targeted Telephone Scams Involving Live Attackers

Ian G. Harris[1], Ali Derakhshan[1], and Marcel Carlsson[2]

[1] University of California Irvine, Irvine CA 92697, USA
harris@ics.uci.edu, aderakh1@uci.edu
[2] Lootcore, Sweden
mc@lootcore.com

Abstract. We present the results of a research study in which participants were subjected to social engineering attacks via telephone, *telephone scams*, in order to determine the features of scams which people are most susceptible to. The study has involved 186 university participants who were attacked with one of 27 different attack scripts which span different independent variables including the pretext used and the method of elicitation. In order to ensure informed consent, each participant was warned that they would receive a scam phone call within 3 months. One independent variable used is the time between the warning and launching the scam. In spite of this warning, a large fraction of participants were still deceived by the scam.
A limitation to research in the study of telephone scams is the lack of a dataset of real phone scams for examination. Each phone call in our study was recorded and we present the dataset of these recordings, and their transcripts. To our knowledge, there is no similar publicly-available dataset or phone scams. We hope that our dataset will support future research in phone scams and their detection.

Keywords: social engineering · telephone scams · attack dataset

1 Introduction

Social engineering attacks, or scams, describe the psychological manipulation of people to convince them to do something that they should not do [8,15]. Social engineers pretend to be some trustworthy entity, or some entity with authority over the victim. Social engineering attacks can be delivered in many ways but electronic communications, such as email or text message are common platforms. Email phishing has been shown to be an effective attack over the years, deceiving a broad range of people [11]. Attackers often gain personal information that affects the victims' personal lives, financial wellbeing, and work environment. Phishing, in all of its forms, is very popular in real attacks. The Verizon 2019 Data Breach Investigations Report [26] states that 32% of all breaches included phishing and 78% of all cyber-espionage which involved state-affiliated actors.

A growing problem is social engineering attacks launched over the phone, or *telephone scams*. Telephone-based attacks can be more effective that email-based attacks in part because the victim is involved in a live conversation, so they feel pressured to respond quickly, without having time to think as would be the case with emails. The sheer volume of phone scams has greatly increased recently. First Orion, a call blocking technology company, estimates that more than 29% of all cellphone calls in 2018 were scams, and expects that almost half of cellphone calls in 2019 will have been scams [24]. The financial losses associated with phone scams are significant. There is a wide variety of common phone scams including the Technical Support scam which Microsoft reports targeted over 3.3 million people in 2015 and cost those people $1.5 million [14]. Nearly 70% of frauds reported to the Federal Trade Commission in 2017 where perpetrated by phone, while only 9% were conducted by email during the same period [17]. The 2019 U.S. Spam & Scam Report from Truecaller, a caller ID and spam blocking company, reveals that Americans lost $10.5 billion to phone scams in the 12 month period before the report was released in April 2019 [14].

The importance of phone scams motivates the need to understand what aspects of phone scams cause people to be convinced by them, so that defense approaches may be developed. Many empirical studies have been performed to understand phishing email scams and their ability to convince victims [7,2,11,6]. However, telephone scams are different from phishing emails in their application and effect. Telephone conversations are real-time, unlike emails, so the victim feels time pressure to respond. Telephone scams can also involve direct human-to-human interaction, which has a different emotional impact on a victim as compared to the receipt of an email. A study of user responses to telephone scams has been presented [25] which begins to shed light on critical aspects of scams, such as the importance of caller ID. However, the study presented in [25] uses only pre-recorded attack calls which have a different impact than calls from live attackers. There are also other aspects of telephone scams which need to be understood, such as the sensitivity of users to different items of information which an attacker may ask them to reveal.

In order to perform research to help protect against telephone scams, researchers need access to datasets of realistic telephone scams which can be studied, and used for training and evaluation of detection approaches. Several large datasets of phishing emails are publicly available [20,21,22], but similar datasets of telephone scams are not available. As a result, there is a large body of previous work on the detection of phishing emails, but very little on the detection of telephone scams [3]. It is difficult to build a dataset of telephone scams due to the legal need for consent of both communicating parties in order to record a telephone call, as is required in most states in the US.

1.1 Social Engineering Study

This paper presents the results of an empirical study on the susceptibility of people to telephone scams. We created a set of 27 different *attack scripts* which were used to scam 186 participants. The attack scripts varied over several different

independent variables including the pretext used and the information requested by the attacker. We present results to show the impact of each independent variable on the success of the attack.

Targeting of the Attack We want to understand the impact of three different aspects of social engineering attacks on the success of the attack. We define the **targeting** of an attack as degree to which the attack is personalized to appeal to a subset of the population. An attack which can be applied to a wide range of people is not well targeted, such as an IRS scam which is generically applicable to any adult in the US who interacts with the IRS. An attack which is targeted to a medium degree would be one in which a caller pretends to be from the IT office of a particular company. Such an attack is targeted towards a smaller set of people, those who work at the company in question. A highly targeted attack might be focus on a single individual by referencing personal information which has been gathered using open-source intelligence techniques.

Targeting is used by attackers because it generally improves the effectiveness of the attack, but the attacker may mis-target an attack because she does not have full information about the target. For example, an attacker may assume that people who use a particular app (i.e. tik tok) are generally young, so an attack against users of the app might be designed to appeal to young people. However, not all users of the app are young, so an attack is mis-targeted when it is launched against a user of the app who is old. To formalize the concept of attack targeting, we define the set P to be the set of people whom the attack campaign is made to appeal to, and the set Q to be the set of people who are actually attacked. We refer to the *targeting accuracy*, a, of an attack as follows, $a = \frac{|P \cap Q|}{|Q|}$. An attacker with limited knowledge of the victims is forced to choose a trade-off between increasing the number of people attacked but reducing targeting accuracy because the attack may not appeal to victims.

We explored the relationship between the targeting accuracy of the attack and the success of the attack. We consider the following hypotheses:

- **Alternative Hypothesis.** $H_{1,1}$:The targeting accuracy of the attack, a, impacts the success rate of the attack.
- **Null Hypothesis.** $H_{0,1}$: The targeting accuracy of the attack has no impact on the success rate of the attack.

Sensitivity of Information We are investigating attacks in which the attacker attempts to gain information from the victim and we want to understand the impact that the choice of information has on the success of the attack. Different types of information have different protection requirements from the perspective of the victim. An email address may not need to be hidden, especially if your email is already publicly available. However, the cost of revealing a social security number is high. We expect that the success of a social engineering attack will depend on the type of information requested.

We define an independent variable *information goal* which describes the private information requested during an attack. We consider the following hypotheses:

- **Alternative Hypothesis.** $H_{1,2}$:The independent variable *information goal* impacts the success rate of the attack.
- **Null Hypothesis.** $H_{0,2}$: The independent variable *information goal* has no impact on the success rate of the attack.

Attack Awareness Over Time A common defense against social engineering attacks is the use of "awareness training" to prepare employees to protect themselves [23]. However, the effectiveness of this type of training is not clear because people may easily forget their training over time. As part of our study, we notify participants that they will be attacked as part of the study, so they have complete awareness that the attack will come. However, we call them between 1 and 3 months after joining the study. We expect that individuals may lose their attack awareness over time, so the success rate of an attack will depend on the *attack delay*, the time between when an individual is made aware of a potential attack and the when an attack occurs. We consider the following hypotheses:

- **Alternative Hypothesis.** $H_{1,3}$:The independent variable *attack delay* impacts the success rate of the attack.
- **Null Hypothesis.** $H_{0,3}$: The independent variable *attack delay* has no impact on the success rate of the attack.

1.2 Social Engineering Dataset

We additionally present the recordings of the 186 attack phone calls, and their transcripts, as a publicly-available dataset for use by researchers studying telephone scams. Each recording was made with the explicit permission of the participant. We hope that this dataset can be used by others as examples of both successful and unsuccessful social engineering attacks. Although each call is based on one of only 27 attack scripts, there are significant variations between calls based on the unpredictable responses of the victims. We provide the audio files of the phone calls in addition to transcripts so that researchers can examine prosodic content of the calls.

2 Telephone vs. Email Scams

This study specifically focuses on telephone scams rather than email phishing scams. Many studies have been performed using phishing emails, and datasets of phishing emails have been compiled. However, these studies and datasets do not adequately represent the properties of telephone-based attacks. Phishing studies have the following limitations in representing telephone scams.

- **Communication Metadata** - Emails contain significant metadata produced as part of the communication protocols used (i.e. headers, footers, embedded URL links) which can be used to detect phishing. Unfortunately, much of this information is different for texting and telephone communication, and entirely absent for in-person communication. This problem is most apparent for the problem of authenticating the source of a communication. Many phishing detection approaches achieve high precision and accuracy by analyzing email metadata to determine that the actual source is not the same as the stated source. These approaches are not applicable to non-email communications however. The availability of non-email social engineering attacks will enable researchers to study the detection of a broader range of attacks.
- **One-way, Context-free Communication** - Phishing emails found in existing databases all show a single communication from an attacker to a victim. They do not show conversations between the attacker and the victim. In most cases of phishing attacks, there is no conversation and the entire attack is composed of a single email. Even in cases of phishing attacks which lead to a conversation between the attacker and victim, the phishing emails found in existing databases are individual with no context given. Social engineering attacks launched via texting, phone, or in-person almost always involve a conversation between the attacker and the victim. The context of the entire sequence of communications can contain information essential to identifying an attack. An examples of the importance of context is the use of dialog designed to alter the victim's mood (i.e. urgency, flirtation, etc.). A mood change change early in the conversation can change the victim's response to a request for private data later in the conversation.
- **Text-based vs. Oral** - Text-based communication depends only on text to transfer information, while verbal communication can use properties of the voice, prosody, to transfer information. As a result, people have developed different approaches for encoding information in text as compared to voice. A simple example is a sense of irony which can be captured in the tone of voice during an oral conversation, but might be captured using an emoji in a text-based conversation.

3 Experimental Procedure

We performed a set of experiments to determine the effectiveness of a variety of telephone-based social engineering attacks. Each participant received a scam phone call within 3 months of joining the study. Each scam phone call requested a single piece of personally identifying information (PII). A scam phone call is considered a success if the PII data was provided, and it is considered a failure otherwise. A call is considered a failure if the participant hangs up before he/she has the opportunity to provide an answer. A call is also considered a failure if the participant asks questions which force the attacker to diverge from the script in a significant way. A divergence from the script is acceptable if the participant

explicitly asks for assistance in providing the requested private data, such as, "How do I find my IP address?". If the participant does not answer the phone then he/she is called again up to five times during the next 5 business days in order to establish contact. If the participant does not answer the phone after 5 call attempts then the participant is dropped from the study and their results are not included in the study.

The main difficulty in designing this experiment is the inherent conflict between the two primary goals of *accuracy* and *ethicality* [12]. In order for the experiment to accurately determine the effectiveness of an attack, deception is required to apply the attack in a realistic fashion against an unsuspecting participant. However, in order for an experiment to be ethical, deception of the participants must be well justified in terms of the needs of the experiment and the benefits of the research [19]. In designing these experiments we have used the advice of the Ethics Feedback Panel for Networking and Security (http://www.ethicalresearch.org/efp/netsec/) which provided us with several ideas on achieving accuracy while maintaining ethicality.

The procedural steps are presented here.

1. **Attract Voluntary Participants**: We advertised for participants in the following ways: posters on campus, announcements in classes, announcements on Facebook pages of campus student groups. The target population was primarily campus students and financial compensation of a $15 Amazon gift card was offered for participation.
2. **Obtaining Informed Consent**: We informed the participants of the deceptive nature of the experiments and we obtained their consent before launching the attack. Specifically, subjects were advised that we would attempt to deceive them and that phone calls would be recorded for analysis. We also advised them that the contents of the phone calls would be edited for PII and then published.
3. **Launch the Attack**: At some point within the three month period after the subject joins the study, we launched the attack. Attacks were conducted via telephone and each attack was recorded.
4. **Debriefing**: Immediately after the attack has been concluded, while the participant is still on the phone, the participant was informed that the preceding conversation was actually an attack conducted as part of the study. This occurred whether the attack was successful or not. In cases where the participant hung up the phone before the completion of the attack, the student was later contacted via email for debriefing.

3.1 Ethical and Legal Concerns

A number of issues arose which were addressed in order to gain IRB approval for the study, and which impact the validity of the results. We have considered these issues and we have structured the study to ensure that it is legal and ethical, while still producing the desired result of evaluating the effectiveness of a set of synthesized social engineering attacks.

Legal The use of deception as part of these experiments requires that we consider state and federal laws prohibiting such deception. The first set of laws which impact our study are generally referred to as *wiretap* laws which define when it is illegal to record communications. Federal wiretap laws are "one-party consent" laws which allow communication to be recorded if a single party has given consent to the recording. Our study is at no risk of violating federal wiretap laws since our student assistants who launch the attacks are clearly giving consent for recording the communication. Many states however, including the state in which the study was conducted, have stronger "two-party consent" laws which require both parties involved in a communication to consent to the recording. Our study is also at no risk of violating these laws because we received informed consent of each subject when they first joined the study.

The pretexting component of a social engineering attack includes the act of "impersonation" as a tool to gain the trust of the subject. We are aware that federal laws prohibit the impersonation of any government worker or officer. We have only used pretexts involving campus officials and the campus IRB has explicitly given us permission to do so.

Protecting Participants from Harm There is a risk of two types of harm to the participants of this study.

- **Material Harm:** This describes the possibility of the participant suffering harm in a physical or financial sense. Physical harm is not likely but the participant may reveal information which could enable theft, such as a social security number or a bank account password. The participant might also reveal sensitive private information which could be used by a malicious actor to perform blackmail against the participant.
- **Psychological Harm:** This describes the emotional distress which the participant might suffer from being deceived.

The risk of psychological harm was considered by our IRB to be very low since participants are immediately debriefed at the end of the conversation. In order to protect private participant data which is learned during an attack, we delete all PII from each recording immediately after the completion of the phone call. Each item of PII is replaced in the audio file with a 440Hz tone of equal duration, completely overwriting the PII data in the audio file.

Subject Attack Awareness It is essential to inform each subject of the nature of the social engineering attacks when they enter the study, but the disadvantage of informing the participants is that it may increase their attack awareness and skew their responses. There is a large body of evidence [5,10] showing that the rate at which information is forgotten is exponential in time. As a result, we expect that the subject's attack awareness will degrade quickly after they have given informed consent. At the beginning of the experiment, the subjects were informed that the attack may occur anytime within the following three months. We varied the time period between when the participant joined the study and

when the participant was called in order to explore how the delay impacts the likelihood of the attack being successful.

4 Attack Scripts

Based on previous work studying the content of social engineering attacks [9,16], we describe the key parts of an attack.

- **Pretext** The act of *pretexting* is the creation of a scenario to persuade the target to either provide the desired information, or perform the desired action. We define the pretext of the attack as the communication which is used by the attacker to present the pretext to the target. The context of the pretext will define a false identity for the attacker which is trusted by the target to some degree. The pretext may be as simple as a false introduction such as, "Hi, I am Joe from the bank", but it may also include a detailed description of a false situation which would justify the communication attempt from the attacker's false identity. For example, the pretext might be, "Hi, I am Joe from the bank and we have detected strange activity on your account. We will need to verify your account information before we can fix the problem". This pretext defines a false, trusted identity, and justifies requests for account information.
- **Elicitation** Elicitation is the process of changing the emotional state of the victim to one in which the victim is more likely to provide information. This is generally done by either building a rapport with the victim to make the target comfortable, or by threatening the victim to make the victim afraid of some negative consequence. An attacker might build a rapport by asking for assistance, such as "I need your help to resolve a problem with our database". An attacker might cause fear by making a threat, such as, "We will have to unregister you from your classes unless you resolve this issue immediately".
- **Information Goal** The culmination of the social engineering attack is to request private information, such as, "Please provide your social security number to verify your identity".

4.1 Structure of a Script

Each script describes the statements that the caller must make during an attack, and the sequence in which the statements must be made. Each script is a sequence of the following 4 steps.

1. Pretext
2. Name Verification
3. Elicitation
4. Information Goal

The *Name Verification* step is a request to verify the name of the caller: "Am I speaking to <first name>, <last name>?". This is required to ensure that the

caller is speaking to the participant who has given consent. Aside from Name Verification, each attack script is built from one Pretext, one Elicitation, and one Information Goal. We used 3 possible Pretexts, and for each Pretext we used 3 Elicitations and 3 Information Goals. The scripts are generated as the cross-product of these three sets, so there are 27 scripts in total.

Pretexts We use the following three pretexts which are all targeted towards campus students.

- Pretext 1: Registrar's Office - "Hi, I'm Jim/Jane from the XXX Registrar's Office and I have a couple of questions about your course registration."
- Pretext 2: Transportation Services (parking) - "Hi, I'm Jim/Jane from XXX Transportation Services."
- Pretext 3: Office of Information Technology (OIT) - "Hi, my name is Jane/Jim and I am with the XXX Office of Information Technologies."

These three pretexts vary in terms of the accuracy of targeting. Notice that pretext 1 and pretext 3 both apply to all students, but pretext 2 only applies to a student who owns or has access to a car. Since our study involves the community of all students at our school, pretext 2 is targeted with lower accuracy than 1 and 3 because pretext 2 is not applicable to all victims of the attack.

Elicitations For each pretext, we use three elicitations which make sense with the pretext. Almost all elicitations are meant to either induce fear in the victim, or build a rapport by asking the victim for help. The elicitations used for each pretext are shown here.

- Registrar's Office
  - Elicitation 1.1: "OK, thanks. We believe that our registration system has been hacked so we need to verify your course registration. You are currently registered for ICS 155 in Spring quarter, is that correct?"
  - Elicitation 1.2: "Great. We received a phone call from a person claiming to be you, who requested that you be dropped from all of your courses. We need to verify that request."
  - Elicitation 1.3: "OK thanks a lot."
- Transportation Services
  - Elicitation 2.1: "OK, thanks. We believe that our computer system has been hacked and some of our information on you might have been altered. We need to verify some of your information, if you have a minute to do it right now."
  - Elicitation 2.2: "Alright. Our parking database seems to have multiple records associated with your name. One record says that you own one car and no pending citations, while the other record says that you have 10 pending citations. We need to get some information from you to resolve the inconsistency."

  - Elicitation 2.3: "Thank you. At this point your citation is overdue by 3 months."
- Office of Information Technology
  - Elicitation 3.1: "OK, thanks. It looks like your computer is infected by a virus of some kind and it is attempting to infect other computers whenever you connect to the campus network."
  - Elicitation 3.2: "Thanks. We have detected multiple attacks on the OIT servers which seem to originate from IP addresses which you were using at the time."
  - Elicitation 3.3: "Thanks. Someone has just attempted to setup a new campus account using your name. This may not be a problem. It may be that another campus member has the same name as you do, but we need to be certain."

Information Goals Across all attack scripts, we use a total of 6 information goals: Postal Address, Social Security Number, Email Address, Driver's License Number, License Plate Number, and IP Address. Each information goal is a data considered to be personally identifying information (PII) by our institutional review board (IRB), but they are expected to have different levels of sensitivity from the participant's perspective. For each pretext, we use three information goals which make sense with the pretext.

- Registrar's Office
  - Goal 1.1: Postal Address, "Can you give me your postal address for verification purposes?"
  - Goal 1.2: Social Security Number, "Please give me your social security number for verification purposes."
  - Goal 1.3: Email Address, "Can you give me your email address for verification purposes?"
- Transportation Services
  - Goal 2.1: Driver's License Number, "Please give me your driver's license number so that I can verify your record."
  - Goal 2.2: License Plate Number, "Can you give me your license plate number so that I can verify your record?"
  - Goal 2.3: Social Security Number, "Please give me your social security number so that I can verify your record."
- Office of Information Technology
  - Goal 3.1: Email Address, "Can you give me your email address for verification purposes?"
  - Goal 3.2: IP Address, "Please give me your computer's IP address for verification purposes."
  - Goal 3.3: Social Security Number, "Can you give me your social security number for verification purposes?"

4.2 Controlled and Uncontrolled Variables

Controlled variables are those independent variables which are held constant throughout the experiment so that their value does not obscure the causal relationships which we seek to identify between the other independent variables and the success rate of the attacks. The main controlled variables in our experiment are the *source phone number* used to place the attack phone calls, and the *accent* of the callers.

All attacks were made from a legitimate campus phone number which would appear on the caller ID of the victim with the same area code and three-digit prefix as any other campus number. The number used was not the actual number of the campus offices used as pretexts, but it is safe to assume that in most cases, just the area code and three-digit prefix were sufficient to convince many participants that the call was from an official campus source. We used a real campus phone number to simulate the process of spoofing a caller ID which is most often done by real attackers to enhance the believability of the attack. All of the callers were American and had neutral accents.

Uncontrolled variables are those which might have an impact on the results but were not explicitly controlled as part of the experiment. The main uncontrolled variable was the *gender* of the caller. Of the total 186 phone calls, 60 were made by a man and the remaining 126 were made by two women. The calls made by men and women were well distributed across the set of 27 attack scripts, but most of the calls were made by women.

5 Study Results

A total of 234 people joined the study and 48 of those, 20.5%, were dropped from the study because they did not pick up their phone after 5 phone call attempts during a week. A total of 186 attacks were completed, and of those, 58 were successful, so 31.18% of calls were successful, overall. On average, 6.89 calls were made using each script, and the standard deviation of the number of calls per script is 2.64.

5.1 Demographics of the Participants

The participants were undergraduate students at the University of California Irvine. Figure 1 shows the age distribution of the participants whose average age is 19.46 years old. Figure 2 shows the distribution of the participants according to the school at the university which contains their major. It is clear that the participants are most concentrated in "ICS" which stands for Information and Computer Science. Note that the sum of all numbers in the table is greater than the 186 participants who completed the study because students with double majors are counted twice if their two majors are in different schools.

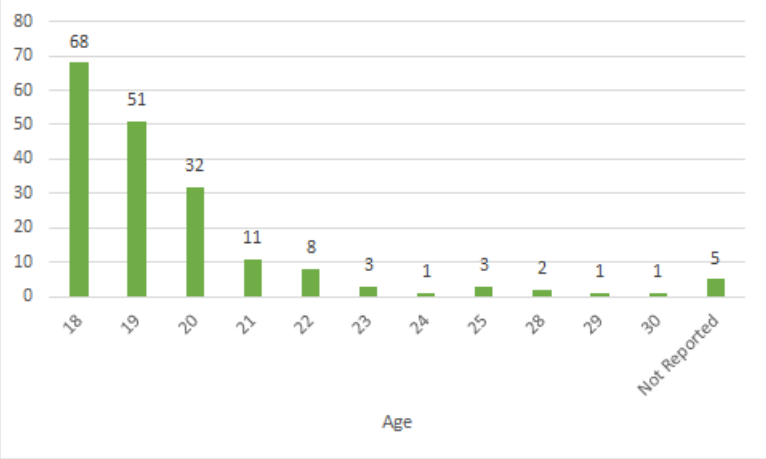

Fig. 1. Age distribution of participants

5.2 Success Rates

To gain insight into what attack features are correlated with success rate, we examine the success rates for different values of the independent variables.

Pretext	**Calls**	**Succ.**	**Succ. Rate**
Registrar's	85	37	43.53%
Transportation	50	3	6.00%
OIT	51	18	35.29%
Total	186	58	31.18%

Table 1. Success rate according to pretext

Table 1 shows the success rate according to the pretext used. Each row, other than the first and last, shows the results for one pretext. The columns show the name of the pretext (**Pretext**), the number of attacks made using that pretext (**Calls**), the number of successful attacks (**Succ**), and the success rate (**Succ. Rate**). It is clear from these results that the Transportation pretext resulted in a lower success rate than the other two. This is probably because some students do not own cars, while almost all students will be registered for classes and have a computer account through the Office of Information Technology (OIT). By questioning during debriefing we found that only 62% of subjects who were scammed using the Transportation pretext owned or had access to cars.

Table 2 shows the success rate according to the information goal. It is clear from the table that victims have some understanding of the sensitivity of information. For example, email address was provided 75% of the time because it

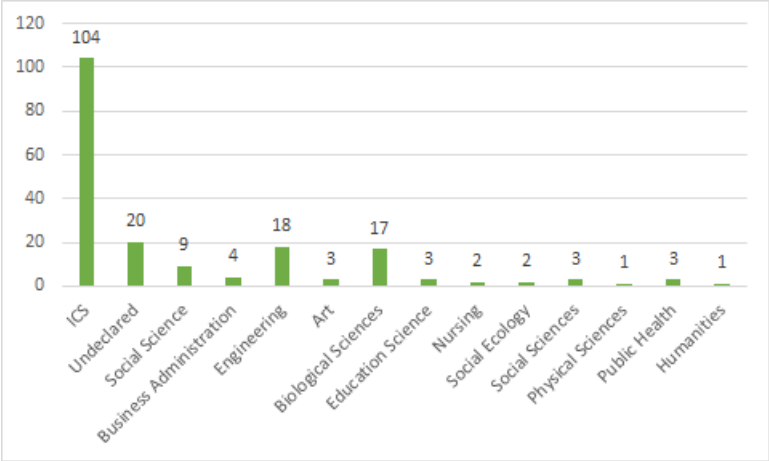

Fig. 2. Major distribution of participants

Info. Goal	Calls	Succ.	Succ. Rate
Postal Address	33	17	51.56%
Soc. Security	58	1	1.72%
Email Address	44	33	75.00%
Driver's License	17	0	0.00%
License Plate	16	3	18.75%
IP Address	18	4	22.22%

Table 2. Success rate according to information goal

is usually easy to determine a student's email address by searching the public campus database. The success rates for Driver's License and License Plate Number are likely to be artificially low because they were only associated with the Transportation pretext, whose success rate is low as shown in Table 1.

Attack Delay	Calls	Succ.	Succ. Rate
1-2 months	87	26	29.88%
2-3 months	99	32	32.32%

Table 3. Success rate according to time frame

Table 3 shows the success rate according to the time frame, the time between when a participant joined the study and when he/she was attacked. When joining the study, participants are informed that they will be attacked, so it is expected that the success rate will increase as the time frame increases, since participants will tend to forget. The success rate does increase, but only by 7.55% between the two time frames.

5.3 Hypotheses Tests

In order to evaluate each hypothesis we computed a binomial logistic regression to test whether the independent variables impact the success rate. Table 4 shows an overview of the estimates of this model. A total of 8 independent variables are used, including 2 dummy variables to represent the categories of the pretext, 5 dummy variables to represent the information goal categories, and 1 variable to represent the attack delay. The 2 pretext variables are defined with respect to Transportation as the baseline condition. The 5 information goal variables are defined with respect to Email Address as the baseline condition. The Attack Delay variable is coded such that 0 represents a 1-2 month delay and 1 represents a 2-3 month delay. The Registrar's and OIT pretexts are considered to have high targeting accuracy since all students register for classes and all students have computer accounts. The Transportation pretext, which is the baseline condition, has low targeting accuracy since only 62% of the victims of the Transportation pretext actually owned or had access to cars.

Independent Variable	β	SE	z value	p value
Registrar's	1.0987	0.477	2.303	0.021
OIT	1.4079	0.684	2.058	0.040
Postal Address	-1.3126	0.559	-2.346	0.019
Social Security	-5.1515	1.111	-4.638	0.000
Driver's License	-3.8505	1.703	-2.260	0.024
License Plate	-1.5787	0.706	-2.237	0.025
IP Address	-2.6593	0.824	-3.228	0.001
Attack Delay	0.0196	0.434	0.045	0.964

Table 4. Logisitic Regression

- $H_{1,1}$: There was a statistically significant positive impact of using the high targeting accuracy pretexts Registrar's (z=-2.303, p=0.021, OR = 3.00, 95% CI [1.18, 7.64]) and OIT (z=-2.058, p=0.040, OR = 4.09, 95% CI [1.07, 15.63]). Based on these results, we can reject the null hypothesis $H_{0,1}$ and accept the alternate hypothesis $H_{1,1}$.
- $H_{1,2}$: The logistic regression produces the following statistics for each information goal.
  - Postal Address, z=-2.346, p=0.019, OR = 0.27, 95% CI [0.09, 0.81]
  - Social Security, z=–4.638, p=0.000, OR = 0.01, 95% CI [0.00, 0.05]
  - Driver's License, z=-2.260, p=0.024, OR = 0.02, 95% CI [0.00, 0.60]
  - License Plate, z=-2.237, p=0.025, OR = 0.21, 95% CI [0.05, 0.82]
  - IP Address, z=-3.228, p=0.001, OR = 0.07, 95% CI [0.01, 0.35]

  All of information goals have a statistically significant negative impact on success as compared to the baseline Email Address. There is also clearly a wide range of odds ratios between 1.0 (for the Email Address information goal itself) down to 0.01. Based on these results we can reject the null hypothesis $H_{0,2}$ and accept the alternate hypothesis $H_{1,2}$.

– $H_{1,3}$: The logistic regression for the Attack Delay variable shows that the CI of the odds ratio contains 1.0 (z=0.045, p=0.964, OR = 1.02, 95% CI [0.44, 2.39]). Although the overall success rate is high, 31.18%, in spite of the fact that participants were made aware of future attacks, there is no indication that their awareness decreased over time. We cannot reject the null hypothesis $H_{0,3}$.

6 Discussion and Limitations

A surprising result was the fact that the attack delay seems to have no impact on the success rate. We assumed that the success rate would increase as participants forgot their "training". However, it is still possible that the time scales that we examined were too large to see the effect. It is entirely possible that an attack delay of a single day, for instance, would be small enough that the study instructions would still be fresh in the minds of the participants.

There were several possible confounding variables which were not controlled for in the experiment. The gender of the participants was not recorded and that may have impacted susceptibility to scams. The manner in which the scams were delivered, the prosody, was not controlled for. Each caller was trained to follow each script when delivering a scam, but it is possible that the manner of speech has an impact on the success rate.

To consider the ecological validity of the experiment, we need to define what the "setting" of the experiment is so that we can consider whether or not the results would generalize to a different setting. One aspect of the setting would be the age of the participants, whose average was 19.46 years. This is quite young and it is reasonable to expect that older people would respond differently to a telephone scam. Another aspect of the setting is the fact that it was college-oriented. The participants were all college students, and the pretexts were all related to college. It is reasonable to assume that college student's reactions to scams might be different than those of people with a non-college background. A further constraint on the participants is that they were all students of a single college, the University of California Irvine. Aspects of the culture specific to UCI could affect the results of the study.

7 Telephone Scam Dataset

We present a dataset comprised of recordings and transcripts of all of the attacks made as part of this study, as well as associated metadata. The repository for the dataset can be found at https://gitlab.com/beatscams/study-on-scam-calls.

The main content of the dataset is contained in two directories, the **audio recordings** directory which contains all of the audio recordings, and the **transcripts** directory which contains all of the transcripts of the audio recordings. Each recording in the audio recordings directory is an mp3 (".mp3" suffix) file whose name is the number of the associated study participant. Participants were

anonymously numbered as they entered the study. Each transcript in the transcripts directory is a Microsoft Word file (".docx" suffix) file whose name is also the number of the associated study participant. Each line in the transcript file is annotated with a time stamp which indicates the start time of the line in the corresponding audio recording. All of the files in the audio recordings directory have been anonymized by replacing PII with a 440Hz tone. All of the files in the transcripts have been anonymized as well.

The repository also contains several files containing metadata associated with each phone call. The metadata is contained in three files, the **CallInfo** spreadsheet, the **ScriptInfo** spreadsheet, and the **ScriptText** file. The **CallInfo** spreadsheet contains one record for each phone call and each record contains the following fields: *Call Number*, *Script Number*, *Time Frame*, and *Success?*. The *Call Number* is the number of the associated participant and the *Script Number* is the number of the script used in the call. The *Time Frame* indicates the time between when the participant joined the frame and when he/she was called. There are two possible values for this field: "0" indicates a time frame between 2 and 3 months, and "1" indicates a time frame between 1 and 2 months. The *Success?* field indicates that the attack either failed ("0") or succeeded ("1").

The **ScriptInfo** spreadsheet describes the contents of each of the 27 scripts used. Each row of the spreadsheet contains a record describing one attack script. Each record contains the following fields: *Script Number*, *Pretext Number*, *Elicitation Number*, and *Information Goal Number*. The *Script Number* is the number of the script, and the remaining fields are the numbers of the Pretext, Elicitation, and Information Goal.

The **ScriptText** file is a Microsoft Word (".docx") document containing a list which associates the Pretext, Elicitation, and Information Goal Numbers with their associated text. This information is the same as the information presented in Section 4.1 of this paper.

7.1 Transcript Examples

Although all of the recordings and transcripts are based on a set of only 27 attack scripts, the responses of the victims cannot be fully predicted. Due to the range of victim responses and the variations from the script which they result in, reveal interesting aspects of effectiveness of the scam and the victim's mindset.

The transcript shown in Figure 3 is a case where the Information Goal, a request for license plate number, did not match this participant since he did not have a car. We used the debriefing to identify the subset of participants who were scammed with the Transportation pretext and did not have access to a car.

Another example transcript is shown in Figure 4 in which the victim is asked for his IP address. The victim is clearly suspicious based on the question, "So what do you need my IP address for again?" on the highlighted line. In spite of that, the victim proceeds to provide the information later in the conversation.

Victim: [0:16] Hello?
Attacker: [0:17] Hi, I'm Jane from UCI Transportation Services. Am I speaking to [censored]?
Victim: [0:26] Sure! Yeah.
Attacker: [0:28] Okay—
Victim: [0:29] This is him. [indistinguishable]
Attacker: [0:30] Okay, thanks. We believe that our computer system has been hacked and some of our information on you might have been altered. We need to verify some of your information if you have a minute to do it right now.
Victim: [0:41] What kind of information?
Attacker: [0:43] Can you give me your license plate number so that I can verify your record.
Victim: [0:47] Oh, I-I'm too poor to have a car actually so I don't have one.
Attacker: [0:50] [small laugh]

Fig. 3. Information goal is not applicable

Victim: [0:00] Hello?
Attacker: [0:01] Hi, I'm Jane and I'm with the UCI Office of Information Technologies. Am I speaking to [Censored]?
Victim: [0:08] Yes.
Attacker: [0:09] Ok thanks. It looks like your computer is infected by a virus of some kind and it's attempting to infect other computers whenever you connect to the campus network.
Victim: [0:18] My computer?
Attacker: [0:19] Yes. Yeah so, so then um can you please give me your computer's IP address for verification purposes.
Victim: [0:30] Alright, I have a Mac book, don't know how it'd get a virus.
Attacker: [0:35] Um, say that again?
Victim: [0:37] I have a Mac book.
Attacker: [0:39] Uh yes, um did you need help looking for your IP address?
Victim: [0:44] No, I'll look, just give me a second.
Attacker: [0:47] Okay.
Victim: [1:30] Hello?
Attacker: [1:31] Yes.
Victim: [1:36] Give me one second, let me find my laptop.
Attacker: [1:39] Okay, no problem.
Victim: [2:40] So what do you need my IP address for again?
Attacker: [2:43] Um for verification purposes.
Victim: [2:48] Okay. [3:14] IP address is in settings right?

Fig. 4. Victim is suspicious

8 Related Work

8.1 Social Engineering Studies

Many studies have been performed in which participant susceptibility is evaluated by evaluating their reaction to receiving a phishing email. Phishing email studies have either asked participants to click on a link or to provide sensitive information, but studies vary in other aspects of the content of the email, such as the pretext used. One study involved a professor sending phishing emails to students in his class requesting their username and password [7]. This attack had a high success rate, 41%, likely in part because the email source was a trusted person, the professor of the course. Several phishing email studies use a trusted email source such as a member of the IT department [2] or a friend identified using open source intelligence [11]. The effectiveness of web browser warning messages has been studied by observing the success rate of phishing emails in the presence of a warning message [6].

Researchers have presented results of full penetration tests against industrial partners which involve phishing emails but also other attack vectors including in-person attacks [18] postal mails [28], and phone calls [28,1]. Another attack vector which has been explored is the use of QR codes which represent links to phishing websites [27].

Rather than launch attacks, some studies have instead asked subjects to judge the veracity of websites [1,4] and emails [13] to identify phishing.

Telephone-based scams have been used in several studies. In [28], student-actors were hired to perform comprehensive attacks which included telephone calls, postal letters, and phishing emails. The contents of the phone calls were not revealed, except to list general classes of pretexts and to say that a "range of persuasive techniques" were used. The attacks in this study used a combination of methods, so there is no way to identify the impact of the telephone calls separately from the other approaches used.

The use of telephone-based scams is described in case study involving employees of a bank [1]. Again, detailed contents of the phone calls are not presented. The authors state that attackers "conducted friendly conversations" with participants before asking for internet banking credentials. Examples of elicitations are given including checking privileges and accessibility and checking account integrity.

A recent study on telephone-based scams involved 3000 subjects, 10 different social engineering attack versions, based on 4 attack scripts [25]. The attack scripts were recorded and an autodialer was used to call the participants. The participants were university staff and faculty who were unaware of the study and whose phone numbers were chosen randomly from the university's internal phone directory. Several variables were evaluated including caller gender, accent, and caller ID shown.

8.2 Social Engineering Attack Datasets

Many datasets of phishing emails have been made publicly available for study [21,22,20]. Collectively, these datasets contain well over 100,000 scam emails of various types which have been contributed. To our knowledge, there does not exist a similar dataset containing telephone scams. One likely reason for this is that the laws in many states prevent the recording of telephone calls without prior consent from both parties involved in the call.

9 Conclusion

We present the results of a study on the effectiveness of telephone scams, and we present a dataset containing the recordings and transcripts of these scams. Telephone scams are under-explored as compared to phishing emails and websites, yet the occurrence of telephone scams is on the rise. Our study explores variables which have not been explored in previous work on telephone scams, including the importance of the pretext, the information goal, and the awareness of the victims. Our study also investigates the effectiveness scams involving live attackers rather than pre-recorded messages. To our knowledge, our dataset of telephone scam recordings is the first of its kind to be made publicly available.

10 Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 1813858. This research was also supported by a generous gift from the Herman P. & Sophia Taubman Foundation.

References

1. Aburrous, M.R., Hossain, M.A., Dahal, K.P., Thabtah, F.A.: Experimental case studies for investigating e-banking phishing techniques and attack strategies. Cognitive Computation **2** (2010)
2. Bakhshi, T., Papadaki, M., Furnell, S.: A practical assessment of social engineering vulnerabilities. In: HAISA (2008)
3. Das, A., Baki, S., El Aassal, A., Verma, R., Dunbar, A.: Sok: A comprehensive reexamination of phishing research from the security perspective. IEEE Communications Surveys Tutorials (2019)
4. Dhamija, R., Tygar, J.D., Hearst, M.: Why phishing works. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (2006)
5. Ebbinghaus, H.: Memory: a contribution to experimental psychology. New York: Teachers College, Columbia University (1913)
6. Egelman, S., Cranor, L.F., Hong, J.: You've been warned: An empirical study of the effectiveness of web browser phishing warnings. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (2008)
7. Greening, T.: Ask and ye shall receive: A study in "social engineering". SIGSAC Rev. **14**(2) (Apr 1996)
8. Hadnagy, C., Wilson, P.: Social Engineering: The Art of Human Hacking. Wiley (2010)
9. Hadnagy, C.: Social Engineering The Art of Human Hacking. Wiley Publishing Inc. (2011)
10. Henry L. Roediger, I., Karpicke, J.D.: The power of testing memory: Basic research and implications for educational practice. Perspectives on Psychological Science **1**(3), 181–210 (2006)
11. Jagatic, T.N., Johnson, N.A., Jakobsson, M., Menczer, F.: Social phishing. Commun. ACM **50**(10), 94–100 (2007)
12. Jakobsson, M., Johnson, N., Finn, P.: Why and how to perform fraud experiments. IEEE Security & Privacy **6**(2) (2008)
13. Karakasiliotis, A., Furnell, S.M., Papadaki, M.: Assessing end-user awareness of social engineering and phishing. In: Australian Information Warfare and Security Conference (2006)
14. Kok, K.F.: 2019 u.s. spam & scam report. Truecaller Insights, https://truecaller.blog/2019/04/17/truecaller-insights-2019-us-spam-phone-scam-report/, Last accessed on 2020-02-17 (April 2019)
15. Mitnick, K., Simon, W.: The Art of Intrusion: The Real Stories Behind the Exploits of Hackers, Intruders and Deceivers. Wiley (2009)
16. Mitnick, K.: The Art of Deception. Wiley Publishing Inc. (2003)
17. Olson, E.: When answering the phone exposes you to fraud. New York Times (December 2018 (accessed June 11, 2020)), https://www.nytimes.com/2018/12/07/business/fraud-robocalls-spoofing.html

18. Orgill, G.L., Romney, G.W., Bailey, M.G., Orgill, P.M.: The urgency for effective user privacy-education to counter social engineering attacks on secure computer systems. In: Proceedings of the 5th Conference on Information Technology Education (2004)
19. Ethical principles of psychologists and code of conduct. Tech. rep., American Psychological Association (june 2010)
20. Scamalot. http://scamalot.com, accessed: 2017-10-11
21. Scamdex. www.scamdex.com, accessed: 2017-10-11
22. Scamwarners. scamwarners.com, accessed: 2017-10-11
23. Scheeres, J.: Establishing the Human Firewall: Reducing an Individual's Vulnerability to Social Engineering Attacks. Biblioscholar (2012)
24. Shaban, H.: Nearly half of cellphone calls will be scams by 2019, report says. The Washington Post, https://www.washingtonpost.com/technology/2018/09/19/nearly-half-cellphone-calls-will-be-scams-by-report-says/, Last accessed on 2020-02-17 (September 2018)
25. Tu, H., Doupé, A., Zhao, Z., Ahn, G.J.: Users really do answer telephone scams. In: Proceedings of the 28th USENIX Conference on Security Symposium (2019)
26. Verizon: 2019 data breach investigations report. https://enterprise.verizon.com/resources/reports/dbir/ (2019)
27. Vidas, T., Owusu, E., Wang, S., Zeng, C., Cranor, L.F., Christin, N.: Qrishing: The susceptibility of smartphone users to qr code phishing attacks. In: Adams, A.A., Brenner, M., Smith, M. (eds.) Financial Cryptography and Data Security (2013)
28. Workman, M.: A test of interventions for security threats from social engineering. Inf. Manag. Comput. Security **16**, 463–483 (2008)