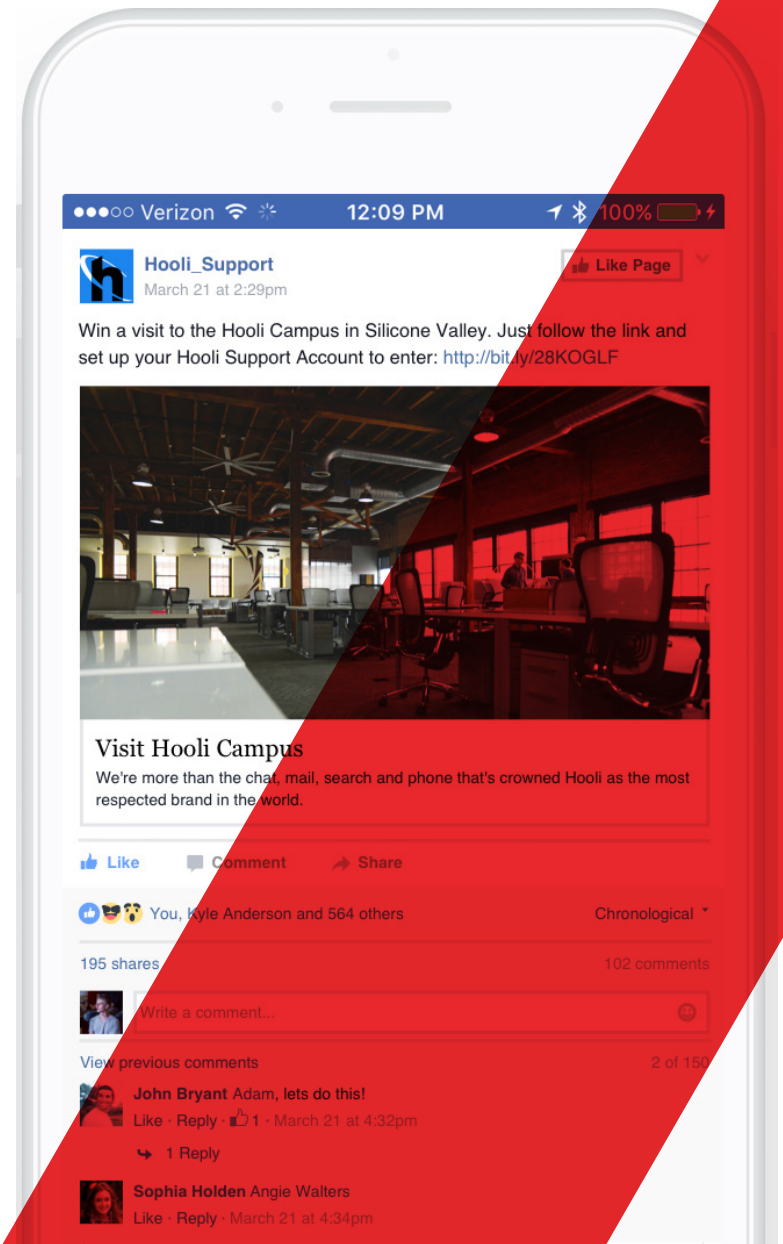
ZEROFOX

RESEARCH

SOCIAL ENGINEERING IN THE SOCIAL MEDIA AGE:

TOP FRAUDULENT ACCOUNT & IMPERSONATOR TACTICS

ZEROFOX RESEARCH



Fraudulent accounts run rampant on social media. But what are they up to? What cyber attacks are they launching? ZeroFOX Research investigates 40,000 fake accounts to find out.

EXECUTIVE SUMMARY

Over the past several years, ZeroFOX has identified and remediated tens of thousands of social engineering profiles and fake accounts impersonating our customers. These accounts spoof a company's brand or executive persona, hijack their logo, and try to mimic the authentic account in order to attack employees and defraud customers. In this white paper, we share some of the trends we have witnessed, delve into the most common and dangerous impersonator tactics, and ultimately try to answer the question: what are all these fraudulent profiles doing?

Fraudulent accounts, also called impersonations, are outrageously easy to create. The easy signup process lowers the barrier to entry for new users, but also makes it easier for attackers to quickly start a campaign. For cybercriminals, conducting their day job has never been more trivial, and, just like they did on email, attackers spoof a brand or its executives to deliver their payload to customers. Today, the social engineer has more tools at their disposal than ever before to create a convincing fraudulent persona and distribute their attack.

In this white paper, we analyze nearly 40,000 identified impersonator profiles to uncover trends over time and the most commonly observed TTPs (Tactics, Techniques, and Procedures) and payloads. We analyze nearly 1,000 of them in depth, often engaging with the cybercriminal to understand their intentions and methodologies. Analysis at this scale is only possible through the use of automation, natural language processing (NLP), image analysis, and machine learning. Without these advanced technical capabilities and a dedicated security research team, organizations are incapable of comprehensively addressing to the threat. Thus, they are blind to the total impact and loss to their business.

The tactics used by these fraudulent accounts are devious and diverse, ranging from traditional social engineering ploys to actually paying money to advertise the scam to reap higher rewards. The networks' attempts to deter this behavior, by providing "verification" to real corporate accounts, has led to a new breed of impersonations and "verification scams." The broader impersonator landscape revealed many tactics meant to lure the user into buying competitor or counterfeit merchandise, providing personal information to unknowing fake recruiters, entering fabricated contests that steal personal information or money, engaging in fraudulent financial scams, and much more.

We've only scratched the surface when it comes to combatting impersonators. While we encountered traditional payloads such as phishing and malware, we found a larger set of threats unique to impersonations on social media. These included unseen scams, fraud, brand abuse, and follower farming. This broader threat landscape extends beyond targeted threats and represents a more systemic issue of risks impacting enterprise security, privacy, and reputation. If allowed to go unresolved, these threats impact the organization's bottom line and damage fundamental customer trust in the organization.

HIGHLIGHTS:

- In the last two years, we've seen the overall number of malicious impersonations increase11x between December 2014 and December 2016.
- Impersonators are most commonly found on Facebook, Twitter, and Google+, though impersonators were also found on Instagram, YouTube, and LinkedIn.
- Verified account impersonators are systemic across the networks and were found on Facebook, Twitter, and Instagram; while also using YouTube to promote their attack.
- Verified account impersonators advertize their payloads through paid, promoted ads.
- Nearly half (48.1%) of all malicious social media impersonators disguise their payload as a fake coupon or giveaway, hijacking the brand to attract promotions seekers.
- Over ⅓ (37.6%) of all malicious social media impersonators send their targets to a phishing page to steal social media account credentials, credit cards, and personal information.
- Impersonators regularly wipe accounts and leave them dormant to avoid detection between attack campaigns.
- Some impersonators create locked accounts to hide their malicious activities, allowing them to take the activity out-of-band through email, direct message, or phone.
- Impersonators crop or modify company images and logos to evade rudimentary image matching and hashing detections.
- Impersonators pivot across networks, posting links to scam impersonations accounts on different channels, making it difficult for the primary network to detect attacks or fraud.

11X INCREASE IN MALICIOUS IMPERSONATIONS
FROM DEC.'14 TO DEC. '16

WATCH OUT FOR **VERIFIED IMPERSONATIONS** ON FACEBOOK, TWITTER AND INSTAGRAM

48% OF MALICIOUS IMPERSONATORS USE **FAKE COUPONS** AS TACTICS

38% OF MALICIOUS IMPERSONATORS DIRECT TO **PHISHING PAGES**

$ SOME IMPERSONATORS ACTUALLY **PAY TO PROMOTE** THEIR ATTACK

TABLE OF CONTENTS

1. Introduction

1.1 Definition of a Social Media Impersonator

2. Impersonations Are On The Rise

3. Methodology

3.1 Text Analysis
3.2 Image Comparison

4. TTP Trends
5. Top Fraudulent Account Tactics

5.1 Verification Phishing
5.2 Malicious Paid Advertisements
5.3 Fraudulent Customer Support
5.4 Dormancy
5.5 E-Commerce Phishing
5.6 Fake Promotions and Fame Farming
5.7 Financial Scams
5.8 Fraudulent Job Recruiters
5.9 Brand Infringement and Counterfeit Merchandise

6. Conclusions

Disclosure

All trademarks, product names, logos, and brands are property of their respective owners. All company, product and service names used in this white paper are for identification purposes only. Use of these names, logos, and brands does not imply endorsement.

All the fraudulent account information and activity shown in this white paper are being or have been reported to the social networks and the businesses.

1. INTRODUCTION

GRANDMA! WHAT ENTICING LINKS YOU HAVE!
ALL THE BETTER TO PHISH YOU WITH, MY DEAR.

If Little Red Riding Hood were on social media, she wouldn't stand a chance.

For someone who couldn't differentiate between her own grandmother and a wild animal in her dear nanny's nightgown, the challenge of parsing the real from the fraudulent on social networks would be nothing short of impossible. The wolves have come out the forest; now they're internet savvy and have more than big ears and big claws and big teeth at their disposal. Once upon a time, they needed to swallow up their victims before slipping into their clothing – now they only need 15 minutes and a coffee shop internet connection.

The art of impersonation, pioneered by the Big Bad Wolf, has come into its own in the age of social media. On Twitter, Facebook, LinkedIn, Instagram, and many more, the barriers to creating a fake persona are trivial. Even non-technical criminals, social engineers, and scammers can create a profile that mimics your favorite retailer, your dentist or, yes, even your grandmother. A simple Google search can give the criminal all the information they need to build a fake account: profile picture, description, interests, and activities. Especially savvy social engineers will simply lift this information from the genuine account of the individual, group, or brand, ensuring their fake account is a near carbon copy.

Impersonators come in many shapes and sizes. Some impersonators imitate individuals, like friends or family, to orchestrate highly targeted social engineering attacks. Some imitate professionals like recruiters, lawyers, or IRS agents to pull off scams. Others imitate big brands and their executives. Each category has its own flavor of associated malicious activity, and each attack leaves behind a trail of digital breadcrumbs that can be followed in order to accurately predict how, when, and where future attacks will be carried out.

This white paper focuses on brand impersonations, in which the cybercriminal imitates a well known company or organization to gain credibility with their victims. In this study, we investigate nearly 40,000 brand impersonators from the past several years across 6 different social networks: Twitter, Facebook, Instagram, LinkedIn, Google+, and YouTube. We use a suite of machine learning, natural language processing, image recognition, and other data science techniques in order to measure the relative similarity between an impersonating profile and the genuine account. These approaches are the foundation of the ZeroFOX Platform, which automatically ingests social media data and analyzes if for impersonations.

These curated datasets allow us to analyze trends over time, among different social networks and more. We take a deep dive on ten commonly observed attacker TTPs to uncover exactly what the impersonations are up to, how they deliver their payload, and how much damage is dealt to their targeted enterprises. We conclude with our recommendations after years of ingesting and analyzing this data.

1.1 DEFINITION OF A SOCIAL MEDIA IMPERSONATOR

Broadly speaking, an impersonation account is anything that pretends to be someone or thing else. The motivations for creating such an account are diverse, ranging from innocuous fan accounts and parodies to outright malicious impersonations slandering a brand

or hijacking their victim's authority to launch scams and phishing attacks. At the end of the day, an impersonator can be seen as an integral step in any social engineering campaign on social media, regardless of the payload or specific tactics. Social engineering is the fundamental, overarching tactic of all impersonations. Specific TTPs are discussed at length in sections 4 and 5.

A note on terminology and usage in this white paper: we use both "impersonator" and "impersonation," the former generally referring to the human perpetrator behind the attack and latter referring to the actual profile. For the sake of style, we also use several variations somewhat interchangeably, such as "fraudulent account/profile," "malicious account/profile," and "fake account/profile."

All social networks provide a definition of an impersonation in their Terms of Service (ToS), highlighting what is permissible and what is not. Anything in violation of the ToS is liable to be warned, blocked, or banned. On Twitter, for instance, parody, fan, and news accounts are permitted only when the account explicitly states their purpose in the user bio; it must literally contain the text "parody," "fan," or "news."

However, there are nuanced variations for the definition of an impersonation among the different social networks. Each has their own unique profile configurability, cultural norms, and means of engagement and expression. Below is how the 6 social networks covered in this report define "impersonations" as pertaining to their ToS:

Social Network	Definition of Impersonation	How to Report
Twitter	• Profiles which portray another person in a confusing or deceptive manner • Exceptions: • Profiles where only commonality is same name • Profiles which clearly state "not affiliated with or connected to any similarly-named individuals". • Parody, commentary, or fan accounts.	https://support.twitter.com/articles/20170142
LinkedIn	• Profiles which post inaccurate content (including employment, qualifications, and affiliations) • Profiles with an image that is not the owner's likeness or head-shot photo • Profiles with pseudonyms or for other people	https://www.linkedin.com/help/linkedin/answer/30200
Facebook	• Profiles using misleading or inaccurate information to artificially collect likes, followers, or shares • Exceptions: Facebook Pages are allowed for pets, organizations, movies, video game characters, and other purposes	https://www.facebook.com/help/174210519303259help/174210519303259
Instagram	• Accounts for anyone other than self. • Accounts with false, inaccurate, stale, or incomplete information • Accounts which impersonate people or entities	https://help.instagram.com/370054663112398
Google+	• Pretending to be someone else • Pretending to represent an organization • Deceiving users into thinking that a page officially represents a person, business, or organization • Using an official logo as your profile picture. • Exceptions: Fan commentaries and parodies if clearly named	https://support.google.com/plus/troubleshooter/1715140
YouTube	• Copying a channel's profile, background, or text, and writing comments to make it look like somebody else's • Using another individual's real name, image, or other personal information to deceive people • Exceptions: Impersonation does not include channels or videos pretending to represent a business.	(Channel) https://www.youtube.com/reportabuse (Individual) https://support.google.com/youtube/contact/impersonation

TABLE 1.
Social networks' definitions of an impersonation.

2. IMPERSONATIONS ARE ON THE RISE

Impersonations can dynamically be created, banned, re-created, re-banned, rinsed, and repeated. Accounts often remain dormant, waiting months or years before coordinating a phishing campaign. Bad actors might create accounts or activate dormant old ones, broadcast attacks, steal credentials and money, then abscond, all in a matter of minutes or hours. One month later, the impersonator can do it all over again with a fresh set of accounts. We've noticed that these attacks tend to coincide with viral trends, temporary promotional offers, or other popular real-life events.

Using the ZeroFOX Platform, we retroactively analyzed alerts going back to October 2014. The platform identified 39,915 impersonations across 6 major social networks (Figure 1). To identify these impersonations, we used a methodology that fundamentally relied on text (name, bio) and image analysis. Text analysis involved text matching, homoglyph identification and natural language processing. Image analysis involved comparison to base images that may have been cropped, superimposed with other content, rotated, resized, rescaled, or underwent other affine transformations. Details of this analysis is covered in section 3.

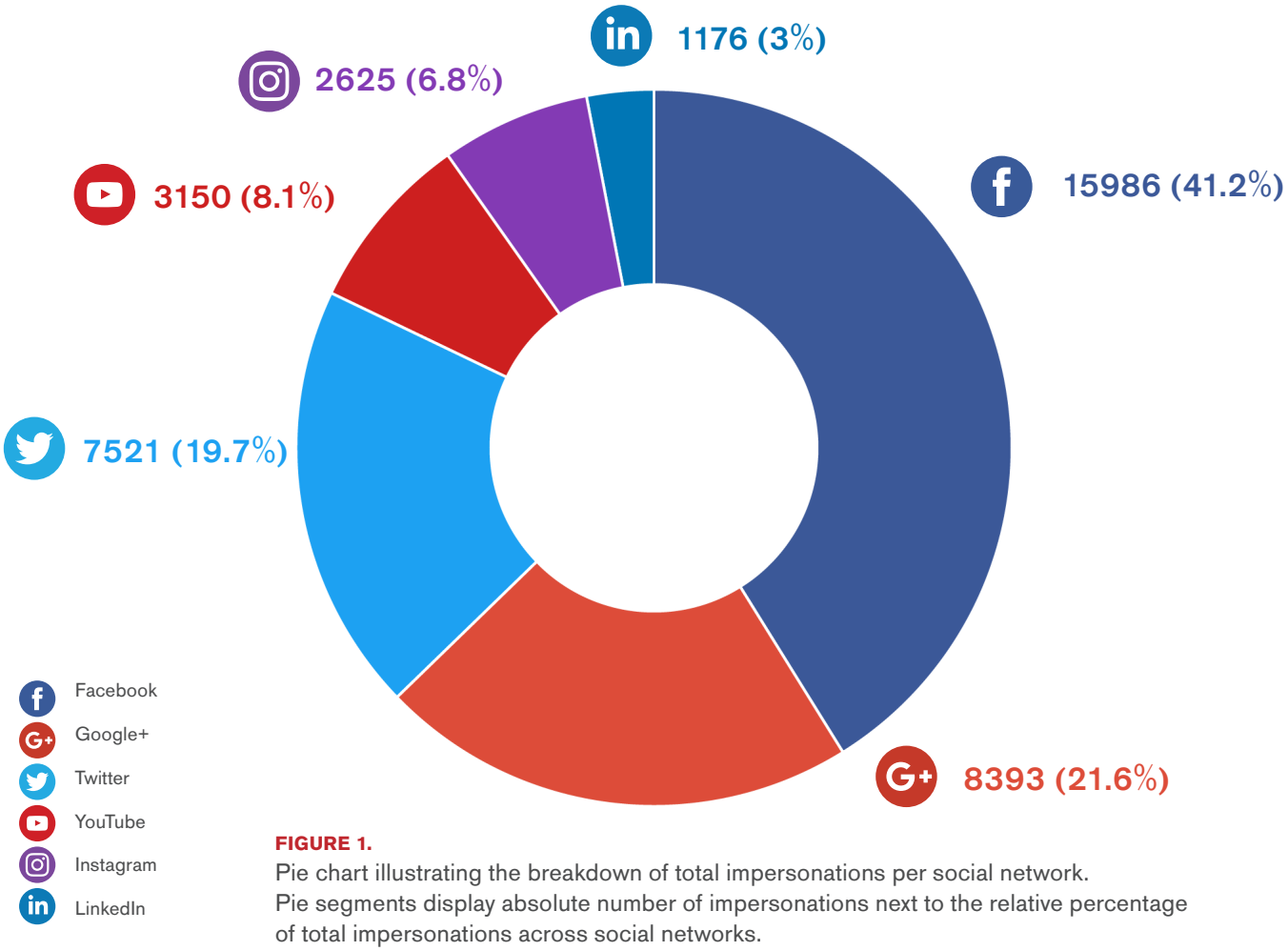


FIGURE 1.
Pie chart illustrating the breakdown of total impersonations per social network.
Pie segments display absolute number of impersonations next to the relative percentage of total impersonations across social networks.

To ascertain the extent to which the impersonator problem was growing or shrinking over time, we bucketed all impersonations according to the time they were detected by the ZeroFOX Platform per social network. Besides a few outlying months, we observed an upward trend in the number of impersonations over time (Figure 2). It is important to note that the numbers below represent only a small sliver of all impersonations in the wild.

Social networks are first-come, first-serve when it comes to username ownership. Late comers often get stuck with usernames that are less desirable or that incorporate numbers appended at the end to make them unique. This is especially the case for companies with names that are also commonly used words (e.g. "@Apple"), people with common names (e.g. "@JohnSmith") and famous people (e.g. "@DonaldTrump"). In fact, even the authentic Twitter account of president Trump himself is famously @realDonaldTrump rather than @donaldTrump. The impersonators have caught onto this behavior as well. From the data above, 1,062 impersonators incorporated credibility-building words like "official," "authentic," "real," "authorized," "actual," and "legitimate" within their names, screen names, and descriptions.

Another common theme observed involves impersonators who target military members and veterans. From the data above, 1,047 impersonators incorporated military-associated words like "military," "navy," "army," "air force," "marines," and "nato" within their names, screen names, and descriptions. Impersonators try to penetrate the social media circles of military members to try to steal personal and sensitive information.

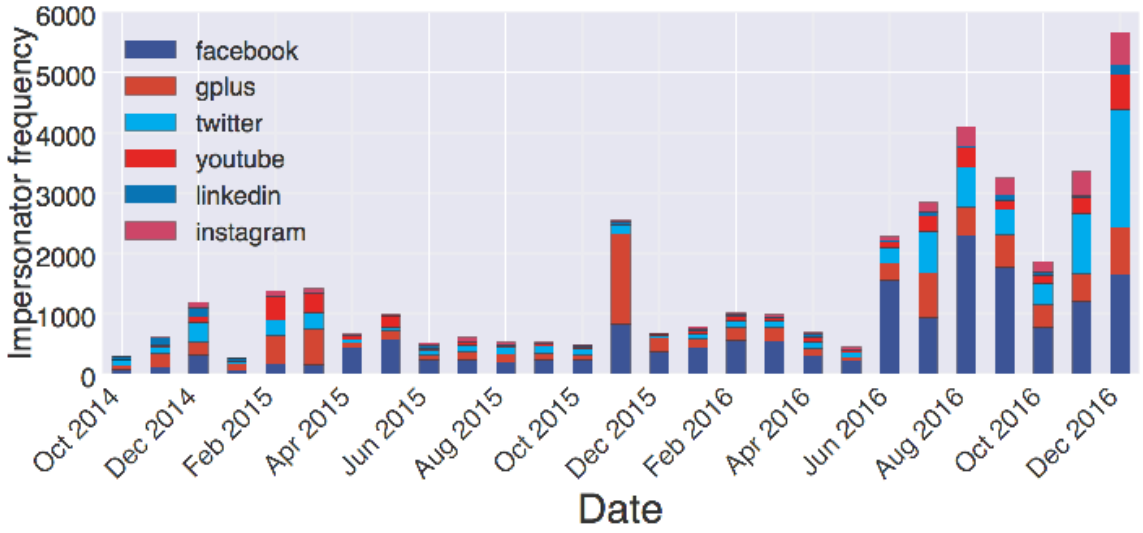


FIGURE 2.
Monthly impersonations are slowly on the rise between October 2014 and December 2016.

3. METHODOLOGY

Impersonations aren't always easy to find. Faithful mimicry and sleeper behavior can make them blend in with the morass of unremarkable profiles that populate the world of social media. Social media users mistake them all the time; if they didn't, cybercriminals wouldn't keep creating them at such a massive scale. Automating the detection of these profiles is harder still, in part because of the large variety of behaviors that impersonators manifest, discussed at length in Section 5. Although these behaviors change, a successful impersonation typically relies on two things: a name and an image.

Other elements, such as biographical information and posted content, are frequently manipulated as well, but without a convincing name and image, an impersonator's chance of success is greatly diminished. These elements are a user's first and often only point of contact with other profiles. For instance, a friend, connection, or follow request generally only contains the name and profile picture of the account, and users rarely take the time to do a more exhaustive investigation.

To detect these patterns of manipulation, the ZeroFOX Platform analyzes two primary data types: text and images.

3.1 TEXT ANALYSIS

The single most important string of text involved in an impersonation attempt is a profile's name. For instance, if someone wants to impersonate Acme Technology Corporation, they will need to create a profile with the same name or some variant thereof (such as Acme Technology, Acme Tech, Acme Corp, ATC, etc.).

There are a number of methods that can be used to compare strings of text, but the concept underlying most methods is called "edit distance." Loosely speaking, this is the minimum number of operations (edits) that are required to transform string A into string B. By this metric, the more operations required for the transformation, the more different the two strings are. The definition of an "operation" varies between algorithms, but the concept is the same.

This metric works well for certain use cases (such as comparing Acme Technology Corporation to Acme Technology Corp), but not for others (such as comparing Acme Technology Corporation to ATC). The edit distance in the latter case is large, despite the symbolic similarity of the two strings. One obvious solution is to include predictable variations on the string in the initial analysis. However, unforeseen variations can still lead to false negatives, and the ZeroFOX Platform uses several other natural language processing techniques to augment its results.

3.2 IMAGE COMPARISON

The ZeroFOX image comparison system uses a variety of techniques to determine if two images are similar. They can be broken down into two broad categories of image analysis methods: similarity hashing and feature detection.

At a high level, similarity hashing enables the transformation of data from one domain (like an image) into another (usually lighter weight) domain while preserving some sort of similarity function. This makes it possible to compute similarity in the lighter weight domain, which can dramatically reduce the cost of computation. This can then be used to convert image pixel data into short text strings that preserve certain measures of similarity relevant to a human's interpretation of an image.

Although efficient, similarity hashing has poor tolerance for significant transformations like large crops, rotations, and other affine transformations. One solution to this is to use feature detection, which can be used to sift through the noise of pixel data in order to find the signal. Instead of working with individual pixels, feature detection works with corners, lines, and other structures that usually represent meaningful parts of objects in the real world.

These strategies make image comparison efficient and tolerant of a wide range of image adjustments, including cropping, scaling, overlays, color changes, rotation, and other affine transformations. This tolerance is particularly useful for detecting impersonators in the wild, as they frequently copy then transform images derived from legitimate entities on social media.

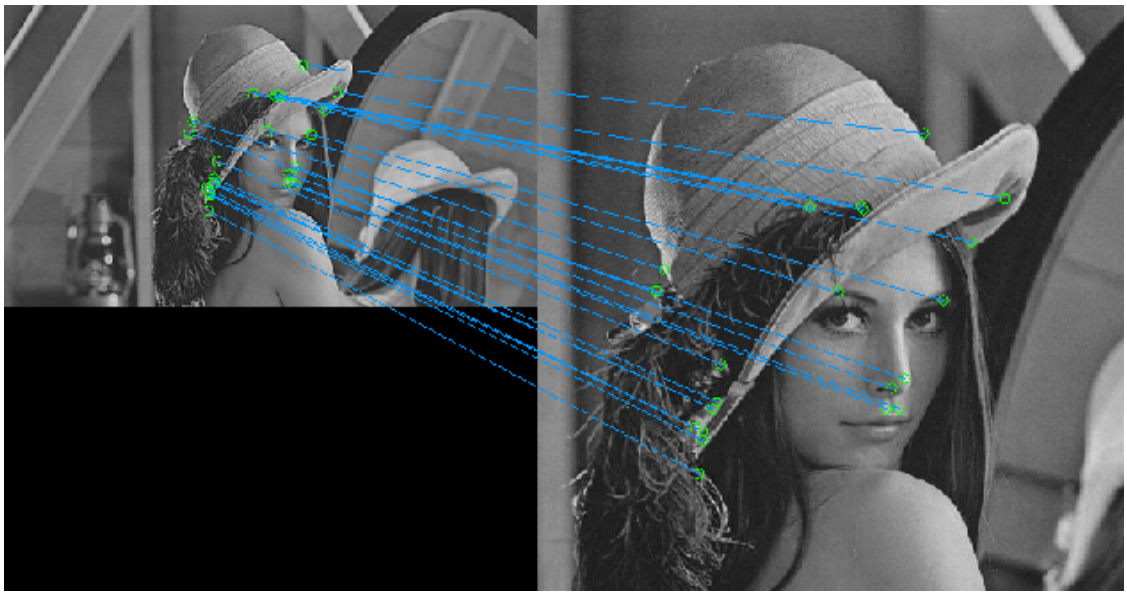


FIGURE 3.
Matched features (circular green endpoints of blue lines) between an image and a cropped, rescaled variant.

4. TTP TRENDS

The next question, and most important, we sought to answer is - what are all those accounts doing? What tactics are they using to lure in their victims and what payloads are they delivering? What is their end goal? To answer those questions, we pulled a sample of impersonations, analyzed their anatomy and created a list of most common and dangerous TTPs (Tactics, Techniques, and Procedures).

Impersonators weaponize their tactics in a variety of ways, but all have a common goal of socially engineering the user into engaging with the impersonator because of an offer, discount, contest, quick return on money, a job, and more. To identify the distribution of these tactics, we took a snapshot of a two week period in early 2017 and conducted deeper analysis on the sample. The accounts were then categorized by tactic and payload to understand the attack flow and distribution of the threats (Figure 4, 5).

Note: "Brand hijacking" is any attempt to capture viewers and then prompt them to explore the products or services the impersonator is selling, often spam sites, unrelated products, or counterfeits goods.

Once a user engages with the impersonator account, a payload is delivered. This delivery mechanism can pilfer personal or credit card information from the victim, infect their device with malware, execute a scam, or simply redirect them to counterfeit merchandise.

It's important to note that in additional to the direct payload, these attacks represent a reputational risk to organizations that tarnishes the brand and erodes customer trust. If customers are subsequently phished, infected by malware, or scammed out of money, they are likely to lose faith in the organization. As such, the lifetime value of a customer is valuable metric when assessing the damage of an impersonation campaign against an organization.

IMPERSONATOR THREAT TACTICS

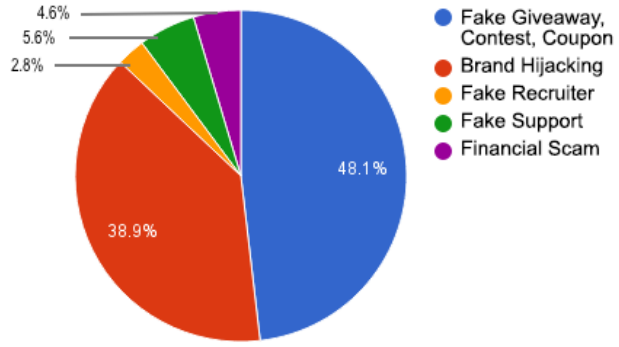


FIGURE 4.
Distribution of impersonator threats by tactic.

IMPERSONATOR THREAT PAYLOADS

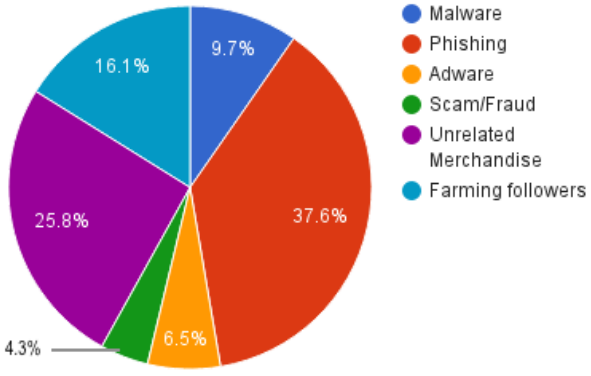


FIGURE 5.
Distribution of impersonator threats by payload.

5. TOP FRAUDULENT ACCOUNT TACTICS

Of the tactics mentioned in the previous section, we decided to thoroughly investigate the most common and damaging tactics. To do so, we created honeypot accounts, engaged with the impersonators, and observed the social engineering attack within a sandboxed environment. This allowed us to reveal the anatomy of the attacks, identify commonalities and differences, and more clearly understand motives.

5.1 VERIFICATION PHISHING

On nearly every major social network, the genuine accounts of popular brands and celebrities will almost always be adorned with a verification badge, usually in the form of a blue checkmark adjacent to the profile picture. The social networks use verification badges to help their users differentiate between genuine accounts and fake accounts attempting to exploit the real account's popularity. Brands and social media influencers understandably scramble to earn the coveted checkmark to boost their authority and encourage user engagement. To satisfy the demand for verified accounts from both users and the popular accounts that they follow, the social networks have established ways to apply for verification. They review the account on a case-by-case basis and decide whether or not to bestow the account with a badge.

As we have learned time and time again, where there are internet users scrambling for popularity or prestige, there will be cybercriminals to take advantage of them. The jockeying for verification, introduced as a means to reduce fraudulent activity, has proven to be a readily exploitable method of cyber attack. Scammers and phishing accounts imitate the networks themselves, claiming to be the authentic verification help account, directing would-be-verified-users to all sorts of malicious payloads. We report some example impersonators pretending to be official support accounts, ostensibly to build credibility for delivering a phishing link (Figure 6).

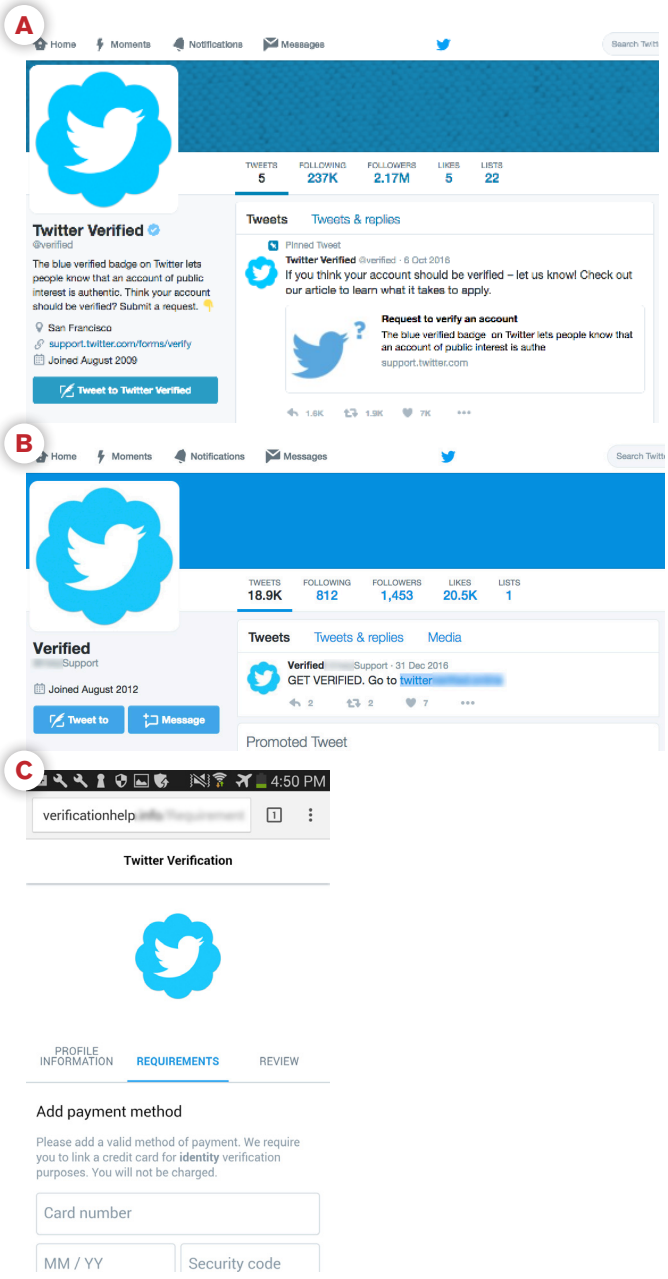


FIGURE 6: The verification phishing scam. A. The real account B. The impersonator account C. Twitter credential and credit card phishing.

The authentic Twitter user @verified (Figure 6A) posts a URL with information about how users can get their accounts verified. Its impersonator (Figure 6B) uses the same default image, similar background, and a deceptive @HelpSupport username with a homoglyph uppercase "i" replacing the lowercase "l." The account laid dormant for 4 years before starting to phish, but now actively engages by posting and "liking" often, following other users, and following back similar accounts spreading malicious URLs. The resolved phishing page can be seen in Figure 6C.

Twitter was not the only social network where this type of behavior was observed. Verification phishing scams proliferated across most popular social networks offering the verified badge feature, including Facebook, Instagram, and YouTube (Figure 7, 9, 10).

On Facebook, verification scams target both "pages" and "profiles." Pages are used by businesses and organizations for marketing purposes while profiles are intended for individuals. The Facebook page example in Figure 7 has the recognizable blue verified badge within its default page picture and background banner picture. The actual verified accounts have the blue badge adjacent to the username. The name of the account is "Get Verified on Your Account," and the banner advertises verification services. The post on the page instructs the victim to download a linked text file with Javascript code.

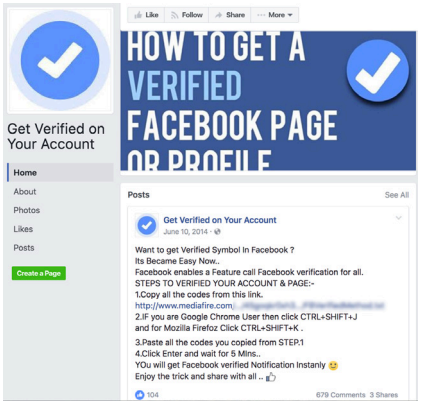


FIGURE 7.
A Facebook Page verification impersonator.

Next, the perpetrator instructs the victim to open the developer console in Firefox while on their Facebook page (Figure 8A). Fortunately, Facebook actually warns of the dangers of using this console to run Javascript. If the user ignores the warning, they are told to paste the Javascript into the console (Figure 8B). The Javascipt begins by capturing the user's Facebook session cookie, a very common technique used for account hijacking (Figure 8C). Additionally, the code uses the hijacked session to "like" multiple pages and lists those accounts later in the script (Figures 8D).

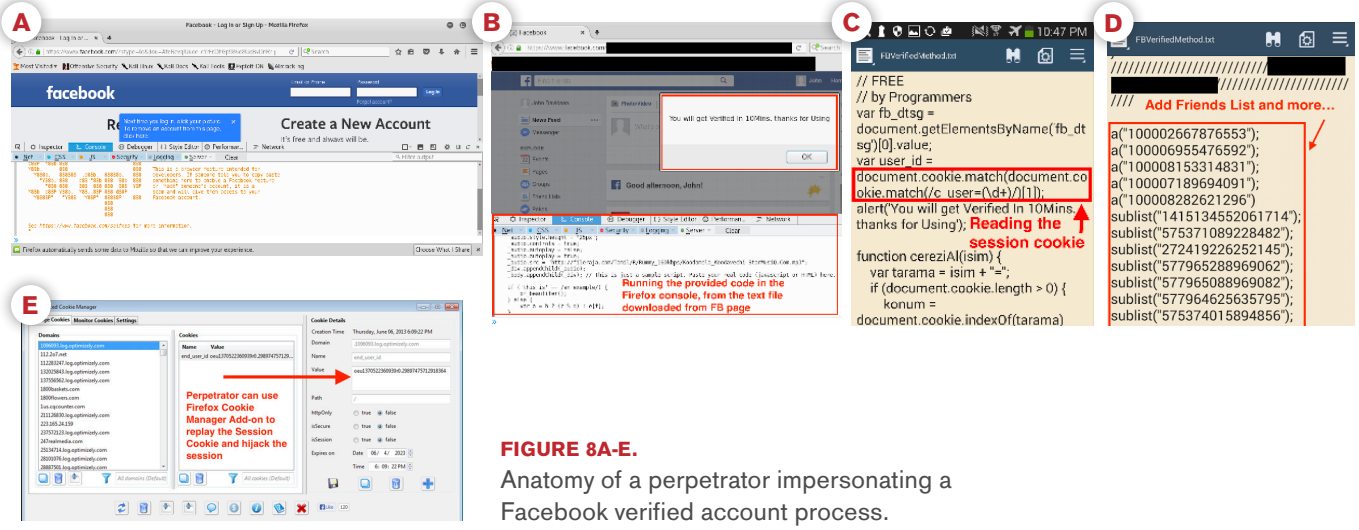


FIGURE 8A-E.
Anatomy of a perpetrator impersonating a Facebook verified account process.

There are a variety of ways of to replay the session cookie by hijacking sessions, including the simple add-on for Firefox, "Advanced Cookie Manager" (Figure 8E). While the session cookie commonly expires when the user logs out, most users do not routinely log out of their Facebook account, therefore the session cookie persists. During this period, the attacker has a window of opportunity to hijack the session and carry out the exploit. The attacker now has access to the account, and they can change settings, post malicious content, attack followers of the account via direct message, follow additional perpetrators, post offensive content, and more.

In Figure 9, an Instagram cybercriminal advertises a method to "Get Verified" by clicking the link in their description. Similar to the Facebook example, this profile also displays a verified badge in the default profile picture. It also has "verifiedbadge" within the username. The link then takes the verification seeker to a fake Instagram login page where their credentials are phished. These accounts attempt to capture as many victims as quickly as possible before getting banned.

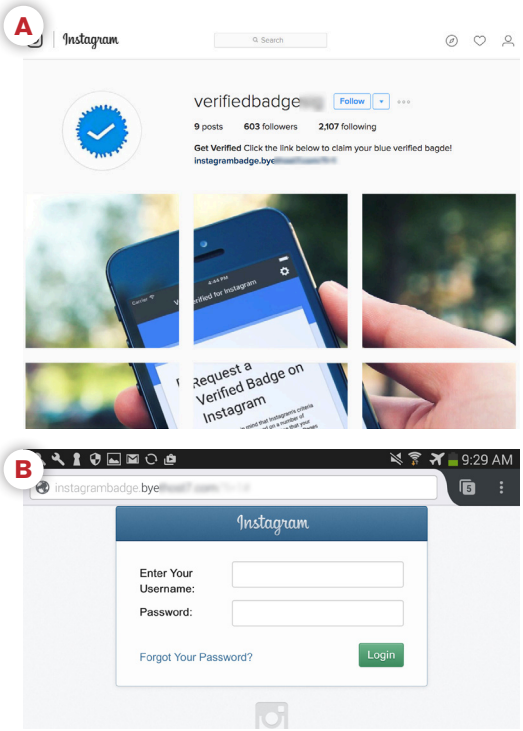


FIGURE 9: A. An Instagram verification impersonator and B. login phishing

YouTube videos are a common method to advertise verification phishing accounts. In Figure10, a video titled "How To Get Verified on Twitter!" shows a screenshare of a Windows desktop that opens Notepad and types out a list of steps to follow in order to get verified, including engaging with a specific Twitter user. By advertising on YouTube instead of Twitter, the attack can fly under the radar and avoid violating Twitter's Term of Service in public view.

The perpetrators of verification phishing inevitably target accounts with a sizeable following – not enough to be already verified, but enough to warrant the user to seek out the verification application. Such accounts are often medium sized business, large businesses late to join social media, social media influencers, or other rising celebrities. This often proves to be the sweet spot for account hijackings, another reason stealing account credentials can be so lucrative. As such, the entire attack, from footprinting to attack to payload to damage can occur entirely on the social network.

These examples show how verification has not solved the problem of impersonating social media accounts.

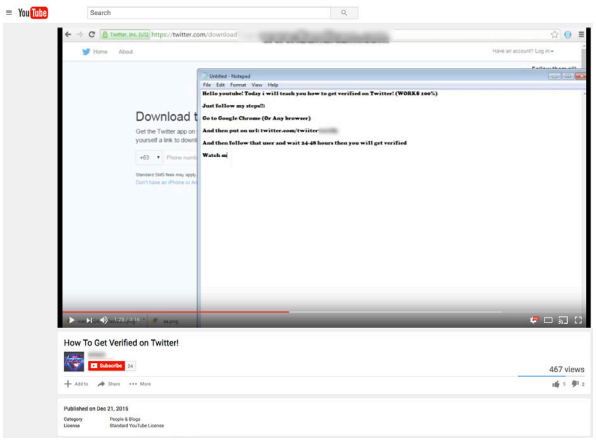


FIGURE 10: YouTube video that instructs viewers to a Twitter verification impersonator.

5.2 MALICIOUS PAID ADVERTISEMENTS

MOST COMMON TARGETS: retail, technology, consumer goods

Another way for a cybercriminal to ensure their attack is viewed by a huge number of potential victims is to use paid promotion, which broadcasts the phishing link to wider audiences. Promotion is a service offered to social media marketers to display an ad to users beyond just their followers, and it is the basis for revenue for most social networks. Scammers using this method take a huge risk; the social networks review ads before they are approved and the scammer may have their entire account banned if the network deems their purposes to be nefarious. Scammers must invest extra time and energy ensuring their promoted content will dupe the network's filters.

Figure 11A shows a verification phisher promoting their tweet. For those cybercriminals that slip through the cracks and have their malicious activity approved by the social networks, the payoff can be huge. In Figure11B, a website offering counterfeit sunglasses at a too-good-to-be-true discount is promoted on Instagram. The website sells fake merchandise despite adopting the real brand's logo. The more scammers are willing to pay, the more the networks will distribute the post.

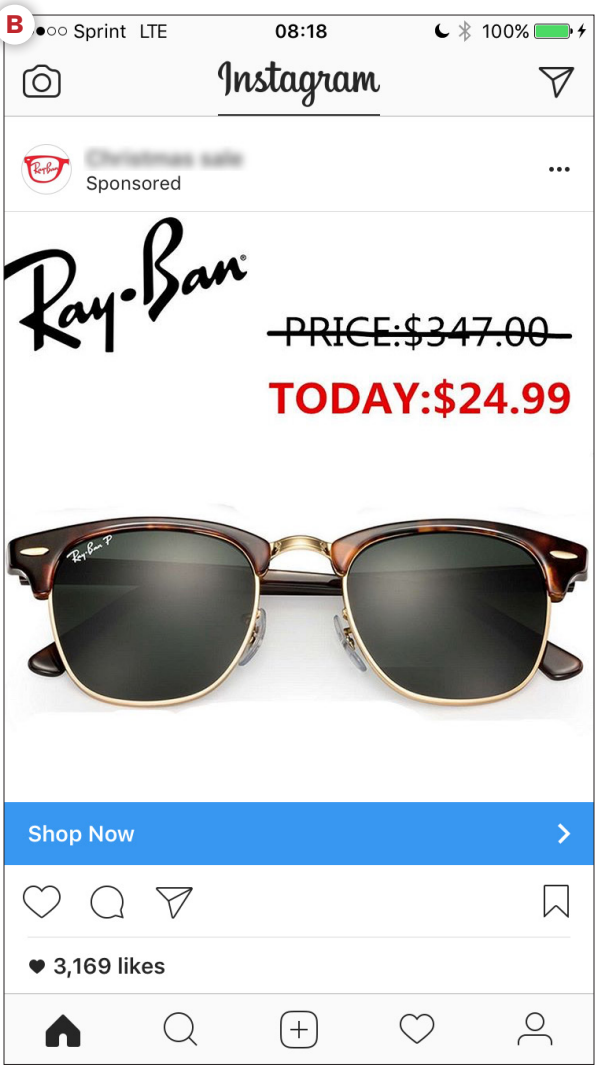


FIGURE 11: Paid advertisement impersonations. A. Twitter paid ads are "Promoted," and this one impersonates the authentic @verified account to broadcast a phishing link, like in Figure 6. B. Instagram paid ads are "Sponsored," and this one impersonates the brand to broadcast a retail scam.

5.3 FRAUDULENT CUSTOMER SUPPORT

The proliferation of social media has revolutionized modern customer support. Gone are the days of waiting on hold over the phone. From product complaints to account security issues to undelivered packages, customers publicly express their discontent by directly mentioning the company's social media account. Companies have responded by forming rapid response teams whose dedicated purpose is to address such customer inquiries. But they aren't the only ones. Impersonators have latched on to the inherent trust that customers place in these support accounts (Figure 12).

Other than the blue verified checkmark, the differences between a bank's real account in Figure 12A and its two impersonators (Figures 12B, C) are negligible to the human eye. Customers with bank accounts self-identify themselves by mentioning the authentic bank's account alongside a personal question, and the impersonator then uses this publicly posted information as a one-stop-shop for victim acquisition (Figure 13A, B, C).

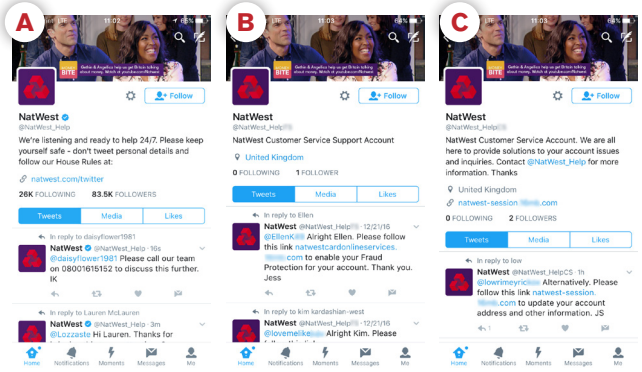


FIGURE 12: Customer support phishing impersonators target a major bank. A. The authentic bank customer service Twitter account. B. An impersonating account that replies to individual customer complaints by delivering phishing links. C. A similar account with the same username pattern "NatWest_HelpXX" that has also contains phishing link within its description.

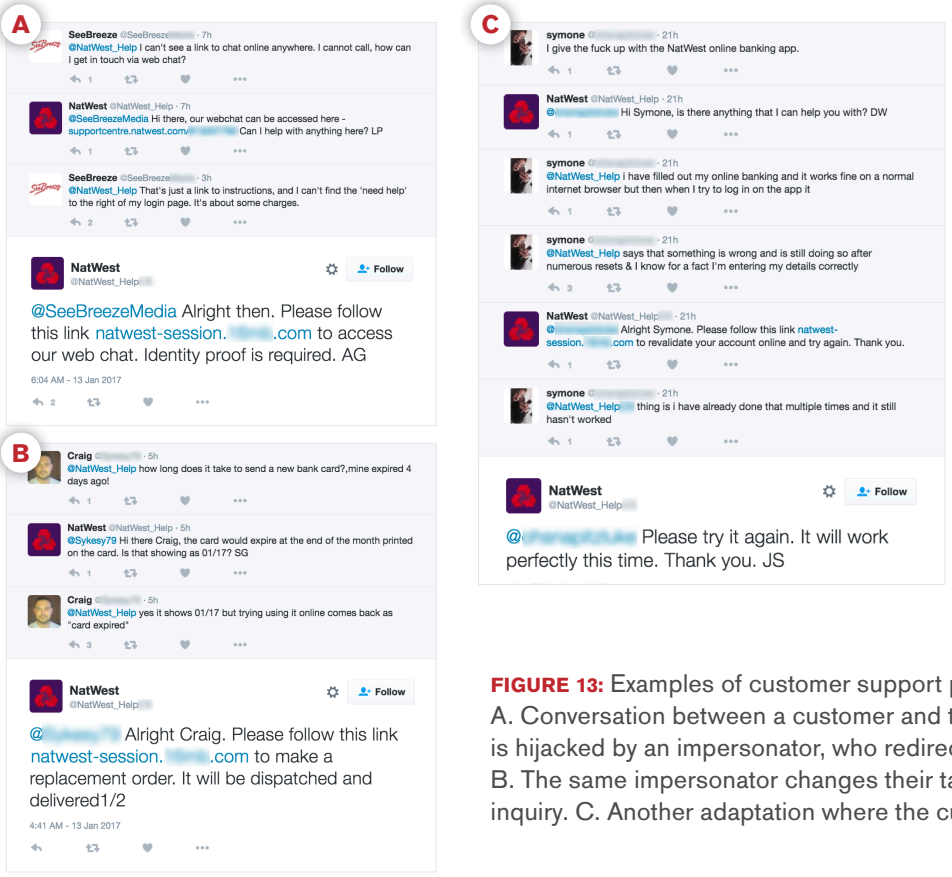


FIGURE 13: Examples of customer support phishing impersonator interactions. A. Conversation between a customer and their bank's official support account is hijacked by an impersonator, who redirects the customer to a phishing link. B. The same impersonator changes their tactics to adapt to the customer inquiry. C. Another adaptation where the customer admits click-through.

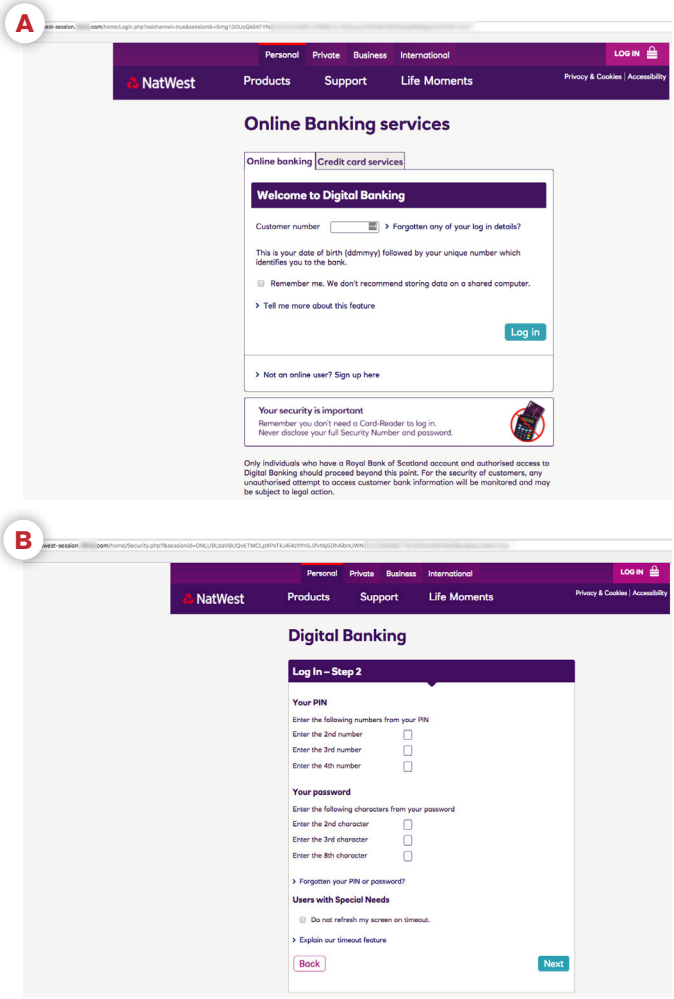


FIGURE 14: Phishing payload click-through redirects to a credential harvesting phishing website. A. The landing page. B. After login, PIN information and other personal banking information is queried by the input forms.

The link redirect destination closely mimics the bank's actual login page (Figure 14A, B). The impersonation extends from the network to the phishing page.

Phishing bank account or credit card information (also known as "carding") has transcended into social media through these bank impersonation accounts, despite many banks having a verified account. To the average user these accounts all look identical with the exception of the blue checkmark.

MOST COMMON TARGETS: media and entertainment, financial services, technology, travel and hospitality, consumer goods, retail

5.4 DORMANCY

Impersonators use a variety of techniques to avoid detection by the social networks. One of the most popular is creating an account but letting it sit dormant for significant periods of time before springing into action. They can return to dormancy just a quickly. The reasons for this are:

1. Older accounts are more credible
For a user doing a cursory check on a potentially malicious impersonation account, the account's age is a good indicator of legitimacy. Users expect the authentic account of well-known brands (Figure 15) to have been around for quite a while. For a scammer, this means "aging" the account makes it more authentic. During this aging process, the account must remain undetected, and thus the perpetrators leave the account dormant and blank.

2. Dormant accounts are more likely to fly under the radar
Cybercriminals regularly wipe the account to avoid detection. Wiping the account helps the attacker cover their tracks between attack campaigns.

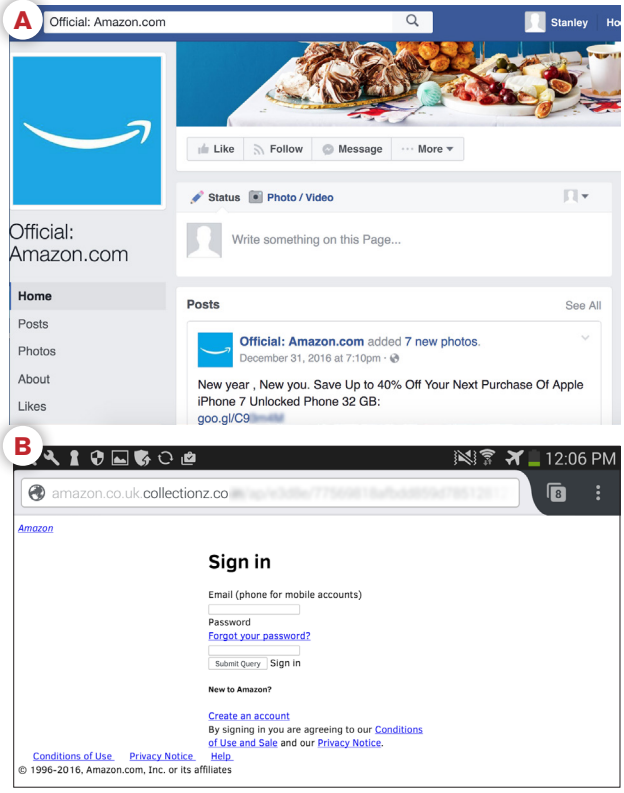
3. The account may have been recently sold
Accounts are bought and sold regularly. Cybercriminals might buy a dormant account with a lucrative handle, perhaps one very similar to that of the brand they intend to impersonate. Once the account has changed hands, it may spring into life and start spreading its attack campaign.

FIGURE 15: A sleeper impersonator lies dormant after conducting an aggressive phishing campaign to steal credentials. A. The perpetrating Twitter account as of January 12, 2017. B. The same account as of December 19, 2016 looks completely different, suggesting the perpetrator is changing its displayed content in an effort to avoid being reported and subsequently banned. C. The phishing website redirected to by the URL in the tweets from B.

5.5 E-COMMERCE PHISHING

Commercial retail transactions have gradually moved from the physical to the digital realm, and not surprisingly, fraudsters have followed suit. As social network platforms become more intertwined with our daily habits, retail companies are hard at work trying to inundate our social feeds with enticing advertisements and hard-to-resist offers for their shiniest products. As this new normal sinks in for the online consumer, malicious actors have stepped in to exploit it.

The Facebook page in Figure 16A contains of a seemingly innocuous shortened link that hides the redirect destination of the ultimate illegitimate phishing website. When the link resolves, it contains the expected login top-level domain and familiar-looking login page (Figure16B). However, there's no website certificate. The perpetrator ultimately seeks to steal login credentials and financial information.



Retailers are also targets for fraud and scams. The fake gift card, coupon, and promotion impersonators in Figure17 can be used to phish information from coupon-clippers, provide discounts codes that bait-and-switch to malware, and even generate usable gift card numbers from fake mobile apps.

Retail scammers distribute links that redirect the user to a page to enter a contest, thus harvesting name, address, email, birthdate, and other PII (Figure17A). Despite following registration instructions, entry confirmation is never received. Instead, the page leads to multiple pop-ups with malware and eventually redirects to a website designed for data extraction (Figure17B). Other impersonator accounts simply request an email address in conjunction with a repost (Figure17C). Once entered, the email is sold to spam lists. The user is typically encouraged to follow steps for providing contact information in exchange for an unfulfilled card (Figure17D). Additionally, a perpetrator can check these against exposed account lists such as *haveibeenpwned*.

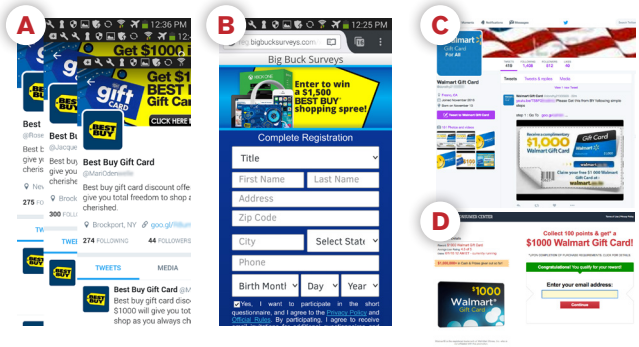


FIGURE 17: A. A fake gift card campaign on Twitter targets shoppers. B. The advertised URL pilfers PII from victims, promotes adware, and posts pornographic content. C. A similar campaign targeting shoppers. D. The advertised URL solicits an email address in order to enter a gift card giveaway.

FIGURE 16: An e-commerce phishing impersonator. A. The perpetrating Facebook page misleads online shoppers by claiming to be "Official" and displaying the official company logo, but its most recent post contains a phishing link. B. The phishing webpage resembles the authentic login webpage.

Perpetrators also stand up survey webpages to phish information from consumers. These are then promoted through social networking accounts impersonating the brand. To circumvent detection, perpetrators will pivot across networks. In the example of Figure 18A, the link first takes the user to Tumblr first, then to the survey infested with adware and malware. Since social networks today typically only analyze direct malware and phishing links, providing a link on another social network circumvents detection (Figure 18B, C).

MOST COMMON TARGETS: retail, consumer goods, media & entertainment

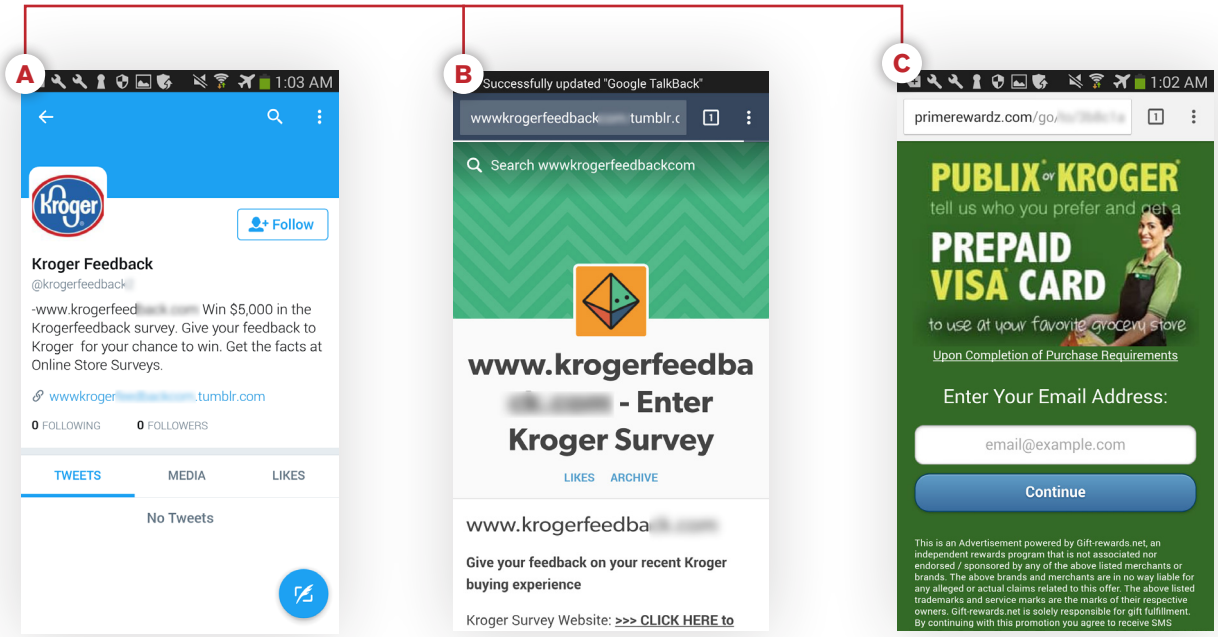


FIGURE 18: A.This feedback account provides a link that takes the user from Twitter to Tumblr. B. This Tumblr account provides a link to the survey. C. The link first requests the user's email address and then takes the user to a second page for detailed information including birth date. Upon completion, it takes the user to more adware and malware.

5.6 FAKE PROMOTIONS AND FAME FARMING

Some impersonators garner followers and likes by promising vouchers, gift certificates, and other fake giveaway promotions. In most instances, they request a @mention or repost of the contest along with an email address or photo. Obtaining followers allows them to inflate their own prominence on social media, a tactic called fame farming.

The value of inflating followers count is threefold:

- **More followers means a more credible account:** there is a feedback loop between offering fake promotions for likes and having a strong following. A strong following increases an account's credibility, and more credibility means more follows. Accounts build this following until they are ready to do something else, almost always something malicious, with the account.
- **Followers now are victims later:** by building a following over time without conducting any overtly malicious activity, the followers are less likely to suspect malicious activity once the account does spring into action. The cybercriminal may begin direct messaging its followers or posting more overtly malicious content, such as phishing links disguised as fake offers or malware in the form of fake contests.
- **Robust accounts can be sold:** scammers and cybercriminals pay a hefty price for accounts with a pre-built following. Building and selling accounts, called "account flipping," is a lucrative tradecraft in the social media cybercrime economy. The recipient of the account may use it for any number of purposes, many of which are covered in this paper.

These impersonator accounts should be a concern to organizations because they hijack brand and trademarked logos, and they target the organization's customers. As with the other examples in this paper, there is a profound element of brand reputation inherent to these cyberattacks that is not part of the traditional cost analysis of an incident. These attacks occur in broad daylight, where marketing teams fight for their brand's share of voice. They target a brand's customer base, especially those that are particularly engaged. Organizations ought to assess these attacks in term of the value of a single customer, not just the direct financial fallout of the attack.

MOST COMMON TARGETS: travel and hospitality, retail, CPG, food and beverage

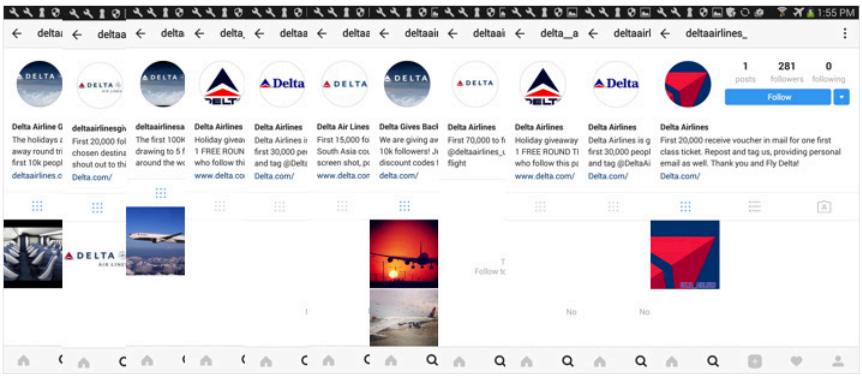


FIGURE 19: A sampling of more than 40 impersonator fame farming accounts for a single company.

5.7 FINANCIAL SCAMS

Financial services are obvious targets for fraud and scams, such as money-flipping scams, work from home scams, card cracking, and more. Financial scammers hijack banks' logos in an attempt to make their services look quasi-official. They monitor legitimate bank profiles on social media and identify when they're followed by a new user. The scammer will then immediately tag them or use an @mention to ask if the user would like to make a quick return on their money. Then the perpetrator takes the conversation with the user to DM to engage off the radar. This activity is not completely hidden; the initial post is public to all including the bank (Figure 20A).

In Figure 20B, the scammer offers a "money-flip" serivce, in which the victim gives the scammer an initial investment and they promise to return it tenfold.

This scammer offered to flip money for a number of banks, going as far as providing their phone number. The bulk of the malicious activity is carried out via private direct messages or off of the platform entirely, making it difficult to detect.

The scammers target victims in dire financial need, often appending hashtags like #help, #debt, and even #singlemom. They also target members of the military and holiday shoppers, who make for lucrative targets. At the end of the day, it's often the banks who eat the costs of these scams, which combined across platforms, could total in the hundreds of millions annually.

MOST COMMON TARGETS: financial services, insurance

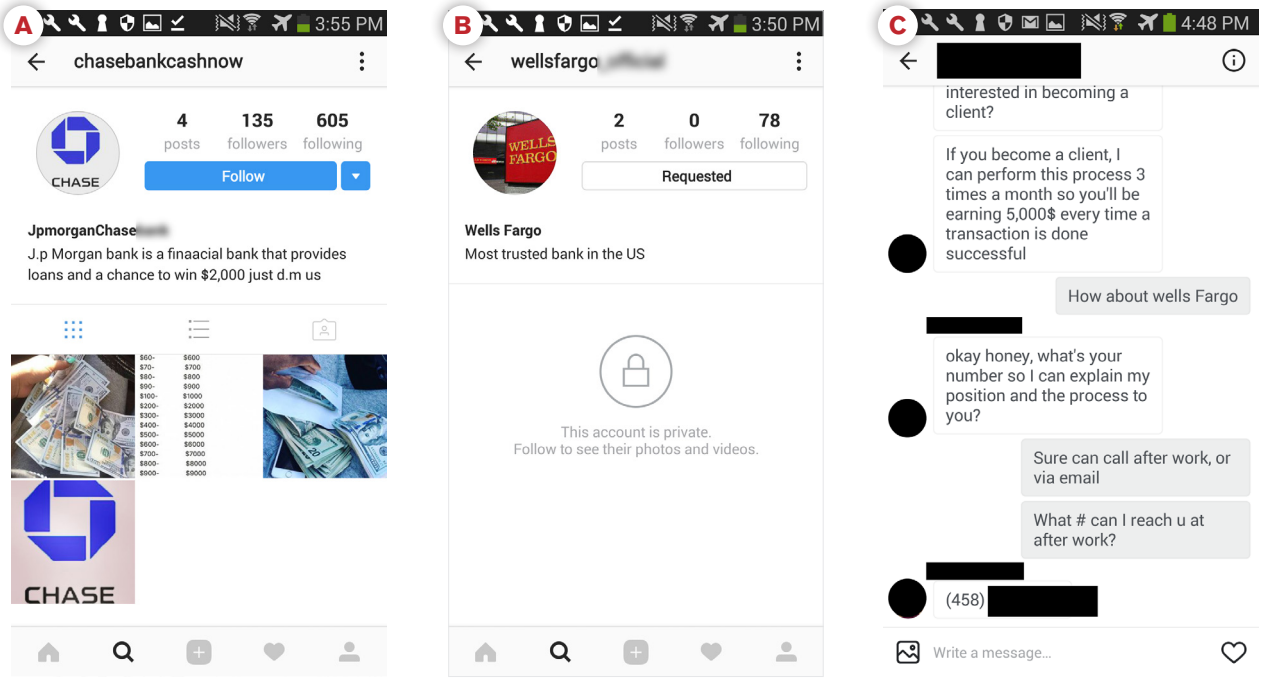


FIGURE 20: A. Financial institutions are exploited to promote money-flipping scams. A. The impersonator advertises money-flipping opportunities in their bio and historical posts. B. Another impersonator locks their account to avoid detection. C. The impersonator from 20B waits for a follower request to begin engaging through DM.

5.8 FRAUDULENT JOB RECRUITERS

Social media also is the new home for job applicants and recruiters. Unless the company has a verified recruiting account, which is very rare, it can be difficult for an applicant to decipher a legitimate account from its impersonator.

Job recruiter impersonator bios commonly contain Gmail, Yahoo, and other free email provider addresses. They encourage applicants to inquire about a job and send their resume, though more advanced scammers can spoof company email domains. Some also include links to official job sites and LinkedIn for follow-up. In most cases, the impersonator uses the company logo to portray themselves as an official recruiter for the company.

The sample of fake recruiters in Figure 21 shows the same perpetrator launching a complex, multi-channel attack across networks on Google+ and Instagram. We can tell it's the same perpetrator because the email and bio are identical on both Google+ and Instagram. The third example is another fake recruiter on LinkedIn.

Once an email is sent to the recruiter impersonator, the perpetrator will either try to extract PII or demand payment for an application fee. Some companies are aware of recruitment scams and have a page on their site asking job seekers to be aware of scammers using unofficial company email addresses.

MOST COMMON TARGETS: oil and gas, financial services, technology, law, executives of all kinds

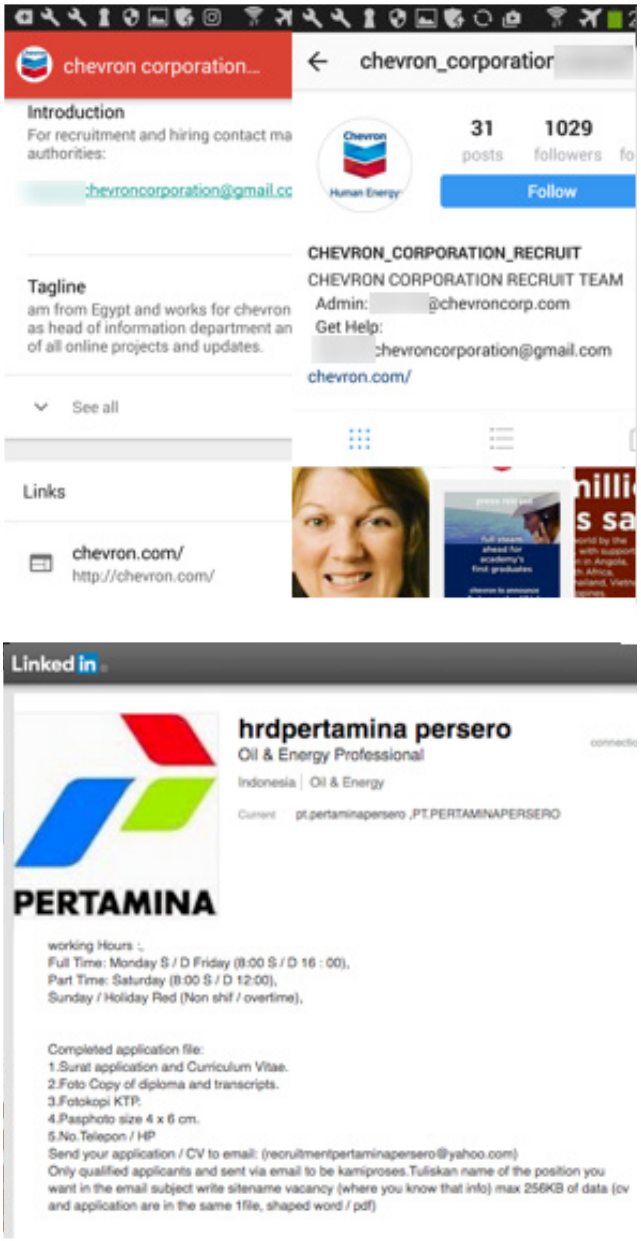


FIGURE 21: Job recruiter impersonators found on Google+, Instagram, and LinkedIn.

5.9 BRAND INFRINGEMENT AND COUNTERFEIT MERCHANDISE

Impersonators were also found to be creating accounts with the company name and logo and using these accounts for promoting competitor products, counterfeits, or other products altogether. For the company, this has a direct impact to the bottomline. Would-be buyers are directed to similar products that they might have otherwise purchased via the official social media account. Additionally, counterfeit merchandise incurs a direct loss to the company and promotes fraudulent activity.

MOST COMMON TARGETS: retail, manufacturing, technology, financial services, food and beverage, pharma

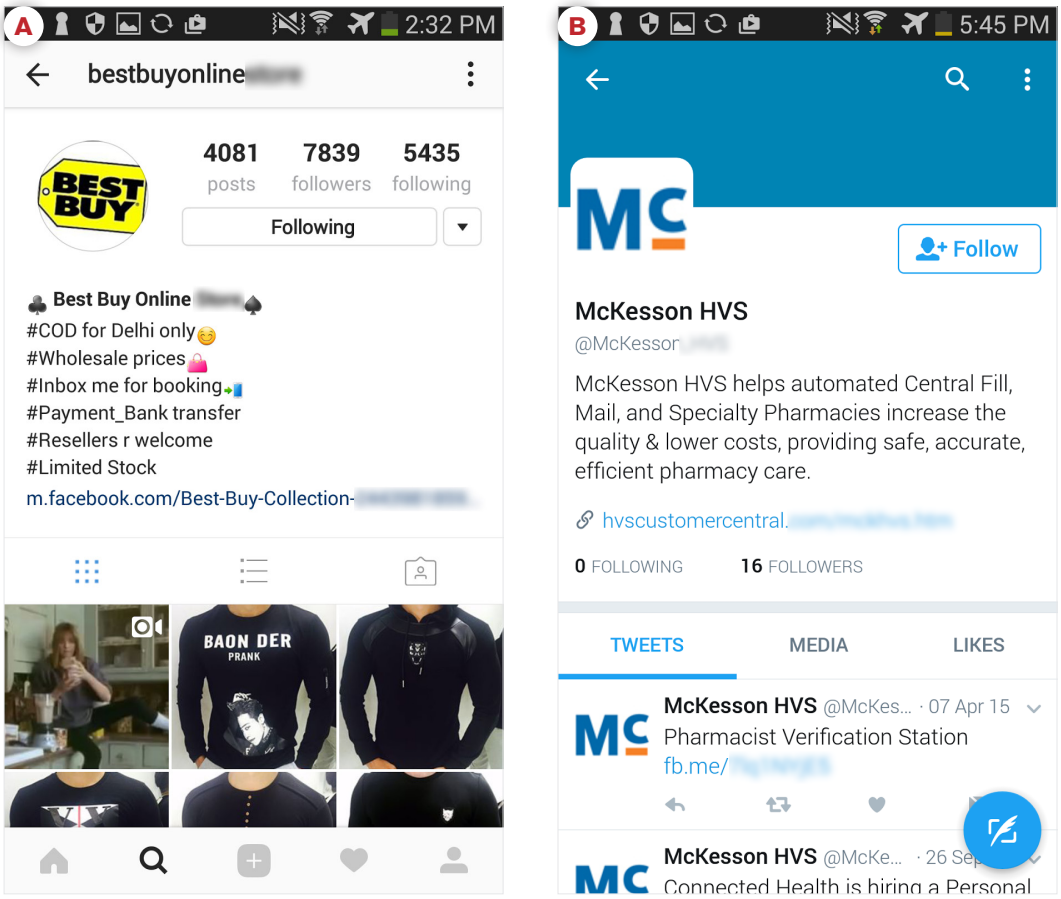


FIGURE 22: A. An impersonator claims to be a retailer but is actually selling unrelated clothing and jewelry. B. An impersonator claims to offer healthcare services, but the link directs users to a website development company that offers unrelated services.

6. CONCLUSIONS

The social networks have taken the first step in combatting the impersonator problem by verifying accounts, indicating to a user that the profile they're interacting with is legitimate and not an imposter. This is similar to websites that are verified using website digital certificates and browsers that highlight the URL in green. But what this approach doesn't provide is any indication of a nefarious account. Social networks rely on abuse reports from their users or manual triage in order to identify and respond to these accounts. This approach cannot keep up with the constant flux of impersonating accounts as they are created and deleted each day.

The problem of fraudulent accounts is systemic across the social networks and the tactics are broad and diverse. Proactively hunting for these accounts requires sophisticated, layered methods using account verification, threat detection, and machine learning (Figure 23).

This approach can be subsequently integrated to allow large-scale, cross-network analysis and improved detection accuracy. Machine learning classifiers can report on these threats targeting an individual or enterprise at incredible scale. Armed with this intelligence, an organization can take a more proactive and timely approach to thwarting threats, requesting account takedowns, and mitigating risk.

Impersonators are an excellent case study for the back-and-forth battle between cybercriminals, social networks, and the users caught in the middle. In our new digital lives, where people are free to assume others' identities and perpetrate malicious activity in their name, brands are increasingly at risk of financial and reputational losses.

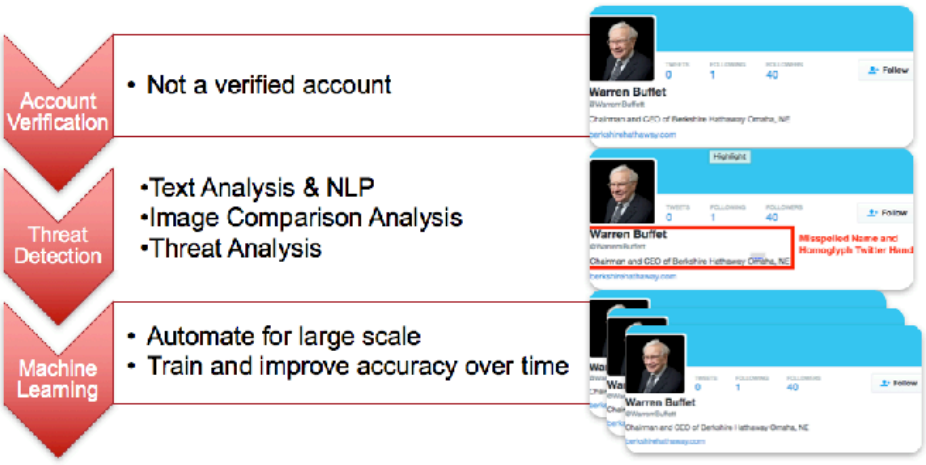


FIGURE 23: New layered approach to combatting impersonators on social media

Disclosure

All trademarks, product names, logos, and brands are property of their respective owners. All company, product and service names used in this white paper are for identification purposes only. Use of these names, logos, and brands does not imply endorsement.

All the fraudulent account information and activity shown in this white paper are being or have been reported to the social networks and the businesses.

ABOUT ZEROFOX

ZeroFOX, the innovator of social media & digital security, protects modern organizations from dynamic security, brand and physical risks across social, mobile, web and collaboration platforms. Using targeted data collection and artificial intelligence-based analysis, ZeroFOX protects modern organizations from targeted phishing attacks, credential compromise, data exfiltration, brand hijacking, executive and location threats and more. Recognized as a Leader in Digital Risk Monitoring by Forrester, the ZeroFOX SaaS platform processes and protects millions of posts, messages and accounts daily across the social and digital landscape, spanning LinkedIn, Facebook, Slack, Twitter, Workplace by Facebook, Instagram, Reddit, Pastebin, Tumblr, YouTube, VK, mobile app stores, the deep & dark web, domains and more.

Led by a team of information security and high-growth start-up veterans, ZeroFOX has raised over $40M in funding from NEA, Highland Capital and others, and has collected top industry awards such as the SINET16 Champion, DarkReading's Top Security Startups to Watch, Tech Council of Maryland's Technology Company of the Year, and the Security Tech Trailblazer of the Year.

ZEROFOX RESEARCH TEAM

The ZeroFOX Research Team is dedicated to investigating malicious activity on social media to better understand how to protect people and organizations alike. Our group is composed of curious and determined scientists, engineers and writers; both techies and storytellers. We are committed to integrity in all aspects of our research process, from data collection to reporting.

All the information in this report is publicly available data collected using the network APIs. No confidential customer information is contained in the report. Additionally, no foxes were harmed in the writing of this white paper.