Socure

The Identities Are Fake, but the Consequences Are Not

Solving the Expanding Problem of Synthetic Identity Fraud

by Jeff Scheidel, Head of Training & Development for Socure

Prologue

It's been many years since people in the security business feared script kiddies, namely younger or less sophisticated hackers armed with downloadable toolkits that allowed even the dumbest of infiltrators to worm their way into other people's networks. That era now seems almost quaint, as a far more intelligent breed has had its way for a long time, creating ingenious methods of penetrating, stealing, and profiting from huge tracts of consumer data.

They pick their targets carefully. As the robber Willie Sutton famously said, he went after the banks because "that's where the money is." In the modern age, the digital gold is the data inside the credit bureaus, retailers, and other hoarders of personally identifiable information (PII). As a result, these organizations need to be laser focused on detecting and preventing an ever-evolving array of data threats.

There are numerous examples of ridiculously vulnerable sources of sensitive data that have been exposed and ransacked, although most are locked down to varying degrees. Still, when these treasure troves of personal information are broken open, they feed the fraud beast exponentially. All that personal data in the wild becomes the fodder for fraud. Social engineering and phishing also generate such data, but it's the breaches that do it on a large scale. It's a two-step crime, in that the data must first be acquired, then made use of, by either the thief or the party to whom he sells the data. And just as criminals evolve how they steal; they also evolve how they leverage what they steal.

The newest incarnation of digital criminal has added more one weapon as he spends utilizes his loot, beyond cunning and technical know-how: patience. In addition to smarts, bad intentions, and a warped talent for IT, a willingness to wait for the opportune moment gives the fraudster yet another leg up.

Organizations practicing advanced fraud detection don't have the luxury of time. They need to up their game, to fight the battle against current fraud techniques as well as those that are still evolving.

Patience that Powers Synthetic Identity Fraud

Although patience may not sound like much of a weapon, consider two particular types of fraud, first party and third party. The former simply requires that an individual not pay his credit card (or other) bills. The latter requires stealing an identity and plugging in one's own contact info.

But synthetic fraud, a broadening threat (and in fact the fastest-growing financial crime in the U.S.), requires taking the time to assemble a fake identity, letting it bake into a reasonable facsimile of a person, then using it as a spearhead to commit major financial crimes. Despite increased vigilance, and the heightened scrutiny due to some very high-profile, high-volume breaches, the amount of PII exposed in the years 2017 and 2018 more than doubled, with almost 450 million more records compromised.

One common target of synthetic fraud: healthcare providers. While Medicare and other provider- based fraud is well-publicized, lesser known are the fraudsters who generate large numbers of phony claims and, in some cases, receive undeserved benefits (including free surgeries) by way of false identities. Thieves may also set up fake identities to file tax returns or receive government benefits.

Regardless of its incarnation, a synthetic identity is a Frankenstein's monster, not a direct identity hijack. There are three primary methods to synthesizing an identity. First, fabrication to create a completely phony identity containing no actual data related to an actual human being. Second, identity manipulation by modifying actual identity attributes of the fake person. And finally, identity compilation, creating that monster from a combination of modified and actual attributes. The ultimate goal is to present a face that can get past a first line of defense with an appearance of legitimacy.

Once an avatar is constructed, the fraudster next might generate a few credit pulls, apply for magazine subscriptions, shop online, or perhaps fund one or more actual accounts with token amounts, enacting minor transactions to put flesh on those phony bones, like an old-time con man fattening up his empty credentials. Establishing an account for a phony ID may require multiple applications, until it is finally accepted. Even rejected applications become part of the history for that identity. Again, this is all part of the gradual, patient process. Some sources break this out into one more stage: repeated application. It's similar to a brute force attack on a password-protected site, in that the perpetrator just keeps applying until an opportunity clicks. Even failed applications add to that synthetic history.

Finally, the criminal leverages that bogus history by using the synthetic avatar to pull off the big one. This chain of events gives the felonious stand-in identity sufficient gravitas to get past an initial scrutiny and appear to be legitimate, leading to the maxing out of credit cards or draining of a funded account. The average cost of "bust out" fraud is fifteen-thousand dollars, hardly a trivial amount when multiplied by the number of occurrences.

A number of years ago, a television newsmagazine sold credit cards online and tracked their usage. Buyers started out purchasing mundane items, such as dog food or shampoo. Once these transactions passed through with no friction, the next ticket items were expensive jewelry. Synthetic fraud is the same kind of exercise in patience, but on a grander and more gradual scale.

Schools and various government bodies have concerned themselves with financial aid fraud, which can take numerous forms, including acquiring funds for classes that are never taken. But in many cases, criminals are bypassing aid fraud and only using schools to acquire edu-based emails and other history, simply for the purpose of later launching synthetic attacks.

Synthetic and first party fraud are often called "victimless" in that the first party won't report himself, and the synthetic person doesn't even exist to report himself. Therefore it may take far more time than in a simple identity theft for the crime to be noticed and called out. But in fact the institutions themselves are the victims, suffering billions in losses that may dampen other investments, and are typically passed along indirectly to consumers. Lenders lost six billion dollars to synthetic fraud in 2016 alone. Merchants who (unknowingly) issue their own store-specific cards to fake persons also take a hit.

Synthetic fraud often leverages actual profile attributes, such as SSNs or other elements of a real person whose financial profile can suffer downstream. This damage is not limited to adults. Over a million minors were the victims of synthetic fraud in 2017, and since they won't apply for various products or services for several years, the theft may not be noticed until they become adults. In addition, there is no telling how many will face additional consequences when they apply for their first jobs, or for college, making the compromise of their data a massive time bomb that is lying in wait until these victims reach adulthood.

But it's not just minors who may face lingering consequences. The leveraging of attributes of others, such as the elderly, deceased, homeless, and incarcerated may not be caught for many years, if ever.

Time Isn't the Only Culprit

There are other factors contributing to the ease with which criminals can build synthetic identities with impunity. In many ways the system is geared toward giving fraudsters a clear path.

First, identity verification in general, across industries, still relies all too heavily on static, compromised data. Social Security numbers, for decades the backbone of validation, have become one of the most commonly stolen assets.

The randomization of SSNs, starting in 2011, was designed to prevent criminals from predicting individuals' numbers. But this random scheme also created an additional opportunity for fraudsters.

Previously, SSNs were generated in ranges (using a combination of geographic and incremental logic), and if a presented SSN was outside an issued range, it was clearly bogus. But with SSNs now generated all over the map, it is now far more difficult for security professionals to instantly identify one as invalid.

With increasing numbers of applications, for loans and accounts and credit, coming in digitally rather than in-branch, criminals are also able to offer up their illegitimate credentials across national borders, without the need to present themselves physically. This anonymous interaction means less fear of getting caught, and therefore less fear of consequences, and subsequently far more attempts. The more attempts, the more chances of success.

Fraud detection tools are simply not up to the task. Meeting KYC requirements is not sufficient. Institutions need to implement advanced models that can recognize evolving fraud types, based on an aggregation of observed fraud. To further power this capability, discovered fraud should also be properly classified. The confusion between first party and synthetic fraud, for example, means subsequent analysis or investigation will be misguided, if it happens at all.

The Consequences for Companies and Customers

Many institutions are required to keep reserves for credit losses, but not fraud losses. This is intended to provide protection in the event of default. Credit losses are a reserved charge-off (and 20 percent of these resulted from synthetic fraud in 2016), while fraud losses are an expense. So they represent an additional danger for at-risk organizations and their customers, who could be left holding the bag if the impact is large enough to constitute an existential threat to the company's long-term health, especially in the event of an economic downturn. And when synthetic fraud has baked long enough to be realized, it may be classified as a credit loss rather than as a fraud loss, which means it's treated as a write-off, rather than a crime worth investigating. This means the bad guys get away without anyone on their tails, and all because those synthetic IDs look and feel like a real person.

Just as many software companies allow their customers to be part of the beta stage of development by catching (and suffering from) bugs, other organizations may allow their own customers who are victims of third-party fraud to help them find their holes. It's the customers whose identities have been compromised that trigger investigations, such as when they receive improper bills or get dinged by collection agencies or other aggrieved parties.

Another sad consequence of the growth in synthetic fraud is that attempts at prevention and detection require tools and the staff to use those tools, representing yet another expanding cost of doing business.

The Synthetic Identity Challenge: Recognizing It as Fraud

Just as some first party fraud is really just otherwise-honest people falling on hard times and struggling to pay their bills, it is not uncommon for individuals with bad credit to use synthetic IDs for basic needs, such as purchasing a car to get to work. The worst perpetrators, however, are repeat offenders, who tend to be extremely calculating and patient and whose goals are clearly not honorable. It's no surprise that many fabricated identities can pass muster in KYC checks, since there are enough data elements or profile attributes to make them look real.

Classifying fraud incorrectly can make it more difficult to create filters based on patterns, since pattern recognition is being trained on the wrong results. Swinging at a target in the dark that you think is taller than yourself will always be futile if that target is in fact shorter. There are also numerous stories of organizations whose fraud teams just do not want to work on synthetic ID cases, since those entities are not real people with real bread crumb trails.

The numbers are frightening. According to the Federal Reserve, between 85 and 95 percent of synthetic applicants are not picked up by standard fraud models.

In order to sniff out synthetics, a fraud model must learn and understand what a normal distribution looks like, for specific data elements, in the context of each element. This requires large training sets of known outcomes and brings us back to two aforementioned issues: the labeling of fraud is often erroneous, and synthetics themselves are often mistaken for other kinds of fraud. But given sufficient, much greater quantities of data, it is possible to recognize the correct patterns, even if not in every instance.

Now that We Recognize it, What Are the Best Strategies to Stop the Synthetic Threat?

One of the strengths of a synthetic identity is that it is composed of very real characteristics. In fact, even if the name or person is fabricated, it may have been so imbued with sufficient history, through the application of actual transactions, that it can pass an identity check. But if there are pieces of other, very real people, those characteristics can also be its weakness. The individual attributes may be legit, but do they relate to each other?

Understanding correlation is one approach to fighting synthetic fraud. The matching of identity elements allows the fraud model to determine whether the identity as a whole is a real person. This is tougher with synthetics versus standard third-party fraud because, while many of the signals are similar, the patient fraudster will make the pattern tougher to detect.

Machine learning is another of those industry buzz terms that has perhaps gotten overused, but it is still a powerful tool. Humans are not capable of processing the massive amounts of data needed in order to recognize the signals of synthetic identities. But automated systems can. If imbued with sufficiently calibrated artificial intelligence, capable of drilling down not only on individual attributes but also the relationships between them, such systems can detect and deflect a majority of synthetic attempts.

Some organizations largely dispense with the use of SSNs in their verification programs because of the many well-known breaches, finding that the correlation between other elements is a better predictor. The Social Security Administration plans to open up its own API to allow verification of name, date of birth, and SSN itself. Therefore, in theory, synthetic fraud is a problem that will be largely remediated by the very organization whose policies have at least partially enabled it. Of course, given enough time, fraudsters will find new attack vectors.

In the meantime, forward-looking risk managers will need to keep constructing models, and feeding those models enough properly vetted outcome data, to battle a virulent form of fraud that injures institutions now and may haunt its compromised victims for many years to come.

ABOUT SOCURE

Socure is the leader in creating high-assurance digital identity verification technology. Its predictive analytics platform applies artificial intelligence and machine-learning techniques with trusted online/offline data intelligence from email, phone, address, IP, social media and the broader Internet to verify identities in real-time.

Socure

For more information, visit socure.com