The University of Texas at Austin
Center for Identity

A Model for Calculating User-Identity Trustworthiness in Online Transactions

Brian A. Soeder
Suzanne Barber

2015

UT CID Report #1505

This UT CID research was supported in part by the following organizations:

The Identity Theft Assessment and Prediction Tool

Jennifer Brenner and K. Suzanne Barber

Your pet's name could be a fraudster's best friend. And in 2013, it has never been easier to get our hands on these "nuggets" of information, really, about anyone. Social Media sites such as Twitter and Facebook make it easy for just about anyone to piece together your life in a mere few clicks. "Despite all the awareness that people have about identity fraud and privacy on social networks, there is a disconnect between what people are disclosing in online space and social environments and what they may be using in other places of their lives," said Thomas Oscherwitz when he served as the chief privacy officer for ID Analytics, a San Diego-based consumer risk management firm. Simply put, announcing on our Facebook pages that we are headed to Europe for a month-long vacation and that our favorite dog Spunky needs to be watched, is detrimental. Not only have we made potential fraudsters aware we will be away from our home for an entire month, we've also provided our pet's name, which happens to be a security validation for many sites. "Even listing daily activities can let strangers know your routine and put you at risk," says Gail Cunningham, spokeswoman for the National Foundation of Credit Counseling. But what exactly are fraudsters doing with the information they find? How are they using your favorite animal's name as a 'in' into ruining your credit…or worse, your life?

The ITAP, or Identity Theft Assessment and Prediction Tool, is a tool that allows computational representation and quantitative measurement to better understand a fraudster's behaviors and inevitably, make connections and visualize patterns of those behaviors. How are fraudsters accessing information, i.e. through what vulnerabilities? What tools are they using in order to overcome security hurdles? What steps are they taking? Behavior patterns and trends identified in the ITAP will answer these and many more questions to include the entry points, vulnerabilities and consequences of fraudsters. According to the U.S. Department of Justice, in 2010, 8.6 million households had at least one member age 12 or older who experienced one or more types of identity theft victimization. Until now, we have been one, two or even three steps

behind fraudsters, only reacting to cases of identity theft after the damages have been done. Across a wide range of market sectors including consumer services, financial services, healthcare, energy, education, government services, national security and law enforcement, proactive prediction of the following is planned:

1. Most common data required to initiate fraudulent behavior capabilities,
2. Most common data taken or acquired by the fraudster,
3. Steps that often work together which result in identity theft or breach of data, and
4. Commonly used resources used to achieve the execute identity theft, fraud and abuse.

The ITAP itself is structured as a database, which takes input "Scenarios" and breaks them down into distinct steps. Scenarios are categorized into different market segments, which allow an analysis where these threats are actually taking place and what industries may be most susceptible to attack. Scenarios are stories or situations where identity theft, fraud and abuse has already taken place. The ITAP describes the situation in full detail, including who the major players were, who the victims were and what exactly took place. Each Scenario is then broken down into "capabilities". Capabilities are the actual steps, or things the fraudster was capable of doing, in order to achieve the overall fraud. For example, the ITAP currently describes a very elaborate scenario of home equity fraud. The first step, or capability, includes the fraudster obtaining a job as a Loan Officer to learn the internal processes involved in processing loans and other such related documentation. We considered this a crucial step in the overall scheme since the knowledge he acquired here inevitably allowed him to pull off the entire fraud. He was able to learn exactly what the proper procedures were in handling loan documentation, how authentications were handled by financial institutions, when clients were calling to check on statuses as well as which types of banks were the easiest targets.

On a more granular level, each of these capabilities, or steps, has a variety of components that help explain how it was carried out and for what purpose. What was the intent? Firstly, each step has data inputs and data outputs. The data inputs are absolutely crucial. They answer the question, what pieces of data were required for this step in the criminal process to be carried out. To elaborate on the previous example, one of the steps the fraudster took in order to commit the home equity fraud was to search ancestry.com for the victim's mother's maiden name. In order to

successfully search ancestry.com, he had four data inputs, the victim's date of birth, previous address, social security number and the name of a relative. Understanding what pieces of information went into a step helped to analyze what fraudsters are actually using to piece together our lives online. In addition to noting the inputs of each step, this research aims to understand the data outputs. What exactly was the fraudster seeking ? It is important not only to think about what the fraudster needed in order to carry out this specific step in the fraud, but also, what were they trying to accomplish? The benefits of this analysis are twofold. Not only do we begin to understand exactly the way the mind of the criminal works, we also begin to see patterns and different ways that fraudsters are using 'x' to gain hold of 'y'.

Within the construct of each "capability", the ITAP model also lists resources the fraudster used to complete the step and of course, the overall fraud. By analyzing the resources, this research can potentially find ways to limit access to common ones whenever possible. These resources can be anything from malware code, a call spoofer or a credit card skimmer to some software, such as Photoshop. Photoshop was in fact one of the many resources that the fraudster used in the Home Equity Fraud described herein. Using Photoshop, the fraudster was able to cut and paste signatures of wealthy couples from publically available loan documents he found online, onto wire transfer forms. He then used these wire transfer forms to transfer money between the victim's account and his own account. Anything the fraudster physically uses can be considered a resource, and unfortunately, the Center for Identity's research indicates that many of these resources are readily available.

In addition to the "capabilities" data and resources ITAP models to describe the criminal behavior, actors are also modeled. Who exactly played a part in completing the step? Was it a hacker? Was it a skimmer? Did someone internationally play a part? It is very important to assess who the key players are in each step of the overall fraud. In doing so, we are able to view the big picture of how many and what type of people were necessary in carrying out the attack. In the home equity case, although there was one major fraudster, he did receive assistance internationally. He worked with an international fraudster and broker when transferring the stolen money outside the US. This allowed him to send the money outside the US and bring it back inside in an attempt to avoid being caught. With the assistance of these two additional people, he was able to launder roughly $7 million dollars every two weeks until he was eventually caught.

Another important feature of the ITAP is the ability to link specific steps together by adding START and STOP conditions. This allows the research to link steps together in the criminal process to indicate that one step is absolutely required to be completed before the next step can begin. We link steps together, if and only if, one would always follow the other. Thus, the ITAP creates true connections between the steps in the criminal process. Upon further analysis, ITAP can throw red flags when similar pairs of steps occur in future instances to signal new instances of identity fraud. For example, in our Home Equity Fraud scenario, the fraudster was able to run a credit report on the victim on annualcreditreport.com with a goal of gaining specific HELOC details. The HELOC details, or Home Equity Line of Credit details, provide information on a specific type of loan the fraudster used to carry out his scam. Because the data input to running the credit report required knowledge of the victim's address, date of birth and social security number, the preceding step, the building of the victim's profile by paying for a background search on a skip-tracing site, is a necessary START condition. Meaning, the fraudster needed to complete the background search prior to being able to access the credit report. We now can make clear connections on how the information is flowing and reveal the dependencies between certain steps. This type of pattern detection will prove invaluable in predicting future identity theft scenarios.

In terms of analytics, this research captures a wealth of actual empirical data about the behaviors of identity criminals in the ITAP model. The data gathered will paint a better "big picture" on how identity fraud is occurring. Currently, we are funneling information into the Identity Threat Assessment and Prediction (ITAP) model to properly categorize threat behavior and detail patterns of behavior. First, the research is detecting the most common data inputs required for any given step within a scenario to occur. What information did the fraudster already have on hand? Did the fraudster have access to your email account? Were they able to view your place of birth or birth date on Facebook? By understanding what the fraudsters have in the beginning, we may be able to change the way we feel about certain elements of data which we, until now, thought to be secure. Perhaps a mother's maiden name is no longer a secure method of identifying someone since it can be fairly easily discovered by viewing family trees on Facebook or Ancestry.com.

Discover patterns of fraudster behavior will provide tools to predict and disrupt this criminal behavior by individual, businesses and government agencies providing and managing this data. Additionally, this research aims to understand exactly what the fraudster is after in each given

scenario. What are the most common data outputs? Although big picture this seems fairly obvious, it is important to ascertain exactly what the fraudster hoped to accomplish at each step. This will allow us to understand exactly *how* the fraudster is getting his/her hands on certain pieces of information. How exactly did the Home Equity fraudster end up with the victim's Mother's Maiden Name? He was able to Google information and then easily use it as a data input on Ancestry.com. Perhaps it is time to consider different data inputs as a method of identifying someone, especially on sites such as Ancestry.com, where the entire purpose of the site is information discovery. This shift in thinking by individuals and businesses may be especially important if the data being acquired could be potentially detrimental to someone's identity.

Furthermore, this research will detect any repetitive groupings that exist between steps in these criminal scenarios. This is important because it could serve as a prediction technique for future identity theft scams. Are there two or three steps that typically go together across multiple scenarios? Could these steps, when completed in conjunction, throw a red flag that identity theft may be amiss? We certainly believe so. Because the ITAP model connects steps in the fraudster's behavior, the research is using the ITAP to see what steps commonly work together in the process of committing a scam. Does A typically follow B? And how can we use knowledge of this to prevent C?

Lastly, this research is seeking to analyze commonly used resources across all the scenarios. What are the most common resources being used to commit identity theft and fraud? Can we limit access to these very common resources or, at a minimum, make the providers of these resources aware that they are in fact being used in an ill manner? Often times, these resources are pinnacle to the completion of the step, and ultimately the theft so any analysis we can do in this area could spring us ahead in identity theft detection and prevention. The research effort will also apply this specifically to market segments to view specific trends within any given industry. What resources are aiding in medical identity theft or in identity theft in financial institutions?

The most unsettling part of the Home Equity Fraud example was that the fraudster didn't begin the fraud with your wallet, access to your bank account or your credit card number. He began with nothing. Through the manipulation of various vulnerabilities, online databases, knowledge of the interworkings of loans and financial institutions, and specific resources, he was able to start with nothing and build profiles on victims until he eventually stole millions. The ITAP

model is analyzed to discover these seemingly random series of steps, resources and data elements and model a deliberate, sophisticated criminal process, explaining how advancements were made. Through further analysis of commonly utilized data elements, resources and pattern detection, results of the ITAP research will thwart future identity theft scams, or minimally, be aware of warning signs.

The University of Texas at Austin
Center for Identity

For more information on Center for Identity research, resources and information, visit **identity.utexas.edu.**