Written Testimony of ID Analytics Corporation

Before the Subcommittee on
Financial Institutions and Consumer Credit

Hearing on H.R. 3997,
the "Financial Data Protection Act of 2005."

Washington, DC
November 9, 2005

Mr. Chairman, Ranking Member Sanders, and other distinguished members of the Subcommittee on Financial Institutions and Consumer Credit, ID Analytics is pleased to submit for your consideration a summary of relevant findings from its forthcoming "National Data Breach Analysis."

We are submitting written testimony, even before the study is publicly released, because we believe the Committee should have the best evidence available as it ponders legislation with respect to how criminals are using (or not using as the case may be) information obtained from data breaches.

By way of background, ID Analytics is a San Diego-based company that provides Identity Risk Management solutions to a number of the nation's largest financial institutions, credit card issuers, and wireless companies. ID Analytics is in a unique position to offer insight into the data breach problem because of our analysis of the "breached files" of three highly publicized data breaches involving hundreds of thousands of identities.

This analysis was conducted using ID Analytics ID Network, a nationwide, cross-industry collaborative fraud prevention system. ID Network Members are organizations that contribute consumer identity elements sourced from their customer management processes, including account applications, requested changes of account information, and tendered payments, in the interest of collectively preventing identity theft and related fraud. Each ID Network Member has agreed that identity fraud prevention requires a new level of collaboration and has entrusted ID Analytics to develop and maintain the technology required for comprehensive and effective Identity Risk Management.

For the purposes of this research, ID Analytics classified each breach as either an "identity level" or "account level" breach. An identity level breach involves the most sensitive data available – names, Social Security Numbers (SSNs), dates of birth, addresses, and other personally-identifiable information. An account level breach involves mostly account data such as credit card numbers and credit card expiration dates.

Summary of Findings:

During the summer and fall of 2005, ID Analytics conducted an analysis of the breach files of three widely-publicized data breaches involving hundreds of thousands of consumer identities. The primary purpose of this analysis was to determine the degree to which identity fraud results following a data breach.

One of the breaches analyzed involved a serious breach of identity-level information on consumer reports. Two of the breaches involved the disclosure of account-level information on credit card accounts. Selected key findings are as follows:

- **ID Analytics' analysis of the identity-level breach, which involved over 100,000 consumer identities, revealed the following:**
  - Misuse of the breached identity information began gradually, spiked around the discovery of the data breach and declined precipitously after the breach was publicly announced.
  - Fraudsters used identity data manipulation, or "tumbling," to avoid detection and prolong the scam.
  - The calculated fraudulent misuse rate for consumer victims of the breach was 0.098%. This rate is less than the annual rate of identity fraud for all Americans reported by the FTC Synovate report in September 2003.
- **ID Analytics' analysis determined that the two account-level breaches did not indicate patterns of new fraudulent activity.**
- **Technologies now exist that can measure the fraud risk associated with breached identities and results are being proven.**

Study of the Targeted Identity Breach:

In mid-2005, ID Analytics was approached by an ID Network Member and asked to explore opportunities to use the ID Network and its associated technology to determine if identity fraud was resulting from a well-publicized data breach of an unaffiliated third party.

The data breach in question was what ID Analytics considers a "targeted breach," meaning there was a deliberate theft—or hacking—of data. It was also what we call an identity level breach as the data stolen consisted of more than 100,000 consumer identities, including Social Security Numbers, dates of birth, names, and other sensitive information. ID Analytics' scientists and fraud analysts set to work to determine if the breached data was being fraudulently misused and, if so, to propose a strategy for preventing any further identity fraud resulting from the breach.

ID Analytics did, in fact, discover fraudulent misuse associated with this major data breach. While we will not go into great detail about the science and analysis used to discover the fraud, we will attempt to explain the basic method.

The underlying theory was that identity information associated with a breached file should not exhibit suspicious patterns and relationships unless that information was being misused in an organized manner, as in the case of a fraud ring perpetrating identity fraud.

Under normal circumstances, any two identities should exhibit subtle, but not suspicious, relationships to each other. For example, husbands and wives cohabitate, and thus share addresses and telephone numbers. Two random individuals can even share the same names, and thus their identity data "relates" in an innocuous manner. Yet two previously unrelated identities should not suddenly begin sharing SSNs, addresses, or telephone numbers. Such suspicious relationship patterns become evident as the identity is asserted on subsequent new account applications; these patterns can be indicative of identity fraud in action. In isolation, many of these patterns appear safe, but with an extremely wide perspective and through millions of repeated observations, sophisticated analytical technology can help interpret suspicious patterns, such as those associated with a data breach that is resulting in identity fraud.

ID Analytics' analysis of the breached file yielded the following results:

(1) Roughly 1 in 1,020 breached identities (0.098%) were used to commit identity fraud. This rate is less than published reports about the annual rate of identity theft affecting the general population.

This rate of fraud in a breach population, called hereafter the "misuse rate," speaks to an important truth about identity-level breaches. Practical constraints, and not the size of an identity-level data breach, determine the amount of identity fraud that is likely to result from a data

breach. It is the fraud ring's available resources that determine how much fraud follows a targeted, identity-level data breach. Fraud rings simply do not have the time or manpower to use hundreds of thousands of identities available to them in their nefarious pursuits.

While initially surprising, these seemingly low use rates from data breaches, upon further consideration, appear rational. Assume the following:

- Five minutes to fully and accurately complete a new account application that is likely to be approved
- One application per unique identity
- Average 6.5 hours per work day, five days per week, 50 weeks per year

Given the above constraints, it would take on individual fraudster over 10 years to fully utilize a breached file consisting of one million consumer identities. Should the fraud ring outsource these tasks at a rate of $10 per hour in an effort to fully utilize the breached file within one year, the fraud ring would have to hire 52 workers and spend over $830,000. This scenario also overlooks other practicalities, such as procuring the applications, logistics around receiving loan instruments (credit cards or loan checks), and the need to launder the proceeds of the fraud. These practicalities imply that there exists a feasibility limit associated with fraudsters committing identity fraud using breached identities.

However, misuse rates could continue to increase drastically over time if the vibrant black market for "identities" remains unimpeded. Today, there is no evidence of a central, thriving, continuously-operated black market, although there is evidence that some stolen consumer data is sold via internet relay chat (IRC) networks and through other internet-based communications channels. By selling any amount of the remaining identities (those not able to be used because of the "feasible limit"), fraud rings could maximize the proceeds from their efforts and exact a far greater degree of harm to consumers, industry and government over time. It should be clear that this scenario calls for consortium-based, real-time, identity-centric technology solutions to prevent ensuing identity fraud.

(2) Fraudsters used identity data manipulation, or "tumbling," tactics to avoid detection

ID Analytics' scientists observed that the fraudsters misusing the consumer identities associated with this data breach were engaging in creative tactics to prolong the scam and avoid detection.

Figure 1: Evidence of Identity Data Manipulation, or "Tumbling" Over Time

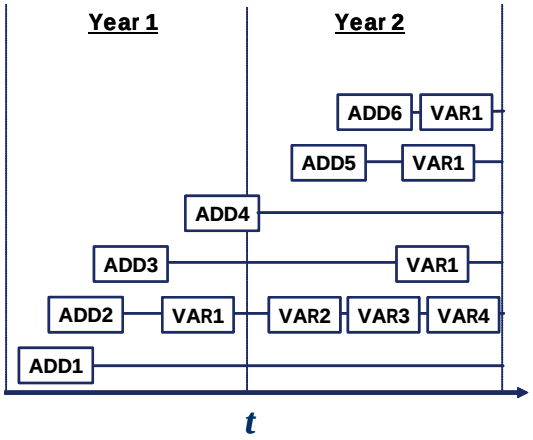
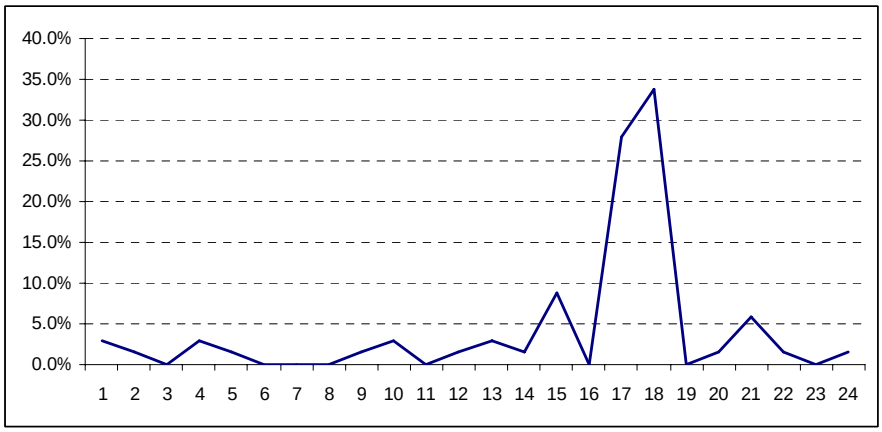


Figure 1 provides evidence of one such technique that has been referred to as "tumbling." The fraud ring in this example chose to manipulate the addresses submitted as part of the account applications over time, resulting in obscure, yet difficult-to-detect variations of the original address. The manipulations illustrated here amounted primarily to changes in apartment numbers or spellings of street names. Interestingly, scientists observed a dramatic increase in these manipulations in the latter days of the identity fraud scam.

(3) Misuse of the breached identity information began gradually, spiked around the discovery of the data breach and declined precipitously after the breach was publicly announced

Over the 24-month observation window for this data breach, there was a 12-month pattern of low rates of misuse followed by a brief 6-month period of a high rate of identity use, and then a steep reduction in identity use after 18 months.

Figure 2: Monthly Identity Use Rate for Selected Data Breach



The increased rate of misuse began around month 14. This elevated rate of misuse lasted for months and dropped off precipitously when the breach was announced around month 22.

ID Analytics can only speculate on this rate of misuse by the fraudsters responsible for this particular data breach. One possible answer is that the fraudsters were using the identities sparingly in order to avoid detection. Around month 14, when the breach was discovered by the commercial entity, the fraudsters may have realized the game would soon be up and tried to maximize the cash value of the data in their possession. Once the breach was announced, the misuse of the identities fell precipitously. We do not know at the time of this study whether the identities will be further misused in the future, but continued monitoring would be required to make such a determination.

Study of Account Level Breaches:

If identity fraud does result following an account level breach, the lasting effect to the consumer involves unwinding this damage through a rigorous series of calls to credit reporting agencies, the issuing lender, and any number of police departments (depending on jurisdiction).

But account-level data breaches can lead directly to credit card transaction fraud. In contrast to identity fraud, where identity elements of a consumer are typically used to perpetrate financial fraud across

numerous cards and accounts, credit card transaction fraud by contrast involves just using a particular credit card to perpetrate fraud. This type of fraud does not burden the consumer as much, nor does transaction fraud persist after the institution takes appropriate measures since either the merchant or the card issuer bear the financial losses resulting from the fraud.

Since account-level data breaches generally involve the disclosure of credit or debit card numbers, expiration dates, and names, when the institution reissues the account number and invalidates the compromised one, transaction fraud is prevented for that victim. Most institutions assume 100% of the liability in these cases of fraud where identity fraud is not a concern because a name alone is not enough information by which to commit identity fraud.

However it is theorized that account-level data breaches can lead to identity fraud if a new account is (or multiple accounts are) opened in the victim's name. Logically speaking, if a fraudster obtained a name and a credit or debit card account number, he could use the internet to "find" the victim and steal the other necessary information (SSN, date of birth, etc.) to perpetrate identity fraud on new account applications or to access existing accounts and defraud the victim.

ID Analytics was approached by an ID Network Member to seek an answer to the following question: **Does a fraudster who accesses an account-level data breach file have the intent or ability to gather additional identity information on the breached identities in an effort to perpetrate follow-on identity fraud?**

This ID Network Member provided ID Analytics with two separate account-level breached files. Both of the breached files originated from US-based retailers' computerized account databases that had been accessed illegally. ID Analytics conducted analyses on both files, but presents results for this report in the aggregate.

While the hackers responsible for the breach did not have consumer identifying information other than name, account number and expiration date, the ID Network Member appended that identity information to the file in order to determine if there was any attempted identity fraud following the breach.

ID Analytics' analysis of the breached file yielded the following results:

(1) No Widespread Fraud Patterns Detected in the Account Level Breach

There was no evidence that the breached file was being exploited by fraudsters to perpetrate large-scale identity fraud scams. While there was one account-level fraud attempted out of 1428 breached accounts (one account level breach above the average misuse rate of 0.07% and one below the average rate of misuse), we found that there was no evidence that follow-on identity fraud had been perpetrated against the two breached account level populations.

(2) Identities from the account-level breach file exhibited an unsuspicious distribution of Social Security Number relationships to reported fraud when compared to a control group.

The table below compares the percentage of fraud hits found by SSN within the ID Network.

Table 1: Social Security Number Relationships to Reported Fraud

Number of SSN Relationships to Reported Fraud	Account Level Breached Group	Control Group
None	99.38%	99.12%
1 fraud hit by SSN	0.61%	0.86%
2 fraud hit by SSN	0.01%	0.02%
3 or more fraud hits by SSN	0.00%	0.01%

As Table 1 illustrates, identities from the account-level breach file exhibited an unsuspicious distribution of SSN relationships to reported fraud when compared to the control group. Both this account-level breach file, as well as the identity-level breach file actually appeared safer than the control group on this SSN-only dimension.

Table 2: Comparison of Anomalous Address Links

Number of Anomalous Relationships by Valid Address	Account Level Breached Group	Control Group
>4	0	1
4	0	0
3	0	1

Table 3: Comparison of Anomalous Telephone Relationships

Number of Anomalous Relationships by Valid Phone Number	Account Level Breached Group	Control Group
>4	0	0
4	0	0
3	0	0

As Tables 2 and 3 indicate, the account-level breach file exhibited no suspicious relationships to either addresses or telephone numbers, indicating an extremely low probability that the affected consumers will become victims of identity fraud.

ID Analytics conducted many other tests on this data set and believes that consumers affected by this account-level data breach will not fall victim to identity fraud in any significant numbers.

Conclusion:

ID Analytics recognizes that the information provided by these three breach analyses is only the beginning of our understanding of how criminals are capitalizing on data breaches to perpetrate fraud.

We appreciate the opportunity to express our views and would welcome the opportunity to more fully brief the Committee on the findings of our study.

Sincerely,

Mike Cook
Vice President
ID Analytics