id:analytics.

15253 Avenue of Science
San Diego, CA 92128

T. 858.312.6200
www.idanalytics.com

Analysis of the BIRLS and HEC Data Using ID Analytics Identity Risk Technology – How Much More at Risk are Veterans than Others of Identity Theft?

U.S. Department of Veterans Affairs

Date of Report: 17 December 2012

This report was processed in accordance with Order Number GST0309DS6024, Task #: R3083865-15

Revision: 1.0
Last Update: 12/14/2012

Table of Contents

1. Executive Overview

This report describes results from a detailed analysis of two special data files for the Veterans Administration: the BIRLS and the HEC file. Generally, the BIRLS file is the population of living veterans and the HEC file contains records for health benefits recipients, likely including veterans and their dependents.

We begin the analysis by doing detailed data statistics on each file. In Section 2 of this report we describe the fields, field populations, and the results of our identity resolution and population overlap analysis. We further identify large numbers of deceased individuals listed in both the HEC and the BIRLS files. Along with this reports we are returning a data set that contains all the flags and scores for each record in both of these data sets, including whom we can find, whom we believe to be deceased and various scores that measure the risk of identity fraud for each person on these files.

Section 3 of this report describes our most recent Data Defense analysis on the HEC file. We note that we performed the same data Defense Analysis and returned the results on the BIRLS file in a separate recent report. In Section 4 we report on what we believe is organized misuse of the identities of individuals on the HEC file. We report on 57 specific individuals whose identities appear to be compromised and misused, and we show explicit detail for 7 of these individuals. We further give recommendations on appropriate actions for the VA to provide remediation and to investigate possible systematic abuse.

Section 4 describes our examination of the risk of identity fraud for each individual in both the HEC and BIRLS file through our MyIDScore. We have scored each of the ~38 million records in the BIRLS and HEC files for personal risk and are returning these scores to the VA, who will then have a measure of personal identity fraud risk for each veteran in their system. We then quantitatively assed the risk of the veterans to the general population and found that the veterans have substantially higher risk than the non-veteran population. Finally we identified specific regions in the country where the veterans are at higher risk than their non-veteran neighbors. These identified risk hot spots are areas of likely systematic identity abuse and are candidates for further investigation.

In Section 5 we use our Consumer Notification Service alerting platform to assess how veterans use their identity in the marketplace. We then determine the risk profile of that activity to understand how often the identity of the veteran is being misused. We then look at veterans enrolled in commercial identity protection services to understand how they react to alerts / activity that they did not authorize, so we can apply the findings to the broader population of veterans and ultimately learn how many veterans are victimized by misuse / identity fraud. The data suggests that hundreds of thousands of veterans are victimized annually.

2. Data Hygiene and Statistics

2.1 Input Data Validity Analysis

All the analysis in this report was carried out on the VA's BIRLS file dated 9/20/2012, a set of 20,502,673 records, and the VA's HEC file received on 11/20/2012, a set of 17,832,405 records that is presumed to be a fairly clean data set of VA benefits recipients.

Each record in the BIRLS data set contains at most the fields:

- Social security number (SSN)
- First name
- Middle name
- Last name
- Date of birth (DOB)

Each record in the HEC data set contains at most the fields:

- Social security number (SSN)
- First name
- Middle name
- Last name
- Date of birth (DOB)
- Email Address
- Address
- City
- State
- Zip Code
- Home Phone
- Mobile Phone
- Work Phone

Table 1 shows the statistics on the BIRLS data fields as received from the VA.

Table 1. Input Validity Statistics/Field Population on the received BIRLS file (Only for successful Identity Resolution Responses).

	SSN	First Name	Middle Name	Last Name	DOB
Empty	0 (0.00%)	1,274 (0.01%)	2,965,161 (14.46%)	72 (0.00%)	531,825 (2.59%)
Partial	0 (0.00%)	57,982 (0.28%)	4,168,390 (20.33%)	603 (0.00%)	446,542 (2.18%)
Frivolous	8 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Invalid	351 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	8,910 (0.04%)
Valid	20,502,209 (100.00%)	20,443,312 (99.71%)	13,369,017 (65.21%)	20,501,893 (100.00%)	19,515,291 (94.87%)

Table 2 shows the statistics on the HEC data fields as received from the VA.

Table 2. Input Validity Statistics/Field Population on the received HEC file (Only for successful Identity Resolution Responses).

	SSN	First Name	Middle Name	Last Name	DOB
Empty	177,466 (1.00%)	32,355 (0.18%)	4,080,681 (22.88%)	25,456 (0.14%)	271,631 (1.52%)
Partial	17,422 (0.10%)	92,424 (0.52%)	7,143,758 (40.06%)	760 (0.00%)	0 (0.00%)
Frivolous	455 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Invalid	5,038 (0.03%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	33,724 (0.19%)
Valid	17,632,015 (98.88%)	17,707,617 (99.30%)	6,607,957 (37.06%)	17,806,180 (99.85%)	17,527,041 (98.29%)

In the tables above we use the following definitions:

- **Empty** – no data in that field, partial or otherwise.
- **Partial** – input value is not complete, for example, only four digits in SSN, missing month or year, middle initial only.
- **Frivolous** – field value is recognized as a known frivolous value, for example, SSN is 999999999 or 123456789.
- **Invalid** – out of range for Date of Birth or a never-issued SSN.
- **Valid** – Field value appears good.

The first entry in each cell of Table 1 and Table 2 is the number of records with that characteristic of that field, and the second number in each cell, in parenthesis, is the percent of all records that have that characteristic, rounded to the accuracy shown.

2.2 Identification of Likely Deceased Records in the BIRLS & HEC Records

It is important to understand the records that exist within the BIRLS & HEC records as we determine relative risk of each consumer. Thus it is important that we use our ID Network visibility to try to better understand who in the BIRLS & HEC records may actually be listed as deceased on the SSA data files. We use the SSA Death Master File to determine which SSNs are reported to be associated with deceased individuals. Table 3 shows the results of our analysis of the number of BIRLS records where we can attach a possible death indicator. Table 4 does the same for the HEC records.

Table 3. Decease flag distribution across BIRLS records

Deceased Flag	# of Records	% of BIRLS Records
N – No	19,448,407	94.86%
U – SSN Only	64,671	0.32%
Y – Deceased	989,595	4.83%

In this table we find 989,595 BIRLS records for people whom we strongly believe to be deceased. For these records we find a match on the SSN, Name and DOB fields from the BIRLS records to the SSA Death Master File. We also see 64,671 unknown deceased BIRLS records where the BIRLS SSN is on the SSA Death Master File but the BIRLS name DOB are not.

Table 4. Decease flag distribution across HEC records

Deceased Flag	# of Records	% of BIRLS Records
N – No	12,354,034	69.28%
U – SSN Only	70,688	0.40%
Y – Deceased	5,230,216	29.33%
E – Empty SSN	177,467	1.00%

In this table we find 5,230,216 HEC records for people whom we strongly believe to be deceased. For these records we find a match on the SSN, Name and DOB fields from the HEC records to the SSA Death Master File. We also see 70,688 unknown deceased HEC records where the HEC SSN is on the SSA Death Master File but the HEC name DOB are not. It is important to note that nearly 1/3 of the HEC file provided to ID Analytics linked to deceased records. In follow up analysis below the deceased records may be withdrawn from consideration.

2.3 Identity Resolution Results

2.3.1 Identity Resolution Status and Confidence

Table 5 shows the resolution status BIRLS records as received from the VA.

Table 5. Resolution Status for BIRLS Records

Resolution Status	Count	% of Records	Cumulative % of Records
R – Resolved	19,184,548	93.57%	93.6%
N – New	1,054,972	5.15%	98.7%
A - Ambiguous	241,369	1.18%	99.9%
I – Insufficient	21,674	0.11%	100%
F – Frivolous	4	0.00%	100%
Total	20,502,567	100.00%	

Table 6 shows the resolution status HEC records as received from the VA.

Table 6. Resolution Status for HEC Records

Resolution Status	Count	% of Records	Cumulative % of Records
R – Resolved	15,765,201	88.41%	88.4%
N – New	1,722,873	9.66%	98.1%
I – Insufficient	202,390	1.13%	99.2%
A - Ambiguous	141,275	0.79%	100%
F – Frivolous	657	0.00%	100%
Total	17,832,396	100.00%	

Table 7 below shows the distribution of the number of records that were resolved at what confidence values. The first column shows the confidence value (in .05 bins), the second shows the number resolved at that confidence level, and the third column is the percent of records resolved at that confidence level.

Table 7. Identity Resolution Confidence Value for the Resolved BIRLS Records

Identity Resolution Confidence Value	BIRLS Records Count	% of Total Records
0.05	937	0.00%
0.10	1,117	0.01%
0.15	177	0.00%
0.20	1,626	0.01%
0.25	3,399	0.02%
0.30	2,284	0.01%
0.35	1,329	0.01%
0.40	15,303	0.08%
0.45	28,180	0.15%
0.50	16,767	0.09%
0.55	5,571	0.03%
0.60	14,745	0.08%
0.65	63,084	0.33%
0.70	79,263	0.41%
0.75	42,838	0.22%
0.80	18,780	0.10%
0.85	259,235	1.35%
0.90	115,087	0.60%
0.95	3,470,547	18.09%
1.00	15,044,279	78.42%

Table 8 below shows the distribution of the number of records that were resolved at what confidence values. The first column shows the confidence value (in .05 bins), the second shows the number resolved at that confidence level, and the third column is the percent of records resolved at that confidence level.

Table 8. Identity Resolution Confidence Value for the Resolved HEC Records

Identity Resolution Confidence Value	BIRLS Records Count	% of Total Records
0.05	391	0.00%
0.10	472	0.00%
0.15	222	0.00%
0.20	960	0.01%
0.25	2,438	0.02%
0.30	1,115	0.01%
0.35	1,448	0.01%
0.40	10,600	0.07%
0.45	16,122	0.10%
0.50	11,340	0.07%
0.55	6,713	0.04%
0.60	14,952	0.09%
0.65	24,718	0.16%
0.70	31,065	0.20%
0.75	16,344	0.10%
0.80	2,665	0.02%
0.85	155,011	0.98%
0.90	56,389	0.36%
0.95	1,620,893	10.28%
1.00	13,791,343	87.48%

Once we have assigned an Identity Number to as many of the BIRLS and HEC records as possible, that is, all the resolved and new records, we can further examine these different populations to determine file overlap between the two files, duplicate records, and possible field replacements within each file.

2.3.2 Record Overlap within BIRLS and HEC files

Table 9 below shows the number of unique identity numbers in the BIRLS and HEC files and how many of those identity numbers exist in both files (overlap). The overlapped records link to the same identity, but keep in mind the HEC file has substantially more information about the veteran than the BIRLS file.

Table 9. Overlap between the BIRLS and HEC files

	Unique Resolved Records
BIRLS	19,094,973
HEC	15,666,464
Overlap	8,377,979

2.3.3 Duplicate Record Analysis

Table 10 below shows how often a unique resolved identity number was observed within the BIRLS file.

Table 10. Identity Resolution analysis results – How often is an identity key seen in the BIRLS records?

Times Seen	Count	% of Records	Cumulative % of Records
1	19,005,978	99.53%	99.53%
2	88,427	0.46%	100.00%
3	558	0.00%	100.00%
4	11	0.00%	100.00%

In this table we find 19,005,978 BIRLS records that appropriately map to a single unique person. We find 2 x 88,427 = 176,854 BIRLS records that map to 88,427 people, each with two BIRLS records. We find 3 x 558 = 1,674 BIRLS records that map to 558 people, each with three BIRLS records. And we find 4 x 11 = 44 BIRLS records that map to 11 people, each with 4 BIRLS records.

We thus find 111,655 people that appear to have more than one HEC record. We don't know why or how this may occur, but given that this is a medical benefit file it may be a cause for concern at the VA.

Table 11 below shows the how often a unique resolved identity number was observed within the HEC file.

Table 11. Identity Resolution analysis results – How often is an identity key seen in the HEC records?

Times Seen	Count	% of Records	Cumulative % of Records
1	17,262,741	99.36%	99.36%
2	109,993	0.63%	99.99%
3	1,469	0.01%	100.00%
4	117	0.00%	100.00%
5	45	0.00%	100.00%
6	13	0.00%	100.00%
7	10	0.00%	100.00%
8	1	0.00%	100.00%
9	2	0.00%	100.00%
10	1	0.00%	100.00%
15	2	0.00%	100.00%
16	1	0.00%	100.00%
17	1	0.00%	100.00%

In this table we find 17,262,741 HEC records that appropriately map to a single unique person. We find 2 x 109,993 = 219,986 HEC records that map to 109,993 people, each with two HEC records. We find 3 x 1,469 = 4,407 HEC records that map to 1,469 people, each with three HEC records. And we find 4 x 117 = 468 HEC records that map to 117 people, each with 4 HEC records, etc…

We thus find 111,655 people that appear to have more than one HEC record. We don't know why or how this may occur, but it is a cause for concern at the VA. We explore these duplicate records in detail in Task 1 below.

In the next section we will examine the results of the Identity Manipulation Score and attributes on the entire 19,094,973 resolved BIRLS records and 17,262,741 HEC records. After the Identity Resolution has been applied we have appended to the 19,094,973 BIRLS, and 17,262,741 HEC resolved identities the Identity Manipulations Score and attributes. We first report statistics by the Identity Manipulation attributes, each considered separately. Keep in mind that these statistics are for the identities in the BIRLS and HEC data files, but the numbers come from activity in the commercial world outside of the VA visibility.

For each resolved identity we have an indicator of how many SSNs are associated with that identity. Many people, unfortunately, have multiple SSNs associated with their identity. In

previous analysis of the entire U.S. population we found that about 6% of the U.S. population has multiple SSNs associated with their identities. About 60% of these anomalies are due to typos and about 40% are deliberate.

2.4 Recommended BIRLS Field Changes

Our commercial activity visibility allows us to see many instances of how U.S. consumers use their personal identifying information in their commercial activities. Our Identity Resolution System allows us to determine who is who, even when the PII is fragmented or confusing. This visibility and toolset allows us to examine the BIRLS & HEC records, resolve the majority of the identities, and then evaluate the fields on the BIRLS & HEC file for their likely accuracy. As a result we are able to recommend which fields could be replaced with which new values. The table below shows the summary statistics of these suggested replacements.

Table 12 & 13 below shows the number of recommended field replacements for the BIRLS & HEC files, respectively. Field replacements are only recommended for values that are non-transient and should not change: SSN and DOB.

Table 12. Summary statistics of suggested field replacements for resolved BIRLS records

Field	Suggested Replacements	% of Total Records
SSN	102	0.00%
DOB	561,240	2.74%

We find 102 SSNs, and 561,240 DOBs that we believe should be replaced with more accurate values within the BIRLS file. Certainly the SSNs and DOBs should be considered here more carefully since they, in theory, should be permanent field values.

Table 13. Summary statistics of suggested field replacements for resolved HEC records

Field	Suggested Replacements	% of Total Records
SSN	34,671	0.22%
DOB	81,773	0.52%

We find 34,671 SSNs, and 81,773 DOBs that we believe should be replaced with more accurate values within the HEC file. Certainly the SSNs and DOBs should be considered here more carefully since they, in theory, should be permanent field values.

Standard output from ID Analytics' Identity Resolution technology includes recommended values associated with a given identity. Therefore, as a result of running the BIRLS and HEC records through this service to uniquely identify individuals it was determined that certain fields in the BIRLS and HEC databases should be updated because they are either invalid or inaccurate. ID Analytics only made recommended changes to DOB, and SSN. ID Analytics only made

recommendations for replacement when there was a high confidence that the input field needed to be corrected.

2.5 Identity Manipulation Score and Attributes

The Identity Resolution System data structure is based on a complex hierarchical representation of all the identity variation we have seen these 313 million people use in their above-described commercial activities. We have detailed visibility into all the different PII variations used, and we can quantify these variations on a continuum from benign and appropriate to deliberately improper and criminal. For example, we see particular individuals deliberately using dozens of SSNs and/or DOBs as they apply for credit products.

With this unique visibility and identity resolution capability we have constructed a set of complex algorithms that examine and quantify the nature and extent of deliberate identity manipulations for each of these 313 million individuals active in the U.S using commercial products and services. We have created an "Identity Manipulation" Score (IM Score) that measures the deliberate manipulation of PII. Specifically, this score examines the variations in SSNs, DOBs, names and addresses used by each individual as they apply for credit products and services. We look for multiple uses and other nontrivial variations, thus eliminating as best as possible simple typos. This IM Score is tolerant to typos in SSN, DOBs, use of nicknames in the first name (John, Johnny, Jack…), last name changes, and common address variations. The IM Score is designed to avoid such typos and false alarms and to focus on the identity variations that are intentional and improper.

Along with this Identity Manipulation Score we also product IM attributes that give insight into what we see going on with this person in his commercial world interactions. The IM attributes are summary characteristics that we see associated with that particular person. Specifically, the IM attributes are

- **Number of SSNs** – includes possible typos.
- **Number of SSN manipulations** – excludes likely typos.
- **Number of DOBs** – includes possible typos.
- **Number of DOB manipulations** – excludes likely typos.
- **Number of SSN and DOB manipulations** – number of instances where both the SSN and DOB are simultaneously manipulated, which can be a strong indication of deliberate intent.
- **Number of first name roots** – a first name root is associated with each set of nicknames, so this attributes indicates the number of essentially different first names used (Steve, Stephen, Stevie are all considered the same first name, but Steve and Fred are different).
- **Number of last names** – number of essentially different last names, ignoring very slight variations.
- **Number of address manipulations** – counts the number of instances where we observe suspicious address variations, such as 123 main street, 123 main avenue, 123 main street unit 1.

These particular attributes are measures of the types of variations that are typically associated with fraudulent identity manipulations.

For the 1,054,972 BIRLS and 1,722,873 HEC records designated as "New", there's not too much further analysis to be done at this time. These are records that look fine as of now but we've never seen them before in ID Analytics' commercial applications. This also applies to all the records that have been identified in tables 5 and 6 as ambiguous, insufficient and frivolous records.

We now turn our attention to the 19,184,548 BIRLS and 15,765,201 HEC resolved (and not new) records. For these records we are able to assign a unique Identity Number, which is in essence a unique person label, and we have a history on these people. The tables below illustrate the number of SSNs, number of manipulated SSNs, number of DOBs, number of manipulated DOBs, number of SSN and DOB manipulations, number of first names, number of last names, and number of manipulated addresses associated with each resolved record.

Table 14. Identity Number analysis – number of SSNs linked to each individual on the BIRLS file.

Number of SSNs	Count	% of Total Records
0	500	0.00%
1	17,535,079	91.40%
2	1,477,429	7.70%
3	140,862	0.73%
4	21,194	0.11%
5	5,278	0.03%
6	1,992	0.01%
7	880	0.00%
8	481	0.00%
9	265	0.00%
10	158	0.00%
>10	430	0.00%

We see in this table that 500 BIRLS identities have no SSN associated with them in our commercial data visibility. We find that 17,535,079 or 91.40% of the people in the BIRLS file have the appropriate single SSN associated with them in our commercial data. After this we find anomalies. It's interesting to note that about 8.6% of the people in the BIRLS file have multiple SSNs associated with them, which is greater than the 6% of the U.S. This is another indication that the BIRLS file has higher than average "badness."

We see 1,477,429 people in the BIRLS file who inappropriately have 2 SSNs associated with them in our commercial data. 140,862 BIRLS people have exactly 3 SSNs in our commercial visibility. More than 588 BIRLS individuals are found to have 10 or more SSNs associated with their name in our commercial data. These are clearly not typos. A rough rule of thumb can come

from the observation that about 6% of the population has an SSN typo. The odds of having two independent SSN typos are 0.06 times 0.06 = 0.0036. Thus if a person has three or more SSNs associated with their identity it's more than 99% (1.0 – 0.0036) likely that it's deliberate.

Table 15. Identity Number analysis – number of SSNs linked to each individual on the HEC file.

Number of SSNs	Count	% of Total Records
0	481	0.00%
1	14,564,573	92.38%
2	1,080,156	6.85%
3	98,513	0.62%
4	14,941	0.09%
5	3,722	0.02%
6	1,321	0.01%
7	613	0.00%
8	301	0.00%
9	171	0.00%
10	117	0.00%
>10	292	0.00%

We see in this table that 481 HEC identities have no SSN associated with them in our commercial data visibility. We find that 14,564,573, or 92.38%, of the people in the HEC file have the appropriate single SSN associated with them in our commercial data. After this we find anomalies. It's interesting to note that about 7.6% of the people in the HEC file have multiple SSNs associated with them, which is greater than the 6% of the U.S. This is another indication that the HEC file has higher than average "badness."

We see 1,080,156 people in the HEC file who inappropriately have 2 SSNs associated with them in our commercial data. 98,513 HEC people have exactly 3 SSNs in our commercial visibility. More than 409 HEC individuals are found to have 10 or more SSNs associated with their name in our commercial data. These are clearly not typos. A rough rule of thumb can come from the observation that about 6% of the population has an SSN typo. The odds of having two independent SSN typos are 0.06 times 0.06 = 0.0036. Thus if a person has three or more SSNs associated with their identity it's more than 99% (1.0 – 0.0036) likely that it's deliberate.

Next we examine a more complex Identity Manipulation attribute, the "Number of SSN manipulations" attribute. This attribute attempts to remove typos and focuses on deliberate SSN variations. We do this by examining the nature of the differences in the SSNs and the frequency of the observations of the different SSNs associated with a single person. For example, if we see a single digit different and only on one occurrence/event, we will call that a likely typo.

The following table and figure show the statistics for the number of SSN manipulations on BIRLS and HEC identities, as seen in our commercial data.

Table 16. Identity Number analysis – number of manipulated SSNs linked to a BIRLS identity, as seen in the commercial world.

Number of Manipulated SSNs	Count	% of Total Records
0	18,492,400	96.39%
1	645,722	3.37%
2	36,816	0.19%
3	5,934	0.03%
4	1,786	0.01%
5	808	0.00%
6	391	0.00%
7	209	0.00%
8	126	0.00%
9	98	0.00%
10	48	0.00%
>10	210	0.00%

Here we see that about 96.4% of the identities in the BIRLS data have the appropriate zero manipulated SSNs associated with their identity. That leaves 3.6%, or close to 700,000 people on the BIRLS files who are seen to have inappropriately manipulated their SSNs outside of the VA as they apply for credit and other products and services.

Table 17. Identity Number analysis – number of manipulated SSNs linked to a HEC identity, as seen in the commercial world.

Number of Manipulated SSNs	Count	% of Total Records
0	15,250,885	96.74%
1	480,695	3.05%
2	26,961	0.17%
3	4,207	0.03%
4	1,223	0.01%
5	516	0.00%
6	237	0.00%
7	139	0.00%
8	87	0.00%
9	71	0.00%
10	38	0.00%
>10	142	0.00%

Here we see that about 96.7% of the identities in the HEC data have the appropriate zero manipulated SSNs associated with their identity. That leaves 3.4% or more than 500,000 people on the HEC file that are seen to have inappropriately manipulated their SSNs outside of the VA as they apply for credit and other products and services.

Next we continue the same analysis for the number of DOBs for each identity on the BIRLS & HEC files.

Table 18 shows the number of different Dates of Birth on each identity in the BIRLS file.

Table 18. Identity Number analysis – number of DOBs linked to a BIRLS identity, as seen in the commercial world.

Number of DOBs	Count	% of Total Records
0	1,554,406	8.10%
1	15,232,086	79.40%
2	2,143,357	11.17%
3	221,128	1.15%
4	26,664	0.14%
5	4,781	0.02%
6	1,224	0.01%
7	468	0.00%
8	184	0.00%
9	96	0.00%
10	57	0.00%
>10	97	0.00%

We find about 8% of the resolved records have no DOB in our commercial data, about 79% have the appropriate single DOB, and about 13% inappropriately have more than one DOB associated with their identity as seen in our commercial data. Keep in mind that these include data entry errors.

Table 19 shows the number of different Dates of Birth on each identity in the BIRLS file.

Table 19. Identity Number analysis – number of DOBs linked to a HEC identity, as seen in the commercial world.

Number of DOBs	Count	% of Total Records
0	2,019,601	12.81%
1	12,032,425	76.32%
2	1,534,911	9.74%
3	154,219	0.98%
4	19,067	0.12%
5	3,427	0.02%
6	867	0.01%
7	351	0.00%
8	137	0.00%
9	72	0.00%
10	43	0.00%
>10	81	0.00%

We find about 13% of the resolved records have no DOB in our commercial data, about 76% have the appropriate single DOB, and about 11% inappropriately have more than one DOB associated with their identity as seen in our commercial data. Keep in mind that these include data entry errors.

Similar to the SSN analysis, we have constructed a special identity manipulation attribute that attempts to remove the many typos and other data entry errors around the Date of Birth. Table 20 and 21 show the statistics around this “DOB Manipulation” attribute.

Table 20. Identity Number analysis – number of manipulated DOBs linked to a BIRLS identity, as seen in the commercial world.

Number of Manipulated DOBs	Count	% of Total Records
0	17,270,957	90.03%
1	1,768,530	9.22%
2	129,714	0.68%
3	12,123	0.06%
4	2,104	0.01%
5	587	0.00%
6	250	0.00%
7	121	0.00%
8	59	0.00%
9	38	0.00%
10	21	0.00%
>10	44	0.00%

Using the "Number of Manipulated DOBs" attribute we find about 9.2% of the BIRLS identities have improper DOBs associated with their identities as seen in our commercial data.

Table 21. Identity Number analysis – number of manipulated DOBs linked to a HEC identity, as seen in the commercial world.

Number of Manipulated DOBs	Count	% of Total Records
0	14,388,215	91.27%
1	1,274,511	8.08%
2	91,398	0.58%
3	8,709	0.06%
4	1,531	0.01%
5	412	0.00%
6	200	0.00%
7	100	0.00%
8	44	0.00%
9	25	0.00%
10	21	0.00%
>10	35	0.00%

Using the "Number of Manipulated DOBs" attribute we find about 8.1% of the HEC identities have improper DOBs associated with their identities as seen in our commercial data.

Another Identity Manipulation attribute looks for very strong evidence of deliberate identity manipulation, and that is the "Number of SSN and DOB Manipulations" attribute. Here we look for events in the commercial world where both the DOB and the SSN are different from what is expected. It is much less likely that this simultaneous difference comes from typos.

Table 22 and 23 show the statistics around this strong deliberate manipulation indicator.

Table 22. Identity Number analysis – number of manipulated SSN and DOB combinations linked to a BIRLS identity, as seen in the commercial world.

Number of Manipulated SSN and DOB Combinations	Count	% of Total Records
0	19,060,442	99.35%
1	113,946	0.59%
2	7,881	0.04%
3	1,397	0.01%
4	441	0.00%
5	175	0.00%
6	91	0.00%
7	51	0.00%
8	28	0.00%
9	28	0.00%
10	24	0.00%
>10	44	0.00%

Table 23. Identity Number analysis – number of manipulated SSN and DOB combinations linked to a HEC identity, as seen in the commercial world.

Number of Manipulated SSN and DOB Combinations	Count	% of Total Records
0	15,677,877	99.45%
1	80,359	0.51%
2	5,400	0.03%
3	962	0.01%
4	286	0.00%
5	127	0.00%
6	54	0.00%
7	32	0.00%
8	27	0.00%
9	22	0.00%
10	19	0.00%
>10	36	0.00%

Next we examine the BIRLS and HEC identities for name anomalies, again looking at our commercial data and not the input record name fields.

In our Identity Resolution System we have incorporated a nickname table that allows the linking and connecting of many common nicknames. For example, we have a class of names for Kathy that include Cathy, Kathleen, Cathie… and many other variations of this first name class. For each class we identify the most commonly used first name and we assign that as the root of that first name class (Kathy, for this example class). We only consider a first name to be different when it is outside of the first name class, thus we don't count common nicknames to be different first names.

In Table 24 and 25 we show the number of first name roots associated with the BIRLS and HEC identities, respectively, as seen in our commercial data. Generally, we would expect each person to use one and only one first name root, which allows for many different first name variations within nicknames or common different spellings of the "same" first name.

Table 24. Identity Number analysis – number of root first names linked to a BIRLS identity, as seen in the commercial world.

Number of Root First Names	Count	% of Total Records
0	7,014	0.04%
1	17,477,495	91.10%
2	1,603,322	8.36%
3	91,361	0.48%
4	4,915	0.03%
5	364	0.00%
6	49	0.00%
7	18	0.00%
8	4	0.00%
9	3	0.00%
10	1	0.00%
>10	2	0.00%

Table 25. Identity Number analysis – number of root first names linked to a HEC identity, as seen in the commercial world.

Number of Root First Names	Count	% of Total Records
0	27,891	0.18%
1	14,298,195	90.69%
2	1,356,379	8.60%
3	78,028	0.49%
4	4,324	0.03%
5	306	0.00%
6	45	0.00%
7	24	0.00%
8	2	0.00%
9	1	0.00%
10	3	0.00%
>10	3	0.00%

In the same way we examine the number of different last names observed in the commercial data that are used by the people on the BIRLS and HEC files. Last names differences are much different than first name differences because, in the U.S in particular, many people undergo last name changes, sometimes multiple times. Thus we need to be much more tolerant to last name variations in our examinations and investigations, and our Identity Manipulation Score takes this fact into account.

Table 26 and 27 show the statistics around the number of different last names used by the identities on the BIRLS and HEC files in their commercial activities.

Table 26. Identity Number analysis – number of last names linked to a BIRLS identity, as seen in the commercial world.

Number of Last Names	Count	% of Total Records
0	1,317	0.01%
1	17,436,441	90.89%
2	1,491,265	7.77%
3	217,627	1.13%
4	32,953	0.17%
5	4,293	0.02%
6	543	0.00%
7	82	0.00%
8	10	0.00%
9	4	0.00%
10	7	0.00%
>10	6	0.00%

Table 27. Identity Number analysis – number of last names linked to a HEC identity, as seen in the commercial world.

Number of Last Names	Count	% of Total Records
0	948	0.01%
1	14,465,863	91.76%
2	1,112,002	7.05%
3	158,646	1.01%
4	23,897	0.15%
5	3,296	0.02%
6	443	0.00%
7	62	0.00%
8	17	0.00%
9	10	0.00%
10	7	0.00%
>10	10	0.00%

Our final Identity Manipulation attribute is the "Number of Address manipulations." Here we look for and count the occurrences of systematic and frequently slight variations in an address that is typical of deliberate address obfuscation, but as much as possible tolerant to typos. In our mainline fraud work we see people making slight variations in addresses to avoid address matching algorithms but in ways that will still allow physical delivery of items in the mail. Sometimes this is done through the addition of unnecessary secondary unit designators (apt, unit, floor, #...); sometimes it is small changes in street address number, for instance to a neighbor's house. In these cases usually the mail is properly delivered but address matching analysis won't see these anomalies.

Table 28 and 29 show the statistics around address manipulation.

Table 28. Identity Number analysis – Number of manipulated addresses linked to a BIRLS identity, as seen in the commercial world.

Number of Manipulated Addresses	Count	% of Total Records
0	14,561,082	75.90%
1	3,583,697	18.68%
2	790,938	4.12%
3	183,448	0.96%
4	46,829	0.24%
5	12,792	0.07%
6	3,812	0.02%
7	1,245	0.01%
8	377	0.00%
9	178	0.00%
10	59	0.00%
>10	91	0.00%

We find about 24% of the BIRLS identities have possible address manipulations in their commercial activities. Many of these are likely benign, but high numbers of observed manipulations is a reasonable additional fraud indicator.

Table 29. Identity Number analysis – Number of manipulated addresses linked to a HEC identity, as seen in the commercial world.

Number of Manipulated Addresses	Count	% of Total Records
0	12,125,045	76.91%
1	2,838,235	18.00%
2	611,891	3.88%
3	140,190	0.89%
4	35,533	0.23%
5	9,833	0.06%
6	2,932	0.02%
7	965	0.01%
8	327	0.00%
9	118	0.00%
10	52	0.00%
>10	80	0.00%

We find about 23% of the HEC identities have possible address manipulations in their commercial activities. Many of these are likely benign, but high numbers of observed manipulations is a reasonable additional fraud indicator.

Lastly we examine the Identity Manipulation Score on the resolved BIRLS and HEC populations. Table 30 and 31 show the Identity Manipulation Score distribution on all the resolved BIRLS and identities, respectively. Generally, Identity Manipulation Scores above 800 are believed to be deliberate and inappropriate manipulators. ***We find about 10% of the BIRLS population to be above this level of badness, almost 2 million people and about 8% of the HEC population, more than 1.3 million people.***

Table 30. Identity Manipulation Score on the resolved BIRLS identities

Identity Manipulation Score Bin	BIRLS People Count	% of Total Records
25	10,486,901	54.66%
75	179,627	0.94%
100	1,123,226	5.85%
125	218,478	1.14%
150	256,958	1.34%
175	26,805	0.14%

200	62,161	0.32%
225	49,997	0.26%
250	272,593	1.42%
275	91,551	0.48%
300	118,464	0.62%
325	178,304	0.93%
350	168,507	0.88%
375	223,821	1.17%
400	242,239	1.26%
425	155,359	0.81%
450	221,501	1.15%
475	263,246	1.37%
500	459,100	2.39%
525	270,254	1.41%
550	203,723	1.06%
575	224,753	1.17%
600	299,452	1.56%
625	323,296	1.69%
650	241,323	1.26%
675	143,536	0.75%
700	186,464	0.97%
725	172,994	0.90%
750	197,066	1.03%
775	216,771	1.13%
800	199,417	1.04%
825	235,660	1.23%
850	281,888	1.47%
875	271,184	1.41%
900	269,706	1.41%
925	234,130	1.22%
950	172,068	0.90%
975	119,839	0.62%
1000	122,186	0.64%

Table 31. Identity Manipulation Score on the resolved HEC identities

Identity Manipulation Score Bin	BIRLS People Count	% of Total Records
25	9,054,083	57.43%
75	164,164	1.04%
100	1,045,297	6.63%
125	163,650	1.04%
150	280,948	1.78%
175	24,282	0.15%
200	55,156	0.35%
225	41,361	0.26%
250	234,421	1.49%
275	69,590	0.44%
300	86,024	0.55%
325	152,416	0.97%
350	123,858	0.79%
375	166,718	1.06%
400	187,441	1.19%
425	121,986	0.77%
450	167,286	1.06%
475	192,057	1.22%
500	348,987	2.21%
525	187,104	1.19%
550	146,216	0.93%
575	163,952	1.04%
600	213,889	1.36%
625	242,377	1.54%
650	164,709	1.04%
675	109,378	0.69%
700	135,250	0.86%
725	125,001	0.79%
750	139,031	0.88%
775	152,561	0.97%
800	140,576	0.89%
825	160,547	1.02%

850	192,949	1.22%
875	183,442	1.16%
900	182,262	1.16%
925	159,719	1.01%
950	119,656	0.76%
975	84,365	0.54%
1000	82,492	0.52%

3. Data Defense Analysis of HEC File

3.1 Executive Overview

ID Analytics has completed the first Data Defense analysis on the Veterans Affairs (VA) member population that exists in the HEC data file. This HEC file was provided to ID Analytics in November 2012 and contained 17,832,405 Veterans Affairs (VA) members. After thoroughly reviewing the file for elements of organized misuse we believe there are fifty-seven (57) new identities potentially being misused across seven (7) new independent cases. Fully detailed and encrypted information pertaining to these fifty-seven (57) identities can be transferred to the VA for further analysis and investigative purposes upon request. The file contains the following application information for VA members involved in potential organized misuse as it relates to the VA:

- Social Security number
- Segment (e.g. Bank Card, Wireless, Retail)
- Application Date (YYYYMMDD)
- Name (Last, First)
- Address
- Zip Code
- Home Phone (potentially a mobile)
- Work Phone (potentially a mobile)
- E-mail
- SSN Issue Year
- SSN Flag (e.g. "Deceased")
- Date of Birth

ID Analytics for Data Defense proactively monitors the Veterans Affairs (VA) data and detects potential misuse by internal and external parties. By analyzing approximately 17.8 million VA records as well as 740 billion aggregated data elements and 2.9 million reported identity frauds within the ID Network®, the nation's only real-time, cross-industry compilation of identity information, ID Analytics for Data Defense develops an integrated view of each VA member's identity characteristics and their connectedness to other VA members. We call this capability Personal Topology™ - a source of identity intelligence that helps leading communication and financial service companies, as well as retailers, government agencies, and health insurers safeguard their business every single day.

The following report outlines each case of anomalous application activity as it relates to each suspicious address and phone number detected by ID Analytics for Data Defense technology. No information pertaining to SSNs, DOBs, account information or card numbers is included within this report.

3.2 Case Analysis Overview

Each case of potential organized misuse is outlined in detail within this report and includes major elements of interest (EOI) such as addresses, home phone numbers, e-mail addresses and

geographic locations potentially used to perpetrate identity fraud. Full address, phone number, and e-mail information is included within the encrypted data file provided separately and is specific to impacted VA members for privacy purposes. Within the report each case is divided into the following sections:

- Case Number
- Activity Statistics (e.g. industries targeted, VA SSNs involved in activity)
- Major Identity EOIs (e.g. phone numbers, addresses, e-mails)
- Potential Indicators of Misuse
- Geographic EOIs (as applicable)

The Data Defense technology identifies addresses, phone numbers, email addresses and SSNs that have been potentially misused in an organized way. These identity elements are stored in the ID Network and come from millions of applications for credit, wireless subscriptions/service and other types of consumer requests for lending vehicles. When fraudsters try to open new accounts using stolen data they leave tell-tale behavioral patterns across the ID Network that ID Analytics technology can quickly identify.

When analyzing the VA data file, ID Analytics for Data Defense technology must take into account the differing dynamics of the veteran population when compared to the US population as a whole. As a group, veterans tend to live and work together at a much higher rate than the overall US consumer population. For instance, manufacturing plants often times have a high number of veterans working at the building. Because of this, there are falsely anomalous linkage patterns that result from these locations because of increased credit and wireless application activity. In reality, these linkages are driven from non-suspicious credit application activity where veterans are either applying for credit at the location, or providing similar work numbers on applications to verify employment. This is also the case with VA hospitals and other VA employment locations, homeless shelters and military facilities. Over the course of the last two years, ID Analytics has developed algorithms that automatically take these locations into account when applying the technology to veterans' identities, thereby automatically reducing false suspects from manual analyst review.

When possible, ID Analytics provides potential locations where the source of the data theft might have occurred. Previous proprietary ID Analytics analysis shows that employees who steal data generally abuse it within a twenty (20) mile radius of the data theft location. Because the veteran population is so large – potentially 20% of the credit active US population – ID Analytics employs "identity ratios" when analyzing cases of potential organized misuse. These identity ratios help pinpoint cases of potential organized misuse related to the VA versus fraud that occurs to the general US consumer population.

For reference, ID Analytics retains the previous six (6) months of inactive cases from prior Data Defense analyses. Each of these cases are labeled as "Inactive" and do not have recent activity. In the event an inactive case becomes active again, ID Analytics will include fully updated information pertaining to new activity.

3.3 GLOSSARY OF TERMS

Element of Interest (EOI): Address, phone numbers, e-mail addresses and geographic locations deemed interesting due to their relationships with suspicious activity within a particular case. These may include, but are not limited to, potential employee home addresses, work phone numbers, a group of homes in close proximity to VA associated facilities (e.g. hospitals, military bases) or an e-mail address linked to multiple VA identities.

Organized Misuse: Harm attributed to a particular data breach event or fraudulent activity associated specifically with identities included in the analysis. ID Analytics for Data Defense isolates potential harm as it directly relates to the VA and filters out fraud occurring to the general US consumer population.

Identity Ratio: In order to accurately determine whether or not suspicious application activity is associated with the VA, ID Analytics reviews the expected distribution of veterans within a particular geographic region and at specific addresses, phone numbers, and e-mail addresses utilized throughout the activity.

Data Defense (DD) Score: A score used to determine the level of harm associated with a case of potential misuse. The score can be used to rank order cases of potential misuse and measure the level of harm as it relates to other instances of misuse.

Confirmed Fraud: Applications that have been reported as fraud by ID Analytics' clients. The ID Network contains the largest identity fraud database within the United States with over two (2) million reported frauds.

3.4 Overview of Activity Statistics

Case[1]	City, State	DD Score[2]	VA IDs (#)	VA IDs (%)	VA Apps	VA Apps (%)	Last Active[3]	Status[4]
1212-001	West Palm Beach,	13.066	14	93.33%	15	93.75%	10/2/2012	New
1212-002	San Jose, CA	7.000	7	100.00%	11	100.00%	8/29/2012	New
1212-003	Elizabeth, NJ	3.000	6	50.00%	21	31.34%	10/15/2012	New
1212-004	North Hollywood, CA	2.314	9	25.71%	22	36.67%	9/14/2012	New
1212-005	Brandon, MS	2.250	6	37.50%	10	30.30%	10/8/2012	New
1212-006	Forrest City, AR	1.829	8	22.86%	17	26.98%	10/26/2012	New
1212-007	Lakeside Park, KY	1.485	7	21.21%	14	24.56%	9/5/2012	New

[1] First 4 digits represent month-year (MM-YY) of the report. Last 3 digits are the case number.
[2] Data Defense score is calculated as: (Identity Ratio) x (Total Number of IDs)
[3] Represents the date of the last application associated with suspicious activity
[4] Cleared: new evidence indicates low risk; Inactive: no new activity since the last report; New: Recent activity indicating high risk

3.5 Case Specific Review

Case Number: 1212-001 – West Palm Beach, FL (New)

Activity Statistics – West Palm Beach, FL (New)

Overall Activity	Details
Data Defense Score	13.066
Number of VA SSNs	14
Total Number of SSNs	15
Identity Ratio	93.33%
Number of VA Applications	15
Total Number of Applications	16

Industry	Applications	% of Total Activity
Bank Card	0	0.00%
Retail	0	0.00%
Wireless	15	93.75%
Other	1	6.25%

ID Analytics Client	Applications	% of Total Activity
Wireless Provider 6	15	93.75%
Other	1	6.25%

Major Identity Elements of Interest – West Palm Beach, FL (New)

Addresses (EOI)	VA IDs	Total IDs	VA Apps	Total Apps (%)
11322 54TH ST N, 33411	11	12	11	75.00%
3200 OLD BOYNTON RD, 33436	4	4	4	25.00%

Home Phone (EOI)	VA IDs	Total IDs	VA Apps	Total Apps (%)
561-333-3333	4	4	4	25.00%
561-444-4444	2	2	2	12.50%

E-Mail (EOI)	VA IDs	Total IDs	VA Apps	Total Apps (%)
Not Applicable	-	-	-	-

Potential Indicators of Misuse – West Palm Beach, FL (New)

- VA members have no history associated with addresses and home phone numbers linked to this activity
- Application velocity significant beginning in August 2012
  - Eleven (11) applications submitted over one (1) day
- Both address EOIs link to oil change facilities, and the phone number EOIs appear to be fake
  - A fraudster may do this to limit their ability of getting caught
- The large majority of the VA members identities likely being victimized here link to deceased SSNs
  - Fraudster may have a list of deceased veteran information provided that ~93% of the activity links to veterans
- Activity focuses primarily on wireless providers

Case Number: 1212-002 – San Jose, CA (New)

Activity Statistics – San Jose, CA (New)

Overall Activity	Details
Data Defense Score	7.000
Number of VA SSNs	7
Total Number of SSNs	7
Identity Ratio	100.00%
Number of VA Applications	11
Total Number of Applications	11

Industry	Applications	% of Total Activity
Bank Card	11	100.00%
Retail	0	0.00%
Wireless	0	0.00%
Other	0	0.00%

ID Analytics Client	Applications	% of Total Activity
Bank Card 1	5	45.45%
Bank Card 2	1	9.09%
Bank Card 4	5	45.45%

Major Identity Elements of Interest – San Jose, CA (New)

Addresses (EOI)	VA IDs	Total IDs	VA Apps	Total Apps (%)
800 HILLSDALE AVE APT 631, 95136	2	2	6	54.55%
800 HILLSDALE AVE, 95136	5	5	5	45.45%

Home Phone (EOI)	VA IDs	Total IDs	VA Apps	Total Apps (%)
408-710-0110	6	6	10	90.91%
408-710-0100	1	1	1	9.09%

E-Mail (EOI)	VA IDs	Total IDs	VA Apps	Total Apps (%)
CKRAMERXXXX@GMAIL.COM	4	4	4	36.36%
TOOLSHEXXXX@GMAIL.COM	2	2	2	18.18%

Potential Indicators of Misuse – San Jose, CA (New)

- VA members have no history associated with addresses and home phone numbers linked to this activity
- E-mail address EOIs (CKRAMERXXXX@GMAIL.COM & TOOLSHEXXXX@GMAIL.COM) are shared across multiple identities involved in the activity
  - Uncommon for individuals to share personal e-mail accounts
- Application velocity significant beginning in July 2012
  - Nine (9) applications submitted over two (2) days
- Activity focuses primarily on bank card issuers

Case Number: 1212-003 – Elizabeth, NJ (New)

Activity Statistics – Elizabeth, NJ (New)

Overall Activity	Details
Data Defense Score	3.000
Number of VA SSNs	6
Total Number of SSNs	12
Identity Ratio	50.00%
Number of VA Applications	21
Total Number of Applications	67

Industry	Applications	% of Total Activity
Bank Card	3	4.48%
Retail	1	1.49%
Wireless	56	83.58%
Other	7	10.45%

ID Analytics Client	Applications	% of Total Activity
Bank Card 1	1	1.49%
Bank Card 4	2	2.99%
Retail Card 2	1	1.49%
Wireless Provider 3	11	16.42%
Wireless Provider 5	3	4.48%
Wireless Provider 6	42	62.69%
Other	7	10.45%

Major Identity Elements of Interest – Elizabeth, NJ (New)

Addresses (EOI)	VA IDs	Total IDs	VA Apps	Total Apps (%)
37 ATLANTIC ST, 07206	6	12	21	100.00%

Home Phone (EOI)	VA IDs	Total IDs	VA Apps	Total Apps (%)
Not Applicable	-	-	-	-

E-Mail (EOI)	VA IDs	Total IDs	VA Apps	Total Apps (%)
Not Applicable	-	-	-	-

Potential Indicators of Misuse – Elizabeth, NJ (New)

- VA members have no history associated with addresses and home phone numbers linked to this activity
- One identity appears to be utilizing VA member SSNs in the creation of synthetic identities
- Address EOI (37 Atlantic St) is a single family home located in Elizabeth, NJ
  - Uncommon for twelve (12) individuals to be living in a single family home
- Activity focuses primarily on wireless providers

Case Number: 1212-004 – North Hollywood, CA (New)

Activity Statistics – North Hollywood, CA (New)

Overall Activity	Details
Data Defense Score	2.314
Number of VA SSNs	9
Total Number of SSNs	35
Identity Ratio	25.71%
Number of VA Applications	22
Total Number of Applications	60

Industry	Applications	% of Total Activity
Bank Card	2	3.33%
Retail	2	3.33%
Wireless	56	93.33%
Other	0	0.00%

ID Analytics Client	Applications	% of Total Activity
Bank Card 2	2	3.33%
Retail Card 2	2	3.33%
Wireless Provider 3	12	20.00%
Wireless Provider 5	9	15.00%
Wireless Provider 6	35	58.33%

Major Identity Elements of Interest – North Hollywood, CA (New)

Addresses (EOI)	VA IDs	Total IDs	VA Apps	Total Apps (%)
5448 HERMITAGE AVE, 91607	2	6	6	25.00%
3628 W 106TH ST, 90303	1	1	8	13.33%

Home Phone (EOI)	VA IDs	Total IDs	VA Apps	Total Apps (%)
469-243-1179	8	30	17	76.67%

E-Mail (EOI)	VA IDs	Total IDs	VA Apps	Total Apps (%)
Not Applicable	-	-	-	-

Potential Indicators of Misuse – North Hollywood, CA (New)

- VA members have no history associated with addresses and home phone numbers linked to this activity
- Home phone EOI (469-243-1179) is an unlisted cell phone in Dallas, TX and is shared by thirty (30) different individuals
  - Uncommon for multiple people to share a cell phone
- Application velocity significant beginning in June 2012
  - Twenty-three (23) applications submitted over eight (8) days
- One (1) of the applications associated with this activity has been reported as confirmed frauds by ID Analytics clients
- Activity focuses primarily on wireless providers

Case Number: 1212-005 – Brandon, MS (New)

Activity Statistics – Brandon, MS (New)

Overall Activity	Details
Data Defense Score	2.250
Number of VA SSNs	6
Total Number of SSNs	16
Identity Ratio	37.50%
Number of VA Applications	10
Total Number of Applications	33

Industry	Applications	% of Total Activity
Bank Card	7	21.21%
Retail	4	12.12%
Wireless	19	57.58%
Other	3	9.09%

ID Analytics Client	Applications	% of Total Activity
Bank Card 1	2	6.06%
Bank Card 2	5	15.15%
Retail Card 2	4	12.12%
Wireless Provider 5	2	6.06%
Wireless Provider 6	17	51.52%
Other	3	9.09%

Major Identity Elements of Interest – Brandon, MS (New)

Addresses (EOI)	VA IDs	Total IDs	VA Apps	Total Apps (%)
1027 WINDROSE DR, 39047	6	16	10	100.00%

Home Phone (EOI)	VA IDs	Total IDs	VA Apps	Total Apps (%)
601-382-3950	1	8	1	54.55%
601-382-2960	2	4	6	24.24%
601-382-7915	3	4	3	12.12%

E-Mail (EOI)	VA IDs	Total IDs	VA Apps	Total Apps (%)
CO****CK101@GMAIL.COM	2	3	4	15.15%
KATDADDY20XXXX@GMAIL.COM	2	3	2	9.09%

Potential Indicators of Misuse – Brandon, MS (New)

- VA members have no history associated with addresses and home phone numbers linked to this activity
- E-mail address EOIs (CO****CK101@GMAIL.COM & KATDADDY20XXXX@GMAIL.COM) are shared across multiple identities involved in the activity
  - Uncommon for individuals to share personal e-mail accounts
- Address EOI (1027 Windrose Dr) is a single family homes located in Brandon, MS
  - Uncommon for sixteen (16) individuals to be living in a single family home
- Four (4) of the applications associated with this activity have been reported as confirmed frauds by ID Analytics clients
- Application velocity significant beginning in December 2011
  - Fourteen (14) applications submitted over three (3) days
- Activity focuses primarily on wireless providers

Case Number: 1212-006 – Forrest City, AR (New)

Activity Statistics – Forrest City, AR (New)

Overall Activity	Details
Data Defense Score	1.829
Number of VA SSNs	8
Total Number of SSNs	35
Identity Ratio	22.86%
Number of VA Applications	17
Total Number of Applications	63

Industry	Applications	% of Total Activity
Bank Card	0	0.00%
Retail	1	1.59%
Wireless	60	95.24%
Other	2	3.17%

ID Analytics Client	Applications	% of Total Activity
Retail Card 2	1	1.59%
Wireless Provider 3	8	12.70%
Wireless Provider 5	11	17.46%
Wireless Provider 6	41	65.08%
Other	2	3.18%

Major Identity Elements of Interest – Forrest City, AR (New)

Addresses (EOI)	VA IDs	Total IDs	VA Apps	Total Apps (%)
424 W BUFORD ST, 72335	6	20	12	57.14%

Home Phone (EOI)	VA IDs	Total IDs	VA Apps	Total Apps (%)
870-270-6527	6	25	9	66.67%
870-270-2868	4	14	8	33.33%

E-Mail (EOI)	VA IDs	Total IDs	VA Apps	Total Apps (%)
Not Applicable	-	-	-	-

Potential Indicators of Misuse – Forrest City, AR (New)

- VA members have no history associated with addresses and home phone numbers linked to this activity
- Address EOI (424 W Buford St) is a single family homes located in Forrest City, AR
  - Uncommon for twenty (20) individuals to be living in a single family home
- Date of births supplied on applications do not match the date of births submitted on previous applications associated with these identities
- Home phone EOIs (870-270-6527 & 870-270-2868) are unlisted cell phones in Forrest City, AR and are shared by at least fourteen (14) different individuals
  - Uncommon for multiple people to share a cell phone
- Application velocity significant beginning in March 2012
  - Twenty-one (21) applications submitted over four (4) days
- Activity focuses primarily on wireless providers

Case Number: 1212-007 – Lakeside Park, KY (New)

Activity Statistics – Lakeside Park, KY (New)

Overall Activity	Details
Data Defense Score	1.485
Number of VA SSNs	7
Total Number of SSNs	33
Identity Ratio	21.21%
Number of VA Applications	14
Total Number of Applications	57

Industry	Applications	% of Total Activity
Bank Card	9	15.79%
Retail	1	1.75%
Wireless	36	63.16%
Other	11	19.30%

ID Analytics Client	Applications	% of Total Activity
Bank Card 1	9	15.79%
Retail Card 1	1	1.75%
Wireless Provider 3	1	1.75%
Wireless Provider 5	18	31.58%
Wireless Provider 6	17	29.82%
Other	11	19.30%

Major Identity Elements of Interest – Lakeside Park, KY (New)

Addresses (EOI)	VA IDs	Total IDs	VA Apps	Total Apps (%)
313 ROCK CRYSTAL LN, 41017	4	19	10	64.91%

Home Phone (EOI)	VA IDs	Total IDs	VA Apps	Total Apps (%)
661-800-7028	4	17	4	35.09%
513-295-2784	0	4	0	15.79%

E-Mail (EOI)	VA IDs	Total IDs	VA Apps	Total Apps (%)
Not Applicable	-	-	-	-

Potential Indicators of Misuse – Lakeside Park, KY (New)

- VA members have no history associated with addresses and home phone numbers linked to this activity
- Six (6) of the applications associated with this activity have been reported as confirmed frauds by ID Analytics clients
- Home phone EOI (661-800-7028) is an unlisted cell phone in Bakersfield, CA and is shared by seventeen (17) different individuals
  - Uncommon for multiple people to share a cell phone
- Application velocity significant beginning in August 2012
  - Eighteen (18) applications submitted over eleven (11) days
- Activity focuses primarily on wireless providers

3.6 Next Steps and Remediation Recommendations

ID Analytics understands the need to provide remediation and follow-up for VA members associated with activity provided in this report. In order to fully understand the breadth and scope of potential misuse, ID Analytics recommends Veterans Affairs pursue the following actions:

- Review VA log files to determine if an employee has accessed a majority of the affected VA members around the date of misuse
- Investigate VA facilities within twenty (20) miles of misuse. These may potentially include military facilities, hospitals, and other veteran-specific organizations (e.g. American Legion)
- Determine if there are common background elements specific to affected VA members (e.g. could the victims all have visited the same VA hospital prior to the date of misuse?)
- Investigate whether or not VA employees reside at – or around – the addresses associated with the activity detailed in this report
- Investigate whether or not VA employees have provided email addresses or phone numbers as contact information that are associated with the activity detailed in this report
- Provide law enforcement with the addresses and phone numbers where potential misuse has occurred
- Match the list of SSNs included in the encrypted report and determine if any one case has identities that were involved in a previous breach-like incident that the VA is aware of. If a large percentage of SSNs from a single case was involved in a previous incident, the likelihood that harm resulted from that incident would be high.
- Compare the lists of SSNs to veterans who have complained they may be a victim of a breach due to Veterans Affairs. If the veteran's SSN is on the list of potentially harmed consumers, the VA could consider increasing the amount of investment made to remediate that veteran's identity.

After this investigation, if there are still VA members who are considered victims due to the activity detailed in this report, ID Analytics suggests the Veterans Affairs consider offering individualized assistance to affected veterans. Such individualized assistance may include credit monitoring, identity monitoring, fraud alerts and credit freezes.

4. Risk Assessment

4.1 Are the veterans at higher risk than the overall population?

The first question we investigate is the relative identity fraud risk of the veterans population compared to the U.S. population as a whole. To make this comparison we use our MyIDScore, which calculates the risk of any particular person to being a victim of identity fraud. To make this calculation in a quantitatively accurate way it is important to have the complete set of PII:

- SSN
- Name
- Address
- Date of Birth
- Phone number

In the BIRLS file we only have SSN, Name and Date of birth, so this file isn't well-suited for this calculation and comparison to background. Luckily we also received the HEC file which has all the needed fields.

We find 15.7 million identified people in the HEC file, with 7,781,339 records overlapping with BIRLS. We wish to do a comparison of risk between these two populations, which is shown in Figure 1 below.

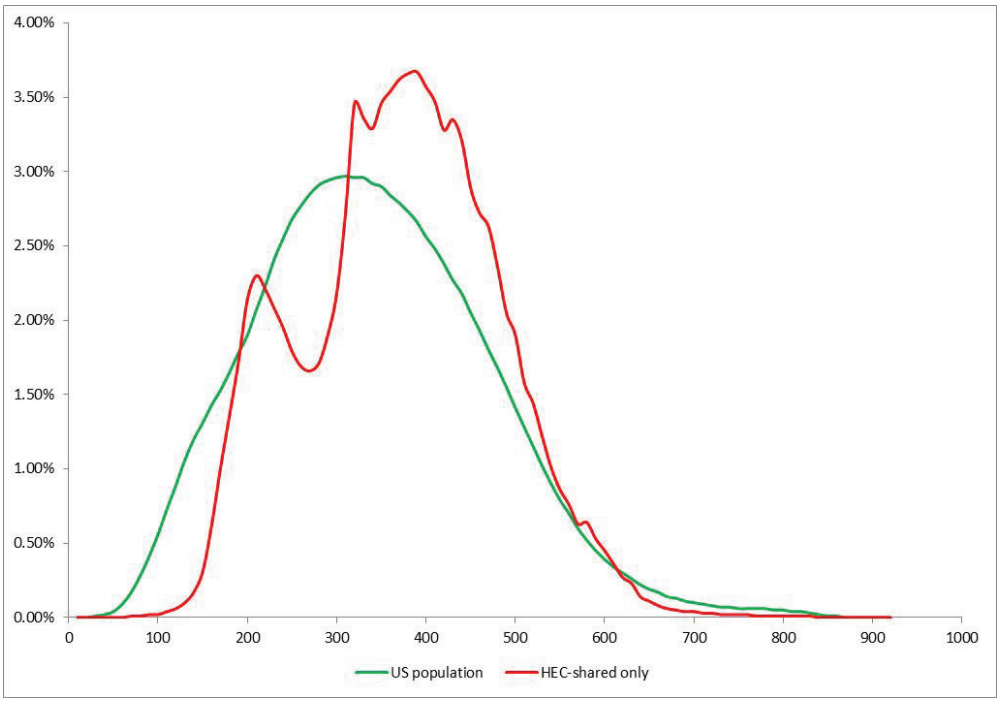


Figure A. Comparison of the distributions of personal risk for two populations: The HEC that are also on the BIRLS file a representation of background U.S. population. We see that the veterans have elevated risk compared to the general U.S. population.

In Figure A we first see that the veteran population's risk distribution is shifted to the right of the background, general U.S. population. We also see an unusual shape on the low risk side of the veterans that remains an anomaly.

4.2 Where are the Veterans at Risk?

In order to answer this important question we use our MyIDScore, which measures the level of risk an individual has with respect to identity fraud. The score considers the personal identifying information of an individual (SSN, name, address, date of birth, phone, email) and examines all other uses of these unique PII components in our commercial visibility, looking for anomalies. We scored the ~20.5 million people on the BIRLS file and we can see explicitly who is at risk and where they are located. The MyIDScore for each veteran is being returned as part of the data delivery for this project.

Using the MyIDScore we can examine where are people at risk for identity fraud. First we ask where the general population is at risk and then we examine the questions for the veterans.

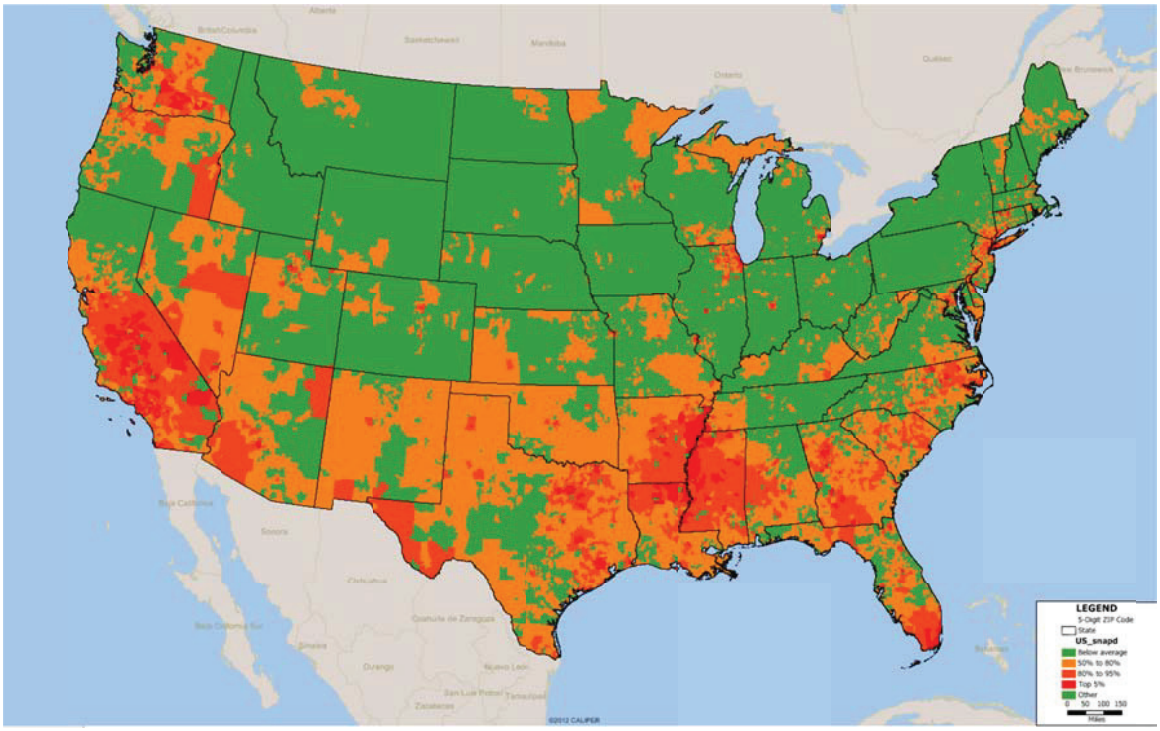


Figure B. Map showing where the general U.S. population is at risk of ID fraud.

Figure B shows the locations in the U.S. where the general population has elevated risk of identity fraud. In this picture the green areas are zip codes where the level of risk is at or below average. The orange represents risk areas between average and the lowest 80th percentile. The darkest red represents the locations of the 5% highest areas, and the lighter red is between the 80^{th} and 95^{th} percentiles in risk. We find risky locations include

- Large portions of southern states: Arkansas, Mississippi, Louisiana, Alabama, Georgia, Oklahoma, Florida, South Carolina, Texas, New Mexico, and Southern California.
- Other states have particular regions at risk: Central Virginia, much of West Virginia, portions of Missouri, Kansas, Arizona, Nevada, much of New Jersey, the border of Oregon and Washington, Northern Michigan
- Many urban areas: Miami, Houston, New York, DC, Detroit, Chicago, St Louis and others.
- We also see broad general areas of below average risk throughout the upper Midwest through upstate New York.

This picture represents where the general population is at risk for identity fraud. Next we ask this same question, but for the veteran population only.

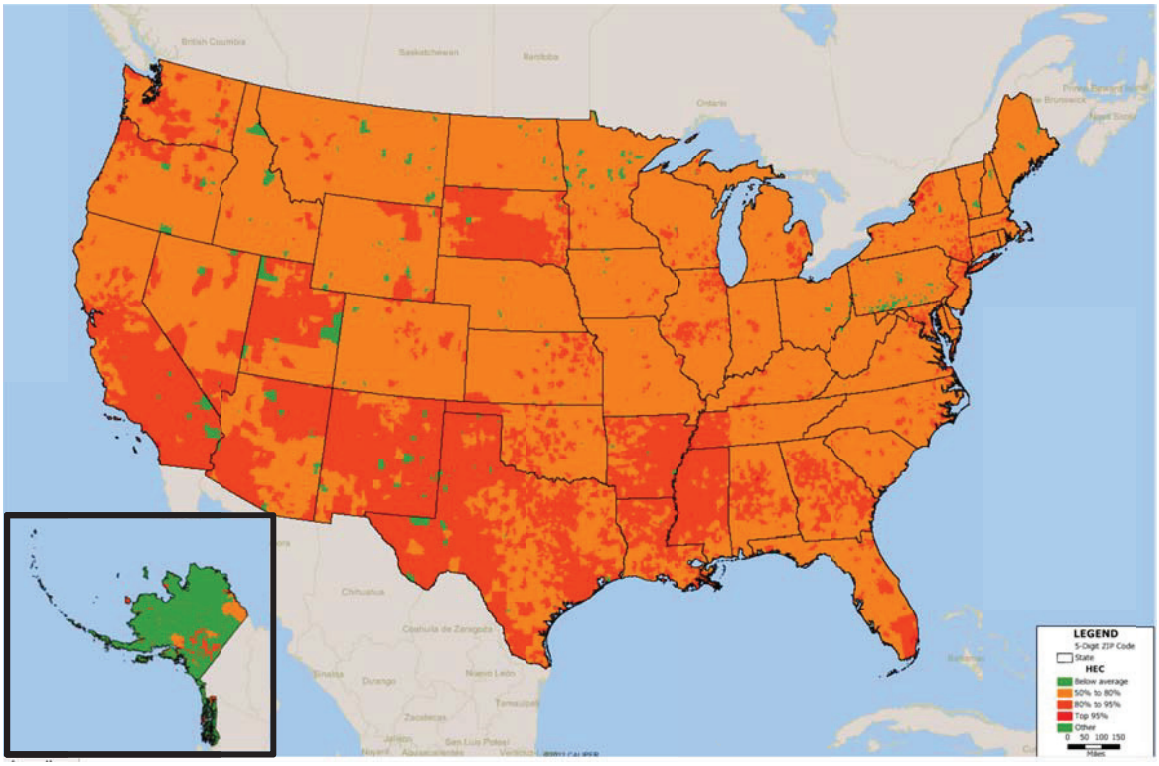


Figure C. Areas where the veterans are at risk of ID Fraud. We find that veterans have higher risk of identity fraud than non-veterans.

In Figure C we show the same risk measure for the veteran population only, using the MyIDScore on the HEC file population that overlaps with the BIRLS file. This picture is using the same scale, and the obvious striking feature is the elevated individual risk level for the veteran population. In Figure B we see the general areas of risk also seen in the general population (Figure B), but for the veterans we also see elevated risk areas in other, less expected areas, where we used the same risk scale as in Figure B.

The next question we ask is where the veterans who have risk levels above their immediate neighbors are. That is, we look at the relative risk of the veterans compared to the general local population, again by zip code. This result is shown in Figure D. Here we find a number of unexpected results.

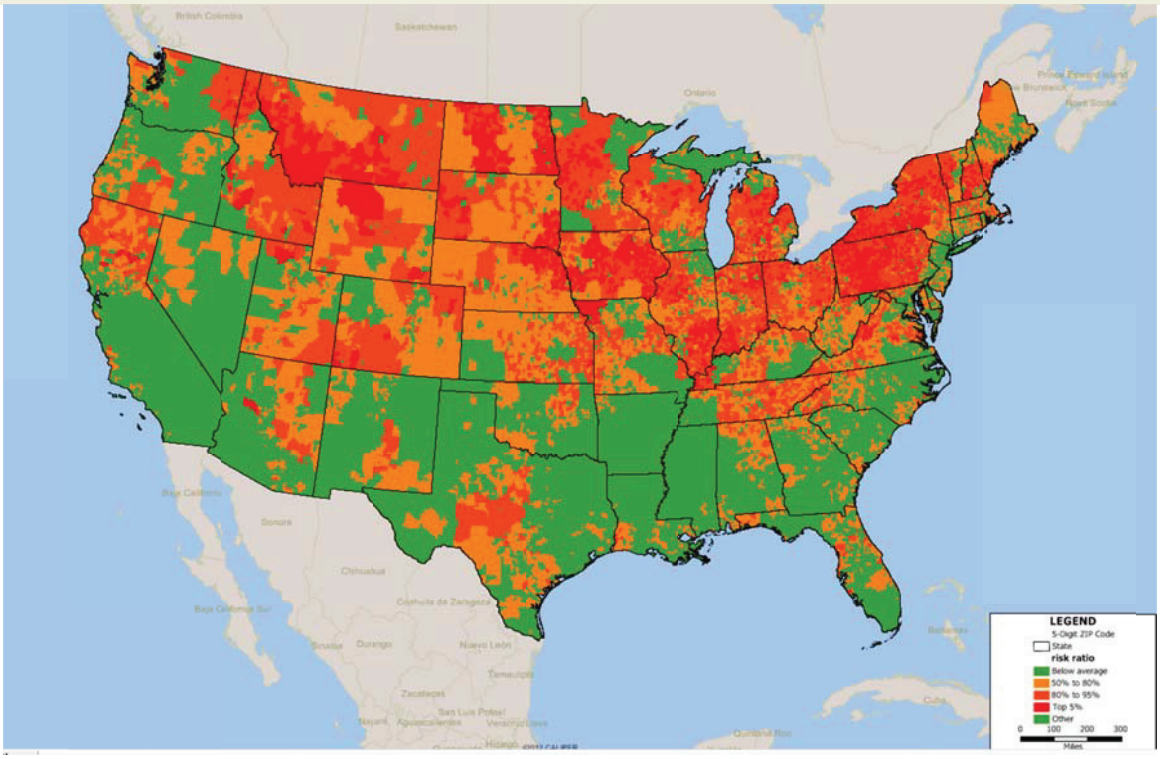


Figure D. Map showing locations where veterans are more at risk than the local population. For each zip code we measure the risk of the veterans and compare this to the risk of the non-veterans in that zip code. This map shows where veterans have a heightened risk compared to similar people, those in the same area.

In Figure D we find that veterans are at higher risk than the background population in much of the upper Midwest, stretching from NY, PA, OH, IN, MI, IL, IA, WI, MN, NB, SD, ND, CO, UT, WY, ID, MT, and Northern California. There are also regions of other states (AZ, NM, TX, LA, MS, GE, FL, TN, KY, NC, VA, WV, MA, CT, VT, NH) that have substantially higher risk than the background population. In Figure 4 we caution that "green" doesn't mean the veterans are safe – only that they are about the same risk level as the local non-veteran population.

In Figure D we can see hotspots where veterans have elevated risk compared to non-veteran peers, that is, people in the same locations. This picture points to possible specific locations of systematic identity theft against veterans. We can further examine the list of the top zip codes that create the hot spots in Figure D, which is shown in Table 32 below.

Table 32. Highest ranking zip codes where the veterans are at significantly higher risk of identity fraud compared to the local non-veteran population. These indicate hot spot locations of possible systematic misuse of veterans' identities.

Zip	BIRLS Risk	Non-veteran Risk	Risk Ratio
59715	614	453	1.36
59718	609	451	1.35
59713	606	460	1.32
59717	606	460	1.32
59710	605	460	1.32
59716	606	461	1.31
59719	606	461	1.31
59806	615	468	1.31
82609	603	460	1.31
59714	604	462	1.31
59808	604	462	1.31
59741	596	457	1.30
59804	606	465	1.30
13329	597	459	1.30
59745	595	458	1.30
59746	595	458	1.30
59747	595	458	1.30
59748	595	458	1.30
59772	595	458	1.30
59740	594	458	1.30
59743	594	458	1.30
59771	594	458	1.30
59735	597	461	1.30
59736	597	461	1.30
59762	597	461	1.30
59022	606	468	1.29
59758	593	458	1.29
59761	598	462	1.29
14004	594	459	1.29
58201	594	459	1.29

Zip	BIRLS Risk	Non-veteran Risk	Risk Ratio
59749	594	459	1.29
82414	594	459	1.29
59739	597	462	1.29
59750	594	460	1.29
12804	590	457	1.29
59807	604	468	1.29
82443	604	468	1.29
59732	596	462	1.29
59759	596	462	1.29
59725	592	459	1.29
59751	592	459	1.29
59760	597	463	1.29
59733	595	462	1.29
52327	591	459	1.29
59755	592	460	1.29
59756	592	460	1.29
59020	601	467	1.29
59602	588	457	1.29
59729	588	457	1.29
59803	588	457	1.29
52302	584	454	1.29
58078	584	454	1.29
59026	602	468	1.29
59088	598	465	1.29
58504	594	462	1.29
59731	594	462	1.29
59875	600	467	1.28
13322	591	460	1.28
57006	610	475	1.28
59025	601	468	1.28

In this table the risk number represents the score location of the top 4% score cutoff, which is a robust and statistically valid measure. The risk ratio is the relative risk between the veterans and the background non-veteran for risk ranking purposes.

Table 32 indicates hot spot zip codes of possible systematic identity abuse against veterans. We observe first that the area around Bozeman, MT has systematic elevated risk to veterans. We also find some other relevant veteran related institutions in other of these elevated zip codes, shown in Table 33 below.

Table 33. Identified veterans' facilities in some of the risk hot spots.

59713 - 59719	Bozeman VA Community Based Outpatient Clinic	Bozeman, MT
59804 - 59808	Veterans of Foreign Wars office; Montana Veterans Affairs Office; Missoula Vet Center; VA Community Based Outpatient Clinic	Missoula, MT
82609	Vet Center; Department of Veterans Affairs - Casper;	Casper, WY
13329	Fort Drum	Dolgeville, NY

Again, the striking feature of Table 32 is the prevalence of zip codes in Bozeman, MT (59715 – 59722) and the surrounding region. Why are so many veterans who live in this area at substantially elevated identity fraud risk compared to non-veterans who also live in this area?

One possibility is the local Bozeman Health Facility, Shown in Figure E below along with the link to this facility from the VA web page.

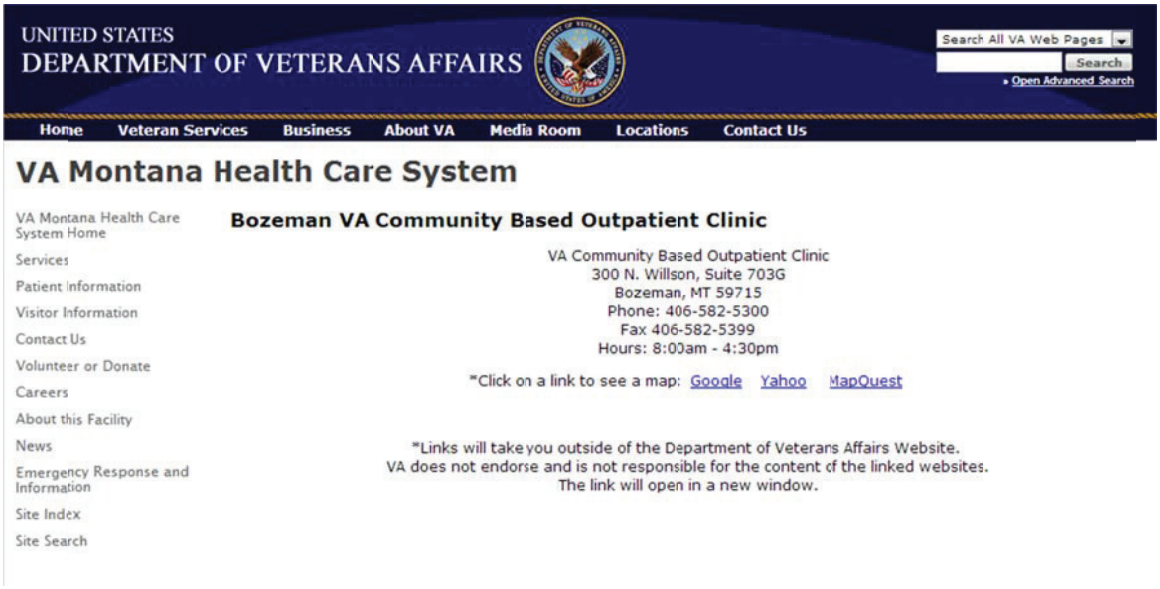


Figure E. Bozeman, MT VA Clinic may be the point of compromise for veteran's identity abuse.

Another way to search for systematic risk in the veteran population is to examine specific zip codes for military facilities. We have a partial list of such facilities and have found the top riskiest ones, shown in Table 33.

Table 33. List of some of the risky military facilities where veterans have higher risk than the surrounding non-veterans.

Zip	BIRLS Risk	Non-veteran Risk	Risk Ratio	Military Institution
99505	620	499	1.24	Joint Base Elmendorf Richardson, AK
13602	630	511	1.23	Fort Drum, NY 13602/13603
80841	625	512	1.22	United States Air Force Academy, Colorado Springs, CO
45433	608	499	1.22	Wright-Patterson Air Force Base, OH
42223	618	516	1.20	Fort Campbell, KY 42223
85613	615	521	1.18	Fort Huachuca, AZ
73503	618	527	1.17	Fort Sill, OK
13603	609	522	1.17	Fort Drum, NY 13602/13603
80840	600	516	1.16	United States Air Force Academy, Colorado Springs, CO
33621	606	527	1.15	MacDill Air Force Base, FL 33621
22211	611	532	1.15	Fort Myer, VA
36362	604	527	1.15	Fort Rucker, AL 36362
32508	618	541	1.14	Naval Air Station Pensacola, FL 32508
32542	620	543	1.14	Eglin Air Force Base, FL

In this table we see a military base in Alaska as a location where veterans have substantially higher risk than the background local non-veteran population. A map of this region is shown below as Figure F, and is a blowup of the previous Figure D.

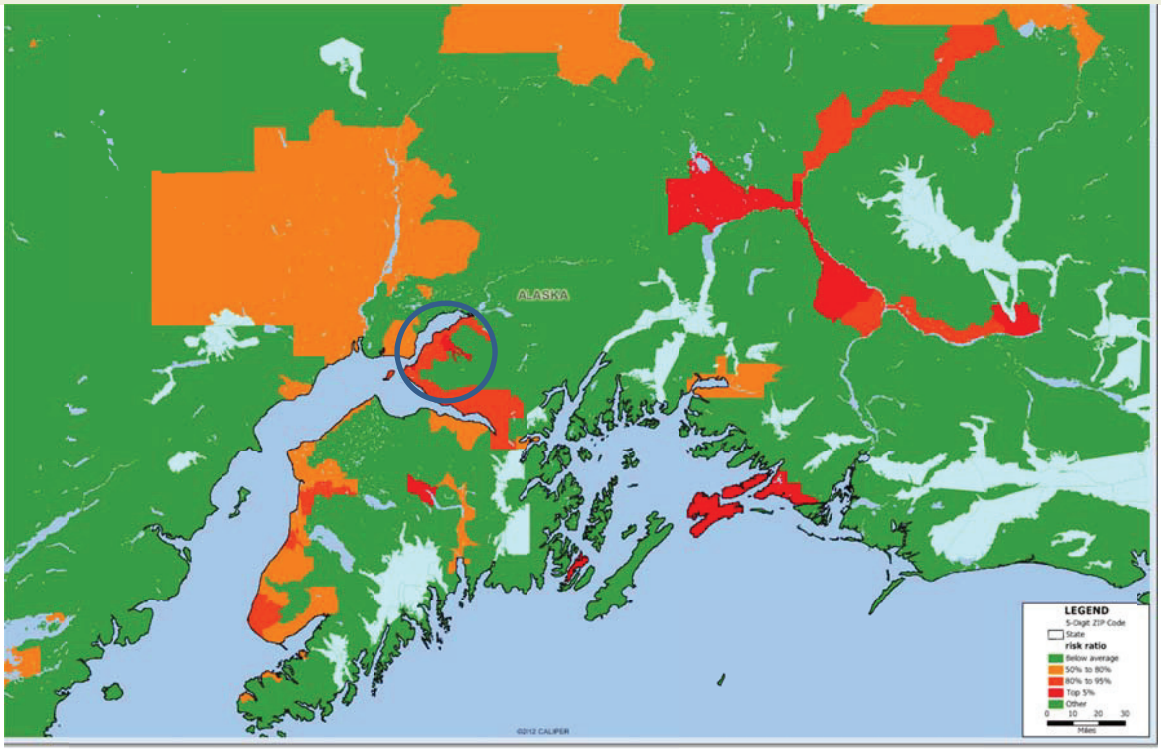


Figure F. Area around the Alaskan Elmendorf Richardson Base outside of Anchorage shows high risk to veterans.

In Figure F we see the area of the Elmendorf Richardson Base inside the blue circle. The risk to veterans for identity fraud attacks is very high here, and in the other regions indicated by red around Anchorage.

The second military installation on this list is Fort Drum in upstate NY. Figure G shows the area around Fort Drum inside the blue circle, which we see has elevated identity risk in a very large area around the military base.

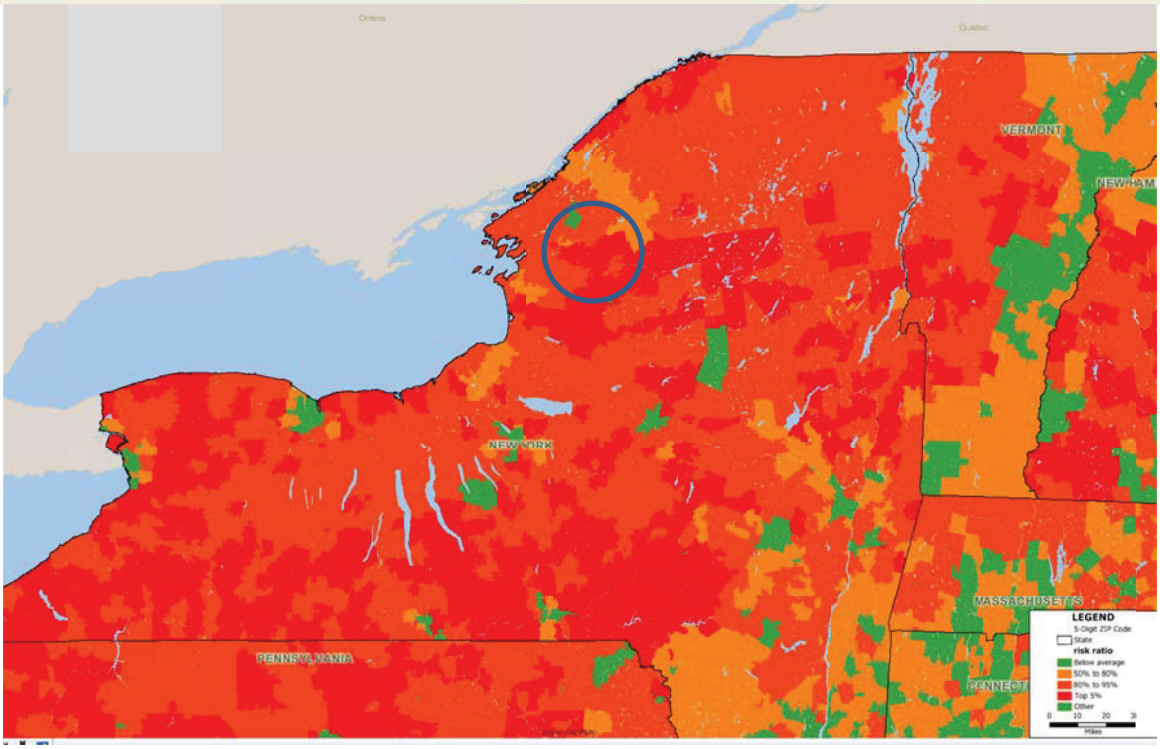


Figure G. The area in upstate NY around Fort Drum shows substantial elevated risk to veterans for identity fraud.

4.3 WHERE ARE THE VETERANS?

A final and interesting question to ask is where are the veterans in general? This is shown in Figure H, where we ask where are the veterans in relation to the general population? In this figure the darker the region the larger percentage of the population is a veteran. We find larger veteran population percentages in both urban and rural areas. As expected, many of the high percentage veteran population areas are around military installations.

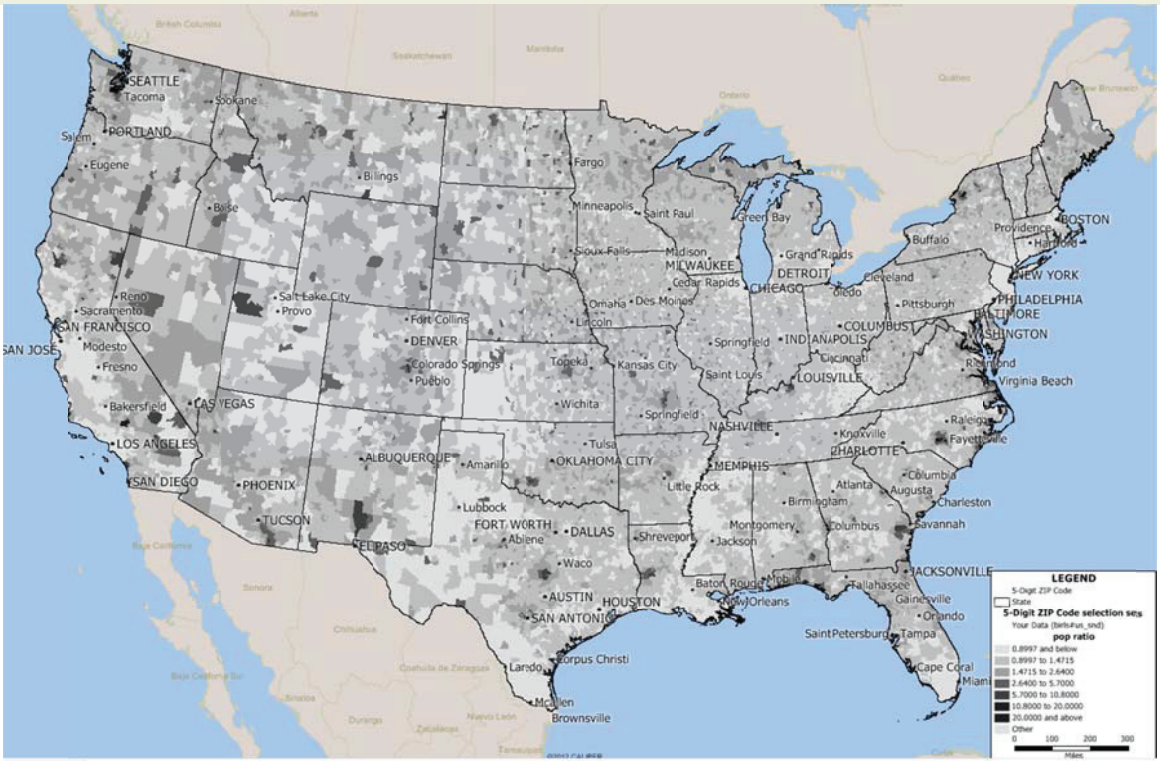


Figure H. here we show where the U.S. veterans live in proportion to the local population. The darker the area the higher percentage veterans living there. There is no concept of risk here, only where they are located, so we use simple dark shading to represent relative veteran population density.

5. Alerting and Activity Study

ID Analytics has performed a retrospective analysis of the BIRLS and HEC data sets using the Consumer Notification Service alerting platform to answer a range of questions regarding actual veteran behavior and activity in the marketplace and the relative risk to the veteran's identity of that behavior.

The Consumer Notification Service continuously scans the ID Network for activity, and generates an alert if new activity matches an enrolled consumer. Alerts are delivered in real-time – often within seconds or minutes. Typical examples of activity that results in a real-time alert include applications for goods or services (e.g. opening a new cell-phone account), requests for changes of address, requests for various forms of financing (e.g. payday loans), and other transactions. The retrospective analysis covered a one year period, from September 2011 to August 2012. The analysis answered questions including:

- The activity of VA members in the marketplace for transactions that asserted the VA member's identity (as measured by alert frequency)
- The industries and companies that drive most of the activity involving the veteran's identity
- The risk profile of the activity – does the activity appear to be misuse of the veteran's identity?
- The frequency at which veterans indicate they have not authorized or initiated activity that is attributed to them
- The amount of identity fraud for veterans who have indicated they did not authorize or initiate activity
- The propensity of veterans to seek out additional identity protection via commercial services

5.1 – Activity & Alert Insights

Figure I displays alert activity from September 2011 to August 2012. Alerts are generated when a Veteran transacts in the marketplace using their identity. As a result alerts are a good proxy for how veterans 'behave' commercially.

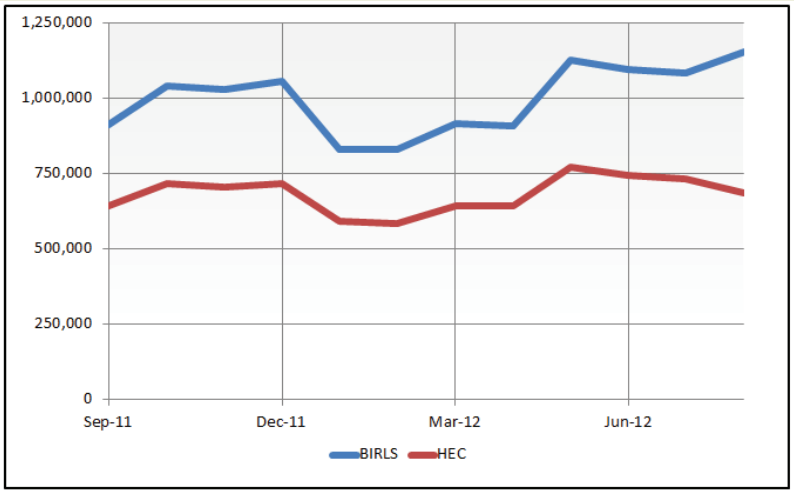


Figure I. Alert Volume September 2011 to August 2012

Figure I illustrates alert patterns for veterans included in the BIRLS and HEC data sets. The trends appear to be typical consumer behavior – alert activity increases over the holiday period (November and December), troughs post holidays, and then rebounds in spring and summer. The cycle then repeats.

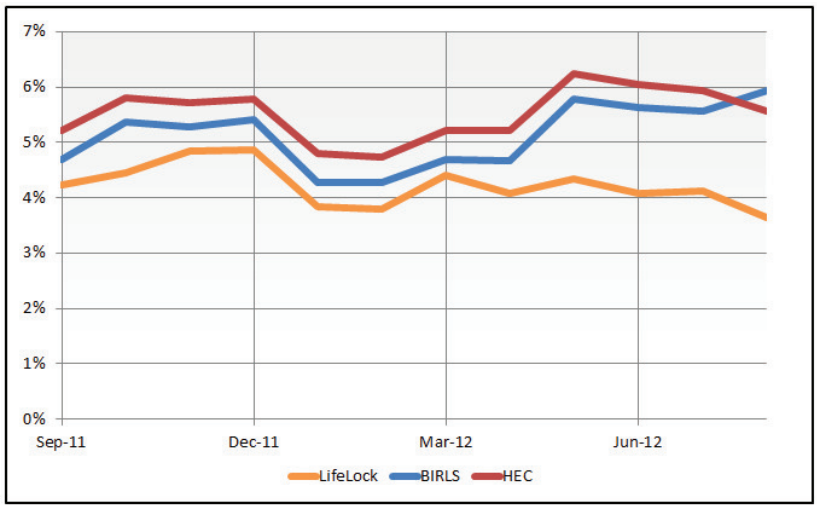


Figure J. Alert Rate of Veterans Compared to LifeLock Subscribers (primarily non-veterans)

Figure J displays the alert rate over the study period. The alert rate is the percentage of veterans who receive an alert in the month indicated. The data compares the rate of activity for veterans (BIRLS and HEC) with the rate of activity for consumers (primarily non-veterans) enrolled with the

LifeLock identity protection service, a leading provider in the marketplace. This data can tell us how commercially active the population of veterans is compared to a non-veteran population.

We can see from the data that veterans (both the BIRLS and HEC files) have substantially higher alert rates than the non-veteran population (represented by LifeLock). This indicates a higher level of activity in the marketplace for the veteran population, which could indicate higher risk of identity misuse.

Figure K, below, displays the types of commercial transactions where the identity of the veteran was used. The predominant usage was to obtain new credit cards, which includes both retail credit (Best Buy, Walmart, Loews, etc.) and bank-branded credit (Chase, Discover, Bank of America, etc.). Telecomm includes both wireless accounts and landline accounts.

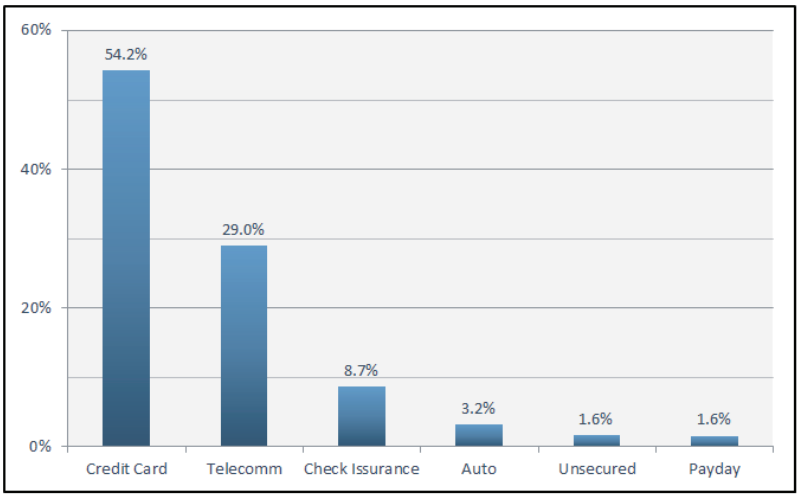


Figure K. Distribution of Alerts across Vertical Industries

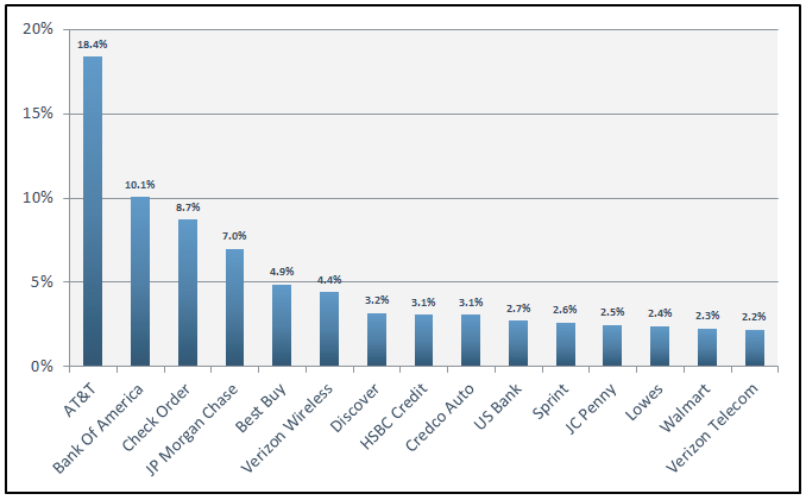


Figure L. Top 15 Companies that Drive Alert Volume for Veterans

Figure L lists the companies that accounted for the majority of the alert / transaction volume. These companies represent the majority of veteran transaction activity from September 2011 to August 2012. The list is dominated by credit card (retail and bank-brand) and telecomm companies.

5.1.1 Alerts, Identity Risk, and Not Me® Cases

ID Analytics is able to assess the relative risk to the veteran for each of the alerts in the study. The risk of an alert is determined by our industry-standard ID Score technology. Our technology uses a complex process involving advanced linking technology, expert variables and state-of-the-art machine learning algorithms to assign a likelihood of fraud to the event containing the asserted personal information. By examining the risk level for each of the alerts, we can identify what transactions were suspicious and likely involved misuses of the veteran's identity.

When a consumer (who could also be a veteran) is enrolled in a service such as LifeLock, they receive real-time alerts that indicate activity involving their identity. After receiving an alert, they have the opportunity to say "that's not me". ID Analytics proprietary and patent-pending Not Me Notification System® is able to connect the consumer's feedback directly with the company where the activity originated. The company can then stop the activity to prevent the fraud or minimize the damage associated with it. Diagram I. illustrates the process.

Diagram I. Not Me Notification System®

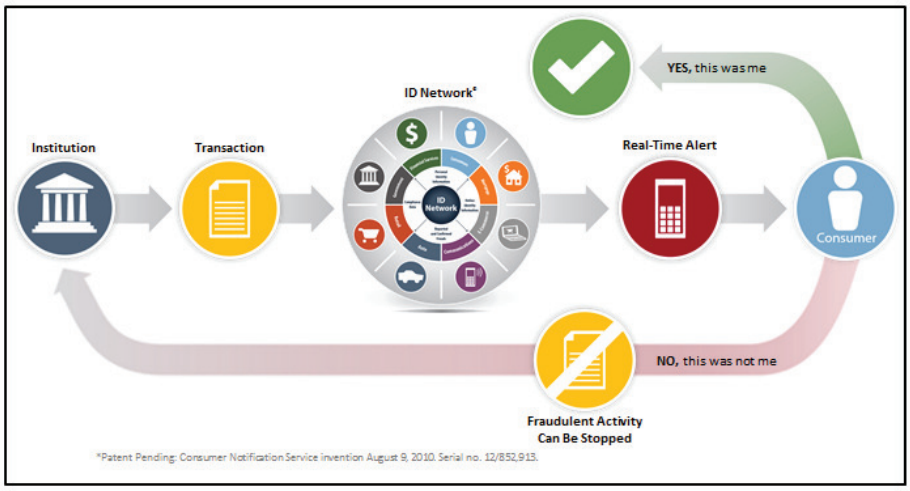


Figure M below shows the rate at which veterans enrolled in LifeLock assert "not me" in response to an alert. The BIRLS and HEC data sets have been compared to the LifeLock population to determine if veterans behave differently than a 'normal' identity protection subscriber. Several conclusions can be drawn from the data. First, the BIRLS and HEC populations are in line with the general rate that LifeLock subscribers assert "not me". Between 3%-5% of all alerts result in a Not Me case for each of the populations. There also appears to be seasonality in the Not Me case rate – there are more cases in the holiday shopping season, followed by a lower rate in the subsequent months (which follows the pattern of alerts).

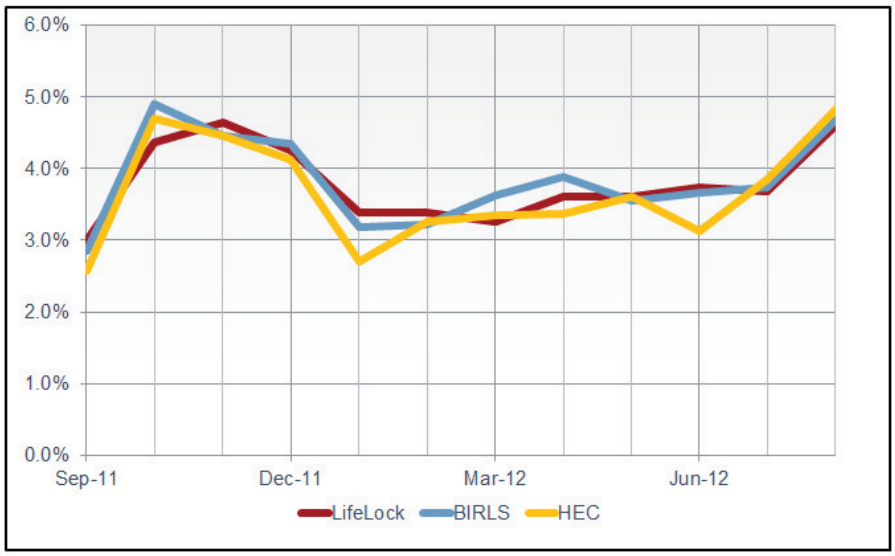


Figure M. Not Me Case Rate for BIRLS and HEC, Compared to LifeLock.

Some alerts are more likely to result in a “not me” assertion. By examining the risk assessment for the alerts delivered and the actual Not Me cases resulting from alerts it is clear that the higher the risk level of the alert, the more likely it is that the veteran / consumer will claim “not me” and a case will be created to investigate their claim.

Figure N shows the risk level of alerts (low, moderate, or high risk) and what percent of those alerts result in a Not Me case. The data clearly shows that alerts categorized as “high-risk” are far more likely to result in a Not Me cases – in fact, high-risk alerts are 14 times more likely to result in a case than a low-risk alert.

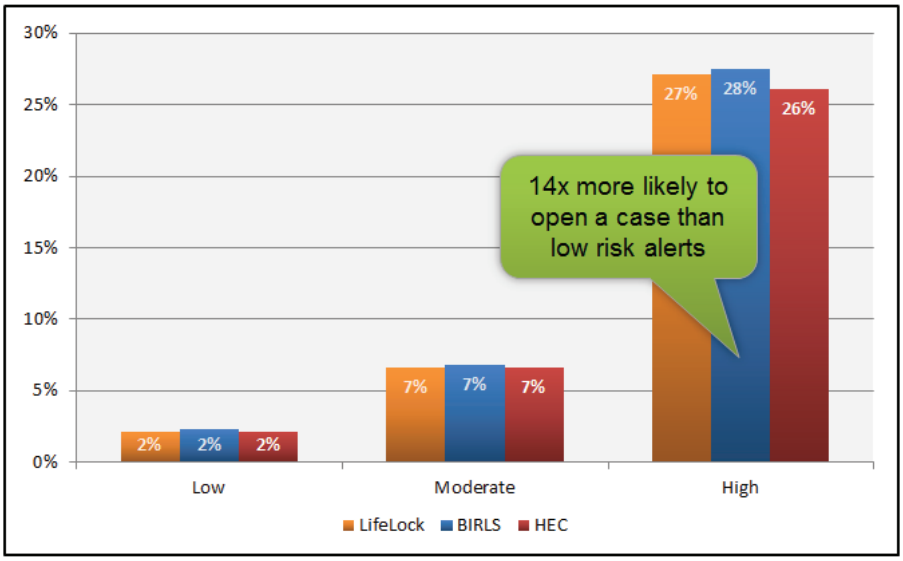


Figure N. Percent of Alerts that become Not Me Cases

Figure O shows the risk distribution for alerts and Not Me cases. As we would expect, Not Me Cases have a much higher risk profile – the median score for cases is nearly 500, versus a median score for alerts of 340.

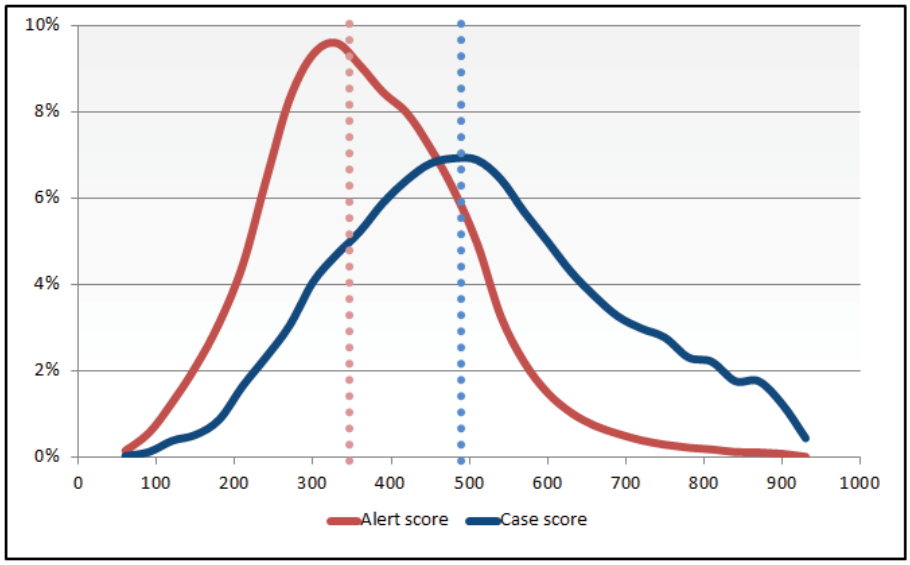


Figure O. Risk Distribution for Alerts and Not Me® Cases

Figure P shows alert risk levels for the BIRLS and HEC populations and also provides a comparison with the LifeLock subscribers. The BIRLS and HEC alerts have very similar risk profiles to the LifeLock alerts. This is interesting, in that the LifeLock population tends to be a riskier population than non-LifeLock consumers (LifeLock consumers often join the service after being victimized).

17% of the BIRLS and HEC alerts (for veteran' enrolled in LifeLock) are in the moderate risk category, while 3% of the alerts are high risk. The 3% of high risk alerts represent over 5,157 BIRLS and 3,399 HEC high risk transactions where the veteran's identity was likely misused.

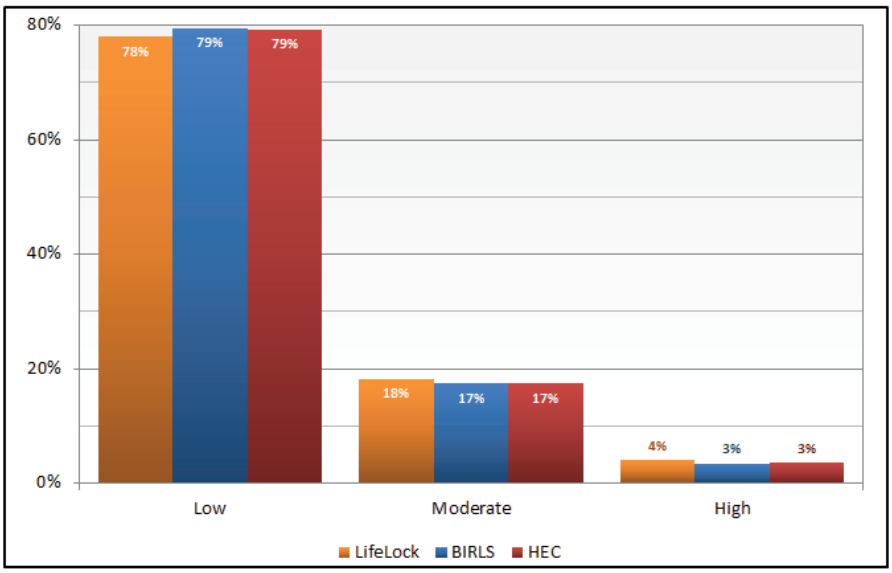


Figure P. Alert Risk Levels for BIRLS, HEC, and LifeLock Populations.

For veterans enrolled in LifeLock, we can examine how often they asserted “not me” after receiving an alert.

Figure Q shows the number of Not Me cases for the BIRLS and HEC populations over the study period. Veterans asserted “not me” 1,419 times within the BIRLS population, and 886 times within the HEC population. It is also interesting to note the seasonality of cases based on the holiday shopping period.

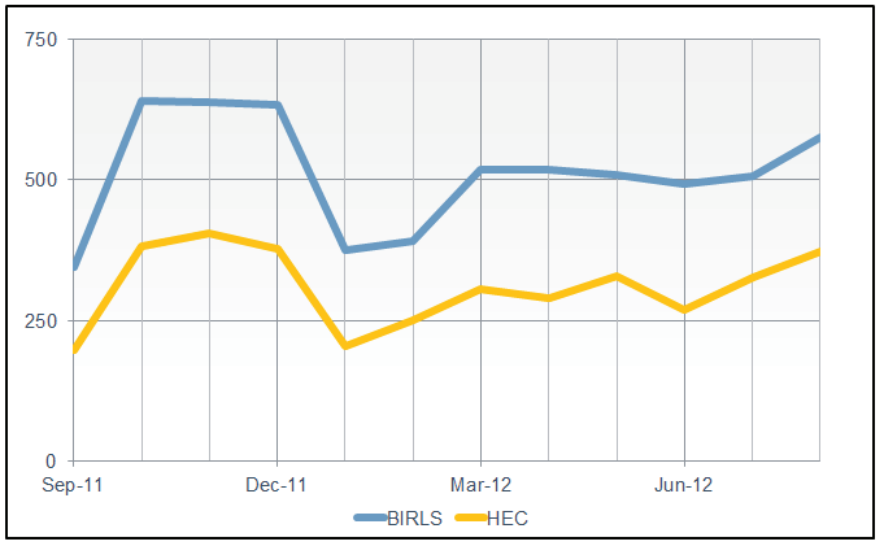


Figure Q. Not Me Cases for Veterans Enrolled in LifeLock

When a veteran asserts “not me”, it is often identity fraud. Figure R shows the percentage of all Not Me cases that are actually identity fraud. Note the increase in fraud rate over the study period. The data shows that between 50%-65% of all cases are identity frauds.

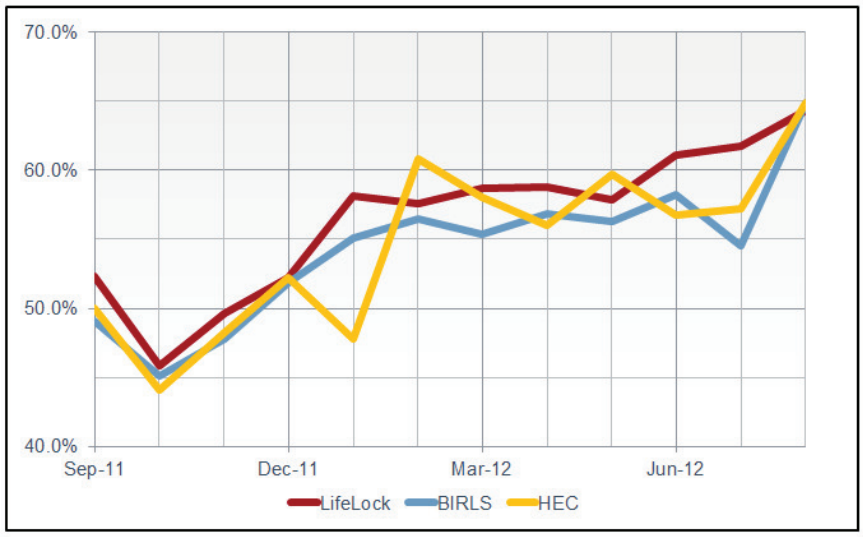


Figure R. Fraud Rate for Not Me Cases

5.1.2 Insights on the BIRLS and HEC Populations

Based on the insights gained from examining alert rates for the overall BIRLS and HEC populations, and the risk levels, Not Me cases, and fraud rates for those veterans enrolled in LifeLock, we can generalize the findings and make assumptions about the veterans in the BIRLS and HEC files.

We know that there were 11,980,697 possible alerts for the BIRLS population, and 8,189,908 possible alerts for the HEC population during the study period. For those alerts, 3.4% of the BIRLS population, and 3.2% of the HEC population are high-risk transactions, which results in 407,344 BIRLS and 262,077 HEC events where a veteran's identity could have been misused.

IF we use the average Not Me case rate (3.9% BIRLS, 3.7% HEC) to determine how many alerts would have likely been escalated as a "not me", we arrive at similar totals. The number of likely cases is 467,247 for BIRLS, and 303,027 for the HEC population, where a veteran could have asserted "not me" in response to the alert – had they been enrolled in a protection service.

Finally, between 50% and 65% of cases are actual identity frauds, which means veterans experienced and had to resolve actual identity misuse between 233k – 304k times for BIRLS, and 152k – 197k times for the HEC population.

5.1.3 Case Studies

ID Analytics has assembled two case studies where we examined activity related to a veteran's identity. In the first case study, the veteran was enrolled in LifeLock and had the opportunity to indicate "not me" in response to an alert. In the second case study, the veteran was not enrolled in an identity protection service and was not able to react to the alerts that would have been provided.

Diagram II displays the case study where Casey C (a veteran enrolled in LifeLock) responded to an alert using his identity and was able to shut down the misuse.

Diagram II. Actual Not Me Case

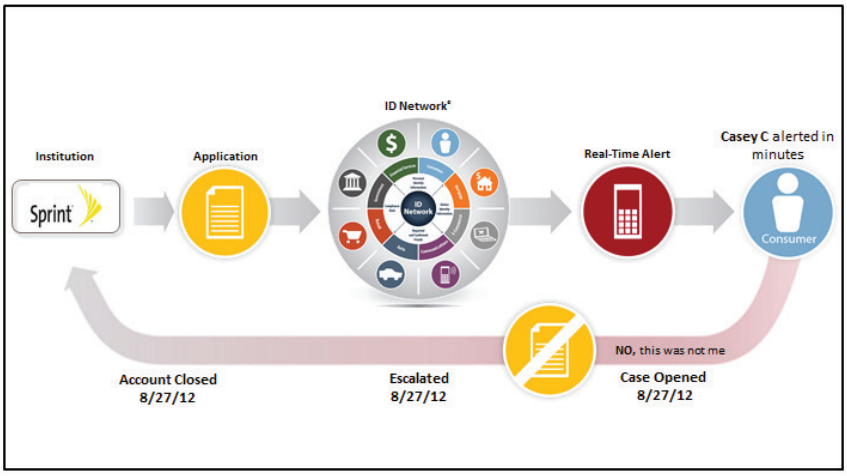
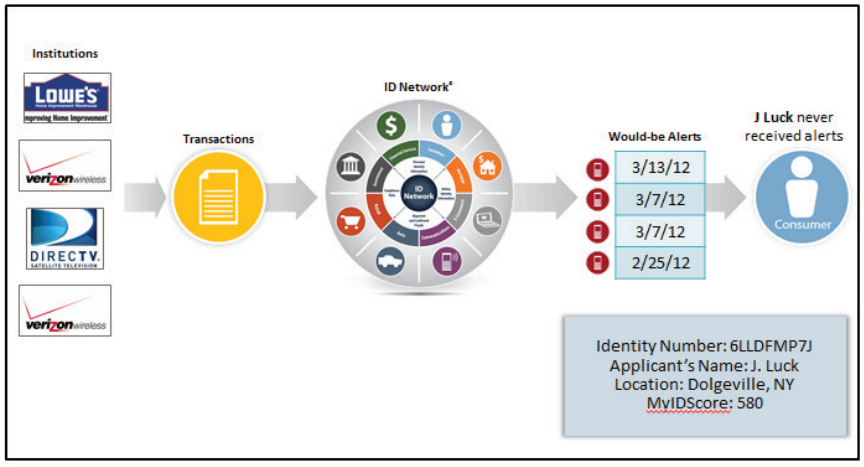


Diagram III displays a case study where the veteran (J. Luck) was not enrolled in a service that provides real-time alerts and the Not Me Notification System. In this case, the veteran's identity was used on applications for a Lowe's credit card, two Verizon accounts, and a subscription to DirecTV.

Diagram III. Identity Misuse Targeting a Veteran



5.1.4 Additional Insights

Figure S displays the age distribution of veterans who received alerts, those who initiated a Not Me case, and those enrolled in LifeLock.

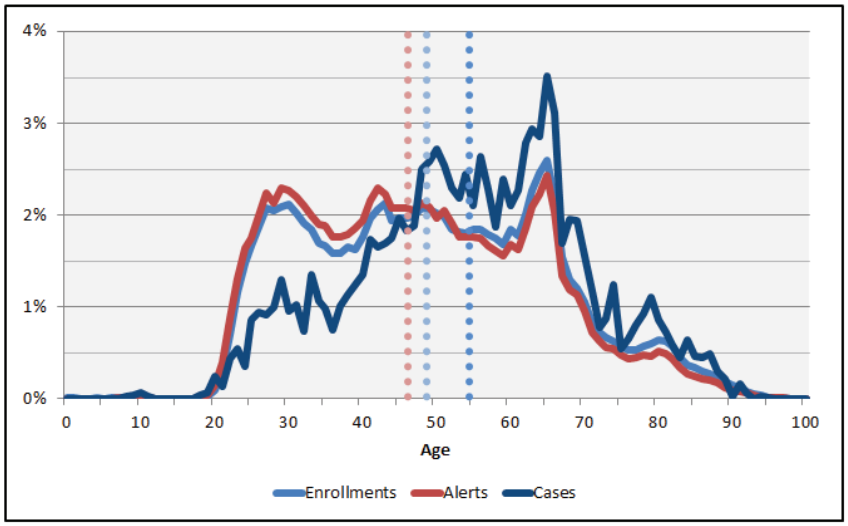


Figure S. Age Distribution.

Figure T lists the states where veterans enroll in commercial identity protection services at a higher rate than then general population.

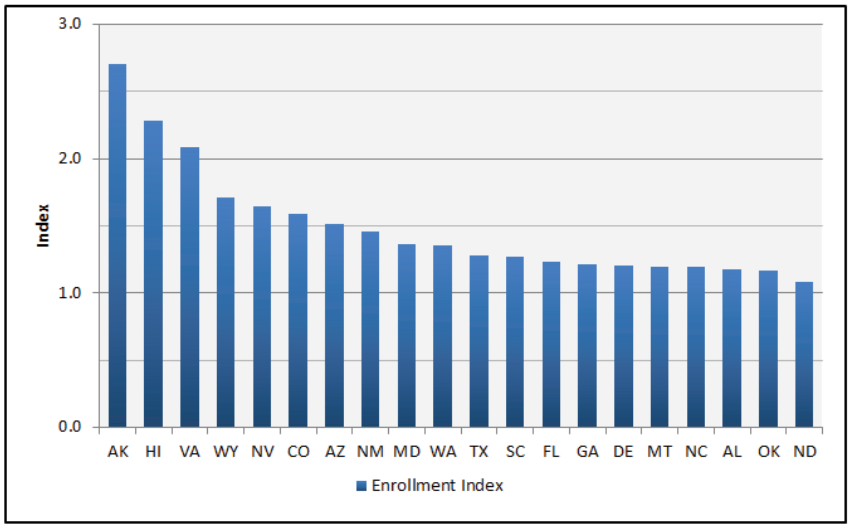


Figure T. Veteran Enrollment in Identity Protection Services, Indexed to the General Population.

5.2 Alerting & Activity Study Summary Findings

- The population of veterans in the BIRLS and HEC files tends to be more commercially active than the general population, as measured by subscribers to the commercial identity protection service LifeLock. This indicates more opportunities where the veteran is asserting their identity and thus more opportunities where their identity could be misused.
- By examining the risk level of the alerts for the veteran populations, we see that 3% of all alerts for identity activity are high-risk alerts as measured by our risk scoring technology. This is in-line with the non-veteran population alert activity.
- For veterans enrolled in LifeLock, we can compare their behavior to non-veteran LifeLock subscribers to understand how often they assert “not me” in response to an activity alert. We find that Not Me case rates are consistent between the populations.
- Key Statistics:
    - Alert Activity Indicating use of the Veteran’s Identity:
        - BIRLS: 11,980,697
        - HEC: 8,189,908
    - Potential Not Me Cases in Response to an Alert:
        - BIRLS: 467,247
        - HEC: 303,027
    - Possible Misuse / Identity Fraud Incidents Suffered by Veterans
        - BIRLS: 233,000 – 304,000
        - HEC: 152,000 – 197,000

6. Glossary of Terms

Identity Manipulator (IM) Score -– A score used to determine how much an identity has been manipulated over time. The score factors in how many variations have been made to the core identity elements like SSN, name, and DOB. The IM Score examines the variations in PII used by a person in the commercial world as he applies for products where by law he is required to use his correct credentials.

Identity (ID) Score – A score used to determine the risk of an individual at account opening and throughout the customer lifecycle. ID Score uncovers many types of identity abuse including family fraud, first-party fraud and synthetic fraud. In addition to detecting suspicious use of traditional identity information such as name, address, Social Security number, and phone number, ID Score integrates online behavior associated with email addresses, IP addresses and geolocation to fine-tune risk assessments

Identity Number – The unique alpha-numeric value assigned to a unique identity; 9 alpha-numeric characters

Resolution Confidence – The confidence that the Identity Number assigned to the record matches the input identity elements provided. 0.0 is low confidence, 1.0 is high confidence. 0.0 is seen as random.

Resolved Record – The input record maps to an existing unique Identity Number

Created New Record – The input identity elements do not match to an existing unique Identity Number; however, there are strong indications that the input identity elements link to one another and represent a unique individual. These records are assigned new Identity Numbers, which can be used to link to other records in the file. Once a record is given an Identity Number, subsequent matches would have a status of R

Ambiguous Record – The identity is ambiguous and can't be resolved. There are indications that the input data links to multiple individuals and as a result a unique individual cannot be identified

Insufficient Record – The identity contained insufficient data to uniquely identity an individual

Frivolous Record – The identity contained insufficient data due to the presence of frivolous input identity elements (e.g. name = abcdefg)

Verified Input – The input data was verified upon resolution

Partial Input – The input data is a partial value of the true value (e.g. SSN = last 4 of SSN match)

Similar Input – The input data is similar to the value that was determined upon resolution

Inconclusive Input – The input data is unable to be verified

Reconstructed Input – The input data did not contain a value; however, upon resolution a value was determined

Date of Death – The date in which the identity was determined to be deceased

id:analytics™

Appendix A

ID Score® 7.2 Reason Codes

Document Last Updated: April 25, 2011

Appendix. ID Score Reason Code Descriptions

ID Score 7.2 reason code descriptions are prefaced with an indicator that describes the source of the underlying reason:

CONSUMER – Indicates that ID Analytics has information directly from the consumer that needs to be conveyed (credit inactive, security freeze, etc.).

EVENT – Indicates that the underlying reason is directly related to one or more data elements provided as part of the risk event itself.

NETWORK – Indicates that the underlying reason is tied to the relationship between one or more data elements provided as part of the risk event and data already in the ID Network.

Code	Description
180	EVENT - Application received through known risky channel
230	EVENT – SSN likely Invalid
231	EVENT – SSN likely Out of Range
232	EVENT - SSN reported as Deceased
233	EVENT - SSN likely issued before DOB
237	EVENT - SSN issue year suspicious
238	NETWORK - SSN associated with risk
500	NETWORK - SSN associated with suspected or confirmed fraud
501	NETWORK - SSN linked to multiple names
502	NETWORK - SSN linked to risky addresses
503	NETWORK - SSN linked to unusual number of addresses
504	NETWORK - SSN linked to unusual number of home phone numbers
505	NETWORK - Unusual number of applications combining SSN with other identity elements
506	NETWORK - Unusual number of historic applications using SSN
507	NETWORK - Unusual number of recent applications using SSN
509	NETWORK - Combination of SSN with other identity elements is generally associated with risk
520	NETWORK - Address associated with suspected or confirmed fraud
521	NETWORK - Address is in a known high-risk area
522	NETWORK - Address linked to multiple phone numbers

Code	Description
523	NETWORK - Address linked to risky SSNs
524	NETWORK - Address type is generally associated with high risk
525	NETWORK - Combination of address with other identity elements is generally associated with risk
526	NETWORK - Unusual number of historic applications using address
527	NETWORK - Unusual number of recent applications using address
530	EVENT - Address is in a known high-risk area
540	NETWORK - Combination of name and date of birth with other identity elements is generally associated with risk
542	NETWORK - Combination of name and home phone number is linked to unusual number of SSNs
543	NETWORK - Combination of name with other identity elements is associated with similar SSN
544	NETWORK - Name and date of birth linked to multiple phone numbers
545	NETWORK - Unusual number of historic applications using name and date of birth
546	NETWORK - Unusual number of recent applications using name and date of birth
560	NETWORK - Combination of home phone number with other identity elements is generally associated with risk
561	NETWORK - Home phone number associated with suspected or confirmed fraud
562	NETWORK - Phone number is in a known high-risk area
563	NETWORK - Phone number mismatch with ZIP code
564	NETWORK - Phone number type is generally associated with high risk
566	NETWORK - Home phone number linked to unusual number of addresses
567	NETWORK - Unusual number of historic applications using home phone number
568	NETWORK - Unusual number of recent applications using home phone number
569	NETWORK - Unable to confirm home phone number in association with other identity elements
580	NETWORK - Combination of elements of the email address generally associated with high risk
581	NETWORK - Unusual number of historic applications using email address
583	NETWORK - email address linked to unusual number of addresses
584	NETWORK - email address linked to unusual number of phone numbers
585	NETWORK - email address linked to unusual number of names
591	NETWORK - Historic usage patterns of identity elements are generally associated with high risk
620	NETWORK - Unusual number of historic applications using IP address
621	NETWORK - Unusual number of recent applications using IP address
622	NETWORK - IP address linked to unusual number of names
623	NETWORK - IP address linked to unusual number of phone numbers
624	NETWORK - IP address linked to unusual number of addresses

Code	Description
625	NETWORK - IP address linked to unusual number of SSNs
626	NETWORK - IP address mismatch with phone number
627	NETWORK - IP address mismatch with zip code
628	NETWORK - Internet connection characteristics generally associated with high risk
701	EVENT – Possible name match with OFAC List
904	NETWORK - Historic usage patterns of SSN are generally associated with low risk
905	NETWORK - Combination of SSN with other identity elements is generally associated with low risk
906	EVENT - No SSN violations found
920	NETWORK - Historic usage patterns of address are generally associated with low risk
921	NETWORK - Combination of address with other identity elements is generally associated with low risk
922	EVENT - Address is in a known low-risk area
923	NETWORK - Address is in a known low-risk area
941	NETWORK - Historic usage patterns of name and date of birth are generally associated with low risk
942	NETWORK - Combination of name and date of birth with other identity elements is generally associated with low risk
950	NETWORK - Historic usage patterns of IP address are generally associated with low risk
951	NETWORK - IP address and zip code match
952	NETWORK - IP address and phone number match
953	NETWORK - Internet connection characteristics generally associated with low risk
960	NETWORK - Historic usage patterns of home phone number are generally associated with low risk
961	NETWORK - Area code and zip code match
962	NETWORK - Combination of home phone number with other identity elements is generally associated with low risk
980	NETWORK - Combination of elements of the email address generally associated with low risk
981	NETWORK - Historic usage patterns of email address are generally associated with low risk
982	NETWORK - Combination of email address with other identity elements is generally associated with low risk
990	NETWORK - Identity elements generally associated with low risk
991	NETWORK - Application received through low risk channel
992	NETWORK - Historic usage patterns of identity elements are generally associated with low risk

Note: We have modified the description of reason code 701 from “EVENT - Name match with OFAC List” to “EVENT - Possible name match with OFAC List”. This description more accurately reflects the output of the OFAC check. Once a possible match is identified, the lending

organization should consult its internal program guidelines for investigating possible matches. The official OFAC website also provides guidance to organizations regarding the evaluation of possible matches, and when it is appropriate to contact the OFAC hotline. (See http://www.treasury.gov/resource-center/faqs/Sanctions/Pages/directions.aspx)