DETECTING ILLEGAL CREDENTIAL SHARING IN VIDEO SUBSCRIPTION

Orly Ovadia-Amsalem, Marcelo Blatt

Cisco, Jerusalem, Israel

ABSTRACT

Today, the video industry faces new types of piracy and threats that cannot be prevented by embedding secure hardware or software in consumer devices. Unlike legacy set-top boxes (STBs), there is no hardware identity built into the second screen consumer video devices. As a result, subscribers can manually enter their account credentials (username/password) and share them, both knowingly and unknowingly, with other non-subscribers. In this paper, we present a method for overcoming this problem, by detecting who is sharing their credentials.

We use machine learning techniques and advanced graph analysis to model different aspects of normal subscriber behaviour: temporal, spatial and watching habits. The models allow us to find anomalous behaviour among subscribers, to set up a threshold, and then to enable service providers to use consequences such as blacklisting devices and suspending sharing accounts.

INTRODUCTION

In this paper, we present our approach for detecting credential sharing in second screen devices, and then helping service providers overcome the credential sharing challenge. Using viewing records from a service provider's logs, we model the typical behaviour of an account, and represent each account as an n-dimensional vector. We use this representation in order to determine a sharing score per account, which reflects the likelihood that an account will share its credentials.

- We implemented machine learning algorithms based on a complex set of statistical, spatial, temporal and behavioural features.
- We performed further analysis on the viewing records using dynamic graph analysis to determine the sharing type.
- We distinguished between two main types of sharing activities, legal and illegal.
    - Under illegal sharing, we observed the cases where the credentials are distributed for profit purposes.

  - Legal sharing included the cases where the credentials are shared with family members or friends. Since we do not have any information about the actual family relation of the subscribers, we incorporate the assumption that family members tend to meet occasionally (e.g., a child living in dormitories) into our algorithms.

The viewing records (logs) we used in our initial trial were captured from second screen devices used by over a million customers and over hundreds of million viewing transactions, all received from a large known service provider.

Since this is an unsupervised problem, we had no training data about the actual sharers, hence, in order to validate our method, we performed post-hoc analysis on our results. This was done with the service provider to validate the shared accounts.

We are now following up our research with a prototype based on real-service provider data, in which data-science and machine learning results are being used in practical actions such as blacklisting devices, tracking activity of suspicious accounts over time, suspending sharing accounts, challenging users in real-time etc. all configurable by pre-defined thresholds.

PROBLEM DEFINITION

Credential sharing has evolved from being a casual pastime to an established industry threat that involves business sharing and stolen accounts. Unlike legacy set-top boxes (STBs) that are manufactured specifically for a service provider, there is no hardware identity built into second screen consumer video devices. This allows subscribers to easily share their credentials (username/password) with other non-subscribers, effectively presenting them with free video services, unbeknown to the service provider.

Recent studies show that credential sharing for video services has led to revenue losses amounting to as much as a billion dollars per year, as well as increased cost of service, and tarnished reputations. As the viewing experience shifts to second screen devices, we expect this problem to increase, especially since sharing is most common among young adults and teenagers.

METHODOLOGY

Our method consists of several phases.

1. First, we processed the data: we removed errors, normalized values, and adjusted time zones, etc.
2. Second, we extracted features from the processed data at the account level. The data of each account over a given time period (3 months in this case) is represented by a single feature vector.

3. We measured the Mahalanobis [1] distance of each vector (account) to a standard account (average of accounts in our baseline data), and normalized a sharing score according to the distance.
4. We then used dynamic graph analysis to differentiate the sharing types (business/stolen and casual).

Data

We conducted our research on viewing-logs from a large service provider. We analysed six months' of logs from the second half of 2016, containing over 1.3 million accounts.

Each account is accessed by a single username and password, and is identified by a unique *account_id*. An account can be accessed via multiple devices; each device is identified by a unique *device_id*.

Each log row, hereafter referred to as a "record", contains the following information: *timestamp, account id, device id,* the way the content is accessed *(linear viewing/VoD), ip,* and *content id*.
We cleaned and processed the logs and then performed geographical enrichment on the IP address using IP geolocation services [2], which translates the IP address to the specific country and city in which the viewing took place. Next, we extracted features from the processed logs.

Features Extraction and Calculation

We partitioned the features into four main groups:

1. Statistical features
2. Behavioural features
3. Temporal features, and
4. Spatial features.

The following sections describe the features extraction process for each of these categories.

1. Statistical Features

These features represent statistical measurements such as counts, averages, and min and max on various log fields.

For example, the number of countries, number of cities, total linear activity, total VoD activity, average number of active days, and the number of devices etc.

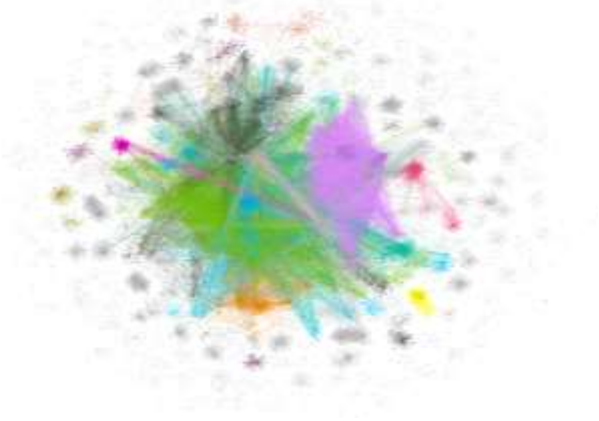

Figure 1

Full network of over 40k VoDs (nodes) and relations among the VoDs (edges). Community detection algorithm is applied on that network. Related nodes (programs) share the same color.

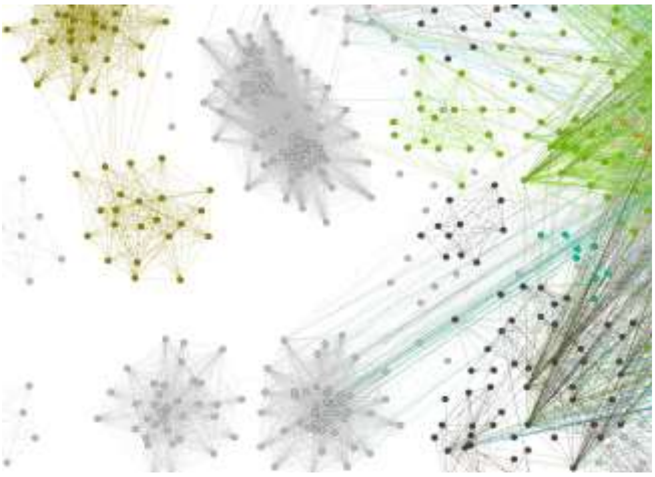

Figure 2

Zoom into the network. Each node is a VoD, each edge is a relation between two programs, each cluster of nodes was discovered by the community detection algorithm, and represents a highly associated group of VoDs.

2. Behavioural Features

This set of features captures the behaviour of an account in terms of time of usage, and taste (genres, etc.). For example: the time of day the account is likely to be more active, the way the content is accessed (linear viewing/VoD) by the subscriber.

For each account, we calculated the number of viewing records per time of day during weekly cycles.

As a result, we got a vector with a dimension of 3 x 3 = 9: three for the beginning of the week, mid-week, the weekend, and an additional three for different times of day: morning, afternoon and evening.

Content Communities

We analysed the content the subscribers watched, in order to model the taste and the viewing preferences the subscribers` choices expressed. We assumed that users tended to watch similar content in terms of show types, genres and other content characteristics.

For each method of accessing content (linear/VoD), we created a graph such that each node in the graph represents a specific program. An edge exists between two content nodes *n1, n2* if a user watched both content *n1* and *n2*. Each edge (*e)* with connecting nodes (n1,n2), was assigned a weight, as the number of occurrences of programs n1 and n2 were watched together. Based on this graph, we looked for strong associations between content and created groups of strongly related items, e.g., all sporting events of the NFL.

Our assumption was that such groups express a set of characteristics that are shared between all items within that group. We ran a community detection algorithm [3] on the graphs, and revealed the underlying communities, i.e., clusters of content items. A community is formed by a set of nodes that are strongly connected to each other, and are not connected much or at all to other nodes.

For each subscriber, we assigned the most dominant clusters according to the content the subscriber watched. We expected to see a few dominant clusters and a stable number of communities per user over time, expressing the user's taste. Exceptionally high numbers of communities, as well as diverse types of communities, may signal that sharing is occurring. Figure 1 illustrates a network of over 40k VoD items. After running the communities detection algorithm, we got as a result 143 meaningful clusters. Figure 2 gives a closer look at the network, focusing on specific clusters.

3. Temporal Features

Features in this group reflect the changes of features over time. For example, we calculated temporal changes in the number of devices as a series of measures of newly used devices, over time within a time interval.

We expected to see that after a short while, there are no more ***new*** devices that are using the same account. If there is an increase over time in the number of devices, the sharing score will increase accordingly. The slope of the series can indicate the sharing-type: a steep slope can imply business sharing when a single account is being shared with many buyers in a short period of time.

4. Spatial Features

Features in this group refer to the spatial aspect of the viewing; in our case, this refers to the geolocation of devices. Spatial features are based on calculations of distances (in miles) between locations (enriched IPs). We expected to see small distances between locations visited by devices in the same account over the time period studied (3 months).

Households

We dedicated a section to this special feature, since this is a key feature in our work. This feature depicts a combination of devices, used by the same account that we believe to belong to the same family, hence, considered to be legal. We used graph analysis to discover households per account from the logs' rows in the following way:

We constructed a bi-partite graph, consisting of two groups of nodes: locations and devices. Per account, we created device nodes for each unique *device_id* at the period of time of the study. Location nodes were created for each unique geo-location we obtained from the IP address. An edge between a device node and location node existed if this specific device had a viewing action from the specified location.

Once this graph was obtained, we checked for the number of connected components on the graph. A connected component is a subgraph where there was a path between any two nodes. In our case, each connected component represented one or more devices active from one or more locations. If a few devices had one or more location(s) in common, we assumed that these devices had been used at the same location within some period of time, hence, belonged to the same household. Figure 3 illustrates a bi-partite graph that forms two connected-components, i.e., two households.

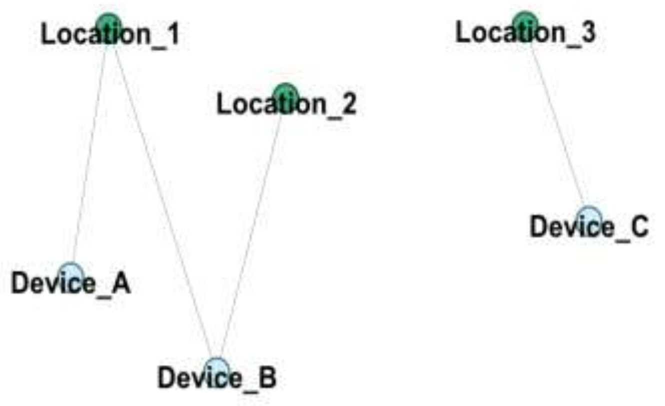

Figure 3

Households graph: Device nodes (blue circles) and location nodes (green circles) form a bi-partite graph. An edge between a device and a location exists if the device had an action from that location. Each connected-component, i.e., a sub-graph where all nodes are connected with edges, represents a household. In this graph, we see two households.

For each account, we expected to find one or at most two households. Thus, accounts with more than two households over a period of time are more suspicious, and likely to have a higher sharing score. As implied, the number of households (i.e., the graph's connected components) was the feature we used in our analysis.

Feature Selection

We extracted and calculated dozens of features. We reduced the number of features to the most relevant to our model. We specifically looked for uncorrelated features, with diverse values. We therefore removed features that were highly non-orthogonal or correlated to other features (i.e., from each block of correlated features, we took a representative feature). We also removed features with very low standard deviation as they didn't contribute to our model. Eventually, we applied principal component analysis (PCA) in order to transform our data to a set of uncorrelated features.

Setting Sharing Score: Anomaly Detection Model

After extracting and selecting the set of features, we determined the final feature vector per account. We associated each account with a sharing score, reflecting the probability that this account is sharing its credentials. For that purpose, we use Mahalanobis distance. The Mahalanobis distance measures the distance between a point *a* (an account) and a distribution *d* (the inferred distribution from all the accounts under research). Mahalanobis distance measures how many standard deviations away an account *a* is from the mean of distribution *d,* this is generalized to n-dimensions. The Mahalanobis distance is given by:

$$\text{E 1. } D = \sqrt{(x-\mu)^T S^{-1}(x-\mu)}$$

Where x is a n-dimensional point, μ is the distribution mean, and S is the covariance matrix.

For each account, we calculate the Mahalanobis distance, and normalize the distance to be in the range of 0-1000.

Differentiating Sharing Types: Rules

We observed two main types of sharing: *casual* and *business/stolen*.

- Casual sharing is the case where credentials are shared within the family or close friends, not for the purpose of profit.
- Business/Stolen sharing is the case where credentials were either stolen or obtained in a fraudulent way, and are used for the purpose of profit. Usually, in such cases, accounts are sold in the black market, possibly, the same account to multiple users.

After detecting sharing accounts, we classified each sharing account into a sharing type, and assigned a “sharing-type” label for each suspicious account. We also fine-tuned the sharing score obtained by Mahalanobis distance.

For the task of assigning the “sharing-type” label, we went back to the *household* feature, and reconstructed it with a reference to the temporal aspect, i.e., as a dynamic graph that is subject to a sequence of updates over time.

Dynamic Graph for Representing Changes in Households over Time

The basic idea was to capture the behavioural changes of the account with respect to time. For example, if more households are added to an account over time, this is very suspicious and may imply business sharing. Also, if a few devices were once connected to the same household, but over time stopped, then, that might imply casual sharing between family members or friends. We model the changes in households as a dynamic graph, i.e., a set of graphs that represents the household’s state at a few points in time. We then extracted features from the dynamic graph to determine the sharing type.

We read the logs as a stream, and processed it in a constant time interval. At each ***time-interval***, we extracted the relevant features for constructing the households graph (Figure 3), and we updated the graph accordingly. The construction and update of the graph is as follows:

- **Initial households graph:** The first time an account appeared in the logs, we constructed the initial graph based on the data in the first interval, the same way the households graph was constructed.

- **Edge weight:** Each edge is assigned a weight, according to the number of transactions found in the logs so far. For example, if account *a* has 200 rows in the logs (i.e., 200 viewing requests), then the correspondent edge will be assigned a weight of 200. Figure 4 illustrates a graph with 1 device that has 2 viewing records from 1 location.

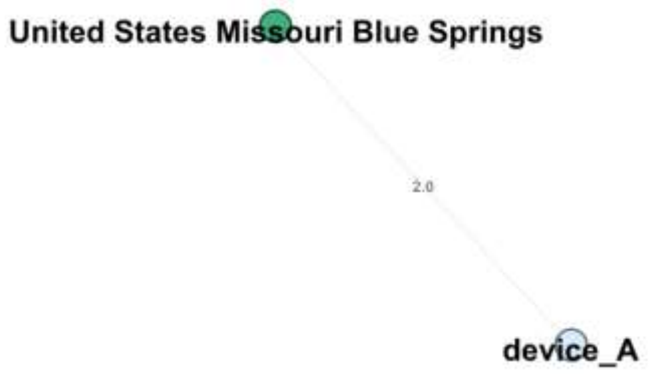

Figure 4

One device at a single location. The graph forms one household, at the first time interval.

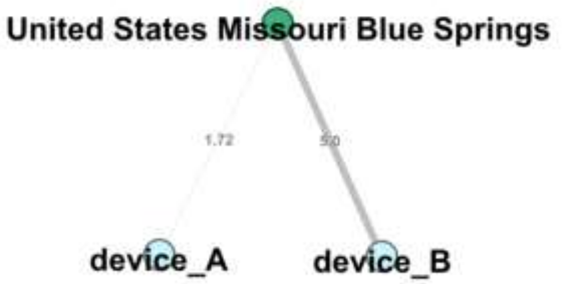

Figure 5

Second device at the same location. The graph still forms one household at the second time interval. Therefore, this account is not suspicious.

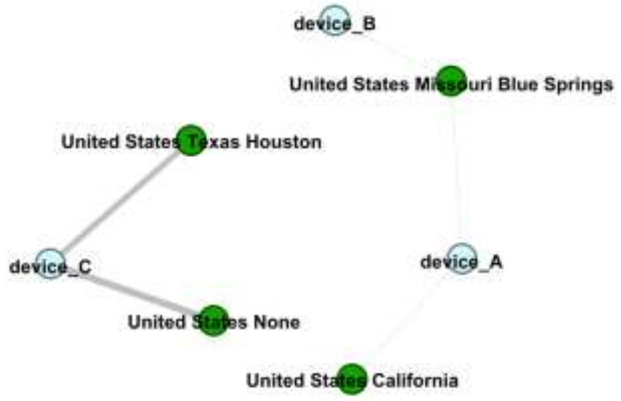

Figure 6

A third device appears in new location, at the third time interval, therefore forming 2 households. This account is starting to be suspicious of sharing credentials.

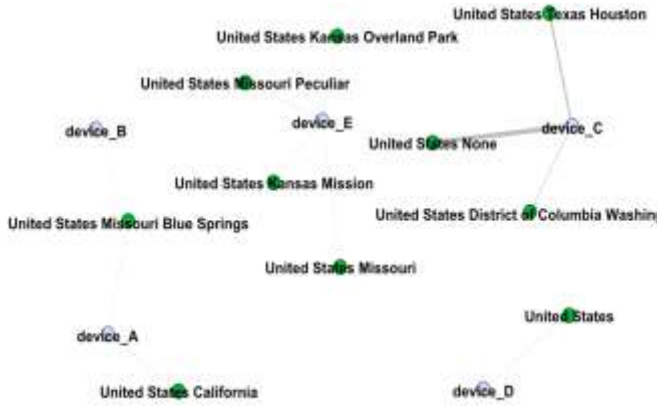
United States Texas Houston
United States Kansas Overland Park
United States Missouri Peculiar
device_B
device_E
device_C
United States None
United States Kansas Mission
United States Missouri Blue Springs
United States District of Columbia Washin
United States Missouri
United States
device_A
United States California
device_D

Figure 7

After the fifth time interval, there are already 5 devices and 10 locations. The graph forms 3 households that are not connected. This account is suspicious of sharing credentials.

Graph Update

The next interval the account appears at, we updated the account’s household’s graph. This update allows the addition of new nodes (device node or location node) as well as the update of an existing edge weight. Each interval, the weight of an edge is reduced by a factor of the time passed. If there was another viewing record from the same device and location, we add this number to the weight after the reduction. The formula for updating the edges weight is given by

$$\text{E 2. } W_{new} = I + W_{old} * \gamma^{time\_passed}$$

where I is the number of viewing records in the current time interval.

This calculation reduces the edge weight such that the resulting weight reflects the importance of the interactions between the device and location, meaning, if a device *d* was used just once from location *l* in the past, over time the edge weight will tend to zero, and we can assume that this edge is not relevant anymore and therefore is a candidate for removal.

As a result of this process, we got a set of graphs, each representing the household according to a specific time interval. Figures [4-8] illustrate dynamic changes in a graph over a time interval, and demonstrate the development of the account over time.

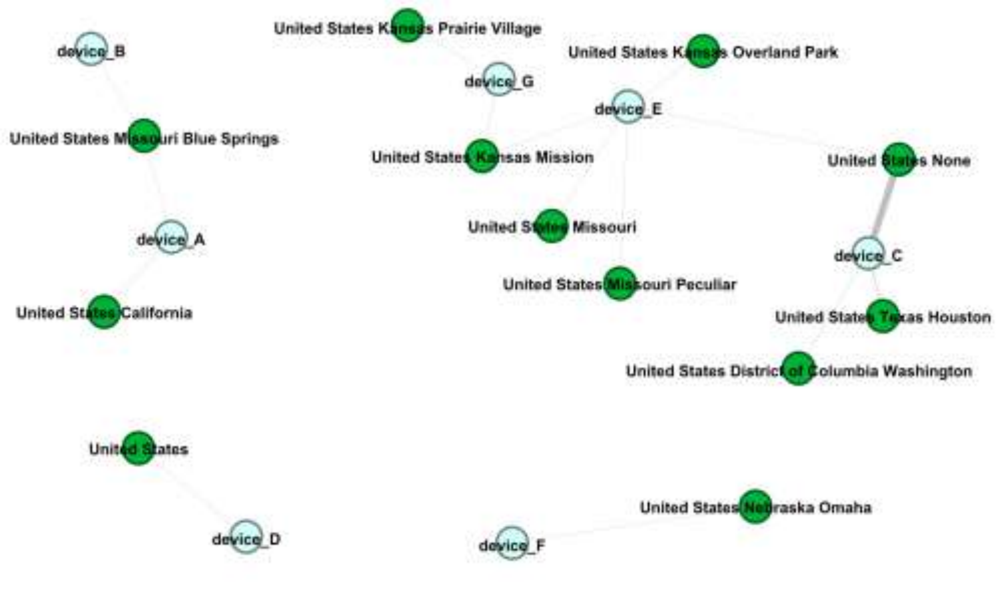
United States Kansas Prairie Village
device_B
United States Kansas Overland Park
device_G
device_E
United States Missouri Blue Springs
United States Kansas Mission
United States States None
United States Missouri
device_A
device_C
United States Missouri Peculiar
United States California
United States Texas Houston
United States District of Columbia Washington
United States
United States Nebraska Omaha
device_D
device_F

Figure 8

After the 11th time interval, there are 7 devices and 11 locations. The graph at this time forms 4 households. Therefore this account is highly suspicious in business sharing.

Extracting Features from the Dynamic Graph

Next, we extracted some features from the set of graphs. These features are used in order to determine the sharing type.

We relied on domain-expert business axioms in order to obtain a set of rules that we used to classify the suspicious accounts. Example of business axioms that are the ground for the assignment of “sharing-types” include:

- A steep increase in the number of households over time implies business sharing for the purpose of profit
- A large geographical distance between households implies illegal credential sharing
- A small geographical distance between households implies legal sharing (could be home/work)
- Two households (or more) that “connect” occasionally and then disconnect (edges are removed) imply legal casual sharing (could be child living in dormitories).

Based on the features we calculated from the dynamic graph and the domain-expert rules, we hand-crafted a set of rules to determine the sharing-type. The features we examined in our decision rules tree include:

- Number of graphs at the period of the time of the study (3 months): a low number weakens the probability an account is sharing its credentials.
- Number of households before and after edge removal: we look for stability of households. A changing household with varied locations and devices is more likely to be suspicious as business sharing.
- Geo-distance between households: We measured the minimum and maximum distance between households. An account with a low number of geographically close households over time is less likely to share credentials. On the contrary, an account with very distant households is likely to be doing business sharing.

We split the decision tree based on the above features, and obtained a tree with over 10 paths to the desired label as a result (Honest, Business/Stolen, Casual).

CONCLUSIONS AND RESULTS

In this study, we presented a data science based scoring method for setting a sharing score per account, which indicates the likelihood of an account to be sharing credentials, as well as a classification method for determining the sharing type (business/stolen, casual) of suspicious accounts. We showed that only a combination of both simple and complex features can successfully model subscriber behaviour and effectively detect suspicious accounts that demonstrate anomalous behaviour. We made use of graph analysis and dynamic graph analysis in our key feature, which allowed us to discover *households* and track the household’s structure over time. We tested our method on real data from a large service provider, and validated our initial results with the service provider and with our security experts.

The practical significance of this work is that the use of our algorithms results in a real-time application that supports actions that can reduce the service providers' loss of revenue and control the credential sharing phenomenon. Actions such as blacklisting devices, tracking activity of suspicious accounts over time, suspending sharing accounts, challenging users in real-time etc. should be available to the service provider, and controllable by a set of configurable parameters.

Results

We ran our algorithm on 3 months of logs. Since this is an unsupervised problem, we set thresholds with the help of a domain expert, which determined the "risk-level" of an account. Figure 9 illustrates the first two projected dimension of our data and the scores (Mahalanobis distances). Figures 10-12 illustrate a few accounts with their associated households graph and normalized score.

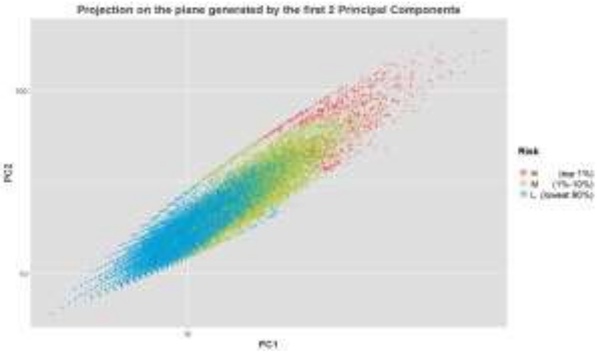

Figure 9

Data projection (PCA), 2 first dimensions. Red dots denote accounts with high risk for credential sharing (top 1% percent), green dots denote accounts with medium risk for credential sharing. Most of the accounts lay on the "safe" area.

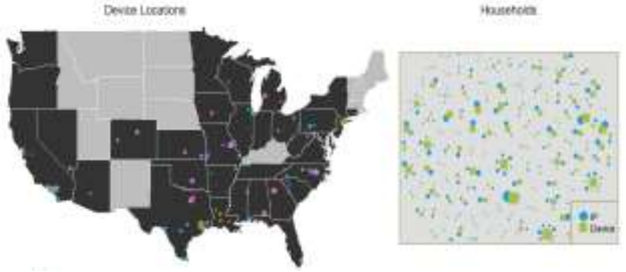

Figure 10

Households graph and device locations of account with sharing score 999 (1st place).

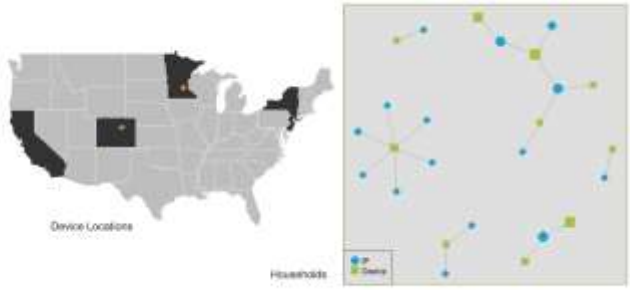

Figure 11

Households graph and device locations of account with sharing score 710 (20k place).

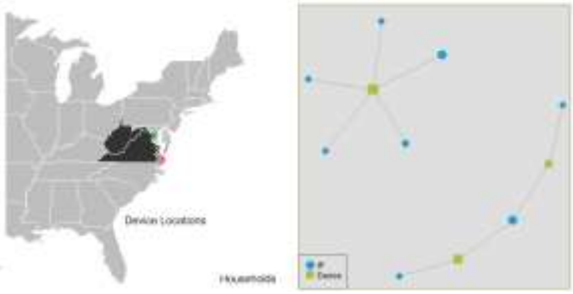

Figure 12

Households graph and device locations of account with sharing score 680 (40k place).

Prototype

We are now developing a prototype based on real service-provider data as part of a wider solution that addresses other security issues relevant to service providers. Figures 13-14 depicts 2 dashboard tabs of the application we developed.

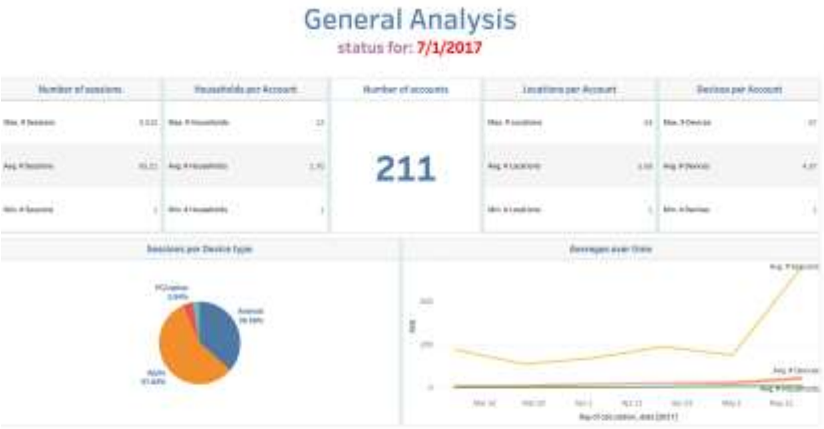

Figure 13

General Analysis tab: This is one of a few descriptive analysis tabs. These screens show general statistics about the data and a subset of the features we extracted.

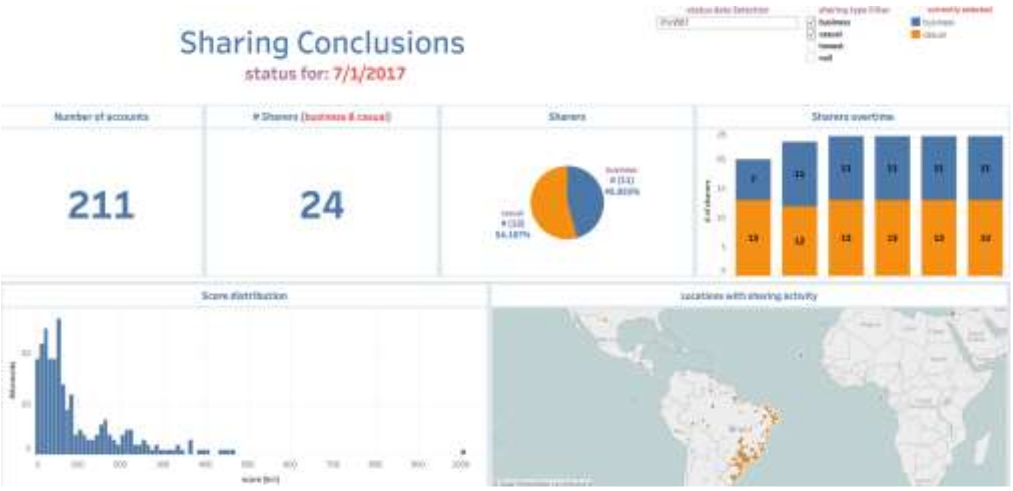

Figure 14

Prototype landing screen: Sharing Conclusions Dashboard, based on the calculations, gives information regarding the current number of sharers in the system, the distributions of the sharing scores (including honest accounts) among the accounts, and the suspicious account locations.

Future Work

In the future, we intend to validate our method on more data sets, and incorporate a feedback loop into our algorithms that will allow us to ingest service providers’ responses and improve our algorithms accordingly.

Endnotes

[1] P. C. Mahalanobis, "On the generalised distance in statistics," in *National Institute of Sciences of India*, 1936.

[2] "IP Geolocation," [Online]. Available: www.maxmind.com.

[3] V. D. Blondel, J. Guillaume and R. Lambiotte, "Fast Unfolding of Communities in Large Networks," *Journal of statistical mechanics: theory and experiment,* 10 2008.