Poster: An Alert System to Avoid Financial Fraud

Tobias Urban, René Riedel, Norbert Pohlmann
Institute for Internet Security, Westphalian University of Applied Sciences
{*urban, riedel, pohlmann*}*@internet-sicherheit.de*

***Abstract*—In the modern information society online transactions are an important part of our daily lives. In this work we present an alert system that determines the current threat level of financial fraud in the internet. We use data from different sources and *off-the-shelf* machine learning algorithms to compute the current threat level. Based on the threat level our alert system issues alerts to raise users awareness of current attack vectors. We tested our approaches with real world online banking frauds. Our preliminary results suggest that this mechanisms can be effectively used to warn users about the current threat situation and therefore help avoiding financial fraud.**

1. Introduction

Online banking and online transactions are a huge part of the modern information society and will even grow in importance. Between 2007 and 2015 the usage of online banking grew from 25% to 46% in Europe [1]. Online banking applications are nowadays successfully attacked by adversaries (e.g. [2]). The total damage caused by these attacks summed up to almost 30 million Euro in Germany in 2014 [3]. Successful attacks on online banking applications (e.g. [2]) are mostly enabled by users who carelessly disclose private information. Hence, the awareness of users has to be raised so that they know about the current attack vectors and how they should act if they are attacked. In this work we present an approach that warns users at times when the threat level for online financial fraud is particularly high. At these times alerts are issued that explain the current threat level to users and how they can protect themselves against current attack vectors. These *in-time* alerts raise user awareness as needed rather than informing them once in general.

2. Identified parameters

If one wants to assess the current threat level it is important to identify the key indicators that influence it. We identified three main categories of indicators for online financial fraud. We used different data sources for each category which are mostly publicly accessible.

- *Phishing* is a broadly used strategy utilized by fraudsters to steal private information from users [4] and has many different manifestations. In this work we consider three of them: (1) Phishing *websites* are a common way to trick users into entering private information on a fake website. Thus, the amount of reported phishing sites (by [5]) is used as indicator for the current threat situation. (2) SPAM *mails* and (3) SPAM *messages* are also common ways to steal private information, for example by luring users onto phishing websites. We used emails (gathered by [6]) and messages that were posted to the Stack Exchange network (provided by [7]).
- *Botnets* and especially banking Trojans still pose great danger (e.g. [8]) to customers of online banking applications. We use the amount of detected banking Trojan infections as an indicator for the current threat level. A large anti-virus vendor provided us with the amount of banking Trojan infections that were detected by their products.
- Publicly known *vulnerabilities* (provided by the National Vulnerability Database (NVD) [9]) are used to indicate the risk that a user might get infected with a banking Trojan.
- To check the accuracy of our developed approaches we use real online banking frauds. A service provider for a banking group provided us with real frauds that have been reported to them.

3. Extracting and weighting the parameters

3.1. Filtering the parameters

Most of the data sources we used do not contain information exclusively related to financial fraud. Thus, the data sources were filtered so that we can conduct further analysis on data that are directly related to financial fraud. At this stage of our work we use a simple list of keywords to filter the parameters. Furthermore, we only consider vulnerabilities that can be used by an adverser to remotely infect a system with malware, based on the information given in the NVD.

3.2. Metric to measure the effectiveness of an alert system

The effectiveness of an alert system S is related to the amount of frauds that are warned about by the system. Obviously a system that publishes an alert every day will always 'warn' about all frauds. Hence, a reduction of the effectiveness of the system is needed for each alerts that is published. Following this approach we use the following formula to compute the systems' effectiveness: Let Ω be the total amount of frauds and T the time span of the test. Let T_{Alert} be the time span in which an alert is active, n days after the alert, with ω the amount of frauds within T_{Alert}. We define the effectiveness of the alert system as follows: $eff(S) := \frac{\omega/\Omega}{T_{Alert}/T} = \frac{\omega \cdot T}{\Omega \cdot T_{Alert}}$. The formula ensures that the effectiveness of the system increases if ω increases and/or T_{Alert} decreases.

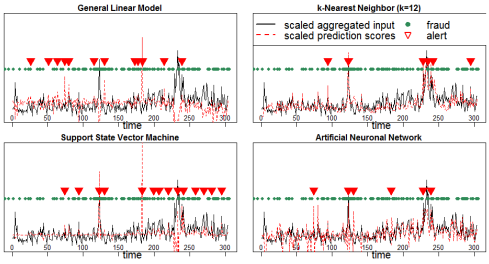

Figure 1. Results of the different approaches.

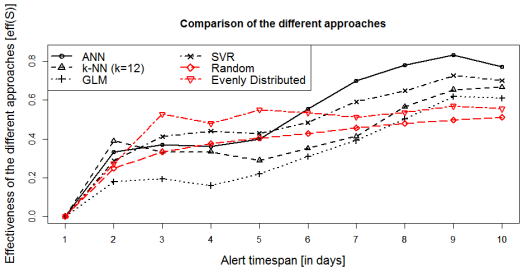

Figure 2. Comparison of the baselines (red) and results of our *off-the-shelf* approaches (black).

4. Determination of alerts

Before utilizing the machine learning methods we aggregated all parameters by day and normalized the weights for the parameters. We used different *off-the-shelf* machine learning methods to determine the times when an alert should be issued. For each approach we computed a threshold that indicates when an alert should be issued. We use $^1/_3$ of our data (test set) to train the different approaches. We compute the threshold by forecasting the training set and use the 95% quantile of the forecasted values as threshold.

4.1. Unsupervised learning

For the unsupervised learning approach we aggregated all different parameter groups to a single value for each day in our test interval (see the black line in Fig. 1). We used the k-Nearest Neighbor (k-NN) algorithm on this time series to detect outliers by computing the Euclidian distance to its k *left* neighbors. A value is considered an outlier if the distance is greater than the computed threshold.

4.2. Supervised learning

Aside of the unsupervised learning approach we used different *off-the-shelf* supervised learning algorithms. The amount of frauds that occurred during the time span of days (n) after a given day (t) is used as dependent variable. We used the following three approaches: (1) A general linear model; (2) Support vector regression (SVR), with a polynomial kernel of degree 3. The hyper parameters of the SVR were optimized using grid search; (3) A (3, 3, 1) feed forward artificial neuronal network (ANN). The network is build performing resilient backpropagation without weight backtracking [10]. Results of the implemented approaches are displayed in Fig. 1.

4.3. Baselines

We compared our mechanisms with two different baselines: (1) We generate 16 alerts at random times and measured the effectiveness of those alerts. We computed the mean effectiveness, over 100 iterations, of these alerts and used the computed value as first baseline; (2) We divided the given timeline into 16 chunks of equal size, issued an alert for each chunk, computed the effectiveness of those alerts, and used it as second baseline. We used 16 alerts because the unsupervised and supervised approaches issued around 16 alerts.

5. Future work

Aside of improving the mechanisms to determine alerts issuances, the design and user acknowledgment of these alerts need to be investigated. We designed different alerts for different communication channels (e.g. email, pop-ups etc.) and are currently conducting a user study to test how subjects react to different alerting channels and alert designs.

6. First results & Conclusion

In this work we presented how *off-the-shelf* machine learning algorithms can be used to compute the current threat level in the online banking business. The effectiveness of our tested approaches are displayed in Fig. 2. For the longest alerting interval (10 days after an alert) each tested approach outperforms our baselines. Our preliminary results suggest that alert systems that use the proposed approaches can be a useful tool to assist and warn users of the current threat situation.

References

[1] Eurostat, the statistical office of the European Union, "Individuals using the internet for internet banking," 2016. [Online]. Available: goo.gl/FJJSha

[2] S. Golovanov, D. Makrushin, and A. Monastyrsky, "Staying safe from virtual robbers," Moscow, 2013. [Online]. Available: https://securelist.com/analysis/user-advice/58328/staying-safe-from-virtual-robbers/

[3] Federal Criminal Police Office, "Bundeslagebild Cybercrime 2014," 2014.

[4] N. P. Singh, "Online Frauds in Banks with Phishing," *The Journal of Internet Banking and Commerce*, vol. 2007, 2007. [Online]. Available: http://www.icommercecentral.com/open-access/online-frauds-in-banks-with-phishing.pdf

[5] OpenDNS, "PhishTank — Join the fight against phishing," 2016. [Online]. Available: https://www.phishtank.com/

[6] Bruce Guenter, "SPAM Archive," 2016. [Online]. Available: http://untroubled.org/spam/

[7] Stack Exchange Inc, "SmokeDetector," 2016. [Online]. Available: goo.gl/uby0j0

[8] N. Etaher, G. R. Weir, and M. Alazab, "From ZeuS to Zitmo: Trends in Banking Malware," in *2015 IEEE Trustcom/BigDataSE/ISPA*, 2015, pp. 1386–1391.

[9] National Institute of Standards and Technology, "National Vulnerability Database," Gaithersburg, 2015. [Online]. Available: https://nvd.nist.gov/

[10] M. Riedmiller and H. Braun, "RPROP - A Fast Adaptive Learning Algorithm."