SPONSORED BY THE

Federal Ministry of Education and Research

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

I. 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. Online banking applications are nowadays successfully attacked by adversaries. The total damage caused by these attacks summed up to almost 30 million Euro in Germany in 2014. Successful attacks on online banking applications 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 alert system 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.

II. 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 (phishing websites, messages, and emails; banking Trojan infections; publicly known vulnerabilities - see Fig. 1). We used different data sources for each category which are mostly publicly accessible. To check the accuracy of our developed approaches we use real online banking frauds.

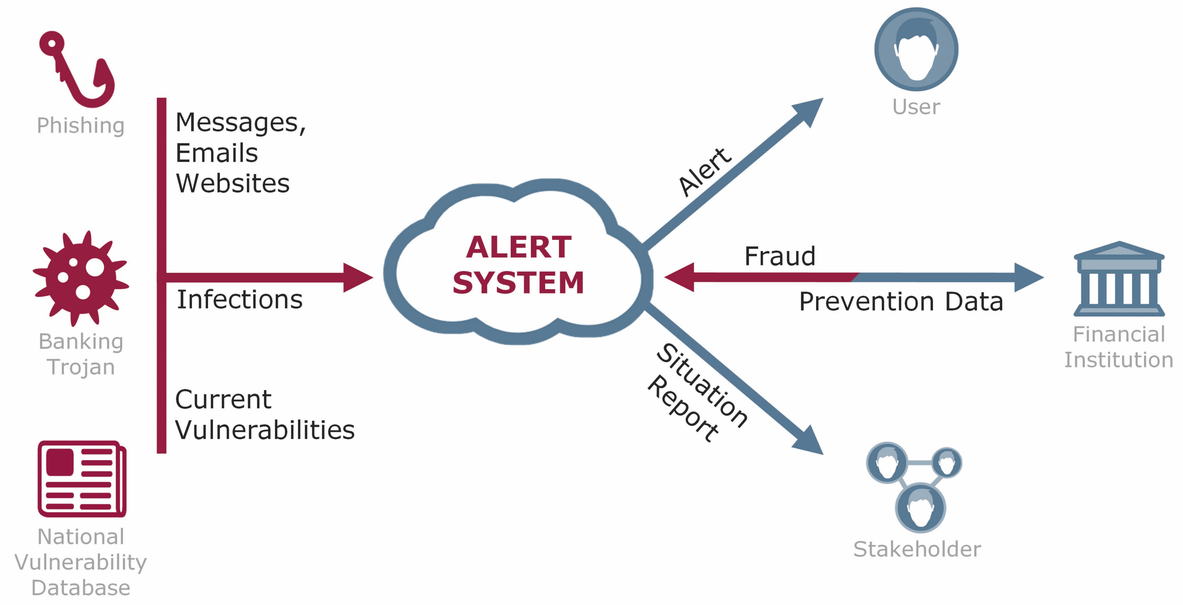

Fig. 1: Overview of the alert system.

III. METRIC TO MEASURE THE EFFECTIVENESS

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:

$$eff(S) := \frac{\omega/\Omega}{T_{Alert}/T} = \frac{\omega \cdot T}{\Omega \cdot T_{Alert}}$$

IV. DETERMINATION OF ALERTS

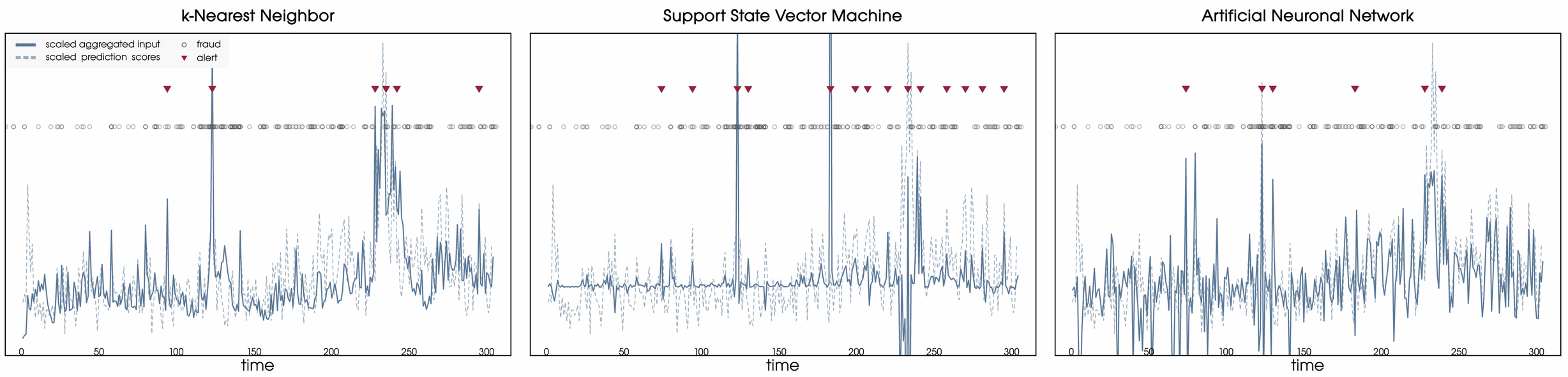

Fig. 2: Results of the different approaches.

A. 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 solid, blue line in Fig. 2). 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 a computed threshold.

B. 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. Results of the implemented approaches are displayed in Fig. 2.

C. Baselines

We compared our mechanisms with two different baselines: (1) We generate 8 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 8 alerts because the unsupervised and supervised approaches issued around 8 alerts.

V. PRELIMINARY RESULTS & FUTURE WORK

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. 3. 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. We designed different alerts (an example is given in Fig. 4) 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.

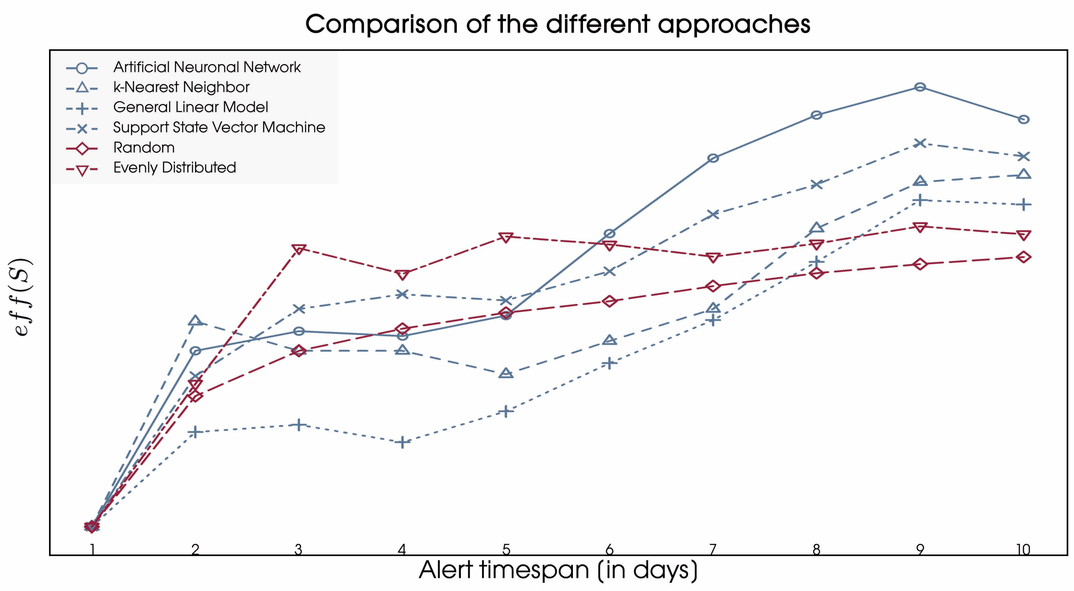

Fig. 3: Comparision of the baselines (red) and results of our *off-the-shelf* approaches (blue).

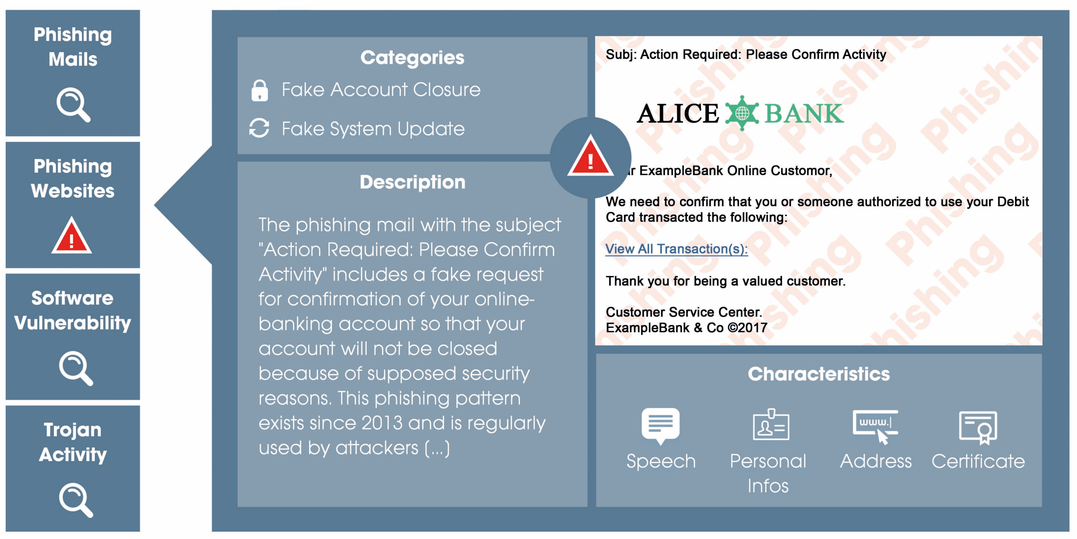

Fig. 4: Example of an alert about high spam activities.