Improvement of User Profiling, Call Destination Profiling and Behavior Pattern Recognition Approaches for Telephony Toll Fraud Detection

Anton Wiens, Sandra Kübler, Torsten Wiens, and Michael Massoth
Department of Computer Science
Hochschule Darmstadt – University of Applied Science
Darmstadt, Germany
e-mail: {anton.wiens | sandra.kuebler | torsten.wiens | michael.massoth}@h-da.de

***Abstract* — Phone fraud attacks cause a massive loss in the telecommunication sector every year. The internet raises new potentials for these attacks. Attackers can use the internet to get illegal access to telecommunication devices such as Voice over Internet Protocol (VoIP) phones and use them for fraudulent calls to expensive destinations in overseas countries. The generated financial damage affects the attacked customers and also the telecommunications companies that provide the telecommunication services. Especially, attacks on small and medium-sized enterprises can threaten their existence. This demands protection for the customers and companies by a fraud detection system with intelligent detection techniques for detection and prevention of these fraud cases. In this work, we present two statistical online user profiling approaches, as well as a call destination profiling approach and a behavior pattern recognition approach for toll fraud detection. All four approaches show good and promising results. Especially, the results of the call destination profiling, for detection of distributed attacks on a single destination, and the behavior pattern recognition, for detection of change in a user's behavior patterns, are promising for future work.**

Keywords—Fraud detection; telephony; CDR; behavior profiling; statistical.

I. INTRODUCTION

This paper is an extended version of [1]. The previous work got improved and combined with other works, which were derived from previous work [2] [3] of the authors for better results. We enrich this paper with new information about the previous work, subsequent improvements and derivations of it. Therefore, an initial user profiling approach, a new user profiling approach, a call destination profiling approach and a behavior pattern recognition approach are presented that try to solve the problems described in the following paragraphs.

Today's voice communication by Voice over IP (VoIP) mostly uses the internet for data transport. There are the drawbacks that the internet can basically be accessed by anyone, and that it links anyone to anyone. For example, it is possible for third parties with criminal intent to access private branch exchange (PBX) systems connected to the internet.

Fraudsters may have multiple options to abuse these systems. Systems that are insufficiently secured may be tapped. Access data that has been saved in these systems could be used to abuse, compromise or even gain full access to the whole PBX. If the PBX system has been taken over, a fraudster will be able to conduct telephone calls to premium rate service numbers or comparable call destinations, generating profit. The resulting cost, on the other hand, will often be charged to the victim or its telecommunication service provider, because of a general rule in telecommunication service providing called "Calling Party Pays".

The Communications Fraud Control Association (CFCA) reports losses of about 46 billion USD caused by telecommunication fraud in 2013, an increase by 15% compared to 2011 [4]. Not only financial damage is a problem caused by fraud attacks. Small providers may also suffer from reputation losses, causing customers to change the provider because of decreased trust and fear of repeated fraud attempts in the future.

To detect and counter these attacks, respectively fraud attempts, fraud detection systems are used. Often, these systems apply methods based on the generation of statistical profiles for each user. User profiles are generated that describe their behavior. These profiles will then be used as input for machine learning techniques, allowing for the detection of fraud [5] [6] [7] [8].

The German company "Deutsche Telekom" reported a huge success in the prevention of fraud cases with potential damages of about 200 million Euro, using an automated fraud detection system [9]. The research project "Trusted Telephony" at the University of Applied Sciences Darmstadt, from which the work at hand originates, pursues the goal to increase security in VoIP telephony, cooperating with a German telecom service provider. A key objective of the project is the development of a fraud detection system.

Recently, fraud cases were caused by security exploits in FRITZ!Box hardware (from the company AVM GmbH), which is often used in Germany [10] [11]. A FRITZ!Box is an integrated, multifunctional routing device, offering internet connectivity, VoIP capabilities and other services in local area networks. Such a unit is very popular in Germany. Because of the large amount of units in use, there is an increased risk in case of security vulnerabilities, especially for private users.

On the other hand, an exploitation of the recently disclosed security vulnerability of such a unit is only one possibility to start such attacks. The security vulnerability has been patched by the manufacturer in the meantime, but in the future, comparable vulnerabilities in similar hardware could turn up.

Therefore, it is important to be able to detect these cases and devise measures to counter them.

Figure 1. Depiction of the generation of call detail records and a detection system using them.

A. Call detail records

The data being analyzed in this work comprises fraud attacks that have been enabled by the occurred and already patched security vulnerability of the FRITZ!Box units [10].

A CDR is a text file containing all parameters of single telephone calls. Each CDR is written by a primary VoIP routing system called TELES.iSWITCH [12] as calls are set up (see Figure 1). CDRs contain information on caller, callee, call duration, starting time, as well as technical network parameters.

B. Structure of the paper

The structure of this paper is as follows: The related work is presented in Section II, followed by an explanation of behavior profiling in Section III, as it is important for the described approaches. Section IV gives an in-depth analysis of attacks on FRITZ!Box units. The following sections except future work and conclusion provide a description of various approaches, each followed by an experimental setup including results and a short conclusion. Section V is about the previous basic user profiling approach, which is being extended by the work at hand. In Section VI, a new basic behavior profiling concept is presented. In order to adapt to the FRITZ!Box incident and to detect these fraud cases, another approach by the authors called destination profiling is described in Section VII. The concept of communication behavior patterns, which also deals with the FRITZ!Box incident, is outlined in Section VIII. An overall conclusion is presented in Section IX, followed by future work in Section X.

II. Related work

This paper is an extended version of [1] and also includes intermediate works [2] [3] that improved the quality of presented techniques greatly.

The first intermediate work presented a behavior profiling different from user profiling to cope with distributed single target toll fraud attacks. The second intermediate work introduces behavior pattern recognition to detect changes in patterns of the users. Patterns can be defined and extended dynamically.

All of the previous and intermediate work is improved in this paper.

In [1], a method for toll fraud detection using statistical user profiling has been described, which can especially be applied when no significant amount of training data is available. Additionally, the method can be run in a mostly autonomous way, requiring only a minimum amount of external administration. The method applies two user profiles, one for a past period of time and one for a present period of time, each containing statistical features. The profiles are used to identify suspicious deviations of the user's behavior, by which toll fraud attempts are detected. In this work, the attacks on FRITZ!Box hardware and the possibility to detect these using the presented method had already been mentioned.

In the work at hand, the method from the preliminary work is adapted more closely to this attack pattern. The new method again uses two profiles of statistical features for each user, but differing in contents and their actual use for the detection of attacks.

Furthermore, other related work also describes different methods of user profiling for the detection and prevention of toll fraud in VoIP telecommunication [5] [6] [8] [13] [14] [15]. In contrast to this work, the previous work [2] does not apply simple user profiles, but a new kind of profile specified as Call Destination Profile. These profiles are used to characterize the behavior of a destination telephone number instead of a user's behavior. It is intended to detect special kinds of attacks this way.

These attacks cannot be detected with user profiling techniques alone and hence would go undetected if the method from [1] was applied.

As a means to visualize user accounts, self-organizing maps (SOM) are used in [16]. This visualization is used to differentiate between normal and fraudulent ones. The features *call destination*, *call start time* and *call duration* are extracted from the CDR data and used for analysis. According to the authors, the method has a true positive rate (TPR) of 90% and a false positive rate (FPR) of 10%.

A framework for self-organizing maps has been developed by Hollmén, Tresp and Simula [17] to cluster probabilistic models. User profiles using data of mobile communication networks have been used for test runs of the system. The output is presented visually, so that the fraudulent calls can be distinguished from normal ones.

The authors of [18] focus on the detection of superimposed fraud using two signature methods, each summarizing a user's behavior. The first presented approach is based on a deviation of the user's current behavior and his signature, while the second is based on a dynamic clustering analysis. In the second approach, a sudden change or "shift" of a user's signature from one cluster to another is the criterion for a classification as fraud. The similarity between a signature and a cluster centroid, which in itself is defined as a signature, is crucial for such a shift. The detection rates of both methods have been estimated: The first one promises a TPR of 75% and the second one a TPR of 91%. Also, a combination of both approaches is examined.

The framework *SUNsHINE*, which is able to detect and prevent VoIP fraud by combining real-time capable components with an offline statistical analysis, is presented in [19]. Multiple data sources,

network traffic data and CDRs, can be used. Different algorithms and techniques are used, e.g., rule sets, profiling, neural networks and clustering. No estimations concerning the detection rate are given.

The intermediate work [3] has been inspired by the concept of clustering algorithms, as the aspect of finding similarities has been adopted, leading to a different point of view in contrast to [2].

III. Behavior Profiling

The term "behavior profiling" describes a technique for differential analysis where the behavior of a given object is represented by a statistical profile. In literature, a distinction is made between absolute and differential analysis.

An absolute analysis examines a whole set of data, trying to identify fraud cases, but does not consider different types of user behavior. A call that may be treated as a fraud case for one user could be no fraud case for another user. For example, one user only makes long calls to his family at weekends and the other user only makes long calls to his family at workdays. If an absolute analysis considers long calls at workdays as fraud cases, the latter user will be considered as fraudulent, just because his normal behavior does not comply with the definition of normal behavior given by the other user. This problem can be avoided by looking at each user and his behavior differently, thus called differential analysis.

Differential analysis is preferred to absolute analysis in most of the related work, e.g., [20] [7] [21] [22]. The main argument is the ability of differential analysis to include the absolute analysis. In other words, a fraud case detected by an absolute analysis can also be found by a differential analysis, but a fraud case detected by a differential analysis cannot always be found by an absolute analysis [20].

In the profile, data from the object is accumulated, which is then used to generate statistics that describe the object's behavior, which are called features. Often, behavior profiles are applied in the form of user profiles [5] [6] [8] [13] [14] [15]. In most cases, a differential analysis is preferred over an absolute analysis. This is because the absolute analysis is a subset of the differential analysis [9].

For example, three variants of user profiling methods are presented in [7]. In this work, the parameters *duration per call*, *number of calls per customer* and *costs per call* are arranged in different ways into the groups *national calls*, *international calls* and *mobile calls.* These are used to generate statistics for the profiles.

User profiles are utilized to describe the behavior of users in the present and in the past, enabling a comparison of behavioral patterns. By this comparison, it is possible to detect suspicious fluctuations. These are analyzed in the next step in order to generate a decision on fraudulent or non-fraudulent behavior.

IV. Analysis of attacks on FRITZ!Boxes

The recent attacks at (and by) FRITZ!Boxes can be divided in two categories. The first category comprises the hostile take-over of a FRITZ!Box by exploiting a security vulnerability in its firmware. The second category comprises possible results of such a take-over, especially secondary attacks that are enabled by then remotely controllable units. Both categories are described in more detail in the following subsections. It is important to note that the initially possible attacks on these units cannot be conducted anymore, since the firmware has been updated by the manufacturer in the meantime [11]. The focus of the work at hand is at the possibility of fraud attacks on telecommunication systems by utilizing taken over secondary hardware, which is not unlikely to happen again in the future, and detecting it.

A. Primary hostile take-over of a FRITZ!Box

The basic idea to perform a hostile take-over of a FRITZ!Box was as follows: An attacker would set up a web site, which is to be visited by potential victims. The attacker would then be able to exploit the known security vulnerability of the FRITZ!Box in order to extract the master password. Using this password, the attacker would be able to access the command shell. Once this is done, the attacker could then deploy system commands, e.g., to make the unit call premium rate service numbers at the cost of the unit's owner [11].

B. Secondary attacks after the take-over

Attack attempts on other systems that had been conducted using taken over FRITZ!Boxes seem to be very similar in their basic approach. For an in-depth analysis, anonymized data on such attack attempts has been provided by a telecom company. The data being used is in accordance to the Federal German Data Protection Act (Bundesdatenschutzgesetz) [23]. All results from this analysis are based on this data and may not represent attack patterns that appeared at other telecommunication providers.

From a single user's view

From the perspective of a single user, an attack attempt may look as follows: An attacker tries to set up a call to a premium rate service number or a comparably expensive call destination, possibly also in another country. This is done multiple times during a short time span. As soon as the attacker has successfully set up a call to a given number, he will try to call this number again, as often as possible, and also in a short period of time. If the call attempts fail (e.g., because the number is not available), the attacker will try another number.

The difficulty to detect such attack attempts lies in the low frequency and the low duration of these calls seen from a single user's point of view. Attackers will avoid a detection using these two parameters by applying an approach described in the next section.

Exploiting multiple users

By exploiting the security vulnerability at multiple victims' FRITZ!Box units, attackers are able to hide their attack attempts neatly. The attack attempts are distributed across multiple taken over units. So, it becomes possible to mask obvious evidence of attack attempts, such as frequency and duration of calls. This will be illustrated by the following examples:

1. Attacker A conducts a hostile take-over of victim C and causes C's unit to start 30 calls to destination number B. The duration of each call is 20 seconds.
2. Attacker A conducts a hostile take-over of victim C and causes C's unit to start 5 calls to destination number B. The duration of each call is 5 minutes.
3. Attacker A conducts hostile take-overs of 30 victims and causes each victim's unit to conduct one call to destination number B. The duration of each call is 20 seconds.

In the first example, the attack at victim C can be detected by the frequency of the calls. In the second example, the attack can be detected by the extraordinarily long duration of the calls. In the third example, the features used before cannot be used again. Figure 2 shows a depiction of example three with just two victims of an attacker calling a premium service number.

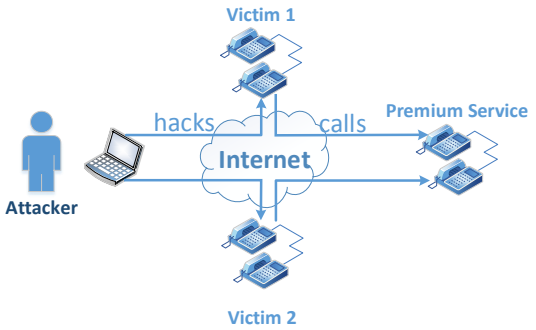

Figure 2. Depiction of example three with two victims of an attacker calling a premium service.

Existing methods often apply user profiling to detect suspicious behavior and potential attack or fraud cases. This way, distributed attacks, as described in the third example, cannot be detected. Therefore, it is necessary to apply a different method for detection.

C. Characteristic traits for detection

From the results of the preceding section, the following characteristic traits for detection can be deduced:

- **Duration of call for a certain user**: The call duration is significantly higher in comparison with the known behavior of that user.
- **Number of calls for a certain user**: The number of calls in a given time span is significantly higher in comparison with the known behavior of that user.
- **Number of calls for a certain destination number:** The number of calls that have been conducted to a given (premium rate service) destination number in a given time span is suspicious.

The first two of these characteristic traits can be detected by applying user profiling if the perspective of a single user is applied. To be able to detect attack attempts using the number of calls, a new method has to be devised. This will be described in Section VI.

V. PREVIOUS BASIC USER PROFILING APPROACH

The description and improved results of our previous user profiling approach will be provided in this section for completeness. The improvements to this concept are described in the sections of the new concepts following this section.

A. Constructing user profiles

For each user, two user profiles exist that represent the present and past behavior in specified time spans. The profile describing the past is called Past Behavior Profile (PBP), and the one describing the present is called Current Behavior Profile (CBP). Each profile uses features, calculated from CDR data, to describe the user behavior in its time span.

Features

Features describe different aspects of a user's behavior. In the profiles, the feature vector shown in Table I was used:

TABLE I. FEATURE VECTOR USED FOR USER PROFILES [7]

Max Calls	Max Duration	Max Costs	Mean Calls	Mean Duration	Std Calls	Std Duration

These are the maximum values (Max) for calls per hour (Calls), the duration of a call and the cost of a call, the mean value (Mean) and standard deviation (Std) for the same CDR information, except the cost.

For those features, the start time, duration and cost information of a CDR are needed. The cost of a call is depending on the user agreement and is not given in a CDR. Therefore, an approximation of costs for a CDR was made, based on country code, number type (mobile or fixed-line) and duration.

These features were used because they delivered the best results in [7]. Many works use standard deviation and mean values of the number of calls and the duration of a call to describe the user's behavior. Some works also differentiate them into national, international or mobile [21] [7] [22].

Profile time span

Each profile P has a length l_P. The PBP additionally has an offset $d \neq 0$, describing the difference in time between the present and the PBP time span (see Figure 3). For a CDR to be included in a profile, it needs to meet the following rules (1) and (2) for the corresponding profile:

$$T_{cdr} < T_n - d \quad (1)$$

$$T_{cdr} \geq T_n - (l_P + d) \quad (2)$$

T_n is the present (n) time, and T_{cdr} is the time of the CDR. If a CDR meets these two rules, it is included in the features of the corresponding profile.

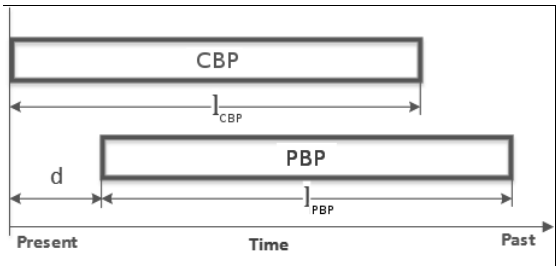

Figure 3. Profile time spans and offset (CBP = Current Behavior Profile; PBP = Past Behavior Profile).

The length (time span) of the profiles and the offset are very important parameters for the detection. The longer a profile is, the more CDRs are represented inside a profile and the statistics have more accuracy and less fluctuations. At the same time, the effects of single fraudulent CDRs become statistically more irrelevant and thus harder to detect. The offset is important for finding fraudulent CDRs that can only be found in groups. It decides how long it takes for a yet undetected fraudulent CDR to be included in the PBP and therefore makes it more unlikely to be found. The length of the offset also affects fluctuations when comparing both profiles. A higher offset causes higher fluctuations, a lower offset causes lower fluctuations likewise.

An optimal tradeoff between the length of the profiles and the offset between profiles needs to be found for best results.

Filling profiles

At first, the profiles need to get filled up for the method to be able to calculate meaningful features. Once the profile contains CDRs for its entire time span, the features can be calculated and used for further analysis. This means that the method has a determined training time for accumulating CDRs that is autonomously done without administration by personnel. In the following, a profile that has been filled up once is called *ready*.

B. Measuring change in user behavior

Once the profiles of a user are *ready*, the change of behavior measured by the profiles can be calculated. This is done by calculating the relative ratio R_F between each feature F of both profiles (PBP and CBP) by (3):

$$\forall F : R_F = \begin{cases} \left(1 - \left(\frac{F_{PBP}}{F_{CBP}}\right)\right), & F_{PBP} \leq F_{CBP} \\ \left(1 - \left(\frac{F_{CBP}}{F_{PBP}}\right)\right), & else \end{cases} \quad (3)$$

This results in a ratio R_F for each feature F, describing the change in behavior for that feature. Each R_F has a range of -1 to 1, with -1 as a maximum decrease and 1 as a maximum increase in behavior measured by that feature.

A ratio R_F for a feature F gives a relative value to the past behavior. It is relative because the severity of a change in user behavior is always relative to the past behavior of the user.

Empty profile

In case of a user not having made calls for a time span greater than the span of all user-specific profiles, one of the profiles of a user can run empty. Once a profile is empty, the calculation of the features is not possible, because they attain a value of zero. Comparing a non-empty profile with an empty profile will result in infinite ratios for the features, allowing for detection of fraud where there is none (e.g., when the PBP is empty and the CBP is not empty). Instead of letting the profile run empty, the last CDR in a profile that is about to become empty is not removed. This prevents the features from getting zero values and keeps user-specific information for fraud detection. Setting the features to a standard value would disregard user-specific behavior and is therefore not done.

Features accepting zero

Features like standard deviation can attain a value of zero, even if the profile is not empty. For example, the standard deviation of the duration attains a value of zero, if all calls in the profile have the same duration. Like in an empty profile, zero values are a problem for calculating the ratios. Therefore, a value ε (depending on the range of the specific feature) is added to the affected feature in both profiles.

C. Detecting fraud

For this approach, fraud cases are to be distinguished by extreme changes in user behavior described by each feature. Thus, for each ratio R_F of a feature F, a limit L_F is introduced. Each ratio R_F is checked if its limit L_F is exceeded, and the number (n) of exceeded limits is checked against an additional limit L_E (E for exceedings). If the limit L_E is exceeded, the CDR is labeled as fraudulent and as non-fraudulent otherwise. The procedure can be described as follows:

1. Set $n := 0$
2. $\forall R_F \in R: (R_F > L_F) \rightarrow (n = n + 1)$
3. $result = \begin{cases} fraud, & n > \mathrm{L_E} \\ normal, & else \end{cases}$

Once a CDR in the CBP is labeled as fraudulent, it is to be excluded from inclusion into the PBP. This prevents the PBP from including fraud cases and obscures potential follow-ups of fraudulent CDRs. This is the first approach chosen for a first experiment. Other approaches for detection using the ratios are discussed in future work.

D. Unexpected fluctuations

Many fluctuations in data and ratios, like weekends and holidays, can be predicted and adjusted for. But there are also fluctuations caused by random events inside the telecom service provider's network, e.g., network, hardware or other failures.

Those fluctuations are hard to predict using user profiles. The idea is to use the relation between absolute and differential analysis. If it is a fluctuation caused by the specific user, the fluctuation is not seen in an absolute analysis. If the fluctuation is global, it will affect all users and will be seen for specific users, too. Therefore, the accumulated behavior of all users has to be measured to detect this kind of fluctuation.

Because the functionality to measure user behavior has already been defined, it can be reused to measure the accumulated user behavior. A global version of a CBP and a PBP is needed for all users. Ratios are calculated the same way as in user profiles. In this case, the ratios are not used for fraud detection, because the source of the fraud cannot be detected by creating profiles for all users. The ratios are used to be included in the user-specific ratios for finding the global fluctuations and removing them from user fluctuations.

The inverse ratios of the global profiles are taken to the power of g and are multiplied with the corresponding ratio of a specific user profile as in (4):

$$newratio = (\mathbf{1} - globalratio)^{g} \cdot userratio \qquad (4)$$

An appropriate value for g is determined in Section V.F. Both ratios have the same scaling and global ratio that describes the change for the user ratio that is still normal. Therefore, the inverse is multiplied by the user ratio. Because the global ratio is much more stable with more samples, it is taken to the power of g. g is dependent on the scaling of $globalratio$ and not on $userratio$.

E. Low usage users

An analysis of the data revealed that on average, each user only makes 6-7 outgoing calls per day. About 47% of the users only conduct 2 calls per day on average. That means a lot of users — and therefore user profiles — include low amounts of calls. Hence, only few samples are available for calculating the statistics, making the statistics inaccurate. A way to handle those fluctuations is to scale the calculated ratios for the user by the number of samples inside the profiles. For the creation of a scaling function $S(x)$, the dependencies of the number of calls in the profiles and the ratios needed to be analyzed.

Before and after scaling a ratio, it needs to be converted to linear space with (5).

$$S(x, y) = 1 - \frac{1}{\left(\left(\frac{1}{1-y} - 1\right) * S(x)\right) + 1} \qquad (5)$$

x is the number of calls in the PBP, and y is the ratio to be scaled. The part $\left(\frac{1}{1-y} - 1\right)$ scales the ratio into linear space, and $1 - \frac{1}{(\dots) + 1}$ reverts it back to the previous space. A full overview of all components and their relationships is shown in Figure 4.

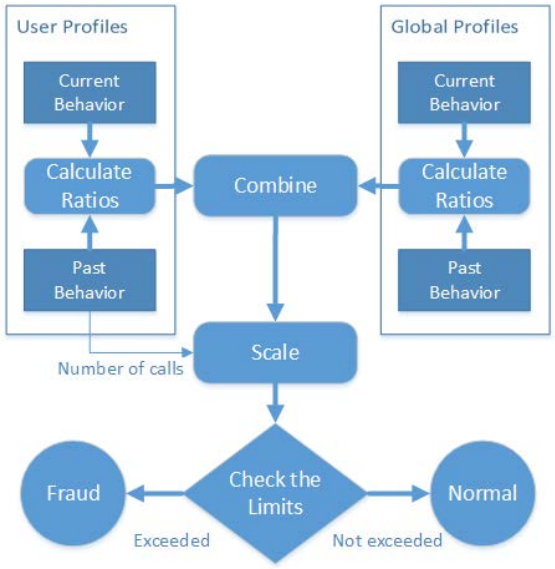

Figure 4. Overview of the components and their relationships.

F. Results of the previous approach

This section describes the test of a prototype implementation of our previous approach in an experiment. The implementation has been done in Java for an existing fraud detection framework of the research project. The data used for the experiment has been generated by a live environment, recorded by a VoIP switching device. The data consists of 76,326 cost impending calls and spans over a time of one month. It has been anonymized in accordance to the German Federal Law on Data Protection.

For the experiment, the whole data set was used, as the system trains on live data with the assumption that fraud cases are rare enough so that the profiles can initially be trained by themselves without greater risks of being manipulated by fraud cases. Assuming the contrary is true and the first data set is containing fraudulent CDRs, the impact would only be that no fraud cases are detected until the fraudulent CDRs are no longer used for the PBP.

For the experiment, profiles of a week's length and with an offset (d) of one day for the PBP are used. In a first run, all occurring ratios are recorded to calculate limits for the ratios, to analyze the parameters for the scaling function and to integrate the global ratios into user profiles. In a second run, the limits were applied and the fraud detection component was enabled.

First results

For the first results, without incorporating the global profiles and the scaling function, the false positive rates (FPR) for different limits were measured. The false positive rate is a very important measure that indirectly determines the expenses due to inefficiency, because administrators need to look at false positives.

Table II shows empirically tested limits for ratios and the number of exceedings. The FPR has been measured from 50,893 samples, where the profiles were *ready*. The limits and the resulting FPRs will be used for comparison with results of the

TABLE II. FIRST RESULTS OF FPR WITHOUT GLOBAL PROFILES AND SCALING FOR DIFFERENT LIMITS WITH PROFILE LENGTH OF ONE DAY AND AN OFFSET OF ONE DAY

Limit for all ratios	Limit for exceedings	FPR
0.25	>0	0.2142
0.25	>1	0.1274
0.5	>0	0.0685
0.5	>1	0.0444
0.75	>0	0.0211
0.75	>1	0.0145

incorporations of global profiles and the scaling function for low usage. These results are for CBPs length of a day and the PBP length of a week.

The following tables show additional information about other time spans for the profiles in comparison to already shown results.

As seen in Table III and in Table IV, the results show that a profile's length of one week with an offset of one day is more promising with much lower FPR than the other profile variants.

TABLE III. RESULTS OF FPR WITHOUT GLOBAL PROFILES AND SCALING FOR DIFFERENT LIMITS WITH PROFILE LENGTH OF ONE DAY AND AN OFFSET OF ONE DAY

Limit for all ratios	Limit for exceedings	FPR
0.25	>0	0.7274
0.25	>1	0.6355
0.5	>0	0.5066
0.5	>1	0.4283
0.75	>0	0.3024
0.75	>1	0.2371

TABLE IV. RESULTS OF FPR WITHOUT GLOBAL PROFILES AND SCALING FOR DIFFERENT LIMITS WITH PROFILE LENGTH OF ONE WEEK AND AN OFFSET OF ONE WEEK

Limit for all ratios	Limit for exceedings	FPR
0.25	>0	0.5964
0.25	>1	0.4467
0.5	>0	0.3096
0.5	>1	0.2239
0.75	>0	0.1309
0.75	>1	0.0386

Global profiles

For the global profiles, the same length and offset was used, because the ratios can be compared better if the parameters are similar. The number of calls was used as the only feature for the global profiles. For the parameter g for scaling the global ratio, see (4), a test value of 1 was used.

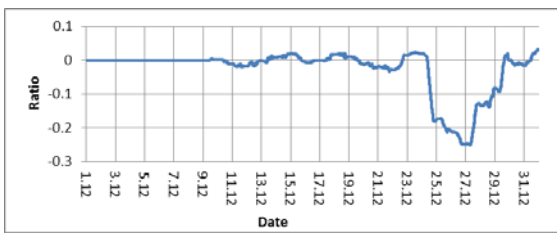

Figure 5. Ratios for number of calls for the whole data in global profiles.

Figure 5 shows the ratios measured for the given data, chronologically sorted. It shows negative ratios during the Christmas holidays in Germany, successfully measuring its effects on the ratios and it can be used to remove those effects from single user behavior. Also, this figure shows when the profiles became *ready*.

The incorporation into profiles of a week's length showed no significant improvements in the FPRs.

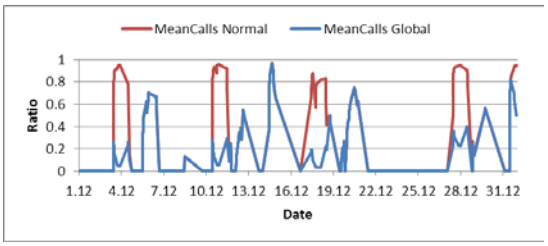

Figure 6. Example incorporation of global ratio into a day-length user profile for feature MeanCalls.

On the other hand, a small scale test of profiles with a day's length showed very good results in removing weekend fluctuations from the profiles. Figure 6 depicts an example for day-length profiles.

The figure shows two curves, MeanCalls Normal showing the ratios of the feature MeanCalls without correction by global profiles and MeanCalls Global with correction by global profiles.

Scaling for low usage

To find an appropriate scaling function, the dependency of the number of calls to the maximum occurring ratios was analyzed. Figure 7 shows an example for four features. It depicts how a low number of samples/calls in a profile can affect the ratios. Therefore, a scaling function was created that scaled the ratios from 0 to 70 calls.

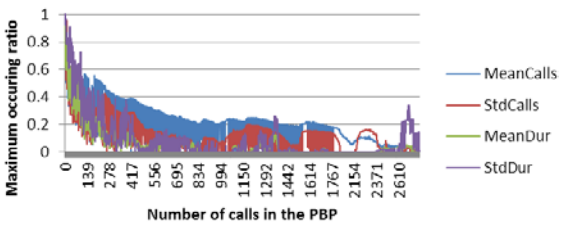

Figure 7. Example for the dependency of max values of the features MeanCalls, StdCalls, MeanDur and StdDur to the number of calls.

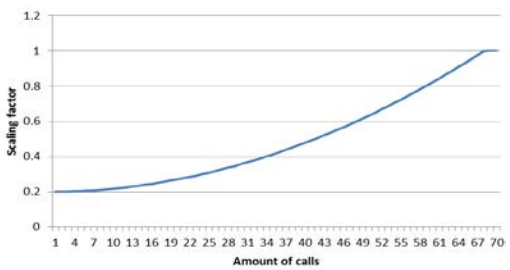

Figure 8. Depiction of the scaling function.

For the scaling function, a simple parable of the form $y = (ax)^2 + b$ was chosen after testing different curves, because it corresponds well to the curve in Figure 7. Using the coefficients $a = \frac{1}{67.1}$ and $b = 0.2$, the scaling begins at 0.2 with 0 calls

and ends at 1 with about 70 calls with a slight increase, as shown in Figure 8. Because about 47% of users only conduct about two calls per day, the scaling function greatly improved the FPRs, as shown in Table V.

TABLE V. CHANGES IN FPR WITH INCORPORATION OF THE SCALING FUNCTION

Limit for all ratios	Limit for exceedings	Old FPR	New FPR	Change in %
0.25	>0	0.2139	0.1684	-21,27%
0.25	>1	0.1272	0.0939	-26,17%
0.5	>0	0.0683	0.0491	-28,11%
0.5	>1	0.0443	0.0290	-34,53%
0.75	>0	0.0211	0.0136	-35,54%
0.75	>1	0.0145	0.0083	-42,75%

Determination of limits

The best way to determine the limits is to optimize the ratio of true positive rate to false positive rate. However, this requires labeled data to be possible. Because of the lack of labeled data, the limits were determined by measuring the 99.5% quantile of all occurring ratios for each feature. The ratios are presented in Table VI. Using these limits, the measured FPR is 1.87%.

TABLE VI. LIMITS FOR FEATURES (99.5% QUANTILE)

Feature	Limit
MaxCalls	0.8247
MaxDur	0.6692
MeanCalls	0.7512
StdCalls	0.8270
MeanDur	0.2985
StdDur	0.5400
MaxCost	0.7387
Mean	0.3835

Final detection rates

Out of the 50,893 analyzed cost impending calls, 1.87% are measured as false positives. Through empirical inspection of the false positives, two users were found with an exceptionally strange behavior pattern. The duration of calls and the number of calls per second was the same in about 200 calls, which is very suspicious. After consultation with the providing telecom company, those calls were considered fraud cases. This shows that the presented approach can detect false positives and reduce the FPR to 1.22%, but does not provide a true positive rate for a decent comparison with related work. Still, 90.23% of the fraudulent calls found in these two users were marked as fraud by the proposed approach. Compared to the approach proposed in [24], which also proposes a statistical, unsupervised method, the approach of this paper has a lower FPR (1.22% to 4.0%). Compared to other supervised techniques, like [13] (with 50% TPR and 0.3% FPR) or [22] (two approaches with 70% and 80% TPR and 0% FPR for both), the proposed approach has a good TPR and FPR and needs no effort for preparing supervised training data.

VI. NEW BASIC BEHAVIOR PROFILING CONCEPT

In this section, a basic concept for behavior profiling is described, consisting of profiles, calculation of features for two different contexts and detection of anomalies in the profiles. The idea for this concept was derived from the previous concept described in Section V. The need for a new concept arose from the problems with low activity users and fluctuations described for the previous approach and its complexity, but also from the potentials for using this approach in a different context. In the following, a description of the new basic behavior profiling concept is given.

A. Profiles

A profile is a collection of historic data in the context of an object. Examples for objects are users or destinations. The historic data, for this work, consists of call information. A profile P holds the information that occurred in a time span with the length P_L with an relative offset P_O to the present time p.

To determine if data with a given timestamp t is inside the time span of a given profile P, the following rules are applied:

$$inside = \begin{cases} true, if\ (p - P_O - P_L) < t < (p - P_O) \\ false, else \end{cases} \quad (6)$$

If data is inside a profile's time span, it is used to calculate features that describe the behavior of an object in that time span. Profiles that describe the behavior of the present will be called Current Behavior Profile (CBP) and profiles describing past behavior will be called Past Behavior Profile (PBP).

B. User profiling

As mentioned in the introduction, we introduce two contexts based on previous work. The first context, as it is common for the related work, is the user's context. The second context is the destinations context.

For the user context, the outgoing call behavior of the user is analyzed to find deviations from the normal or present user behavior to the past behavior. These deviations or anomalies are used to detect fraud.

The new approach has not as much problems to be considered as the old approach. The idea of global user profiling from the old approach for global fluctuations in user behavior, e.g., due to seasonal reasons, got adapted to the new approach and is described in the following subsections.

Features and profiles

The most used features for describing past user behavior used in related work are statistics over the amount of calls and the duration of each call a user makes, e.g., in [5] [7] [20]. Therefore, the standard deviation and the arithmetical average of the call amount and the duration of calls are used in this work.

In contrast to the previous approach, the maxima are not used in this concept, because an analysis

showed weak influence in detection by these features.

The statistics for the duration will be calculated on per call basis and the statistics for the call amount on a per hour basis.

Table VII shows the used feature vector for describing the past behavior. Mean stands for arithmetical average, Std for standard deviation, DpC for Duration per Call, and CpH for Calls per Hour.

TABLE VII. FEATURE VECTOR USED FOR DESCRIBING PAST USER BEHAVIOR

MeanCpH	StdCpH	MeanDpC	StdDpC

From the CDRs, both the timestamp of the call connect and the duration are needed for calculating the features.

For the present user behavior, only the call amount and the average duration for the latest hour are used as features, resulting in the following feature vector described in Table VIII.

TABLE VIII. FEATURE VECTOR USED FOR DESCRIBING PRESENT USER BEHAVIOR

MeanCpH	MeanDpC

The CBP has a length P_L of one hour and no offset. Therefore, MeanCpH stands for the number of calls in the profile. The PBP has an offset P_O of one hour, due to the length of the CBP, and will have a length of one week as discussed in Section V.A and shown with good results in Section V.F.

Detection

The detection of fraud is done by comparing the PBP with the CBP features. For this, we calculate a limit for the mean duration and number of calls of the CBP. The limit is calculated as described in (7):

$$Limit = Mean + Std * r + a \qquad (7)$$

a is an additional absolute part that removes the need to handle users with low amount of calls and empty profiles. It needs to be small enough for not affecting users with a high call count and still protect users with a low call count from unnecessary fraud alerts. The techniques for these cases described and used in the previous work are therefore no longer needed. The Mean and Std part scales with the amount of calls the user does and is therefore a scaling for users with a higher amount of calls. The additional scaling r is for adjustment of the relative part.

Adaption of global user profiling

In the previous approach, we used a global user profiling for analysis of fluctuations in the data like:

- Holidays
- Seasonal fluctuations
- Weekly fluctuations (weekends)
- Daily fluctuations (work/after work)
- Unexpected fluctuation (network problems)

We used global user profiling with the assumption that the whole user base as a single entity provides stable enough statistics that show the global fluctuations but not the single user's fluctuations. This also requires that a single user does not make more calls than all other users together.

In this concept, because of the profiles' length of one week, we only want to look at holidays, seasonal fluctuations and unexpected fluctuations, e.g., caused by network failures. The other fluctuations get evened out by the statistics over one week.

In contrast to the previous approach, we will also use the user profiling provided in this concept for global user profiling. This will make it easier for changes in the global user profiling to be factored in the individual user's profiling.

For the PBP and CBP, the same lengths and features are used as for user profiling.

We measure ratios R for each feature F between the CBPs and PBPs. The ratio is calculated by dividing a feature of the CBP with the respective sum of the mean and standard deviation feature.

The ratios must then be applied to the detection of a single user's profiling. The relative part of the limit formula (8) is adjusted by the respective ratio:

$$Limit = (Mean + Std * r) * R + a \qquad (8)$$

C. First Results

A first analysis in "normal" data showed for single users that the limit is appropriately high enough. This was enabled by the incorporation of the global ratio. Figure 9 shows the difference between normal ratio of a single user and his ratio adjusted with the global ratio for the MeanCpH limit. The figure shows the information of increased activity during working hours that is normally lost in the user's profile due to the smoothing of the statistics over one week. This confirms that we can successfully reinsert an approximation of this information via the global ratio.

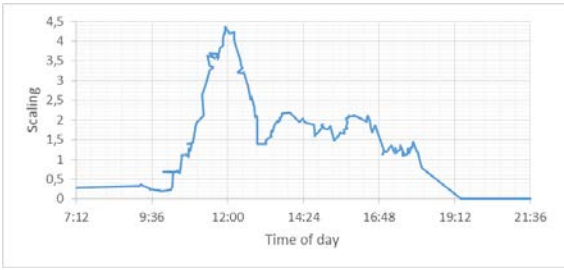

Figure 9. Depiction of the difference in limits with incorporation of global ratio for the MeanCpH limit.

This shows that the new approach is much simpler than the old one, but still handles the problems of the previous approach and promises good results for detection rates (false positive rate and true positive rate).

An extensive analysis over prepared and labeled data for acquiring detection rates of this new user profiling approach is still required.

VII. DESTINATION PROFILING

After the FRITZ!Box attacks happened, as described in Section IV, the previous approach was applied to a data set containing the attacks. The previous approach could not detect the FRITZ!Box attacks because the context of the profiling was that of a single user and not of a destination. Attacks on a single destination distributed over many users could therefore not be seen by user profiling, because of the relatively small effects on the single user's profile. This led to the idea of using a different context for profiling and detection of distributed attacks on single destinations, as we call it Destination Profiling.

In another previous work [2], Destination Profiling was then implemented and good results were generated for detecting this kind of attacks. The results are shown in Section VII.B after the description of the approach for detection is given in the following paragraphs.

Because of the good results in the previous work with profiling the number of distinct users calling a single destination in a defined time span [2], the approach is improved by adding distinct callers as a feature for the Destination Profiling. After the results of the previous work are shown, the new results with the improvement are shown and compared to the previous results.

A. Features, profiles and detection

For Destination Profiling, the same methods as for the User Profiling are not only reused, but adjusted. The duration of a call is not representative for a destination, because different callers have different behaviors. Only the number of calls, that is not per call basis, and the number of distinct callers can give important information about fraudulent usage of the destination.

Only the number of distinct callers can be used, because the number of non-distinct users will lead to the same result as with user profiling. If only one caller is doing a high amount of calls, this can also be detected with user profiling, as there is only one user context doing the calls.

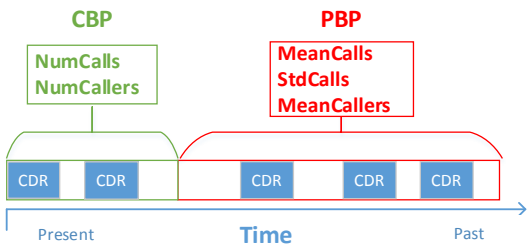

Figure 10. Destination Profiling: Depiction of the current behavior profile and the past behavior profile in relation to time.

For the PBP, the mean and std for calls per hour and callers per hour are used. For the CBP, the number of calls and the number of distinct callers are used. The profile parameters and the detection are the same as for User Profiling. Figure 10 depicts the profiles.

B. Previous prototype

In this section, results from a prototype implementation of the previous Destination Profiling approach are described and analyzed empirically.

Used Data Set

To evaluate the prototype implementation, real life traffic data (CDRs) provided by a local telecom company has been used. The data comprises calls from a time span of two weeks containing about 3.5 million calls. Only the portion of the data with outgoing calls was used, because incoming calls are not relevant to the analysis. The outgoing calls amount to about 470,000. Table IX shows the distribution of the calls for the regions national, mobile and international and are split into connected and unconnected calls.

TABLE IX. DESTINATION PROFILING: NUMBER OF CDRS FOR EACH REGION

REGION	AMOUNT
CONNECTED	325,947
NATIONAL	274,205
MOBILE	42,669
INTERNATIONAL	9,073
UNCONNECTED	153,330
NATIONAL	112,476
MOBILE	24,570
INTERNATIONAL	16,284
TOTAL	479,277

In the first week, no attack attempts (fraud) were contained. This part of the data was applied to initialize the behavior profiles, building the features. In the second week, normal call traffic is contained, as well as about 20,140 fraudulent calls following the typical FRITZ!Box attack pattern. The second week has been used to test the detection abilities.

Experimental Setup

First of all, the relevant thresholds had to be determined, because this is a necessity for high-quality detection results. To accomplish this, a single run of the method, without the fraud detection, is conducted with the first week of the data and every feature value at the time of each call is recorded. The thresholds are estimated by analyzing the resulting values of the CBP for fraud and non-fraud cases and for each region (national, mobile, international). The 99%-quantiles of the number of calls from the CBP, for connected and unconnected calls, as well as national, international and mobile calls each, have been recorded and used as the absolute threshold A_R for each region. The parameter G_R, representing the relative threshold, has been set to $G_R = 1$, for testing purposes.

Finally, a test run with the activated fraud detection and the previously measured thresholds is done and the detection quality is evaluated by

comparing the detected cases to the known cases of fraudulent behavior.

The approach can be described with the following steps:

1. The detection method is deactivated at first
2. The profiles are initialized using the data set of the first week
3. Thresholds are calculated from CBP values as described before
4. The detection method is now activated
5. The data set of the second week is now used as input
6. The results from the detection method are compared to the known cases of fraudulent behavior

Detection results

Thresholds have been determined for successfully connected, as well as unconnected call attempts, each for national, international and mobile calls. Also, the profile values have been calculated and recorded.

The arithmetic mean and the standard deviation both represent valid values to generate relative thresholds. An adjustment with the parameter G_R is only necessary in individual cases.

Under these testing conditions, the detection method achieved a false positive rate of 0.7% or 3,355 false positives (see Table X). Of the known attacks in the data, the detection method was able to identify all attacks, resulting in 100% detection rate or true positive rate. However, there is the possibility that not all attacks are detected because some may still be unknown to the provider of the data. An estimation of a true positive rate of about 95% would be more appropriate.

TABLE X. DESTINATION PROFILING: DETECTION RESULTS

	AMOUNT	RATE
FALSE POSITIVE	3,355	0.7%
TRUE POSITIVE	20,140	100%

Compared to the results achieved in comparable related work (see Table XI), which utilizes unsupervised user profiling, with a FPR of 4% and a TPR of 75% [6] and our previous user profiling approach with a FPR of 1.22% and a TPR of approximately 90% (see Section V), these measurements are as good or even better.

TABLE XI. DESTINATION PROFILING: COMPARISON OF FPR AND TPR

	TPR	FPR
THIS WORK	95%	0.7%
PREVIOUS APPROACH	90%	1.22%
RELATED WORK [6]	75%	4%

On the other hand, no direct comparison is possible, because the detection method itself is partially different, applying a modified approach of user profiling.

Improved prototype

The addition of the number of distinct users as a feature to the profiling allowed the reduction of the limit for number of calls while maintaining the true positive rate. This also affected the false positive rate and reduced it by 0.2%, resulting to a FPR of 0.5%.

VIII. CONCEPT OF COMMUNICATION BEHAVIOR PATTERNS

The concept of communication behavior patterns in one of the intermediate works [3] was developed based on the analysis of the FRITZ!Box incident and experiences with user profiling from the previous work. It differs in the *point of view* from the concept of Destination Profiling [2], utilizes pieces of information from user profiles and adapts the principle of clustering algorithms (unsupervised learning) [25], which is used to find similarities between objects.

Behavior patterns are created in order to reflect a behavior of a user in a specific context. To associate with a behavior pattern, each shall have its own criteria, where similar patterns possess similar criteria. In order to describe the behavior concerning a distinct aspect of a user or a group of users, the calls of a user profile are matched against predefined behavior patterns. A user is able to have matches to several behavior patterns.

To search for behavior patterns using user profiles as objects and to obtain an indication for thresholds, clustering algorithms implemented in WEKA [26] were used. Used algorithms were k-means, EM and an implementation of a SOM (self-organizing map) as a clustering algorithm.

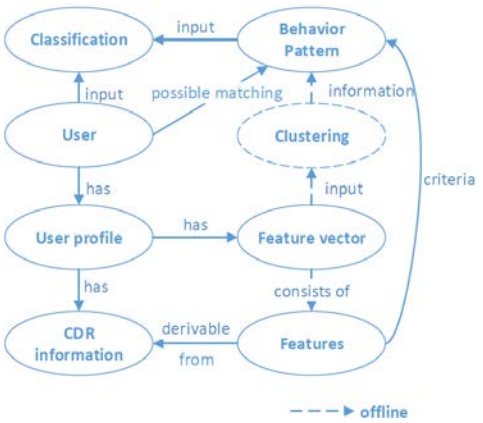

Figure 11. Overview of the relations between the components.

Figure 11 shows an overview of the relations between the components of the concept of behavior patterns. For every user, a user profile gets constructed. Each user profile consists of prepared data retrieved from CDRs, as well as a feature vector, which are defining criteria for behavior patterns. Clustering algorithms providing pieces of information for behavior patterns are used offline, as the gathered information has to be evaluated, as well as they potentially slow down the whole classification process. At last, using a user's behavior patterns and the belonging matching

information, a call can be classified as fraudulent or non-fraudulent.

A. Data preparation for user profiling

The concept of behavior patterns utilizes information retrieved from user profiles. The attributes A_1 to A_4 are extracted from a CDR for a user profile:

A_1 User ID
A_2 Timestamp of the call
A_3 Duration of the call
A_4 Destination number

The first attribute is used for a unique identification of a user. A_2 is used to obtain information whether the call occurred on a weekend and if the call has been made during work hours (7:00 am to 18:59 pm) or after hours (19:00 pm to 6:59 am) with a time span of 12 hours each.

A_3 is used to know whether the call was a call attempt or call connect. The information obtained from A_4 is further categorized into its call region *national*, *mobile* and *international*.

B. Behavior patterns

A behavior pattern reflects a behavior in a specific context of a user. An example is the behavior pattern "International Calls" to which a user gets assigned if he conducts calls to international destinations often. It is possible for a user to *match* one or more behavior patterns.

Features

Every behavior pattern has its own defining set of features *F* called *feature vector*, with comparable behavior patterns having similar defining features. These features are highly dependable on the criteria of a behavior pattern. Therefore, providing an overall definition for a feature vector is not possible. The features are derived from the data contained in a user profile and are grouped in two types, numeric and Boolean (true/false).

Examples for two behavior patterns, their criteria and therefore feature vectors are:

- "*International Calls After Hours"*: The criteria for this pattern are: The call has to be connected, the call region is *international* and the call is made *after hours.*
- "*Weekend Calls"*: The only criterion is for the calls to be made on a weekend.

The numerical value is the accumulation of the respective calls during a time span t_{BP} (in this case, one hour), which applies for both behavior patterns.

Criteria for a behavior pattern match

For a *match* to a behavior pattern, every feature of a feature vector, depending on its type, has to meet its criteria:

- Numeric: A statistical or numeric feature has to pass a *threshold*.
- Boolean: A Boolean feature has to have the value *true*.

For every defined behavior pattern, the criteria are tested. This way, it is possible for a CDR of a user to lead to a match to more than one behavior patterns.

Metric for a match

All calls matching a distinct behavior pattern are stored in respective lists. Over time, the length of such a list (*grade* of a match) can diminish or grow, being further denoted as a *growth* of a match to a behavior pattern.

The *growth G* of a match to a behavior pattern over a time span is measured as:

$$G = \frac{C_L}{\bar{x}\,(C_P)} \tag{9}$$

C_L denotes a list of all connected calls during the current (latest) hour and C_P a list of all connected calls in the past. For both C_L and C_P, calls from the list of matches are used. $\bar{x}$ denotes the arithmetic mean over the respective list.

Change of a match

The *growth G* of a match to a behavior pattern described above is further used as a criterion to mark a current call as *fraudulent*, as it is defined in the following case differentiation:

$$Fraud = \begin{cases} true, G > T_{BP} \\ false, otherwise \end{cases} \tag{10}$$

T_{BP} denotes a threshold for the growth of a match to a behavior pattern. If T_{BP} is passed, the current call, which had been causal for *passing* the threshold, is the first call to be considered fraudulent. All subsequent calls, which are still triggering *true*, are considered fraudulent as well. To regulate how much a growth of a match influences the assignment of a call as fraudulent, each behavior pattern has been given a *weight*, leading to an enhancement of the case differentiation shown in (10) to (11):

$$Fraud = \begin{cases} true, G \cdot w > T_{BP} \\ false, otherwise \end{cases} \tag{11}$$

C. Prototype

Used data

In case of the prototypical implementation of the concept of behavior patterns, real life traffic data over a time span of seven weeks provided by a local telecom company has been used. The first week has been used for the initialization phase, where user profiles, as well as behavior patterns of a user, are constructed, as this week did not contain known fraudulent activity. Out of the seven weeks, there is at least one week included with definite fraud attacks having the pattern described in Section IV. The rest of the data shows partial signs of the FRITZ!Box fraud attack pattern as well. The data set comprises 10,401,547 CDRs. As only outgoing calls, as well as successfully connected calls (call connects) are of importance, 2,749,860 CDRs were left.

Experimental setup

Two simple behavior patterns have been defined:

- *IntCallsPattern*: All connected calls having an international destination match the behavior pattern.
- *IntCallsAfterHoursPattern*: All connected calls having an international destination and having

been conducted in the after hours match this behavior pattern.

The thresholds for the statistical features for both behavior patterns, as well as indications about the thresholds concerning the change of a match to a behavior pattern have been derived using clustering algorithms from WEKA. The applied clustering algorithms were k-means, EM and an implementation of a SOM as a clustering algorithm. For the prototypical implementation, the behavior patterns possess parameters, which can be defined via a XML configuration. Table XII shows the definition of both behavior patterns, including the parameters. The parameter *type of pattern* defines whether the behavior pattern checks for the number of calls (*calls*) or for the sum of duration (*duration*). The possible values for each parameter after "type of pattern" are:

- call type (all / call attempts / call connects),
- destination (all/national/international/mobile),
- timeslot (all / workhour / after hour) and
- weekday (all / workday / weekend)

If no further distinction is made, a parameter gets initialized with *all*.

TABLE XII. OVERVIEW OF THE DEFINED BEHAVIOR PATTERNS AND THEIR PARAMETER VALUES

	BP_1	**BP_2**
name	IntCalls	IntCallsAfterHours
weight	0.9	0.7
threshold (matching)	25.2	8.4
threshold (growth)	0.5	0.4
type of pattern	calls	calls
call type	call connect	call connect
destination	international	international
timeslot	all	after hours
weekday	all	all

Results

An approximation concerning the TPR was possible due to the analysis performed on the data retrieved during the FRITZ!Box incident. It is highly possible that not all fraudulent data has been known during the evaluation of the prototype. The following steps have been applied on the data set:

1. Apply the thresholds and weight values retrieved from clustering algorithms and given from experience, respectively.
2. Run the prototype with the defined two behavior patterns.
3. Analyze the results utilizing the knowledge derived from the analysis of the data, as well as from the local telecom company.

In total, 17,110 fraud cases were reported and analyzed. During the analysis, one customer was noticeable in his behavior to conduct calls to foreign destinations very often, even not during the timeframe of the FRITZ!Box incident. Due to these findings, as well as other aspects found in our analysis, the aforementioned customer can be considered being a call center. As such customers are likely to be added to a whitelist, all calls belonging to such a customer can be ignored, which leads to a total of 13,503 reported fraud cases. A TPR of *98.4%* and a FPR of below 0.01 % have been measured. At this point, it has to be said that surely not all fraud instances of the FRITZ!Box incident could be found. This can be said even though not enough labeled data existed, as valuable time – and therefore, CDRs – passes in order for a user to match a behavior pattern and be associated with the described behavior. Additionally, a threshold concerning the growth of a match has to be passed, resulting in an equivalent to a "settling-in phase". Thus, it is possible that not all fraudulent instances were detected.

IX. CONCLUSION

This paper presents an already simple statistical way of detecting fraud in telephony by analyzing user behavior and finding anomalies. It has flaws concerning the complexity, because of problems with special cases of users and fluctuations. These problems and the need for a technique for detecting the FRITZ!Box attacks led to three new techniques.

With the new user profiling concept, the complexity of the previous approach is reduced and the problems of it are handled, by combining the global profiling with a different statistical approach. The effect of the global profiling is shown for a single user. Still, more features are needed to describe user behavior as it is shown in related work.

The FRITZ!Box attacks led to an adaption of the new user profiling approach to a new context for enabling the detection of these attacks, called Destination Profiling. Destination Profiling allows to detect attacks on a single destination from multiple sources. The developed approach shows promising detection rates and was improved with small changes in this paper.

The concept of behavior patterns utilizes the grouping aspect of clustering algorithms, leading to behavior pattern recognition using information retrieved from user profiles. Pieces of information from a user profile are matched against predefined behavior patterns, which depict the behavior of a user in a specific context. A match to a behavior pattern can grow and if a significant growth in a short time frame has been observed, a call is considered fraudulent.

Overall, the three new approaches that were derived from our previous work show promising results for a combined detection in an online analysis tool.

X. FUTURE WORK

The most important future work is an extensive analysis of all new approaches presented in this work. The analysis needs to be done on a large enough prepared and labeled data set. A very important task of this analysis is to find correlations between the approaches to create an optimized combined detection result.

In particular, the new user profiling approach needs more features for the description of user behavior. Features used in related work need to be further analyzed and adapted to this approach. The incorporation of the global profiling needs to be fine-tuned as well.

As there is little related work for destination profiling, a more detailed analysis of possible new features for the presented approach needs to be done.

The behavior pattern recognition only covers a little amount of possible user groups. The authors see a huge potential in finding new user groups and analyzing them for finding new approaches for fraud detection. Including information given by call attempts and call termination cause codes can further improve the detection result. They can provide insight whether a fraudulent attack is currently prepared or conducted. Additionally, “normal” behavior patterns - e.g., “National Calls” - have to be considered as well, as they can provide further indications on a sudden change in a user’s behavior.

ACKNOWLEDGMENT

We would like to thank the state of Hesse, Germany for supporting this work by providing the necessary funds by the development program LOEWE. Also we would like to thank a German telecom company that provided the necessary data, essential support and knowledge for research and development of practical fraud detection methods for this work.

REFERENCES

[1] A. Wiens, T. Wiens, and M. Massoth, "A new unsupervised user profiling approach for detecting toll fraud in VoIP networks," in The Tenth Advanced International Conference on Telecommunications (AICT 2014) IARIA, 2014, pp. 63-69.

[2] A. Wiens, T. Wiens, and M. Massoth, "Approach on fraud detection in Voice over IP networks using call destination profiling based on an analysis of recent attacks on FRITZ!Box units," in The Sixth International Conference on Emerging Network Intelligence (EMERGING 2014) IARIA, 2014, pp. 29-34.

[3] S. Kübler, M. Massoth, A. Wiens, and T. Wiens, "Toll fraud detection in Voice over IP networks using communication behavior patterns on unlabeled data," in The Fourteenth International Conference on Networks (ICN 2015) IARIA, 2015, in press.

[4] Communications Fraud Control Association, "Global Fraud Loss Survey," October 2013. [Online]. Available from: http://www.cfca.org/pdf/survey/CFCA2013GlobalFraudLossSurvey-pressrelease.pdf 2014.06.23.

[5] M. Taniguchi, M. Haft, J. Hollmen, and V. Tresp, "Fraud detection in communication networks using neural and probabilistic methods," in Proceedings of the 1998 IEEE International Conference on: Acoustics, Speech and Signal Processing, vol. 2, 1998, pp. 1241-1244.

[6] P. Burge and J. Shawe-Taylor, "Detecting cellular fraud using adaptive prototypes," in Proceedings AAAI-97 Workshop on AI Approaches to Fraud Detection and Risk Management, AAAI Press, 1997, pp. 9-13.

[7] C. S. Hilas and P. A. Mastorocostas, "An application of supervised and unsupervised learning approaches to telecommunications fraud detection," Knowledge-Based Systems, vol. 21, no. 7, pp. 721-726, 2008.

[8] H. Grosser, P. Britos, and R. García-Martínez, "Detecting fraud in mobile telephony using neural networks," in Proceedings of the 18th international conference on Innovations in Applied Artificial Intelligence, Bari, Italy, Springer-Verlag, 2005, pp. 613-615.

[9] heise online, "Report: Deutsche Telekom analyzes call data of several calls," [Online]. Available from: http://www.heise.de/newsticker/meldung/Bericht-Deutsche-Telekom-wertet-Verbindungsdaten-saemtlicher-Telefonate-aus-1933436.html 2014.06.23.

[10] AVM GmbH, "Security notice: suspected phone fraud," 06 02 2014. [Online]. Available from: https://www.avm.de/de/News/artikel/2014/sicherheitshinweis_telefonmissbrauch.html 2014.06.23.

[11] R. Eikenberg, "Hack on AVM routers: Fritzbox breach disclosed, millions of routers at risk," 07 03 2014. [Online]. Available from: http://www.heise.de/security/meldung/Hack-gegen-AVM-Router-Fritzbox-Luecke-offengelegt-Millionen-Router-in-Gefahr-2136784.html 2014.04.04.

[12] TELES AG, official homepage. [Online]. Available from: http://www.teles.com/en/teles.html 2014.06.23.

[13] Y. Moreau, H. Verrelst, and J. Vandewalle, "Detection of mobile phone fraud using supervised neural networks: a first prototype," in Proceedings of the 7th International Conference on Artificial Neural Networks, Springer-Verlag, 1997, pp. 1065-1070.

[14] T. Kapourniotis, T. Dagiuklas, G. Polyzos, and P. Alefragkis, "Scam and fraud detection in VoIP networks: analysis and countermeasures using user profiling," in 50th FITCE Congress, 2011, pp. 1-5.

[15] T. Fawcett and F. Provost, "Adaptive fraud detection," Data Mining and Knowledge Discovery, vol. 1, no. 3, pp. 291-316, 1997.

[16] D. Olszewski, J. Kacprzyk, and S. Zadrożny, "Employing Self-Organizing Map for fraud detection," in The 12th International Conference on Artificial Intelligence and Soft Computing (ICAISC 2013), 2013.

[17] J. Hollmén, V. Tresp, and O. Simula, "A Self-Organizing Map for clustering probabilistic models," in Ninth International Conference on Artificial Neural Networks (ICANN), vol. 2, 1999.

[18] R. Alves et al., "Discovering telecom fraud situations through mining anomalous behavior patterns," in Proceedings of the DMBA Workshop on the 12th ACM SIGKDD, 2006.

[19] D. Hoffstadt et al., "A comprehensive framework for detecting and preventing VoIP fraud and misuse," in International Conference on Computing, Networking and Communications (ICNC), 2014, pp. 807-813.

[20] P. Burge, J. Shawe-Taylor, C. Cooke, Y. Moreau, B. Preneel, and C. Stoermann, "Fraud detection and management in mobile telecommunications networks," in European Conference on Security and Detection, 1997, pp. 91-96.

[21] H. Grosser, P. Britos, and R. García-Martínez, "Detecting fraud in mobile telephony using neural networks," in Proceedings of the 18th international conference on Innovations in Applied Artificial Intelligence, Bari, Italy, Springer-Verlag, 2005, pp. 613-615.

[22] M. Taniguchi, M. Haft, J. Hollmen, and V. Tresp, "Fraud detection in communication networks using neural and probabilistic methods," in Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 2, 1998, pp. 1241-1244.

[23] "Der Bundesbeauftragte für den Datenschutz und die Informationsfreiheit: Bundesdatenschutzgesetz (BDSG)," [Online]. Available from: http://www.bfdi.bund.de/SharedDocs/Publikationen/GesetzeVerordnungen/BDSG.pdf?__blob=publicationFile 2014.06.23.

[24] P. Burge and J. Shawe-Taylor, "Detecting cellular fraud using adaptive prototypes," in Proceedings AAAI-97 Workshop on AI Approaches to Fraud Detection and Risk Management, AAAI Press, 1997, pp. 9-13.

[25] N. Grira, M. Crucianu, and N. Boujemaa, "Unsupervised and semi-supervised clustering: a brief survey," in A Review of Machine Learning Techniques for Processing Multimedia Content, Report of the MUSCLE European Network of Excellence (6th Framework Programm), 2005.

[26] WEKA, Machine Learning Group at the University of Waikato, official homepage. [Online] Available from: http://www.cs.waikato.ac.nz/ml/weka/ 2014.12.15.

[27] D. Wang, Q.-y. Wang, S.-y. Zhan, F.-x. Li, and D.-z. Wang, "A feature extraction method for fraud detection in mobile communication networks," in Fifth World Congress on Intelligent Control and Automation, vol. 2, 2004, pp. 1853-1856.

[28] P. Burge, J. Shawe-Taylor, C. Cooke, Y. Moreau, B. Preneel, and C. Stoermann, "Fraud detection and management in mobile telecommunications networks," in European Conference on Security and Detection (ECOS 97), 1997, pp. 91-96.