A Smart Blacklist Filter for the Android System

Tai-Fu Yu and Jui-Chung Hung

Department of Computer Science,Taipei Municipal University of Education
Taipei 100, Taiwan
juichung@gmail.com

Abstract

In this paper, we present an efficient blacklist filter to address fraudulent calls on an Android system. In general, fraudulent calls are almost never included in blacklist databases of smartphones. Therefore, we combined blacklist databases and adaptive neuro-fuzzy inference systems (ANFIS) to reduce the chances of being cheated by a fraudulent caller.

The smart blacklist filter consists of two steps. First, the system checks for whether a call matches the blacklist databases. If it does, the call will be filtered. If not, the call moves on to the second step. Second, an ANFIS method is used to assist in assessing the degree of risk for mismatched calls. From the experimental results, smartphones will have increased security when considering the calls used by ANFIS.

Keywords: Adaptive neuro-fuzzy inference (ANFIS), Android system, Blacklist filter, Fuzzy theory

1. Introduction

In recent years, smartphones are becoming more and more popular. People's dependence on smartphones is also increasing steadily. In addition, Android's market share is very high, so this study used Android as the study platform.

Fraudulent calls are currently one of the main ways of new fraud practices. It is important, when an unknown number rings, to give the user enough time to determine the degree of risk. Currently, filter software in Google Play, such as Call Blocker, BLACKLIST and Calls Blacklist, have blacklist and whitelist functions, both of which require the user to input the phone numbers one-by-one. Some whitelists can be entered via phonebooks. The blacklist function only blocks the number and the whitelist function only allows the phone to answer any number on the list. Although both of these have a filtering function, for any number that's not on the list, they can neither make a judgment nor perform real-time searches. It is a negative approach and does not update known blacklists. We present an application that can help the user determine the degree of risk of the unknown calls, so they will not be deceived by fraud. This study uses a two-stage filter. In the first stage, the filter determines whether the call is in the database of the blacklist; if it is on the blacklist, then the filtering will be completed. If it is an unknown call, i.e., not in the blacklist, it will enter the second stage. In the second stage, artificial intelligence uses the known information to determine the risk that the unknown phone may also be fraudulent, and then informs the user. This structure of artificial intelligence is complicated. Therefore, the ANFIS[1][2] architecture is applied for improved system performance.

The main structure of ANFIS is based on a fuzzy inference system combined

with the characteristics of a neural network. To solve the fuzzy inference system, we need to rely on expert knowledge to constantly adjust the membership function and fuzzy rules to achieve optimization.

Fuzzy theory has been widely used in various fields. The fuzzy inference system is constructed of fuzzy theory and uses fuzzy IF-THEN rules for human knowledge and reasoning processes to perform qualitative description and analysis that simulates human logical thinking ability. Therefore, ANFIS applies learning and computing functions of neural networks to fuzzy theory. It confers the fuzzy inference system with the ability of uncertainty and imprecision processing and adjusts the parameters by organizing training from itself to achieve the best performance.

2. Proposed method

The filter phone software (blacklist application) currently uses most of its database for filtering; however, in general, fraudulent calls are not in the known database. Then, we aimed to implement an intelligent blacklist system on the phone in this study and compensate for the lack of existing filter phone software using a two-stage filter. This blacklist system is divided into two steps; the first step involves determining whether the number is in the existing database; if not, then the second step is initiated. The system builds an intelligent blacklist with the ANFIS structure using expert knowledge [1] to determine the degree of fraudulence in a call. The expert knowledge of the system uses the number of discussions from web pages and the regional database distribution to establish the criteria for the smart blacklist system. The system is implemented on an Android, using Google technology and the artificial intelligence structure to establish the intelligent blacklist system. The details are outlined in the following sections.

2.1. Android

Android is a semi-open operating system based on Linux, using the authorization of Apache License 2.0,that is, to provide free to use. Due to the research to be done blacklist by phone automatically hang up the phone that involves the Android system function. Therefore, we need to use the Android Interface Definition Language (AIDL) to achieve it. AIDL is an Interface Definition Language, using the characteristics of the neutral interface to achieve mutual bridging with different programming languages. This system is required to bridge with Android system-level service. Its methods are as follows:

```
TelephonyManager tm = (TelephonyManager)context
                .getSystemService(Service.TELEPHONY _SERVICE)
      Method m = Class.forName(tm.getClass()
                .getName()).getDeclaredMethod("getITelephony").   (1)
```

We get the Message relating communication by m, but getITelephony that we used is been protected. Therefore, we use the method “Method.setAccessible” to unlock the protection of the system. It can only allow the program to hang up incoming calls directly after unlock the protection. Its syntax is as follows:

```
m.setAccessible(true)
com.android.internal.telephony
.ITelephony telephonyService = (ITelephony)m.invoke(tm)
telephonyService.silenceRinger()
telephonyService.endCall().                                       (2)
```

2.2. ANFIS

ANFIS[1] is a system that can self-adjust parameters. It combines fuzzy inference systems and the neural network of artificial intelligence; therefore, it is suitable for real-time systems. We use it to determine the degree of risk in the second stage of this intelligent blacklist system. The ANFIS neural network is generally

divided into five layers. The input values correspond to the respective membership functions in the first layer and uses a Gaussian function in this system. The formula is as follows:

$$O_{1,qw} = \mu_w\left(x_q\right) = e^{\frac{-\left(x_q - m_{qw}\right)^2}{2v_{qw}^2}},\ q = 1,2,\ldots,N,\ w = 1,2,\ldots,K \quad (3)$$

where $\mu_w\left(x_q\right)$ is the q-th input value corresponding to the w-th membership function, m_{qw} is the average of the q-th input value corresponding to the w-th membership function, v_{qw} is its standard deviation, N is the total number of inputs and K is the number of corresponding membership functions. According to empirical rule, ANFIS input values using internet custom search items and the cumulative number of telephones that correspond to regional numbers in the blacklist database used in this thesis. Then,

$$x_1 = Custom\ Search's\ count$$
$$x_2 = DataBase\ Left(\mathrm{X})\ count. \quad (4)$$

In the second layer, the value from the first layer will be input to the corresponding fuzzy sets. We use the standard intersection operator [10] expressed as follows:

$$O_{2,ij} = w_{ij} = \min\left(m_i(x_1), m_j(x_2)\right),\quad i = 1,2,\ldots,K,\ j = 1,2,\ldots,K \quad (5)$$

where i is the first input value corresponding to the i-th membership function and j is the second input value corresponding to the j-th membership function.

The third layer is the normalization. The value is compressed to a normalized value between 0 and 1. It is expressed as follows:

$$O_{3,ij} = \bar{\omega}_{ij} = \frac{\omega_{ij}}{\sum_{m=1}^{K}\sum_{l=1}^{K}\omega_{ml}}.\quad i = 1,2,\ldots,K, j = 1,2,\ldots,K. \quad (6)$$

The fourth layer will express the output as a linear function of the input and is expressed as follows:

$$O_{4,ij} = \bar{\omega}_{ij}\ f_{ij},\ i = 1,2,\ldots,K,\ j = 1,2,\ldots,K,$$
$$f_{ij} = \alpha_{ij}x_1 + \beta_{ij}x_2 + \gamma_{ij}. \quad (7)$$

The fifth layer adds up all the values of the fourth layer and is the output value, y, i.e., the degree of danger. It is expressed as follows:

$$O_5 = y = \sum_{i=1}^{K}\sum_{j=1}^{K}\bar{\omega}_{ij}\ f_{ij}. \quad (8)$$

3. Simulation

The system uses Android as a test platform. The system setting screen is shown in Figure 3, and the incoming call screen is shown in Figure 4. This experiment uses 300 calls: 200 calls for training and 100 calls for testing. The blacklist proportion is 20%, 40%, 60% and 80%. The training convergence is shown in Figure 5. As shown in Figure 5, convergence of the training algorithm is very quick when there are few blacklist events. For example, when there are 20% blacklist events, convergence is completed in approximately 100 generations. The error rate is shown in Table 1.

The results in Table 1 show that a very high accuracy rate was achieved with the training 4 database using the smart blacklist system. It performs much better than the general database for filtering fraudulent calls. When a user is attacked by a large number of fraudulent calls, the system performs better than the general method, which returns approximately 70% correct. In the training 1 database system, our method also performed better than the general method, which returned approximately 18% correct. Therefore, the system is very suitable for filtering fraudulent calls.

The unknown whitelist may misjudge when the database has more blacklist items. However, even with that potential

error, it provides users with better security than the general method.

4. Conclusions

Recently, crimes committed using fraudulent calls have been increasing. The majority of blacklist applications on the market uses a database to filter calls and cannot determine anything about the unknown calls. In this study, we use a two-stage filter. In the first stage, it uses phone books and databases in the system to make a judgment. If it is an unknown call, the filter will enter the second stage. In the second stage, the filter uses ANFIS to access the Internet in search of expert experience using region numbers as conditions and determine the risk factor. When attacked by fraudulent calls in training 4 conditions, there is a very high accuracy rate (90%). The database of this system can be even more enhanced because of the users' shared knowledge and increased amounts of information in the future.

Table 1: The error rate using the blacklist and proposed method

Test data		Training 1	training 2	training 3	training 4
Blacklist events		20%	40%	60%	80%
Proposed method	Training error	7.03%	17.01%	1.19%	18.56%
	Test error	1.50%	3.25%	5.50%	9.25%
Database-based filter	Test error	20%	40%	60%	80%

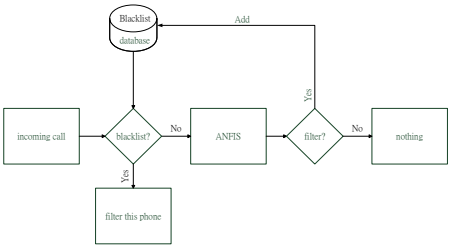

Fig. 1: Smart blacklist system program flowchart

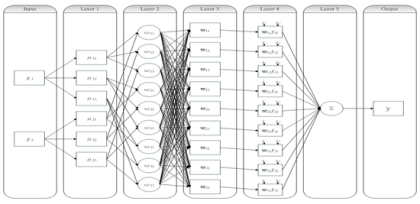

Fig. 2: ANFIS structure diagram

Fig.3: The home screen of the smart blacklist filter

Fig. 4: The view of an unknown caller

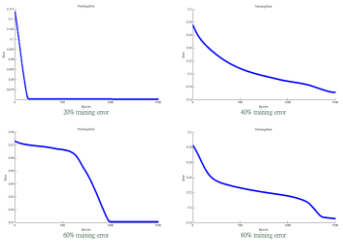

Fig. 5: The training error of the proposed system

5. References

[1] J. S. Jang, "ANFIS: Adaptive-Network-based Fuzzy Inference Systems", *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 23, Pages 665-685, May 1993.

[2] J. C. Hung, "Fuzzy switch ANFIS GARCH model applied to forecasting volatility of stock market", Information Sciences, vol. 179, Pages 3930-3943, 2009