# Detecting False Emergency Requests Using Callers' Reporting Behaviors and Locations

Mahdi Daghmehchi Firoozjaei
*Department of Electrical and Computer Engineering*
*Sungkyunkwan University*
*Suwon, Korea*
*Email: mdaghmechi@skku.edu*

Jaewoo Park
*Department of Software Platform*
*Sungkyunkwan University*
*Suwon, Korea*
*Email: bluereaper@skku.edu*

Hyoungshick Kim
*Department of Computer Science and Engineering*
*Sungkyunkwan University*
*Suwon, Korea*
*Email: hyoung@skku.edu*

***Abstract*—This paper introduces a security framework to detect false emergency requests, which might significantly disturb actual rescue operations for emergency services. We design a threshold-based false request detection algorithm named FRDA to effectively filter out false requests, which is based on a reasonable scoring formula to automatically evaluate the trustworthiness of emergency requests with the caller's reporting behavior and location. This study will be a basis for developing mitigation techniques to deal with false emergency requests.**

## 1. Introduction

Emergency events can be categorized as human-caused and natural disasters, but they are identically unpredictable. Generally, frequent emergencies cause a series of tragedies to the society and need proper emergency planning [1]. Due to the influence of emergency events, the emergency service facilities should be well prepared for any unexpected situation, regardless of time or circumstances. The loss or disability of emergency service capabilities would notably impact the nation's security, public safety, and morale. Although advances in technology enable emergency services to adequately prepare for and effectively recover from terrorist attacks, natural disasters, and other catastrophic incidents, false emergency requests have a destructive effect on their practicalities [2].

The false emergency requests are expensive problems for emergency service providers in the aspect of time and resources. These requests exhaust the emergency services and guide them to be away from people who may be in life-threatening situations and who need urgent help. In this condition, service providers are forced to multiply their resources to assure they are not being overloaded with inappropriate calls and therefore may not be able to respond to real emergencies [3]. Generally speaking, the false alarms can be generated based on good intents (e.g., inappropriate judgments of emergency situation, mistake call or fault generated automatic emergency alarm) or malicious intents. Although the good intent false emergency requests are inevitable, detecting and preventing the malicious emergency requests is a critical issue. In this view, we suggest a framework to evaluate the trustworthiness of the emergency request and detect the reality of the request based on the sender's historical behaviors and the reported information.

To have an effective informing system, the model provides a dataset of critical information related to the callers. Health-related or special service provider needs access to proper data, especially in emergency situations, maintaining the availability and privacy of data is an essential requirement [4]. Every caller, depend on its requirement, shares the required information. For instance, if a dedicated medical service is required, in the emergency case this service will be prepared by the emergency service. Furthermore, if the living location of such a caller is that no vehicle can close thereto, the rescue team should be equipped enough to cope with this obstacle. Our model makes it feasible for emergency service providers to achieve critical information of callers. It is not needed to get this information from caller and every required information is automatically available.

In this model, all emergency requests go to a central server to analyze their trustworthiness. After confirming, depend on the request type, they will be forwarded to the proper emergency service center (e.g., emergency medical service, police, and firefighter service). We define a false request detection algorithm (FRDA) to scrutinize the reliability of the emergency request. The FRDA is a scoring-based algorithm which monitors callers' behaviors and scores them based on the result of the emergency service report. To evaluate the reliability of a request from a suspicious sender, the FRDA performs a verifying function based on the reported information. The key contributions of this paper are summarized as follows:

- We conducted an in-detail study to matching the trust formula in the business community to scrutinize the reliability of the emergency request.
- We conducted the analysis of the trustworthiness of the emergency request based on the reported information and the caller's historical behavior.
- To implement the model and defining the trust parameters, we exploited a real dataset of emergency request of London Fire Brigade.

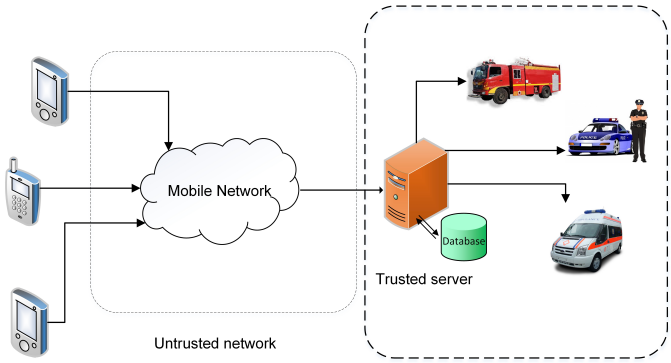


Figure 1. The overview of the emergency informing system

This paper is organized as follows. Section 2 introduces the basic concepts of our model. The FRDA, the reliability verifying function, and defining the parameters of trust are described in detail at Section 3. Section 4 concludes the paper and shows our plan for future work.

## 2. Basic concepts

Basically, this model is based on a made-for-purpose application that should be installed in advance on the caller side. All emergency requests are conducted to a server at the central emergency center. The model makes it possible to benefit the service anonymously based on caller's ID. To avoid any misuse, the application should verify caller by a real phone number (or email address). This verification leads to detecting any duplicate registration and moreover the anonymously malicious request. The server is accessible to callers who are identified by unique caller IDs. Every caller as a client is assigned a profile which consists of the basic information. In this profile, client optionally can share some information about himself/herself which are useful for emergency services. For instance, health situation (e.g., suffering from heart problems), accommodation's condition (e.g., in which floor he/she lives), or being under threatening record which is useful for police emergency service. The overview of the emergency informing system is shown in the Fig. 1.

Depending on the requested emergency service and based on the client's profile, the server adds some information and forwards it to the relevant emergency service provider. In this case, the critical required information will be automatically available for emergency services to provide proper service. As the benefit of this preparation, no time is wasted to get the information from the caller. In the view of privacy preserving, we consider a trusted server with adequate privacy protection and no information leakage. Moreover, the server side is considered to be well equipped which causes no delay and deficiency. The clients' evaluation is performed based on their profiles on the server side. Each client is assigned an index to evaluate its trustworthiness and is updated after every emergency request. The trustworthy evaluation is executed by FRDA on the server side. Detecting any cheating and misuse of the model is another responsibility of the server, which should be performed at the first stage of request analyzing.

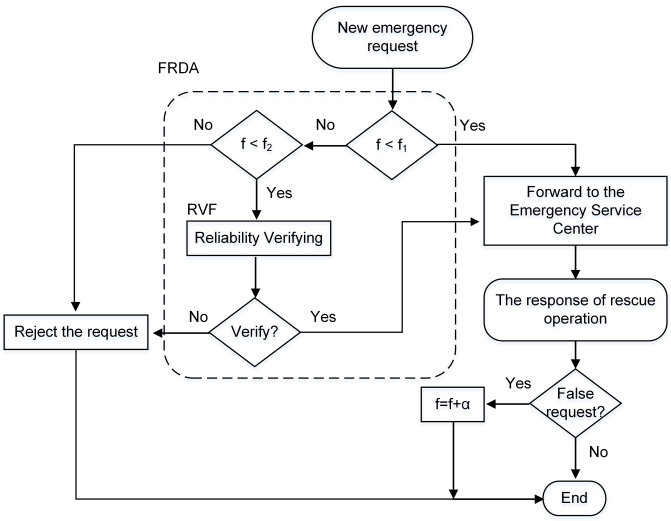


Figure 2. FRDA and indexing an emergency request

## 3. False request detection algorithm-FRDA

To cope with the false emergency service request problem, we define FRDA to scrutinize the reliability of an emergency request. The FRDA is a scoring-based algorithm which relies on the reported information and history of the caller's behavior. Based on the result of the emergency operation which is replied by the emergency service center, a request is considered a false or a real emergency request. We assign a point ($f$) to a caller if its request has been reported as a false emergency request. Figure 2 shows the structure of FRDA in conjunction to the scoring mechanism.

Due to the importance of the responsibility of the emergency services which can mean the difference between life and death for someone in trouble, we should not ignore any request. To reduce the possibility of false request, the model assigns a false index; $f$, to each client and defines two levels of threshold; $f_1$ and $f_2$. The threshold values of $f_1$ and $f_2$ should be defined based on the experiences of the emergency service centers. It should be noted that these values should be set according to experimental reports.

To manage the public service resources, we classify users based on their behaviors. By using a binary classification, we will have two categories: malicious and normal callers. In this case, defining an exact threshold for decision is very difficult. Furthermore, what will happen to an emergency request reported by a caller who has been classified wrongly as a malicious one? To avoid these problems, we turn to ternary classification. By ternary classification, it is possible to quickly offer a service to genuine clients and give more chance to suspicious callers to prove their requests. In this classification the rate of wrongly rejection will be decreased. By experimentally selecting proper threshold

levels for these two levels, we can manage the public service resources optimally.

In comparison to the defined thresholds, the clients fall into three categories:

- Normal client; $f \leq f_1$
- Suspicious client; $f_1 < f \leq f_2$
- Blocked client; $f_2 < f$

The request of the normal client is immediately forwarded to the emergency service center. If FRDA detects that the client is a blocked one, its request will be rejected. To verify the reliability of a request from a suspicious client ($f_1 < f \leq f_2$), FRDA runs reliability verifying function (RVF). By this function, the reliability of the request will be measured based on the reported information and the client's false index. By satisfying the RVF, the client's request will be forwarded to the proper emergency service center. Otherwise, it is considered as a fake client and its request will be rejected.

After the rescue operation, the emergency service center replies the result to the emergency server. If the result shows that the request was a false request, the false index of that client will be increased by $\alpha$, ($f = f + \alpha$) and in the next time the new index will be considered. To distinguish between the false request types, the $\alpha$ is assigned different values. In this condition, a malicious request leads to bigger value of $\alpha$ in comparison to a false request made by mistake. We initially set $\alpha = 1$ but it can be conceded to the emergency service facilities to set it based on their experiments.

### 3.1. Reliability verifying function

To evaluate caller's trustworthiness, the FRDA performs a reliability verifying function. Basically, this function is based on the trust formula enhanced and documented by Charles Green in his book, The Trusted Advisor [5]:

$$T = (C + R + I)/S \tag{1}$$

Where:
$T : trustworthiness$, $C : credibility$, $R : reliability$, $I : intimacy$, $S : self - orientation$

Based on this definition, four primary components have direct impact on the trustworthiness, words, actions, emotions, and motives. The $credibility$ refers to the words and claims of the person and $reliability$ is achieved based on his actions or level of dependability. The $intimacy$ indicates the security one feels when entrusting a person. In a mutual relationship, the $self - orientation$ covers anything that keeps one part to focus on himself/herself more that the other part. By assigning values to the trust equation's factors, we can assess the trust level of a relationship [5]. The trust formula shows mathematically how certain behaviors or attitudes impact on the level of trust that clients have in a business or relationship.

To implement the trust formula, a real dataset of London Fire Brigade is used. This dataset contains 355796 records of the details of every incident responded to since January 2012 in London City. Information is provided for when and where the incident happened and the type of incident dealt with. Generally, the false request can be generated by malicious or good intents and automatic emergency alarm systems (e.g., automatic fire alarm (AFA)). Unlike the first two groups, the false requests caused by automatic alarm systems have no human origination. According to this dataset, $49.17\%$ of total emergency request is false request. Table 1 shows the detail of the false requests among the dataset of London Fire Brigade.

TABLE 1. THE RATE OF FALSE ALARMS IN THE DATASET OF LONDON FIRE BRIGADE

| False requests | Percentage |
|---|---|
| False-AFA | $36.45\%$ |
| False-Good intent | $11.22\%$ |
| False-Malicious | $1.49\%$ |
| Total false alarm | $49.17\%$ |

In these records, $63\%$ of the fire service requests generated by AFA were false. This situation specially occurred in emergency requests for non residential categories. On the other hand, the good intent false requests are usually generated due to misapprehend or incomplete information about the event. In the next section the parameters of the trust formula of emergency request are defined based on this dataset.

### 3.2. Defining the trust parameters

According to the trust formula, four parameters; credibility ($C$), reliability ($R$), intimacy ($I$), and self-orientation ($S$) should be initialized. In this view, we analyzed the information of the dataset to model the trust formula by the emergency request's parameters. By analyzing the records, we categorized four parameters of records to model the trust parameters. These categories consist of property category, address qualifier, request origination, and the user's behavior (false index $f$).

The property category is used to model the credibility ($C$). The analysis shows that the reported property has convincing relation to the kind of emergency request. For instance, the rate of true request in the outdoor category is more while the false request occurs more in the nonresidential categories. In this view, the property category is divided into four groups. Table 2 shows the rate of total, false, and true requests for each group. This information is presented as true requests of fire (T-Fire) and special services (T-Srv) and false requests of AFA (F-AFA), good intent (F-good), and malicious (F-Mal).

Generally, the information related to an event mentioned by an emergency requester helps us to evaluate his reliability ($R$). To model the reliability, the address qualifier of emergency request is used. If the reported event happens at the same place or at a nearby address of reporter leads to different rates of true and false requests. Therefore,

TABLE 2. THE PROPERTY CATEGORY'S GROUPS (%)

| Property type | Total | T-Fire | T-Srv | F-AFA | F-good | F-Mal |
|---|---|---|---|---|---|---|
| Dwelling | 45.86 | 13.41 | 42.40 | 30.29 | 12.17 | 1.72 |
| Nonresident | 30.92 | 8.29 | 10.53 | 72.80 | 7.08 | 1.28 |
| Outdoor | 15.01 | 65.62 | 16.68 | 0.21 | 15.85 | 1.62 |
| Vehicle | 8.19 | 24.63 | 61.37 | 0.09 | 13.06 | 0.85 |

TABLE 3. THE ADDRESS QUALIFIER'S GROUPS (%)

| Qualifier | Total | T-Fire | T-Srv | F-AFA | F-good | F-Mal |
|---|---|---|---|---|---|---|
| Correct incident address | 54.83 | 14.67 | 27.23 | 46.1 | 10.29 | 1.53 |
| Within same building | 20.78 | 8.58 | 35.63 | 46.92 | 7.56 | 1.21 |
| In street | 14.01 | 38.64 | 42.38 | 2.84 | 14.13 | 1.95 |
| Near address | 6.92 | 47.42 | 17.88 | 14.54 | 18.75 | 1.38 |
| others | 3.45 | 57.8 | 18.88 | 0.57 | 21.26 | 1.47 |

the reporter's awareness of the event's address has direct relation to his reliability. For instance, $81.02\%$ of emergency requests is true in the street qualifier, but this rate falls to $44.21\%$ for reported events within the same building. We define five groups of address qualifier which Table 3 shows the rate of total, false, and true requests for them.

Theoretically, in the trust formula, the intimacy ($I$) refers to the security that one feels when entrusting a person. Based on this, we refer to history of malicious false request rate of the event's location. To this end, we classify the boroughs of the service area of London Fire Brigade, according to the rate of malicious requests. Totally, the service area consists of 34 boroughs and they are scored based on the statistic of the malicious request related to each one. For instance, Newham has the highest rate of malicious request while city of London has the lowest rate. Based on this classification, an emergency request reported from the Newham borough has the least intimacy in comparison to other boroughs.

Finally, the behavior of a caller is considered as the resource to value his self orientation ($S$). Basically, the factor of self orientation has to do with the focus of the person in question [5]. To this end, the false index ($f$) of the caller, which is available at the input for RVF, is exploited. Based on the trust formula, higher value of $S$ causes to lower trustworthy. Due to the intent of emergency services for life rescuing or saving the properties, we should not ignore any emergency request. Regarding to these facts, we initialize $S$ to $f_1$. Therefor, we set:

$$S = f_1 + f \tag{2}$$

In this view, a suspicious caller with higher false index has a lower chance to be verified.

### 3.3. Request reliability verification

The RVF evaluates a request's trustworthiness ($T$) based on the reported information and reporter's false index. To this end, a trust threshold value ($TTV$) is defined to verify the request reliability. A request's reliability is verified, if its trustworthiness ($T$) is bigger than the $TTV$. Based on the trust formula, the trust parameters are scaled of 1 to 10 [5]. Therefor, regarding to possible amount of trust parameters, a proper value should be defined for the $TTV$. The TTV directly depends on the experimental information on request dataset.

In the trustworthy evaluation, the parameters of property category, address qualifier, and event's borough are directly evaluated based on the reported event. While, to note the caller's personally behavior as self orientation, we refer to his false index. The reliability of a request from a caller with lower false index has more chance to be verified in comparison to the request from someone with higher index. In this view, to verify a request's reliability the reported information is evaluated based on the behavior of the caller. Therefor, a higher false index caller should report more precise information in order to be verified.

## 4. Conclusion and future work

The problem of the false emergency request is the inevitable fact of all emergency service centers. In this paper, we proposed a framework to score the callers of emergency service according to their behavior. To evaluate the request reliability of a suspicious caller, we defined RVF function to analyze the reported information. Based on the trust formula, the RFV evaluates the reliability of the request from a suspicious caller. The function analyzes property category, address qualifier, and the location of the event based on the caller's false index. To value each parameter of the trustworthiness, the model exploits a real dataset of emergency requests. Empirically, a threshold value is defined for trustworthy evaluation based on the real emergency request dataset.

Practically, the model is built on caller's behavior history. Therefore, the RVF function is performed only for suspicious callers. For the future work, we are developing the verifying function to evaluate the reliability of all emergency requests. Accordingly, the verification is performed based on the reported information on the event independently of the reporter and his behavior.

## Acknowledgments

This research was supported by:

- The MSIP (Ministry of Science, ICT and Future Planning), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2015-R0992-15-1006) supervised by the IITP (Institute for Information & communications Technology Promotion)
- The National Research Foundation of Korea (NRF) grant funded by the Korea government (No. 2014R1A1A1003707)
- The ICT R&D program of MSIP/IITP [2014-PK10-28, Standard Development of Network Security based SDN] (partly)

## References


[1] Han Fuyou, Zhang Hailong, and Dong Liyan. Research on evaluation model of emergency response plans. In *Mechatronics and Automation, ICMA 2009. International Conference on*, pages 5117–5122, 2009.

[2] Emergency services sector-specific plan, department of homeland security, usa, 2010.

[3] False emergency calls, eena operation document- false emergency calls, european emergency number association, 2011.

[4] Kalpana Singh, Jian Zhong, Vinod Mirchandani, Lynn Batten, and Peter Bertok. Securing data privacy on mobile devices in emergency health situations. In AndreasU. Schmidt, Giovanni Russello, Ioannis Krontiris, and Shiguo Lian, editors, *Security and Privacy in Mobile Information and Communication Systems*, volume 107 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pages 119–130. Springer, 2012.

[5] David H. Maister Charles H. Green, Robert M. Galford. *The Trusted Advisor*. The Free Press, 2000.