Application of Social Relation Graphs for Early Detection of Transient Spammers

RADOSLAW BRENDEL and HENRYK KRAWCZYK
Electronics, Telecommunications and Informatics Department
Gdansk University of Technology
Narutowicza 11/12 Gdansk
POLAND
radoslaw.brendel@eti.pg.gda.pl, henryk.krawczyk@eti.pg.gda.pl

Abstract - Early detection of social threats and anomalies is a real challenge in today's dynamic societies. The people form many complex social relations that can be shown by various types of graphs in which the nodes would represent the subjects (individuals or groups of people) and the links would indicate specific relations between them. The analysis of these constantly changing relations can point out specific social threats that are imminent. Observing the tendency of changes in the social relation graphs, such threats can be early detected and adequate preventative steps can be taken. In the paper we present how this approach can be efficiently used to early detect imminent threats of spam to a local e-mail society and isolate groups of spammers before their messages reach the users inboxes.

Key-Words: - social graphs, imminent threats, graph patterns, spam detection, security

1 Introduction

Human societies developed specific social relations in many fields of their activities. Much research has been done so far in analyzing complex relations between individuals and organizations in real world [2][8][11]. We can easily observe them in one of their possible representations – social relation graphs. In such graphs nodes usually represent individuals or whole organizations whereas links between the nodes have the meaning of relations of different types that we want to study. Social relation graphs are characterized by several specific properties. One of them is the small world phenomenon. It means that two nodes in the graph are related with each another by relations of small amount of other intermediate nodes. For instance, an experiment made in 1967 by psychologist Stanley Milgram showed that the chain of social acquaintances connecting one arbitrary person to another arbitrary person required on average only five intermediates (the experiment was carried out on US individuals). Moreover, the analysis on social relation graphs usually allow us to distinguish a certain set of nodes (representing "social" objects) that play main roles in the studied community [1].

Observing the behavior of a certain community for a long enough time and collecting the data about relations between its members, it is possible to create a group of pattern graphs representing typical and abnormal relations in the community. After that we possess the knowledge of what are the typical and abnormal relations in the community being analyzed and any exceptions to these patterns can be distinguished and measured in some way. Detection of these exceptions is crucial from security point of view because their presence usually indicates some anomalies in relations between the members that in consequence cause threats to the community.

In the paper we will show how the method of early detection of incoming threats can be effectively performed taking into account the properties of social relations graphs and the assessment of congruity level between these graphs and relations pattern graphs. We will also present how the method can be used to efficiently detect spammers in one of the most famous social network - Internet e-mail community. Effectively, every e-mail user will be identified as a spammer or a regular user.

2 Social Relation Graphs

In order to perform detection of imminent threats to a certain community we defined two categories of graphs. The first category is compound of directed social relation graphs that describe the current activity state of all the members of the community. The second category consists of pattern graphs describing typical and abnormal behavior of the members.

The social relation graphs are the structures that represent actual relations between community members after a specified period of time. The relation graphs are dynamic in the sense that they are constantly changing over time to reflect the current state of relations established between the members. This dependence on time will be indicated in many beneath formulas by subscript k, i.e. D_k would indicate digraph at a certain moment of k or after the k-th event. Now, let's assume we consider a social network that can be represented in form of the digraph $D_k(V_k,R_k)$, where $V_k=\{v_1, v_2,..., v_s,.., v_p\}$ represent a set of nodes (actors in a real social network) and $R_k \subseteq V_k \times V_k$ representing the relations $\Re$ between the nodes (actors). Assuming binary relations between the actors the digraph can be described by the following corresponding binary matrix $R_k = [r_{i,j}]_{p \times p}$ where:

$$r_{i,j} = \begin{cases} 1 & if\ v_i \Re v_j \\ 0 & otherwise \end{cases} \qquad (1)$$

Of course depending on the social network sometimes it may be important to express in some way the strength of the relationships between the actors. In such cases $r_{i,j}$ would take as values real numbers (most often nonnegative natural numbers).

The *neighborhood of the node* v_s (i.e. the relations between the actor and its neighbors in the digraph can be described by another matrix R_k^s that in fact is a submatrix (block) of the matrix R_k. Assuming that the set of nodes $V_{s,k}^N = \{v_i : v_i \Re v_s\} \cup \{v_s\}$ represents all the neighbors of the node v_s and v_s itself, the neighborhood matrix $R_k^s = [r_{i,j}^s]_{t \times t}$ where $t = |V_{s,k}^N|$ can be defined as follows:

$$\forall v_i, v_j \in V_{s,k}^N \quad r_{i,j}^s = \begin{cases} 1 & if\ v_i \Re v_j \\ 0 & otherwise \end{cases} \qquad (2)$$

We should underline that a relation graph is changing over time thus the relations between community members in the graph are different depending on the moment in which we decide to draw it. Let us look at Fig.1, where it is shown an example of relation graphs of a sample community in three different moments (t_1, t_2 and t_3). We assumed that members of the community represent mobile phone users, the relations we consider are made phone calls between the members and the weights of the relations will indicate how many times one user made a phone call to another (if a user made more than one phonecall).

We can notice that after time t_1 several relations were established (phone calls made) but we have not got clear image of typical relations in the community yet. We cannot distinguish yet any group of people that form a local community (members that maintain very close relations in respect to others). After time t_2 one local community was formed (indicated by circled dashed area). Finally after time t_3 we see that in the studied community three different local groups of members developed very closed relations between their members. Generally it is difficult to say for how long we need to observe a community in order to get a compete image of the relations in it. It just depends on the activity of its members.

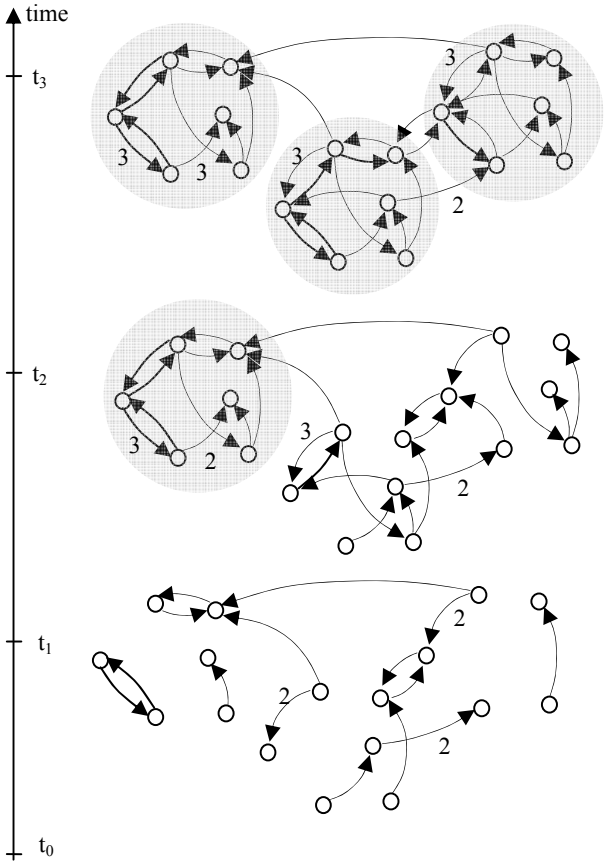

Fig. 1 – Changes a in sample relation graph over time; t_1,t_2 and t_3 are the moments of observations ($t_1 < t_2 < t_3$).

3 Relations Pattern Graphs

We mentioned that to effectively detect imminent threats to a community it is required to define a set of special graphs that we call relations pattern graphs, shortly pattern graphs. Their goal is to describe typical and abnormal behavior of the community members. We are able to create them only after getting familiar with the community

especially deeply studying the relations between its members.

Of course these pattern graphs will be different for every community they describe. For example, considering the Internet e-mail community an activity of one member (one e-mail user) that started hundreds relations with other members probably indicates abnormal behavior typical for spammers, spreading huge amount of advertising e-mails.

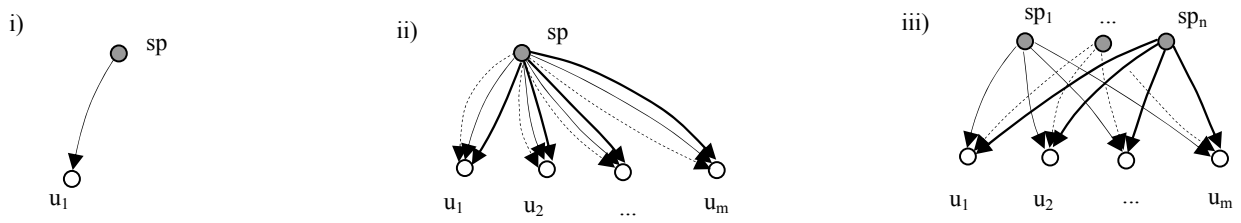

Fig. 2. Relations pattern graphs for three classes of spammers (sp): i) OO-OF; ii) MT-OF; iii) MT-MF

Below, as an example, we present possible relations pattern graphs that can be distinguished for Internet e-mail society. The proposed pattern graphs describe typical and abnormal behavior of e-mail users in regard to the problem of spam in the Internet community. These two types of behavior (typical and abnormal) have been naturally "assigned" to spammers and so-called regular e-mail users accordingly. The nodes of the pattern graphs are e-mail users, whereas the links represent relations started by sending an e-mail between two members (e-mail users).

Considering pattern graphs for spammers we distinguished three types of them representing three different classes of spammers. These are:

- OOOF – Only-Once One-Face; spammers of this class come up in the e-mail network only once and with only one identity;
- MTOF – Multiple-Times One-Face – spammers of this class come up in the e-mail network several times during a period of observation every time having the same identity;
- MTMF – Multiple-Times Multiple-Face – spammers of this class come up in the e-mail network several times during a period of observation every time changing its identity. It is a typical class of *transient* spammers.

Fig. 2 presents three types of pattern graphs corresponding to each class of spammers. In Fig. 2i we see a situation where a spammer sent only one e-mail to one regular user and never came up in the network again. Fig 2ii shows another type of spammer that sent e-mails to a set of recipients (u_1,...,u_m) every time sustaining its identity (sp). In the last Fig. 2iii we see a group of nodes (numbered sp_1,...,sp_n) that sent e-mails to a fixed group of users (u_1,...,u_m) from which nobody has responded (did not want to start relation with the senders). Probably all these sender nodes represent same spammer that every time he sends an e-mail he changes (forges) his appearance. Of course with only a certain probability we can state it but taking into account other attributes of the e-mails (further analyzing the headers of the e-mails for instance) can ensure our assessment.

Fig. 3 presents two non-spam pattern graphs. With high probability it can be stated that nodes ru on both graphs represent regular e-mail users (non-spammers). In Fig. 3i node ru sent an e-mail to a closely tied group of users (u_1,...,u_m). Most of them exchanged at least one e-mail with another user so they must know well each other forming a local small community (like in a real society). Assuming that spammers do not know local relations between e-mail users, the e-mail directed to a closely tied group of users allows us with high probability to classify the sender as a regular user.

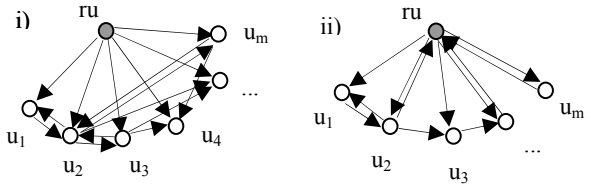

Fig. 3. Relations pattern graphs for regular users (ru): i) ru sent e-mails to a well-known group of users; ii) e-mails sent by ru have been reciprocated.

In Fig. 3ii we see another pattern graph where nodes u_2, u_3, u_4 (recipients) answered to the e-mail sent by node ru. Usually we do not answer to e-mails sent by spammers (especially that their e-mail addresses are often forged). Thus this act of responding to the e-mail lends credence to the sender classifying it with high probability as a regular user.

4 Imminent Threats Recognition Approach

In most cases the efficiency of reaction to anomalies in community members behavior strongly depends on the moment in which the anomalies will be detected. Usually sooner the anomalies are detected, more efficient the taken countermeasures will be. So it is extremely important to develop a solution for early recognition of imminent threats to the community.

The proposed recognition approach begins from the build of the relation graph that describe the current community members activity. The graph is build from historical data collected for a long enough time letting all the typical relations between the community members to be established. We prepared a collection of perl scripts and programs that analyze e-mail logs generated on mail servers systems[4][5]. After that we need to distinguish relations pattern graphs describing typical and abnormal behavior of the community members. In general we can create different set of such pattern graphs according to existing real threats. Having these two types of graphs prepared we are ready to implement the threat recognition procedure.

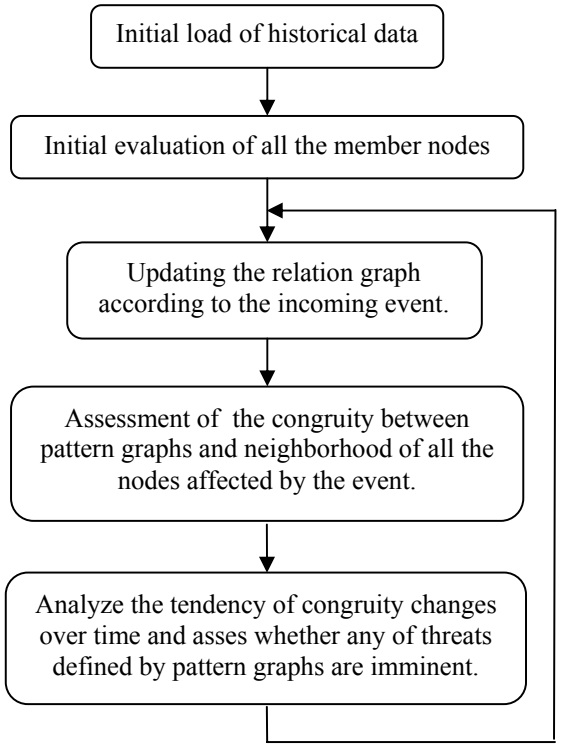

Fig. 4. General imminent threats recognition approach

The recognition procedure is based on the following assumption. If we simply want to state whether a certain community is under one of the previously defined threat (in the form of a relations pattern graph) we could perform a search to check whether any of the pattern graphs is already included in the relation graph. Having found any of the pattern graphs in the relation graph it would indicate that the community is already a victim of this specific threat (for example, in the case of email communication graphs it would mean that spam messages have already reached the users mailboxes). However, our goal is to find a method of early recognition of imminent threats. That is why, it is necessary to predict the creation of such pattern graphs in the relation graph before they appear in it in full form. We can achieve it by creating special *pattern graph similarity functions* calculating the congruity between the pattern graphs and parts of the relation graph. Such functions are specific to each pattern graph. The congruity evaluation can take place on regular basis or on every event that changes the neighborhood of any node representing a community member in the relation graph. Taking into account all the evaluations (after last event and from the past) it is possible to indicate the tendency of changes in evaluation results. Every time when a certain part of the relation graph becomes more and more similar to one of the pattern graphs representing actual threat we can state with certain probability that the threat is imminent. Formally this approach is realized by the *imminent threat recognition algorithm* that we present in the next chapter. However, the general imminent threat recognition approach is shown in Fig. 4.

5 Imminent Threats Recognition (Roles Indentification) Algorithm

Before introducing the imminent threat recognition algorithm let us assume that we already defined a social community according to the formula (1) and that it is also defined a set of relations pattern graphs describing good (positive) and bad (negative) behavior of one actor in the social network. Good patterns are indicated as G^i where i=1...n and bad patterns are indicated as B^j where j=1..m. Then for each G^i and B^j we define the *pattern graph similarity function* that evaluates the similarity between the neighborhood of one node and the corresponding relations pattern graph. It is defined as follows:

$$\rho^{G^i}\left(G^i, R_k^s\right) \ where \ \ i = 1...n \qquad (3)$$

$$\rho^{B^j}\left(B^j, R_k^s\right) \ where \ \ j = 1...n$$

and has the following feature:

$$\forall G^i : \rho^{G^i} \in \langle 0;1 \rangle$$

$$\forall B^j : \rho^{B^j} \in \langle 0;1 \rangle$$

The patterns graphs can be defined in many forms but one of the simplest and most natural would be to define them by means of a set of matrices. The form in which the pattern graph is defined is not important. The only requirement is that the similarity function has to be able to „handle" the form and make the evaluation between the corresponding relations pattern graph and the neighborhood matrix R_k^s.

The above definition of the pattern graph similarity function allows to calculate the similarity of one node's neighborhood in regard to one of the predefined pattern graphs. However there is also a need to define another similarity function that estimates the overall similarity level between one node's neighborhood and all of the relations pattern graphs. Such a function, taking into account all the relations pattern graphs, would answer the question: is the neighborhood (the relations) of the considered node v_s more similar to a group of „good" or „bad" relations pattern graphs. In other words, we want to assign a role to the actor. If the actor represented by the node v_s behaves regularly (conforms to positive patterns) we can say he plays a positive role in the social network and analogically if the actor behaves rather abnormally (having the relations more like in negative patterns) his role is negative. This function, called the *node's neighborhood similarity function* is defined as follows:

$$l_{s,k} = \frac{\max\left(\rho^{G^i}\right) - \max\left(\rho^{B^j}\right) + 1}{2} \qquad (4)$$

for i=1..n and j=1..m.

The above formula is defined in such a way that it conforms the following restriction:

$$\forall s,k : l_{s,k} \in \langle 0;1 \rangle$$

Now we are ready to introduce the imminent threat recognition algorithm that in practice identifies roles of actors of the considered social network. Practically two types of roles of actors are identified: positive and negative ones depending on which type of the relations pattern graphs an actor's relations are look like. The algorithm uses the node's neighborhood similarity function (formula (4)). However, it might happen that the formula (4) returns a value that does not allow to definitely identify the role played by the actor. In this case the role of the actor remains unclassified (unidentified). Generally, the algorithm during the roles assignment process uses several levels of node's neighborhood similarity function that are shown in fig. 5.

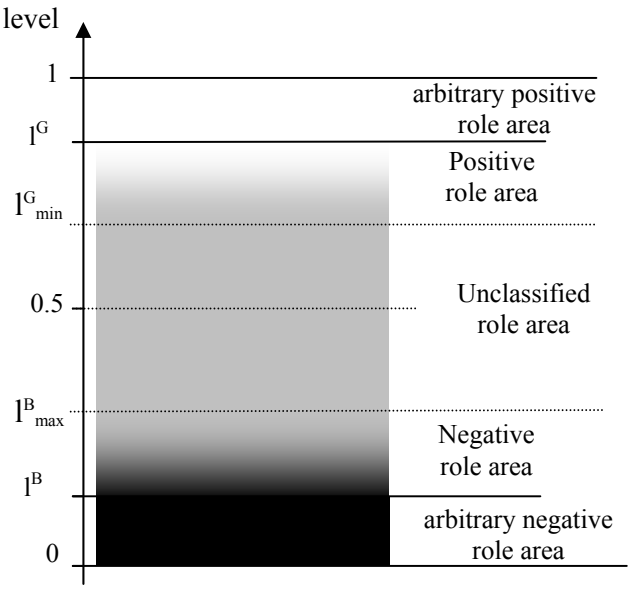

Figure 5. Various levels of similarity used by the roles identification algorithm.

The algorithm is defined in pseudocode and is shown if fig.6. As an input it takes the following values:

l^G - minimum level of similarity that has to be reached by the similarity function $l_{s,k}$ for letting the corresponding node v_s to be arbitrary identified as representing positive role;

l^B - maximum level of similarity that can be reached by the similarity function $l_{s,k}$ for letting the corresponding node v_s to be arbitrary identified as representing negative role;

l_{min}^G - the minimum level of similarity that has to be achieved by the similarity function $l_{s,k}$ for letting the corresponding node v_s to be considered as behaving normally, i.e. if the similarity function $l_{s,k} \in \langle l_{min}^G; l^G \rangle$ the node v_s *can* be classified as behaving normally depending on the tendency of past values of the node's neighborhood similarity functions;

l^B_{max} - the maximum level of similarity that can be achieved by the similarity function $l_{s,k}$ for letting the corresponding node v_s to be considered as behaving abnormally, i.e. if the similarity function $l_{s,k} \in \langle l^B; l^B_{max} \rangle$ the node v_s can be classified as behaving abnormally depending on the tendency of past values of the node's neighborhood similarity functions;

Δl^G - the maximum increment of similarity level „towards good behavior" for node v_s that can be observed during the last events; the level is calculated in such a way that its values are always greater than or equal to 0 (zero). The exact procedure that calculates values of Δl^G is presented below in Fig. 7;

Δl^B - the maximum increment of similarity level „towards bad behavior" for node v_s that can be observed during the last events; the level is calculated in such a way that its values are always less or equal to 0 (zero). The exact procedure that calculates values of Δl^B is presented below in Fig. 8.

```
CALCULATE l_{s,k} , Δl^G and Δl^B
IF l_{s,k} > l^G THEN
  node v_s is behaving normally
ELSE IF l_{s,k} < l^B THEN
  node v_s is behaving abnormally
ELSE IF l_{s,k} ≥ l^G_{min} AND Δl^G > l^G_{min} THEN
  node v_s is behaving normally
ELSE IF l_{s,k} ≤ l^B_{max} AND Δl^B < l^B_{max} THEN
  node v_s is behaving abnormally
END IF
```

Figure 6. Roles identification algorithm

```
IF (Δl^G + (l_{s,k} − l_{s,k−1})) > 0 THEN
   Δl^G_{new} = Δl^G + (l_{s,k} − l_{s,k−1})
ELSE
   Δl^G_{new} = 0
END IF
LET Δl^G = Δl^G_{new}
```

Figure 7. Calculating the value of Δl^G

```
IF (Δl^B + (l_{s,k} − l_{s,k−1})) < 0 THEN
   Δl^B_{new} = Δl^B + (l_{s,k} − l_{s,k−1})
ELSE
   Δl^B_{new} = 0
END IF
LET Δl^B = Δl^B_{new}
```

Figure 8. Calculating the value of Δl^B

As an example of how the roles identification algorithm works, let us assume that in a certain social structure we wanted to identify the role of one node, let us assume v_7. Let us also assume that we have created four relations pattern graphs: two for "positive" roles (G^1 and G^2, normal behavior) and two for negative roles (B^1 and B^2, abnormal behavior). We observed how the relations of the node v_7 were changing during three consecutive events (k=1,2,3) and after each of them we calculated pattern graph similarity functions $\left(\rho^{G^1}, \rho^{G^2}, \rho^{B^1}, \rho^{B^2}\right)$ and of course the neighborhood similarity function $\left(l_{s,k}\right)$. The hypothetic results are presented in Tab. 1. The graphical representation of the results is shown in Fig. 9.

Table 1. Sample evaluations of similarity functions for the hypothetic node v_7 after three consecutive events.

# of events (k)	ρ^{G1} (G^1,R^7_k)	ρ^{G2} (G^2,R^7_k)	ρ^{B1} (B^1,R^7_k)	ρ^{B2} (B^2,R^7_k)	$l_{7,k}$
0	0	0.00	0.00	0.00	0.50
1	0	0.00	0.00	0.00	0.50
2	0	0.17	0.13	0.75	0.21
3	0	0.11	0.25	0.83	0.14

In Fig. 9 apart from values of $l_{7,k}$ (k=0,1,2,3) there are also indicated values of Δl^B_{new} and Δl^B after the *k=3* event that allowed the roles identification algorithm to identify node v_7 as a actor playing probably negative role in the considered network. Precisely, the algorithm classified the user identified as node v_7 as behaving abnormally because its node's similarity function $l_{7,3} < l^B_{max}$ and $\Delta l^B_{new} \geq \Delta l^B$. The property of this algorithm is that it is able to identify the role of an actor even if not all of the typical properties of the role he or she tends to play have not been revealed yet.

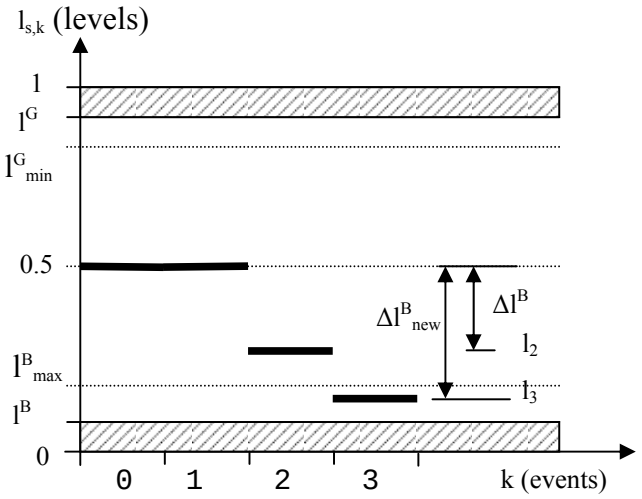

Figure 9. Early identification of actor's negative role.

6 Spammers Identification Using Imminent Threats Recognition Approach

As an example of practical implementation of the proposed recognition approach we decided to use it in identification of spammers. The goal was to identify potential spammers in an e-mail network to prevent regular users inboxes from being flooded with spam messages.

Research on e-mailing in Internet indicated that e-mail graphs have properties of typical social relations networks [9][12]. Thus, groups of people (e-mail users) tend to form local communities, where their members are tightly connected between each other. However, occasionally some anomalies can be observed that disturbance the regularities in the e-mail graphs. These disturbances can be the first symptoms of unusual e-mail traffic that we used to call spam. In the war of spam the probability of a victory is highly related to the early detection of spammers activities. That is because spammers tend to act by surprise and dynamically. If the detection time of spammers' activity is not short enough, they will reach the goal, i.e. their e-mails will arrive at their destinations.

Following the recognition approach described in the previous paragraph we established a five-steps algorithm that let us identify roles of e-mail users and classify them into one of the following lists: SL (Spammers List), RL (Regular Users List) and UL (Unidentified Users List). It consists of the following steps:

1. Initial load of historical data – to start the roles identification of e-mail users we need to load historical data collected for a certain period of time.
2. Initial identification of all the nodes – it serves as a reference point for all the further evaluations.
3. Add the incoming e-mail to the graph – every e-mail that comes is added to the relation graph and the e-mail's sender clustering coefficient (CC [9]) value is updated (CC is often used in pattern graph similarity functions).
4. Calculate adequate similarity functions for each sender node whose neighborhood has been changed by the event – here comes all the evaluations of similarity functions associated with each relations pattern graph;
5. Identify roles of all the users whose neighborhood has been changed by the event using roles identification algorithm.

In our case the relation graph represent e-mails exchanged by the users for a certain period of time. The time has to be long enough to let all the typical relations between the users to be established. In our tests we observed e-mail traffic related to one of the faculties of the Gdansk University of Technology. It turned out that one-month period of observation of e-mail traffic between local users and outside world was enough. Further collecting of data did not change the graph significantly. It means that after one month major relations between e-mail users were already established.

Next, we distinguished typical spammers and regular users behavior and described them in the form of relations pattern graphs. Finding right pattern graphs is probably the most crucial part of the whole process. It also points out how well we know the community we want to protect and whether we are able to describe precisely the threats against which we intend to protect the community. Of course a set of the pattern graphs is dynamic in the sense that once isolated graphs can change over time and new once can come out. For every pattern graph it is required to create a pattern graph similarity function $\left(\rho^G \text{ or } \rho^B\right)$ that will calculate a congruity between the pattern graph and an indicated part of the relation graph.

Below we present a sample pattern graph similarity function that can be associated with one of the spammers pattern graph presented in Fig. 2ii. This type of spammer we classified as MT-OF (Multiple-Times One-Face) and its characteristic property is sending e-mails to many regular users always being presented under the same identity. The other property of this kind of spammer is that it directs his messages to a loosely coupled group of

users that usually do not know each other. The similarity function uses clustering coefficient value of a node calculated according to the formula proposed by Watts and Strogatz [9]. Thus, the sample pattern graph formula can be of the form:

$$\rho^{B^1}\left(B^1, R_k^s\right) = \begin{cases} 1 - \dfrac{CC_k^s}{CC_{\min}^B} & for \ \left|V_{s,k}^N\right| \geq 5 \wedge CC_k^s \leq CC_{\min}^B \\ 0 & otherwise \end{cases} \quad (5)$$

where

CC_k^s - clustering coefficient for v_s after k-th event,

$CC_{\min}^B$ - minimum value of clustering coefficient required for the node to be classified as non-spammer. Let us assume $CC_{\min}^B = 0.3$.

In Fig. 10 it is presented a part of the relation graph illustrating how the neighborhood of the node v_s was changing over time becoming more and more similar to the defined pattern graph. Let us assume that the sender represented by the node v_s sent e-mails at three different moments that corresponds to events marked in Fig. 10 as 1, 2 and 3. Evaluating the pattern graph similarity function (ρ^{B^1}) after every event we obtain:

$$\rho^{B^1}\left(B^1, R_1^s\right) = 1 - \left(\frac{2}{12}\right) \quad / \quad 0.3 \approx 0.44$$

$$\rho^{B^1}\left(B^1, R_2^s\right) = 1 - \left(\frac{2}{20}\right) \quad / \quad 0.3 \approx 0.66$$

$$\rho^{B^1}\left(B^1, R_3^s\right) = 1 - \left(\frac{2}{42}\right) \quad / \quad 0.3 \approx 0.84$$

The last calculation indicates that the node v_s in the relation graph represents a spammer with the level of similarity equal to 0.84. We can confirm in our belief looking at the tendency of changes of the similarity values over time.

Of course to make our assessment more trusty we have to take into account results for all the similarity functions associated with relations pattern graphs, both spammers pattern graphs and regular users pattern graphs in respect to the node v_s. Analyzing how the similarity values change over time for all the pattern graphs for every node we can early point out the nodes that tend to behave like spammers and with a certain probability mark incoming e-mails as spam messages.

Our simple example showed how we can identify one node to be a spammer by estimating how his behavior is different from (or close to) a typical behavior described by a specific pattern graph. According to this evaluation one node can be classified to a certain group of users where each group would represent a level of 'suspicion' of its members. If we make the calculation over all the nodes at the certain time, it is also possible to asses the overall security level for a whole community. Of course the formula calculating this level is tied with the specific community taking into account not only the similarity levels calculated for each node but also, for instance, the importance (or power) each node has in the community.

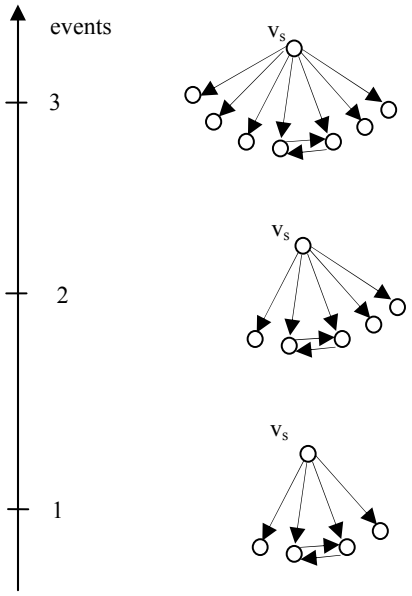

Fig. 10. The neighborhood of node v_s becomes more similar to spam pattern graphs over time.

7 Tests and results

Using the roles identification algorithm presented above we made a test with a local community consisting of e-mail users of one of the faculties of Gdansk University of Technology including other external e-mail users communicating with this faculty members. Initial load of the relation graph was made on the basis of system servers log files with registered e-mail traffic related to the users of the faculty.

After that we made an initial roles identification of all the nodes representing senders we are going to classify over time. Initial evaluation serves as the first referencing point when analyzing tendency of changes in values of the pattern graphs similarity functions.

After the initial one-month period, we collected all the incoming e-mails for one-week making the on-line roles identification of e-mail users. Almost 10.000 messages were collected containing spam

and regular messages. When an e-mail arrived the sender was first classified as spammer or regular user according to the results obtained from the node's neighborhood similarity function. Because the roles identification was based on values representing levels of similarity of one node to be a spammer or regular user, it was necessary to establish some thresholds used then in roles identification algorithm. Finally, every e-mail user has been classified on one of three lists: RL (Regular Users List), SL (Spammers List) and UL (Unrecognized Users List). The results of the test are shown in Fig.11.

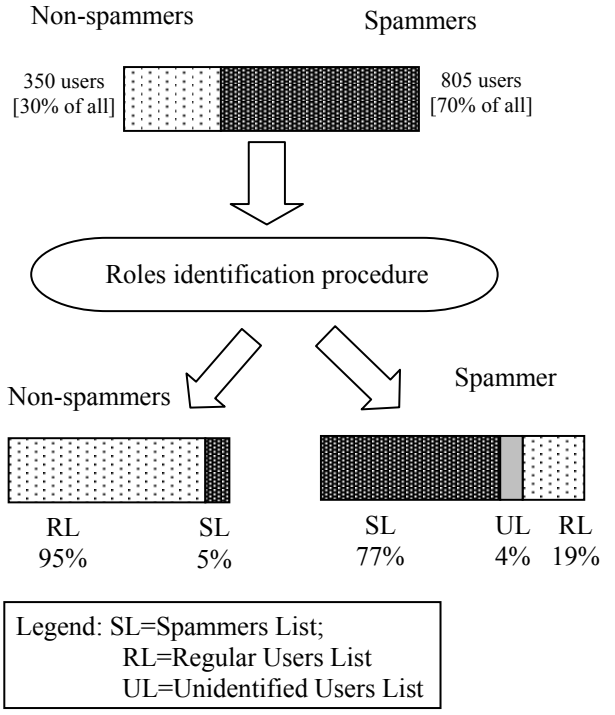

Fig.11. The results of roles identification process.

During the test we classified the amount of 1155 e-mail users of Internet e-mail community. 30% of them were regular users, the rest represented a group of spammers. Considering regular e-mail group of users, 95% of them were correctly identified as non-spammers and only 5% was misclassified as spammers. Contrarily about 19% of spammers were identified as regular ("positive") users and 77% has been assigned the correct role of spammer. Only 4% of users was left unclassified (the procedure could not uniquely assign any role to them).

8 Conclusions

The paper presented the general recognition approach of imminent threats to a community using social relation graphs. The critical part of this approach refers to the process of building relations pattern graphs that describe what is typical (acceptable) and abnormal behavior of the certain community. We need special monitoring techniques and identification algorithms to define such graphs. Moreover, the pattern graphs can also change over time what makes behavior classification much more difficult. We want to underline, that our proposition is flexible to consider the most difficult to detect class of spammers called transient spammers. It allows for taking into account the dynamic aspects of relations between actors that tend to change over time.

From the security point of view a state in which every part of the relation graph matches one of the pattern graphs representing typical relations can be treated as secure one and any exception to this state can indicate security compromise. Of course exact matching happens rarely so it is more practical to measure the level of congruity between pattern graphs and the relation graph. The graphical representation of such situations (see Fig. 9) allow an expert to take final decision.

We showed how the recognition approach can be used in roles identification process. We identified each e-mail user of a sample Internet e-mail community as spammer, regular e-mail user, some of them leaving them unidentified, however. Non-spammers users were classified efficiently, only 5% of them were misclassified. Unfortunately a significant part (about 20%) of spammers were classified as regular users (roles identification failed). The main reason of it was that spammers were able to reach local communities of e-mail users. It means that although this method is very efficient in recognizing typical activity of users, abnormal behavior requires additional pattern graphs to be created. These could take into account domain names of e-mail senders, IP addresses of e-mail servers and the possibility that generally the sender can forge their names.

In respect to e-mail users identification the proposed recognition strategy can give especially good results when we combine it with other techniques like content-based antispam tools. The proposed method correctly recognizes about 80% of all the users. After further improvement in spam classification this method can become a very effective tool in the initial spam recognition process leaving only a small part of all messages to be classified finally by heavy load content-based analysis tools.

References:

[1] L. A. Adamic, "Zipf, power-laws, and Pareto – a ranking tutorial" www.hpl.hp.com/research/idl/papers/ranking/ranking.html.

[2] P.O. Boykin, V. P. Roychowdhury, "Leveraging social networks to fight spam", *IEEE Computer*, April 2005

[3] U. Brandes, T. Erlebach, *"Network Analysis. Methodological Foundations"*, Springer, 2005

[4] R.Brendel, H.Krawczyk, "Spam classification methods based on users' e-mail communication graphs", *Proceedings of The Second IEEE International Conference on Technologies for Homeland Security and Safety*, Kadir Has University 2006

[5] R.Brendel, H.Krawczyk, "Detection Methods of Dynamic Spammers' Behavior", *Proceedings Of International Conference on Dependability of Computer Systems*, Szklarska Poreba, Poland 2007,

[6] P. J. Carrington, J. Scott, S. Wasserman, *"*Models *and Methods in Social Network Analysis"*, Cambridge University Press, 2007

[7] R. Diestel, "Graph Theory", Electronic Edition 2005, Springer-Verlag Heidelberg, New York 1997, 2000, 2005

[8] H. Ebel, L-I. Mielsch, and S. Bornholdt, "Scale-free topology of e-mail networks, *Physical Review* E 66, 035103(R) (2002)

[9] Z. Gyongyi, H. Garcia-Molina, "Spam: it's not just for inboxes anymore", *IEEE Computer*, October 2005

[10] M.E.J. Newman, "The structure and function of complex networks", http://arxiv.org/abs/cond-mat/0303516, March 2003

[11] M.E.J. Newman, S. Forrest, and J. Balthrop, "E-mail networks and the spread of computer viruses", *Physical Review* E 66, 035101(R) (2002)

[12] B. Whitworth, E. Whitworth, "Spam and the social-technical gap", *IEEE Computer*, October 2004

[13] S. Wasserman, K. Faust, *"Social Network Analysis.* Methods *and Applications"*, Cambridge University Press, 2007