# Using Complex Networks for Mining Malicious Activities in a Collaborative Map

Carlos Caminha, Vasco Furtado

Universidade de Fortaleza, Brazil
{carlos.o.c.neto, furtado.vasco}@gmail.com

**Abstract.** Collaboration with content sharing via digital maps is a type of application that is characteristic of the context of the social web. A malicious activity that is difficult to detect in this interactive context is the generation of a false trend on the map as the result of a plot in which several false reports by more than one person are done. In this paper, we describe how modeling complex networks of crime reported on a collaborative (or crowd) map can help identify regularities, and therefore show deviations arising from malicious activity. The idea here is to model a network comprised of users who reported crimes and the locations where such crimes were reported (e.g.: a census tract). Starting from a bipartite network model in which the vertices are individuals and census tracts, we projected a monopartite network of users in which the edges indicate the strength of connection between them. This connection strength indicates the degree of co-relatedness of the reports of crime made by these two users in a particular place. By characterizing this, we were able to observe that the relationships of non-hub users among themselves are typically no stronger than the relationship between such non-hub users and the hubs. If this happens, the evidence of malicious activity becomes clear. Simulation of malicious activities in this dataset has allowed evaluating the contributions and limitations of our approach.

Categories and Subject Descriptors: H.2 [**Database Management**]: Miscellaneous; H.3 [**Information Storage and Retrieval**]: Miscellaneous; I.7 [**Document and Text Processing**]: Miscellaneous

Keywords: Complex Networks, Data Mining, Security on the Web

## 1. INTRODUCTION

There has recently been an explosion of interest in using the web to create, assemble, and disseminate geographic information provided voluntarily by individuals. Crowd mapping, combining the aggregation of a Geographic Information System and crowd-generated content, flourishes daily on the Web [Rouse et al., 2007]. Sites such as Wikimapia (http://www.wikimapia.com), Click2fix (http://www.click2fix.co.sa), Crowdmap (http://www.crowdmap.com), and OpenStreetMap (http://www.openstreetmap.org) are empowering citizens to create a global patchwork of geographic information, while Google Earth and other virtual globes are encouraging volunteers to develop interesting applications using their own data. In crowd map applications, the digital map works as a blackboard for accommodating stories told by people about events they want to share with others typically participating in their social networks.

One of the prominent exemplars of this kind of system is WikiCrimes [Furtado et al., 2010] (http//www.wikicrimes.org). The idea behind WikiCrimes is to provide a common area of interaction among people so that they can report and monitor the locations where crimes are occurring. WikiCrimes allows users to access and to register criminal events on the computer directly in a specific geographic location represented by a map. Alarms that indicate the most risky places and heat maps are example of services produced by the website to the people in general.

In crowdmap there is a need to keep in balance the trade-off between diminishing the constraints imposed to the users with the intention to increase the number of participants in the system, and the rigid control that can be imposed to avoid unwanted behavior, such as the reporting of false information [Caminha et al., 2010]. For this reason, little information about the users is available, which ultimately makes difficult to apply methods that relies on the reliability of the source of the reports. Content analysis is also very difficult here, because it is hard to attest that a particular report made by someone is false or not. Related work on that line would be the analysis of review content such as [Ott et al., 2011] but no study on the domain of crime reporting has been conducted yet. Such approach would have to face the challenge coming from the fact that a crime report does not contain clues, such as adjectives to highlight the quality of a product, which might indicate fake activities.

Moreover, some malicious activities are not done exclusively by the report of a single user, which could be detected by anomaly detection techniques [Breunig et al., 2000] (e.g. LOF). False trends that are more difficult to identify are those caused by a group of users (which actually can be done by a single person with several accounts) that report crimes in a certain place with the aim of highlighting it in comparison to others. These malicious actions cannot be captured only through analysis of reports or of users individually; they require an investigation from the perspective of the relationships among users. Finally it is important to point out that this problem cannot be approached by supervised machine learning algorithms because there is no dataset with historical examples of malicious activities.

This motivated us to consider the exploration of complex networks modeled after the information of the users, the reports, and the locations where the reports were made. This model makes use of patterns identified in previous work [Cançado, 2005], [Furtado et al., 2009b], showing that the distribution of crimes by census tract follows a power law. It is verified in this context that there are few places that concentrate many crimes, and many places that concentrate few crimes. On the other hand, the literature on collaborative systems has shown that people's participation in systems such as crowdmaps also has a skewed distribution, which is popularly called the 90-9-1 rule [Rouse et al., 2007]. Many users participate little, and few users participate very actively.

Our approach in this paper is first to characterize the data described in WikiCrimes, which led us to investigate the existence of power law distributions suggested in the literature. Then, we were able to identify new regularities that are evident, particularly with regard to the correlation between users who report crimes in certain places. Starting from a bipartite network model in which the vertices are individuals and census tracts, we propose an innovative way to project this network into a monopartite network in which the vertices are only the users. The innovation relies on a heuristic defined to measure the strength of connection between the two users that is based on the degree of co-relatedness of the reports of crime made by these two users in a same place.

Based on this modeling and on information obtained by the characterization of the data such as the distribution of crime per census tract and the distribution of reports from users, we were able to find a kind of regularity within the context of WikiCrimes. This regularity refers to the fact that hubs have a high geographic coverage (i.e. they report crimes in the majority of census tracts) therefore demonstrate a well-defined behavior with regard to their connections with non-hub users. By characterizing this, we were able to observe that the relationships of non-hub users among themselves are typically no stronger than the relationship between such non-hub users and the hubs. If this happens, the evidence of malicious activity becomes abundantly clear by using typical anomaly detection methods. Simulation of malicious activities in this dataset has allowed evaluating the contributions and limitations of our approach. The generalization of the approach was prospected from the analysis of data from Amazon about the review of products. The representation of these data as complex network has shown that they share similar features with WikiCrimes, in particular the relationship between hubs and non-hubs.

## 2. MODELING COLLABORATION AS A COMPLEX NETWORK

### 2.1 Representation of Reports of Crime in Census Tracts

On collaborative maps, users in different geographical regions mark bits of information. Specifically in WikiCrimes, users report occurrences of various types of crimes anywhere in the world on a digital map. Hence, our complex network model is based on information from users, reports of crimes made by such users, and the locations where the reports refer to, represented here by census tracts. This model is based on a bipartite graph and its projection.

The directed bipartite graph $G^b(U, S, E^b)$ has - as vertices - the users $u$ $(\in U)$ and the census tracts $s$ $(\in S)$. An edge, $e^b$, represents the fact that there was a report by a user $u$ in a tract $s$. The weight of the edge $e^b(\in E^b)$ is obtained from the number of crime reports made in $s$.

In order to have a representation that indicates the strength of the connection between the users, we projected the bipartite graph on to a monopartite graph in which vertices are the users. Thus, users who reported crimes in the same tract will have an edge that joins them. We have defined a heuristics to compute the weight of this edge that is based on the number of crimes these two users reported in common in that location. In other words, if users report crimes in more than one tract, the number of crimes reported by them in common in these tracts is added to the weight of the edge.

Formally, the monopartite graph $G$ $(U, E)$ has on its vertices the users $u$ $(\in U)$ and the edges $e$ $(\in E)$. The weight of an edge $e$, $w(e_{u,u^1})$, between two users $u$ and $u^1$ $(\in U)$ is calculated based on the weight of the edges of $u$ with $s$ $(\in S)$ and of $u^1$ with $s$ $(\in S)$ in the bipartite graph $G^b$. More specifically, $w(e_{u,u^1})$ is is the sum of the minimum number of crimes reported by $s$ $(\in S)$ and $u^1$, in all tracts, according to the formula below:

$$w(e_{u,u^1}) = \sum_{i=0}^{n} (min(w(e^b_{us_i}), w(e^b_{u^1 s_i})))$$

Where $n$ is the number of census tracts, $s_i$, in which both $u$ and $u^1$ reported crimes.

The strength of the relationship between users, represented by the weight of the edge $w(e)$, indicates that the higher the weight, the more those users reported crimes in the same locations.

Figure 1 illustrates the steps of the process of construction of this graph where four users report crimes in three census tracts. In Step 1 we see the bipartite graph of user and census tract in which repeated edges represent the several reports done by the users in the tracts. In Step 2 the graph is transformed in a way that the weight of each edge is the number of crimes reported by the user in the tract. In Step 3, the census tracts are eliminated and users become connected directly together if they have reported a crime in the same tract. In Step 4, the weights of the edges $w(e_{u,u^1})$ representing the relationship between the users, are calculated.

### 2.2 Characterizing WikiCrimes data

Modeling user interaction as a complex network allows one to extract the main properties of such network to better understand how users relate to one another. The analysis was segmented by cities, that is, we generate networks that represent the reports of users in a particular city. This is due to the fact that malicious activity carried out by a group of users and that is more harmful to WikiCrimes is generating a false trend of violent area (typically represented by a *hot spot* [Mollenkopf et al., 2000]). Our premise is that the context usually operated by a user is that of a city (usually the home town). Note that this choice does not affect the generality of our approach. It only facilitates the semantic

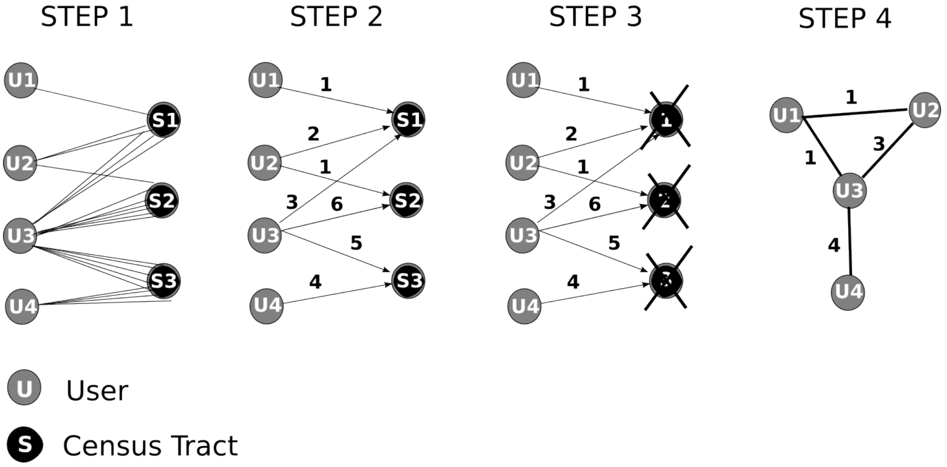

Fig. 1. Transformation of the bipartite graph into a monopartite graph in four steps

analysis of the results and allows us to work with a significant data in statistical terms, which would be difficult for low granularities.

Finally, it is noteworthy that WikiCrimes has very uneven levels of participation among different cities. The segmentation allowed us to focus the analysis in Fortaleza, a city with more than 2 million inhabitants and in which participation in WikiCrimes is intense. 14.63% of users come from there (1667 out of 11,394). The bi-partite network built to Fortaleza is fully connected and includes 416 users who registered 12,984 crimes in the period from 21/11/2007 to 25/04/2012 in the 3018 census tracts of Fortaleza. These data are available in `http://www.wikicrimes.org/wikicrimesapi/crimes\_per\_census\_tract.zip`.

The analysis indicates that the process of forming this monopartite network does not seem random. On the contrary, the characteristics extracted from the bipartite network show a preferential connection process of the nodes, where new vertices inserted tend to connect with the hubs of thenetwork. The number of crime reports per census tract follows a power law with exponent of 2.88 (see Figure 2b). This finding confirms previous work on the distribution of crime occurrence per census tract on urban metropolis [Cançado, 2005], [Furtado et al., 2009b]. The distribution of crime reports per user, the out-degree of the bipartite graph, confirms the 90-9-1 rule [Whittaker et al., 1998], which states that active participation in collaborative sites is dominated by 1% of the users. The majority of them actuates with very low frequency. It follows a power law of exponent 2.19 as the plot in a log-log scale of the degree distribution of Figure 2a illustrates.

As for the monopartite graph generated from the bipartite one according to the description in Section 2.1, the main properties are the following. The network hub has degree 357, i.e., this vertex has nearly 20% of all the network's edges, which total 1590. It is also interesting to see that it connects with 357 vertices, and this value represents more than 90% of the total number of nodes on the network. This stems from the fact that the hub reported crimes in almost all the census tracts (more precisely: 82% of the tracts). We are going to go back on that issue later.

The network has a low density, with only 2.35% of the maximum density that a network of 393 nodes can attain. This value shows the great distance between the degree value of the graph's hubs and the degree value of the vertices with fewer adjacencies on the network. The shortest path between all vertices is, at most, four leaps.

The hubs seem to play an important role in maintaining this property at such a low value. The fact that they report crimes in many places makes the shortest paths between users that report in few places remain very small. The clustering coefficient to the network is 0.93. This high value is due to the projection we made from the bipartite graph to a monopartite one, with complete sub-networks, because all the users who report crimes in the same location have edges connecting them all together.

Ranking the hubs allow us to verify their scope, that is to say, the number of census tracts they cover. This proved to be particularly relevant information because the largest hub of the monopartite graph connects with 90% of other users. This follows from the fact that in the bipartite graph, the user with greater out-degree crimes recorded in 82% of tracts. We decided to analyze what this process means from the original bi-partite graph. We saw how, in general, users who report more crimes (those who have high out-degree), do so in various census tracts, thus demonstrating a high coverage of them. Due to this characteristic, we found that, in the case of data from WikiCrimes few hubs users can have an almost complete coverage, i.e., can have records of crimes in almost all the census tracts.From a more general form, this fact comes from the value of the exponent of the power law that characterizes the out-degree of the reports from users. In the case of WikiCrimes, the exponent was 2.19. In this case, the first 6 hubs had made at least one record of crime in more than 95% of census tracts. Figure 2c shows the relationship between the non-hub users with the hubs to WikiCrimes and Amazon data. For example, observing the dotted line representing WikiCrimes data, we can verify that 97% of non-hub users have an edge with at least one of the sixth first hubs (1,5% of users).

## 3. MODELLING ANOMALIES FROM CORRELATIONS

### 3.1 Overview

The properties extracted from the monopartite network show that hubs are very connected to non-hub users, the result of comprehensive activity that exists in the reporting of events in various census tracts.

The explanation for this fact is that in collaborative systems of crime reports, where official data and data coming directly from the population are mixed, the hubs are typically entities or people with good reputations, such as government agencies, and that possess information that is mapped on a wide geographical area. In WikiCrimes, when analyzing data from Fortaleza, we were informed by the WikiCrimes administrator that the hubs are the so-called "certifier entities" [Furtado et al., 2010], project partners that hold a large volume of information on crimes, such as insurance brokers, police officers, specialized media, etc. Because they are aware of several crime occurrences, they make their reports in various census tracts in the city.

This scope of the hubs has an essential role in the behavior pattern of users and their reports of crimes, and seemed to be the key for detection of activities that may indicate fraud. It is noteworthy

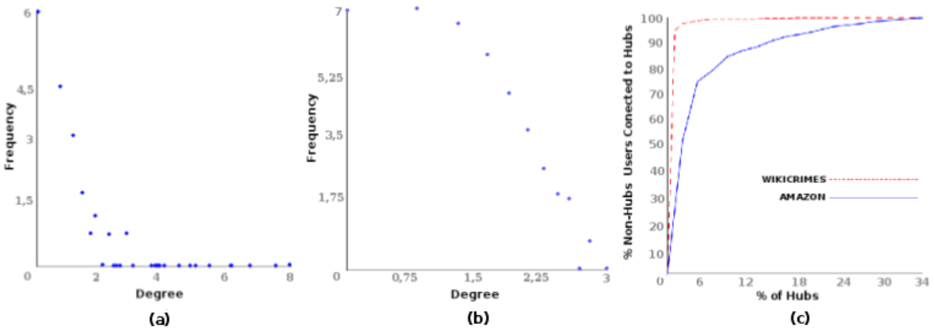

Fig. 2. (a) Plot log-log of out-degree distribution of the bipartite graph of crimes report per ranking of users; (b) Plot in a log-log of in-degree distribution of crime reported per ranking of census tracts; (c) Relation between hubs and non-hubs in WikiCrimes and Amazon(We are going to explain Amazon data in Section 6. Here our analyses will concentrate on WikiCrimes)

that in the monopartite network of users, the weight of the edge determines the degree to which two users, connected by that edge, report events in the same places. Analyzing the relationship of these users, we noted that, in all of them, the edge with the highest weight is one that refers to its direct connection to a hub. In the case of WikiCrimes, we're talking about a set of only 14 users. Here we saw a type of regularity that is evident of this type of scenario, i.e., non-hub users report crimes in areas where hubs also do. In other words, the coincidence of reports in a given area is much more likely to occur with reports made by hubs rather than with reports made by other, non-hub users.

This relationship pattern between non-hub users and hubs proved to be worthy of verification, because the absence of this pattern in the behavior may indicate suspicious relationship between users, whereby people who do not have a volume of reporting sufficient to generate a false trend map, start to have this power by adding their entries to those of other users at that location.

Formally, verification of this pattern can be described as follows. Let $U_h(U_h \subseteq U)$ be the set of users $u_h$ with a high degree $d$, $(d(u_h) > \theta$ where $\theta$ is a system parameter) and a triple $T$ formed by a hub $(u_h \in U_h)$ and by two users $u_1$ and $u_2$ $(\in U)$. The correlation factor $\rho(u_1)$ of a vertex of the triple is a ratio of the weight of the edge $w(e_{u_1,u_2})$ of this vertex with another user by the weight of the edge with the hub, $w(e_{u_1,u_h})$, as expressed in the following equation:

$$\rho(u_1) = w(e_{u_1,u_2})/w(e_{u_1,u_h})$$

Based on the correlation factor of users in these triples, one can see the limit for a given vertex (user) to be considered anomalous in its activity of reporting events on maps. We verified that, for the *Fortaleza Network*, it is normal to expect values of $\rho \leq 1$. The value of $\theta$ has to be chosen from an analysis as that we have shown in Figure 2c. As $\theta$ determines the number of non-hubs users who will be explored, it works like a confidence value determining the error that a priori it is known to have with the method. In WikiCrimes we have chosen to work with 1,5% because we had the information that all these hubs were trusted. It means that only 3% of users are not going to participate of the computing of the correlation factor because they are not connected to the hubs.

### 3.2 An Example in a Hypothetical Scenario

Below we will detail our explanation with an example in a simplified scenario where three users will attempt to generate a false trend in a region. To do so, they will observe where the map shows a high density of events, and will then report false information in another area with the intention to make this other region a hot spot as well.

Given the following situation: in a particular region we have a number of sectors (based on census tracts), one with 40 reports of crime and all the others with a number of reports close to 10 events. Assume also that in this network, formed by this collaboration, it is normal to expect that the vertices have a relation of greater weight with the hubs than other with network users (a characteristic identified in WikiCrimes). Of course, the sector with 40 events will be shown as the hot spot. Then three malicious users decide to make a sector with a low density of events (for example, 10 reports) a dangerous sector. They must report at least 30 entries to "equalize" the danger in this new region. So they divide among themselves the 30 reports with each reporting 10 events. We will detect the odd behavior when we realize that those three users will have a stronger relationship between themselves than their relationship with the hub, straying from the normal pattern of the network. Figure 3 shows the bipartite (a) and monopartite graphs (b) before and after a malicious activity provoked by three users.

Note that in Figure 3(b), after the malicious activity, the vertices(other users) have links with themselves with weight of exactly one while with the hub the weight is greater than or equal to

1. Consequently, their is less than or equal to one. As for the malicious users, they have among themselves links with weight of 10 while the weight of the links with the hub is 6. Thus the $\rho$ for these three users is 1.66(10/6). The higher the value of $\rho$, the more extraneous/abnormal the vertex will be. But in order for an element to be considered extraneous, the limit of $\rho$ will depend on the level of participation of the hubs in the evaluated network. In networks with strong participation of the hubs, when such hubs report numbers of events close to 90% of the total amount of information on the network, it is normal to expect values of $\rho$ below 1, but to the extent that the participation of hubs decreases, the expected value of $\rho$ increases.

## 4. IDENTIFICATION OF MALICIOUS ACTIVITY IN WIKICRIMES

In general users who jointly make numerous reports in few places are suspicious. When only one user reports numerous crimes in one place, it's easy for a network administrator to identify malicious activity. However, when several people report the same number of crimes, but divide the total number of events among themselves, identifying the problem becomes a difficult task, especially if we imagine that several people are reporting crimes at the same time.

We applied our method for detecting anomalies in WikiCrimes data (for all data from the Fortaleza Network) and verified that the only users who have behavior different from the majority in this network are the hubs. This does not mean that they are anomalous, as we already expected distinct behavior from this type of user because it reports a lot of information in many places.

There are several techniques for classifying anomalous elements in models such as this. As this type of analysis involves two variables (number of places where users reported events and the correlation factor) that are calculated and plotted on a 2D graph, we can (for example) use a method that ranks the points in terms of distance to the other points; our model tends to place points representing anomalous users far from the others($\rho > 1$)), because they report events in a few places and have a high correlation factor. Hubs also tend to be distant from the other users, but this does not represent an anomaly in our model, as we already expected them to behave differently because they report crimes in many places.

Figure 4a shows the plot of each user's correlation factor by the number of places where such users reported crimes in Fortaleza. Since hubs report crimes in several places, they appear at the highest points of the y-axis. The largest squares occur when many values of x,y are repeated, i.e., the more the coordinates are repeated, the larger the square around them will be. The colors vary as a function of the degree of abnormality of each point; this level was calculated using proximity- based anomaly detection techniques [Tan et al., 2006] with k = 10. The color chart ranges from dark gray (Low) to

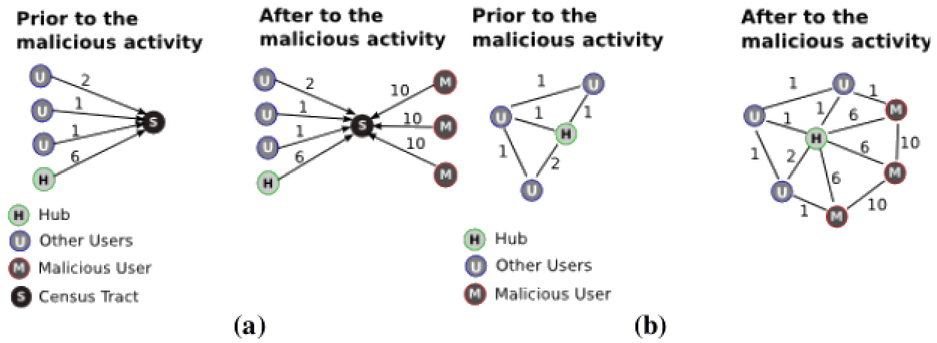
Prior to the malicious activity
After to the malicious activity
Prior to the malicious activity
After to the malicious activity
H Hub
U Other Users
M Malicious User
S Census Tract
H Hub
U Other Users
M Malicious User
(a)
(b)

Fig. 3. In (a) we see two bipartite graphs that represent users reporting crimes in a census tract.There is a hub who has reported six crimes and other users with less reports. Also there is the graph after the report of the three malicious users. In (b) we see the monopartite projections with the relationship between users

light gray (High). Note that only the hubs had a high degree of abnormality.

The final analysis of the data from Fortaleza indicates that at present there is no trend generated from malicious activities, but such analysis does not allow us to assess whether there has been no malicious activity during the four years of recorded data. We decided to carry out a temporal analysis by applying the method at intervals of time. We conducted this analysis by defining pauses in the process of building the network, thus we would be able to detect malicious activity that existed at a particular time, but afterwards was attenuated by the collaboration that continued happening over time. The strategy we used to analyze the WikiCrimes network temporally was to define pauses in time every month, using October 2008 as a starting point. We chose that date to start our analysis because we understand that the network "stabilized" in that month. In this period almost all of the major hubs were formed, and all the properties found in the end network already existed at that point.

The month-by-month analysis of the construction of the monopartite graph of WikiCrimes in Fortaleza allowed us to detect a suspicious relationship between two network users in December 2008. Upon verifying the correlation factor $\rho$ of users in that month, we noticed one of them (the user with id = 911 in the WikiCrimes dataset) with a value of $\rho = 2$, which is odd for the pattern detected by us. This occurred because such user reported two crimes in a census sector in which only one hub had reported only one crime. Additionally, the user with id = 1404 had reported crimes in several places, but two of these crimes were in the same sector in which user 911 reported a crime. This formed a triple between two users (one non-hub and one hub), enabling detection of the anomaly.

At this point, we took the case to a WikiCrimes administrator for assessment, who did not consider it as a matter of malicious activity. This is because user 1404 was a WikiCrimes partner who had just been registered as a Certifier Entity. This also explained why, in the following month, the alert of malicious activity was not reproduced. This Certifier Entity, responsible for creating the suspicious relationship with user 911, reported 121 crimes in 111 sectors, becoming one of the hubs of the network. Thus a triple was not formed between two non-hub users and one hub user, being impossible to calculate for user 911.

The case of malicious activity detected in December 2008 in WikiCrimes proved to be very interesting because, although it was subsequently interpreted as a false positive, at that time it was an anomaly that warranted an alert to the system administrator.From the WikiCrimes' administrator we have heard that the inexistence of abnormalities is probably because malicious activities in WikiCrimes are not so elaborated and done by individuals rather than groups. Also, reporting is in

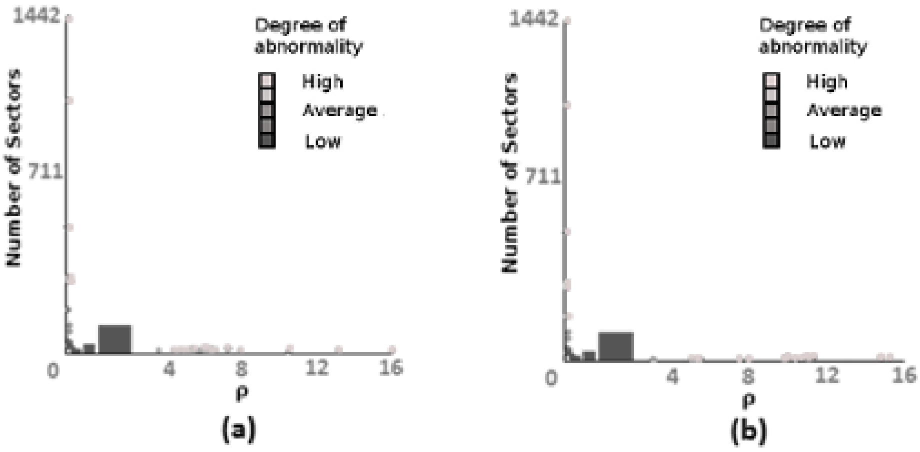

Fig. 4. Plot of the correlation factor of each vertex by the number of sectors they reported as having crime events for (a) Fortaleza and(b) Curitiba

majority dominated by certifier entities that are trusted by the system.

In addition to the city of Fortaleza, we tried the same analyses in other Brazilian cities. Unfortunately, Rio de Janeiro and São Paulo did not have enough crimes to form a monopartite graph of users. Rio de Janeiro has 10,552 census sectors with only 259 crimes distributed within them, which makes the network very disjointed when we performed the projection of the bipartite graph to a monopartite graph. In the case of São Paulo, the monopartite network becomes even more disjointed, because we have 18,285 sectors with only 286 crimes.

The dataset of Curitiba-PR, despite having a much smaller mass of crimes than Fortaleza, has very similar properties in its network structure. Curitiba has 2,387 census sectors with 878 crimes distributed in these locations. The projection in its bipartite graph generates a monopartite graph with 36 users, where its hub has 22 connections. This monopartite graph has a diameter of 3 with a degree of centrality 0.552 and clustering coefficient 0.894. Another noteworthy factor is that, in this city, the hubs also have great coverage, more precisely 85.44%. This assured near total connectivity between hubs and non-hubs of the monopartite network.

We applied in Curitiba the same techniques used to detect malicious activity of groups in Fortaleza. First we evaluated the behavior of users on the complete network, then we carried out a month-by-month temporal analysis starting from October 2010, but no strange or malicious behavior showed up in any of the tests. Figure 4b illustrates the graph with the correlation factor of each user by the number of census sectors where such users reported crimes, this time with data from Curitiba from December 2007 to April 2012.

## 5. SIMULATING MALICIOUS ACTIVITY IN WIKICRIMES

As it is impossible for us to guarantee that the WikiCrimes data are fully produced by non- malicious users (which, incidentally, precluded the application of supervised methods), we decided to evaluate our approach based on a simulation environment where we could represent malicious and non-malicious users.

### 5.1 Methodology

The simulation was executed on the WikiCrimes data already described for the city of Fortaleza. In order to operate in this environment, groups of users that aimed to report crimes were added to the data. The groups considered malicious were those in which the likelihood of a report of crime made by someone in the group strongly depends on previous reports made by users of the same group.We assume that a user deemed as malicious wants to generate a trend in a certain place and therefore, in his choice of reporting a crime, takes into account the reports of crimes and the places that other users in his group have chosen. On the other hand, users who do not wish to deliberately generate a trend, report crimes without any such bias.

In our simulations, each agent has a desire to carry out malicious activity, which varies as a function of the probability $p$ of reporting crimes in places where reports made by members of the group have been made. We assume that potentially malicious agents have a very high value of $p$, whereas non-malicious agents have low value of $p$. Therefore, for each simulation, we created 5 user groups. The number of members of each group varied uniformly between 2 and 5. The number of reports of crimes for each user group also varied uniformly between 5 and 20.We also created two procedures to choose the location where the agent to register a crime in case it wants to generate a trend. In the first procedure, called *random preference*, the agent chooses randomly among the places that have had crimes recorded by someone in its group. In the second, called *preferential attachment* procedure,it follows a strategy of preferential attachment and records preferably in places that had more entries by group members.

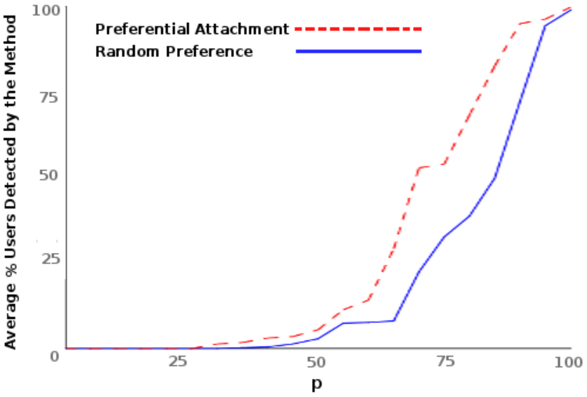

Fig. 5. Relationship between the probability of a user wishes to create a trend per the percentage of those users detected by our method

For each procedure, we ran 10 simulations independently of one another for each value of p ranging from 5% to 100% in intervals of 5%. After each run, we calculated how many members of groups were detected by our method.We expect the detection rate is reduced for low values of p and grows while $p$ increases.

### 5.2 Results

The results are depicted in Figure 5. As expected the rate of detection grows for high values of $p$. When $p$ is equal to 90% the accuracy of the method is greater than 90% for the two types of procedure of reporting. The cases of (potentially) false positives were rare, and the only happens when $p$ was greater than 35%. Even so, in this scenario, only 0.056% (1/177) of the users that made reports were implicated as malicious, but probably were not. Detection is higher when the agents use the preferential attachment procedure. This happens because such a procedure favors the concentration of reports in a place what is easily identified by the method. The results are positive in terms of detecting false positives. Typically, regular users report without considering previous reports of friends. In fact, in WikiCrimes, the regular user does not have in the system the information of who reported where,what makes unlikely to act in this form.

A particular case that warrants separate remarks is that of two agents that reported four crimes each in two census sectors, but no hub had reported information in these sectors. This prevented detection of the activity by our method, because it does not formed a triple between these two users and a hub. This fact is very unlikely, considering that the hubs have a wide coverage as seen in WikiCrimes (we will return to this issue below).

In Figure 6, one can see the behavior of the correlation factor for two of the ten simulations we ran for a specific scenario in which $p$=35% indicating potentially non-malicious users and $p$=95% indicating potentially malicious ones. In these graphs, we can see the plot of the correlation factor $\rho$ by the number of sectors that each user reported crimes in WikiCrimes, but now with the inclusion of the users generated in these simulations. In (a) we see the WikiCrimes data together with the users generated in simulation 1.In this scenario,theu sers considered malicious are shown in light gray and have a value of $\rho$ greater than 1. Only one user that could potentially be considered with a malicious profile was not detected.In (b), the results of another specific simulation are shown, the only one in which false positives occurred. In this simulation, a user from group 4 had $\rho = 3$.

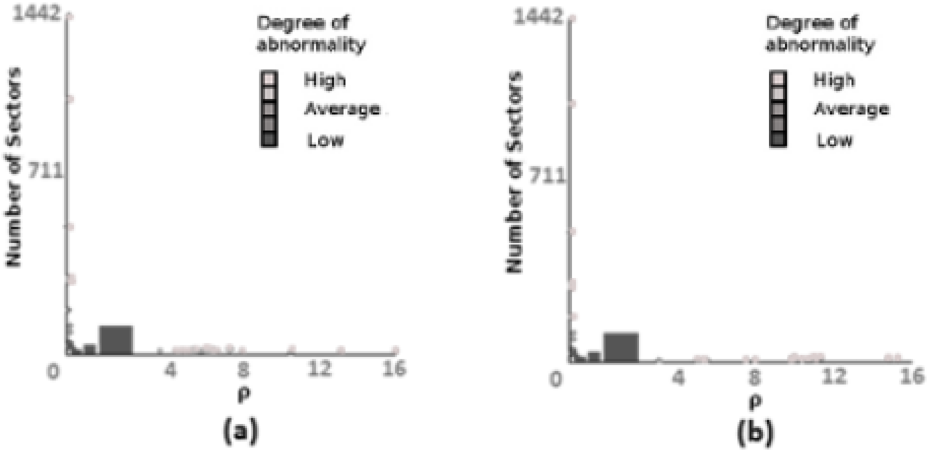

Fig. 6. Density plots with the results for two simulations

### 5.3 Understanding the Scope of the Hubs

As we aforementioned, our approach is sensitive to the scope of the hubs with respect to the number of places they report crimes. When they record in several places more possibilities of forming triples with non-hubs occur.We then decided to evaluate this for other values of exponent. For this same scenario, in which the number of users is 397 and the number of census tracts is 3018, we built a generator of a bi-partite graph inspired by [Guillaume and Latapy, 2006]. Knowing that the distribution of degrees of users who report crimes follows a power law and that the distribution of crimes reported in each census tract also follows such a distribution, we generate graphs from the following procedure:

(1) Generate a distribution with exponent of -2.88 (the same of WikiCrimes) to represent the in-degree distribution of vertices representing census tracts;
(2) Generate distributions with exponent between 1.1 and 3 to represent the out-degree distribution of vertices that represent the users who report crimes;
(3) Connect these two sets of vertices randomly.

In order to understand the impact of the value of the exponent alpha in the coverage of hubs, we varied its value from 1.1 to 3. Figure 7a illustrates the behavior of scenarios obtained varying the number of reported crimes (10,000 to 100,000 reports).

When the *alpha* value increases, the coverage of hubs increases and the proportion of hubs that cover most of the census tracts decreases. In general, when *alpha* is greater than 2, the range of 95% of the census tracts is obtained with only 1% of vertices with higher connectivity (hubs).

Briefly, the high coverage of the hubs of users in collaborative maps is related to the size of the exponent of the power law. It was precisely what WikiCrimes data from Fortaleza and Curitiba showed. A generalization of this assumption does not seem too hard to be induced because that's been evidenced by the literature in the area that participation in collaborative systems follows a power law at scales larger than one.

### 5.4 Understanding the Number of Members in the Group

The simulations performed led us to understand the impact of group size in the calculation of $\rho$ for each of its members. We realize that when we fix the number of crimes of a group and we varied only the amount of membership, while maintaining the probability of registration in the same locations in

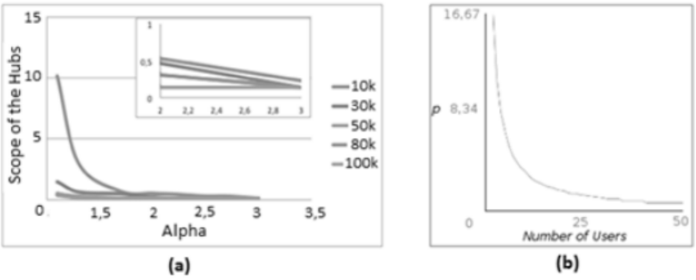

Fig. 7. (a) Correlation between alpha of the out-degree distribution of reports of the user and percentage of hubs that cover 95% of the census tracts; (b) Number of users by the correlation factor ($\rho$), where $\rho$ is calculated for a simulation of a false trend of 200 crimes in WikiCrimes

95%, the method has a sharp drop in their level of accuracy. What happens is that the larger the group the more difficult it is to detect anomalous correlation factors because as each user of this group reports a few crimes, relations with hubs will be stronger than with the group members themselves.

If we define that 200 crimes have to be inserted in a tract by only 2 users (100 each) the correlation factor would be $\rho$=16.67. Slightly increasing the community of users that would equally report all 200 crimes, we arrived at a total of 34 people, where each user would report an average of 6 crimes, implying $\rho$=1 for such users, which would place them near other users who do not generate a false trend on the network. We can see that with these three scenarios, as more people work together in order to generate a false trend in this type of network, the more the value of $\rho$ for such users tends to drop, and consequently the more difficult it is to identify the anomalous behavior. In Figure 7a, we illustrate the behavior of $\rho$ when more people divide among themselves the 200 crimes of the previous scenarios. The value of 1 is obtained with 34 people.

The limit of 34 people for the previous scenario, illustrated in Figure 7b, may actually be much higher, because the division of crimes equally among users is unreal. It is more common to imagine the division of crimes among users following a normal distribution. Assuming this, we have a much higher limit of users for the malicious activity of 200 crimes, because some users would probably report an above-average number of crimes and therefore would be identified by our method.

## 6. TOWARDS A GENERALIZATION OF THE APPROACH

In order to study the generality of the method proposed for WikiCrimes we decided to investigate how it could be applied in other domain. Our intent here is solely to show that our approach has wide applicability and to support some of the assumptions we made in our simulations. Assessment as for the accuracy of the method will be the subject of future work up because they require unavailable semantic information about the domain.

We model the data from the Amazon product reviews of 1999[1] as a bipartite graph. The vertices represent reviewers and products, which are in turn connected by revisions. The weight of the edge is a product between the number of revisions and the evaluation, ranging from 1 to 5, given in the review.

We concentrate on the giant component that owns 50% of the vertices of the complete graph. This component has 57,561 reviewers and 43,717 products. Altogether there are 118,283 edges representing revisions made. The second largest giant component has only 51 vertices. A malicious activity in this area and recently explored by [Ott et al., 2012] is the appearance of groups of reviewers who together make positive reviews that may influence the evaluation of a certain product.

[1] Available in `http://www.cs.uic.edu/~liub/FBS/fake-reviews.html`

The in-degree distribution and the ratio of out-degree of the graph, as in WikiCrimes, follows a Power Law with exponent -2.08 and -2.51 respectively. In terms of range, we found that 10% of hubs review 71% of products, which means a connection (in the monopartite graph) with 87% of regular users. Figure 2c shows the behavior of hubs in relation to the connectivity to other reviewers. One can see that, although not as pronounced as in WikiCrimes, the relationship of a few hubs connecting with many non-hubs users also happens here. If one want to work with 100% coverage of non-hubs users should choose the 34% of the first hubs.

We apply our method to the data by selecting 1.5% of the largest hubs as we did in WikiCrimes. This meant that 77% of regular users connect to the selected hubs. Figure 8 shows the behavior of the correlation factor after the application of the anomaly detection method based on a distance (k = 200).

In this scenario 254 reviewers (0.3% of the reviewers) were indicated as anomalous (in light gray). As in WikiCrimes, hubs some considered anomalous (26 out of 864) because have made a large number of reviews. On the other hand regular users with high $\rho$ (228 reviewers that have $\rho>6$) are also considered anomalous due to their relationship with hubs. A deep analysis of these anomalies is beyond the scope of this work, but some cases deserve mention. We have seen that there are groups of reviewers working together in the same products and always providing good ratings (from 4 to 5). A large group of 17 people acted together in nine products from March until December 1999, all members of the group had $7 \leq \rho \leq 9$ by the strong relationship between themselves and weak relationship with the hubs. This weak relationship with the hubs took the fact that three out of the nine hubs evaluated some products, but always with very low grades (rating from 1 to 3). Another group detected that deserves mention was one of only two reviewers. They evaluated only one product, but three times each, always rating at the maximum (5 stars). A hub only rated this product giving it one star. This led the two users to receive a $\rho = 15$.

## 7. RELATED WORK

Wikis, in general, are based on the concept of radical trust; i.e., it is believed that individual participation, for the most part, includes correct information. Nevertheless, the identification of attempted fraud or vandalism is necessary. The challenge imposed on WikiCrimes and collaborative maps in general is to assure the credibility of the information recorded on the map, and requires the study of different approaches.

One approach to minimizing the problem is to assign a value to the credibility of the information

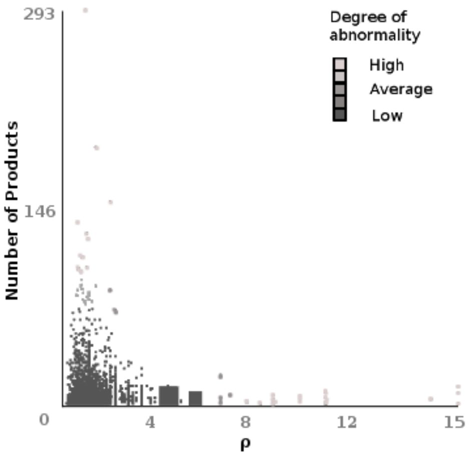

Fig. 8. Density plot indicating anomalous users with $\rho$ greater than 6

based on a model of user reputation and trust. The analysis of the users' social network, for example, was one of the ways we used to achieve this [Furtado et al., 2010] and [Wu et al., 2006]. The reputation of a user depends on the reputations of his/her friends. Propagation occurs through the social network. Reputation models, however, lack the level of granularity to capture malicious activities such as generation of a false trend that can come about with an excess of false reports made by various people because may not exist any explicit friendship connection between a group of malicious users. Identifying evidence of these problems through data mining is another recommended approach.

Furtado and colleagues [Furtado et al., 2009a] have developed an algorithm that – based on reported events – tries to identify patterns that indicate excesses or abuses coming from an individual or group of individuals. The basic idea is to identify the existence of communities [Appel and Junior, 2011] on the social network and verify if there is one such community that dominates the reporting of events (reported for a hot spot, in particular). In order to do this, they used algorithms for identifying communities developed in the context of social network analysis. In addition to considering the structure of the community in terms of connection density, it was necessary to consider the participation of users (represented in the nodes) in the formation of hot spots. Since the formation of hot spots varies with the zoom level, the authors proposed identifying communities for each one of the social networks related to the hot spots at each zoom level.

Although the first results obtained with this approach were satisfactory, they soon showed that a high number of false positives could occur as a greater participation in the system starts to occur. Groups of users can report information and generate false trends in a region without forming a community among themselves. Assuming that most of the reports are true, one must try to find this type of anomaly, if any.

In anomaly detection, the goal is to find objects having behavior that is very different or extraneous in relation to others. In the context of WikiCrimes, these objects can be users, whereby the intent is to extract characteristics of normal behavior thereof, then identify extraneous elements [Breunig et al., 2000].The task of detecting anomalies brings many challenges, as it may be necessary to use several or just one attribute to detect one of these elements. We have showed that such methods can perfectly be used in conjoint with our method by exploring the relationship between hubs and non-hubs as we have highlighted.

Machine learning is complementary approach rather than an alternative to our proposal. Supervised and semi-supervised approaches cannot be applied here because there is no data with classification. Methods for discovering frequent substructures such as [Yan and Han, 2007] and [Zaki, 2004] would help in generating candidates for analyses, for instance, users that frequently report conjointly. However, we are proposing a method based on the frequency of co-occurrence of structures but also in the relationship between "few with many" identified on data.

Similar work has been seen in the context of malicious activity in mining consumer reviews on the web [Mukherjee et al., 2012], [Ott et al., 2012]. The similarities relate to the fact that some of these activities are caused by user groups seeking to establish trends (positive or negative) for certain products and thus influence the opinion of buyers. The approaches applied to solve this problem assume that you can apply supervised learning algorithms to discover a pattern that indicate malicious activity, based on the characteristics of the group of reviewers and reviews. [Mukherjee et al., 2012] for example, made use of experts to assess the review of certain groups and say which of them could be declared as malicious. Try something similar in the context of reports of crimes would be very hard due to the difficulty of characterizing a group and, especially, what comes to be a false report.

Regardless of the technique chosen to detect extraneous elements, such technique will always require the selection of variables, attributes or classes as input for the use thereof. One of the main contributions of this study is to propose a way to represent the collaboration of crime reports in places such as a complex network, by modeling the relations between users in such a way as to make it possible to use divers methods to separate potentially malevolent users from users with desirable behavior.

## 8. CONCLUSION

Our scientific research has sought to represent the participation in crowdmaps through a complex network. By doing so, we can better understand the patterns that form based on the relationships between users who report events on the maps. Thus, we were able to formally measure the relationships between users and identify patterns of the behavior thereof within the network. The hubs of this network are the key to detecting anomalies. This stems from the fact that the hubs are usually entities with a high reputation and very often refer to government agencies. These are users who participate actively in the reporting of events and do so in several places, exposing a clear pattern of relationship between them and other users.

In particular, by following this approach, we show that with a very simple method without any knowledge about the domain we were able to detect a type of malicious activity that is difficult to identify in the context of crowdmaps. We were able to identify that relationships between non-hub users among themselves are typically no stronger than relationships between non-hub users and the hubs. When this occurs, the possibility of malicious activity becomes strongly evident. We formalized how this can be evidenced by means of a measurement of correlation between non-hub users and hubs. This measure can be input for machine learning and data mining algorithms becoming complementary to them.

From simulations we show that agents with low are a good characterization of the users who have no interest in carrying out malicious activity. We believe that p cannot be too low because there is not totally discarded a coincidence of reporting between persons of the same group in a certain place. Not so in the interest of forcing a trend, but that they may have knowledge of common crimes to a certain region (near their homes or workplaces). It would be possible to estimate this value? As future work, we plan to investigate WikiCrimes data in order to define communities and determine what percentage of users in a community recorded crimes in the same places.

Specifically in the case of WikiCrimes, it was found that the hubs were mostly trusted users. We also saw that they report the vast majority of crimes in various locations, which makes them relate to other network users. What are the validity of our method and the behavior of $\rho$ when users are not trusted hubs? We believe that if you can identify a few trusted users with high scope, our method can be applied. Characterization of data in the context of data about review sindicated a different threshold for $\rho$ compared with that identified in WikiCrimes. Future investigations will seek to formalize the intuition that less participatory trusted users would increase the values of the threshold of $\rho$ in this type of network as also identified in Amazon data.

## REFERENCES

Appel, A. P. and Junior, E. R. H. (2011). Centaurs - a Component Based Framework to Mine Large Graphs. *JIDM*, (2)1: 19–26.

Breunig, M., Kriegel, H., Ng, R., Sander, J., et al. (2000). Lof: identifying density-based local outliers. *Sigmod Record*, 29(2):93–104.

Caminha, C., Furtado, V., Vasconcelos, E., and Ayres, L. (2010). Uma ferramenta de autoria para criação de mapas colaborativos para aplicações em egov 2.0. In *Anais do XXX Congresso da Sociedade Brasileira de Computação*, Belo Horizonte, Brazil.

Cançado, T. (2005). Alocação e despacho de recursos para combate à criminalidade. Master's thesis, Dissertação de mestrado, UFMG, Belo Horizonte.

Furtado, V., Assunção, T., de Oliveira, M., Belchior, M., and D'Orleans, J. (2009a). A method for identifying malicious activity in collaborative systems with maps. In *International Conference on Advances in Social Network Analysis and Mining*, pages 334–337, Athens, Greece.

Furtado, V., Ayres, L., de Oliveira, M., Vasconcelos, E., Caminha, C., D'Orleans, J., and Belchior, M. (2010). Collective intelligence in law enforcement–the wikicrimes system. *Information Sciences*, 180(1):4–17.

Furtado, V., Melo, A., Coelho, A., Menezes, R., and Perrone, R. (2009b). A bio-inspired crime simulation model. *Decision Support Systems*, 48(1):282–292.

Guillaume, J. and Latapy, M. (2006). Bipartite graphs as models of complex networks. *Physica A: Statistical Mechanics and its Applications*, 371(2):795–813.

Mollenkopf, J., Goldsmith, V., McGuire, P., and Sara, M. (2000). Identification, development and implementation of innovative crime mapping techniques and spatial analysis. *US Department of Justice*, page 27.

Mukherjee, A., Liu, B., and Glance, N. (2012). Spotting fake reviewer groups in consumer reviews. In *Proceedings of the 21st international conference on World Wide Web*, pages 191–200, New York, NY, USA.

Ott, M., Cardie, C., and Hancock, J. (2012). Estimating the prevalence of deception in online review communities. In *Proceedings of the 21st international conference on World Wide Web*, pages 201–210, New York, NY, USA.

Ott, M., Choi, Y., Cardie, C., and Hancock, J. T. (2011). Finding deceptive opinion spam by any stretch of the imagination. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, volume 1, pages 309–319, Stroudsburg, PA, USA.

Rouse, L., Bergeron, S., and Harris, T. (2007). Participating in the geospatial web: collaborative mapping, social networks and participatory gis. *The Geospatial Web*, pages 153–158.

Tan, P., Steinbach, M., Kumar, V., et al. (2006). *Introduction to data mining*. Pearson Addison Wesley Boston.

Whittaker, S., Terveen, L., Hill, W., and Cherny, L. (1998). The dynamics of mass interaction. In *Proceedings of the 1998 ACM conference on Computer supported cooperative work*, pages 257–264, New York, NY, USA.

Wu, B., Goel, V., and Davison, B. (2006). Propagating trust and distrust to demote web spam. In *Workshop on Models of Trust for the Web*, volume 190, Edinburgh, Scotland.

Yan, X. and Han, J. (2007). *Mining graph data*, chapter Discovery of frequent substructures, pages 97–115.

Zaki, M. J. (2004). Efficiently mining frequent embedded unordered trees. *Fundam. Inf.*, 66(1-2):33–52.