Collective intelligence in law enforcement – The WikiCrimes system

Vasco Furtado\textsuperscript{a,d,*}, Leonardo Ayres\textsuperscript{a}, Marcos de Oliveira\textsuperscript{c}, Eurico Vasconcelos\textsuperscript{b}, Carlos Caminha\textsuperscript{a}, Johnatas D’Orleans\textsuperscript{a}, Mairon Belchior\textsuperscript{a}

\textsuperscript{a}Universidade de Fortaleza – UNIFOR, Av. Washington Soares, 1321, Fortaleza-CE, Brazil
\textsuperscript{b}Pontifícia Universidade Católica – PUC-Rio, Rua Marquês de São Vicente, 225, Gávea, Rio de Janeiro-RJ, Brazil
\textsuperscript{c}University of Otago, P.O. Box 56, Dunedin 9054, New Zealand
\textsuperscript{d}ETICE – Empresa de Tecnologia do Ceará, Av. Pontes Vieira 220, Fortaleza-CE, Brazil

1. Introduction

The general public is making ever greater use of collaborative systems. This momentum of collaboration is primordially leveraged by the Web 2.0\textsuperscript{[18]}, in which the difference between information producers and consumers decreases significantly, since several applications emphasize the production of information by any Internet user. A culture of sharing induces people to externalize their feelings, opinions, experiences and even their goods. With the mass production of content, collaboration mechanisms such as Wikipedia (http://www.wikipedia.org) came almost naturally.

Despite the huge success of Wikipedia, finding good and useful causes that are capable of involving thousands or even millions of people operating with little or no coordination is still a big challenge. A slight variation of the wiki concept illustrated by Wikipedia has been dubbed crowd sourcing\textsuperscript{[10]}. In that context, mass collaboration occurs via human resource outsourcing based on awards (often financial) or even punishments that “motivate” people’s participation in large campaigns. In the context of public and government areas – where social problems would be a potential target to be treated by means of mass collaboration – crowd sourcing is particularly difficult, since few investments are available to leverage people’s participation.

\* Corresponding author. Address: Universidade de Fortaleza – UNIFOR, Av. Washington Soares, 1321, Fortaleza-CE, Brazil. E-mail address: vasco@unifor.br (V. Furtado).

0020-0255/$ - see front matter © 2009 Published by Elsevier Inc.
doi:10.1016/j.ins.2009.08.004

Please cite this article in press as: V. Furtado et al., Collective intelligence in law enforcement – The WikiCrimes system, Inform. Sci. (2009), doi:10.1016/j.ins.2009.08.004
In the Law Enforcement domain, besides the aforementioned difficulties, another particular challenge must be faced. In that context, finding the equilibrium between people’s participation and information credibility is crucial. Anonymous mass collaboration is the easiest way to receive information; however, the credibility of the received information is depreciated, because the source of information is unknown.

It is in that particularly complex domain that we have initiated a project called WikiCrimes (http://www.wikicrimes.org), which is driven by three goals: (i) to give more transparency and publicity to criminal information, (ii) to provide means for citizen crime prevention, and (iii) to reduce the phenomenon of crime under-reporting (crimes that are not notified to authorities). These goals have been on the political agenda of several countries around the world, particularly those in which the population suffers with high rates of violence.

WikiCrimes aims to offer a common interaction space among the public in general, so that they are able to notify criminal facts as well as keep track of the locations where they occur. The goal is to obtain collaborative individual participation for generating useful information for everyone. It is based on the principle that the ones who hold information about crimes are the citizens. If they want to make such information public, they can. Here, we are appealing to the feeling of sharing that is so usual in victims of violence. When someone is a victim of any type of crime, it is usual to tell someone about the fact. Typically, those who are close to the victim are the first to be informed. What we intend with WikiCrimes is to provide an environment to be that “global blackboard” of stories told by people about crimes, in order to help alert other people on a scale larger than their closest social contacts. In other words, if there is active participation, crime mapping starts to be done collaboratively, and everyone will benefit from having access to information about where crimes occur.

In this article we describe the main goals, motivations and challenges we face during the development and implementation of WikiCrimes. We particularly concentrate on the ways we have created to keep the “participation vs. reliability of information” trade-off at good level. We also discuss and analyze our experiences and the obstacles that we face.

2. Collaborative systems

We understand collaborative systems as computer tools that support the interaction (direct or indirect) among a certain number of people with individual and/or collective goals. Research into collaborative systems is divided into several areas and assumes different yet inter-related terminologies such as Groupware, Computer-Mediated-Communication (CMC), Computer-Support Cooperative Work and Learning (CSCW/CSCL) [7]. “Computer-Mediated Communication” can be seen as the vast area that encompasses computer-mediated communication research, including collaborative systems. “Groupware” refers to technology applied to computers and networks designed to enable work in groups. This technology can be used to communicate, cooperate, coordinate, troubleshoot, compete and or negotiate. Examples of groupware technologies are e-mail, newsgroups, wikis, video conferences and chats [9]. CSCW and CSCL refer to the areas of study that will examine the design, adoption and use of groupware tools in the context of working and learning.

All these kinds of collaborative systems have been strongly influenced by the recent trends of the Web. The Web can be seen in terms of two distinct phases, and one way to distinguish these phases is to look at the number of producers and consumers of information. In the traditional Web (or Web 1.0), there are few producers and many consumers. The main producers are specialized persons (web designers), companies (both public and private), and traditional media. The typical Internet user only accesses the information made available by these producers. In Web 2.0, the difference between producer and consumer of information decreases significantly, since several applications emphasize the production of information by any Internet user. The recent wave of blogs, photologs, wikis, online communities, social networks, etc. is an example of this context. With the mass production of content, cooperation mechanisms such as Wikipedia came almost naturally. Another type of application that is beginning to emerge is aimed at the integration of information from different sources (“mashup”).

3. WikiCrimes: motivation and goals

The veracity and accuracy of information about where crimes occur, as well as the information on the characterization of such crimes, has always been on the agenda of discussions about Public Safety in Brazil and in various other countries. Traditionally, this information is monopolized by law enforcement agencies and is therefore characterized as a highly centralized mechanism. This monopoly ultimately creates tension in the relationship between such agencies and society at large, because it is commonly opposed to the precept of disclosure and transparency of information required by a democratic regime. Allied to this context are the crises that have characterized the daily routines of law enforcement agencies as well as their limitations to provide a quality public service, which tend to diminish citizens’ trust in those agencies. These factors encompass some of the reasons for the growing problem of under-reporting – the low rate of reporting crimes – that has occurred [12]. It has become common for one to hear someone who has been mugged say that they did not file a police report because they thought it would not bring about any effect. Polls conducted with crime victims in several Brazilian states show that under-reporting may, in densely-populated areas, reach 60% for certain types of crimes. The result of this can be disastrous in terms of formulation of public policies and particularly in the planning of police actions, since the official crime mapping may be reflecting a trend that is somewhat different than what is actually occurring in real life.

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. It is based on the principle that the ones who hold information about crimes are
the citizens. If they want to make such information public, they can. Thus, individual participation, in a collaborative manner, can generate collective intelligence. In other words, if there is active participation, crime mapping starts being done collaboratively, and everyone will benefit from having access to information about where crimes occur. WikiCrimes is a typical Web 2.0 application. It allows users to access and to register criminal events on the computer directly in a specific geographic location represented by a map. That is one of the phases of the crime mapping activity, in which an analyst collects, maps, visualizes, and analyzes crime incident patterns.

4. Main architecture

The open and participatory characteristics identified in WikiCrimes makes it susceptible to abuses or fraud attempts. It is important that as many people as possible collaborate with the system, contributing to the growth of its data records. But it is equally important that the information registered in the system be reliable, so that the system can become a trustworthy source of information. In saying this, our claim is that an open collaborative system must be viewed as a kind of open multi-agent system, where a number of human and/or artificial agents interoperate pursuing their individual or common goals. These agents can contribute positively or negatively to the organization and goals of the system as a whole, and the agents are free to join and leave the system as they wish, as long as they obey certain rules that must be observed concerning the management of the system.

The management of interoperation among agents is a complex task, and robust techniques and methodologies for the development of reliable and open Multi-Agent Systems (MAS) have been studied in academia [1,2,5,8,14,16,20,24]. Such techniques are aimed at the modeling and implementation of features that give openness to those agents, allowing them to have the ultimate choice of obeying regulations or dealing with possible sanctions imposed by the MAS norms. After all, agents are autonomous entities and the biggest challenge is to have a coordination system where the agents can be free to decide what to do, but at the same time be encouraged or enticed to obey the regulations of the artificial society they are entering.

4.1. WikiCrimes as a multi-agent system

We have therefore modeled WikiCrimes as an Open Collaborative System [17], which is comprised of the following components:

WikiCrimes = \{A_\text{sys}, A_\text{context}, R_\text{ext}, O_\text{domain}, O_\text{system}, O_\text{reputation}, P\_\text{protocols}, N\}.

where, \( A_\text{sys} \) agents playing system roles; \( A_\text{context} \) agents playing context roles; \( R_\text{ext} \) a set of external roles that are played by agents that interact with the system; \( O_\text{domain} \) domain ontology (here identified as crime ontology); \( O_\text{system} \) system ontology; \( O_\text{reputation} \) reputation ontology; \( P\_\text{protocols} \) Set of interaction protocols; \( N \) set of norms represented as commitments among agent roles listed in \( R_\text{ext} \).

Agents playing system roles (\( A_\text{sys} \)) are the Institutional Agent, the Monitor Agent and the Reputation Agent. The Institutional Agent is responsible for registration of agents playing external roles. This agent manages a workflow of the system, which defines the participation of every agent that is part of the system in the whole interactive process. The Monitor Agent is responsible for punishing or rewarding agents that break or follow the norms of the system, respectively. This agent manages the life cycle of commitments that represent the system norms defined in \( N \).

The Reputation Agent is responsible for keeping updated a database of reputation owned by the agents registered in the system. The reputation is calculated based on the agent interactions and acceptance of the system norms defined in \( N \).

Agents playing context roles (\( A_\text{context} \)) are the referral agents responsible for keeping a Social Network formed by agents that play the roles of Registered User, Invited User and Certifier Entity. The social network is formed when agents playing the Registered User and Certifier Entity roles register a crime and indicate or refer another agent. The System Administrator and the Database Administrator Agents are responsible for managing administrative tasks and keeping the criminal facts updated in the database, respectively.

The external roles played by the agents that interact with the system (\( R_\text{ext} \)) are basically the role of registered user that represents a typical user of WikiCrimes, who is able to register, confirm and disconfirm crimes, browse the environment, denounce abuse, leave messages for the Institutional Agent about the system, and indicate other agents to confirm crimes. The Invited User Role is played by the agent that is indicated to confirm a crime. The Certifier Entity Role is a special kind of role played by agents that have a respected position in the system. Agents playing this role have the reputation of a System Agent, since they represent institutions with high credibility as sources of crime information. The Browser User Role is played by users that only browse the institutional environment, basically seeking information in the system. The Institution Agent will try to encourage this kind of user to assume the role of Registered User.

The main interaction protocols (\( P\_\text{protocols} \)) are registration, deregistration, assuming a role in the system, changing a role in the system, breaking a commitment, fulfilling a commitment, registering a criminal fact, indicating another agent to confirm a crime, confirming a crime positively, confirming a crime negatively, denouncing abuse, commenting on a registered criminal fact, leaving a comment for the system, registering an observation area, deregistering an observation area, sending alerts, receiving alerts, asking for a safe route, querying the criminal database, and querying the reputation data base. The set of norms \( N \) represents the following commitments between the agent roles listed in \( R_\text{ext} \):

Please cite this article in press as: V. Furtado et al., Collective intelligence in law enforcement – The WikiCrimes system, Inform. Sci. (2009), doi:10.1016/j.ins.2009.08.004
Registered Users accept registering criminal facts that will have positive confirmations and that won’t be denounced as abuse.

Invited Users will register with the system and confirm positively or negatively criminal facts associated with the invitation.

Certifier Entities will exercise dominance over the registration of criminal facts in specific geographic areas.

The commitments are represented by objects that have a specific life cycle and will influence the rewarding or punishment of reputation points for the agents that fulfill or break (respectively) such commitments.

4.2. Basic ontologies

Two ontologies are represented in WikiCrimes for the representation of crimes and reputations. The crime ontology is used by most WikiCrimes agents and is depicted in Fig. 1. Basically it represents two concepts: crime and report of crime. A crime has a type, a time, an address, and type of weapon used. The report of crime has information about the reporting itself, such as whether or not the fact was communicated to the police.

The reputation ontology was inspired by the work of [19]. However, we specialized some of the concepts to cope with the particular features of WikiCrimes. Fig. 2 depicts the classes we have defined in WikiCrimes (isa relationship) and some instances in that context (the io relationship). Entity is any element of the system that is susceptible of being evaluated or having an active part in the generation or diffusion of evaluations. Entities being evaluated play the social role of Target, in WikiCrimes represented by all the agents playing the Re (External Roles). Entities evaluating or generating evaluation are represented by the Source class. In WikiCrimes all the agents playing As (System Roles) are seen as Source. The Institutional Agent and Monitor Agent are sources for the computation of the Acquired Reputation-AR (see Section 5 for details). The Institutional Agent, Referral Agent, and the System Administrator Agent are sources for the computation of the Propagated Reputation-PR (also, see Section 5 for further details). The Reputation Agent is source for the computation of the combined reputation of AR and PR. Agents that spread the evaluations play the role of Gossiper entities and are represented in WikiCrimes by the Reputation Agent. Agents that receive the evaluation and observe it for further interactions in the system are represented by the agents that play Er (External Roles). The Focus class represents the context of the evaluation, the goodness or badness of some Entity’s Evaluation is towards to a specific norm, standard or skill [19], see Fig. 2. The class Value represents a value that express how good or bad is the target entity in the context represented by the Focus element and is generated by the Source entity. In WikiCrimes the norms represent a set of rules that must be obeyed by the agents while interacting in the system. The value of the reputation in WikiCrimes is represented by a real number belonging to the

![Fig. 1. Crime ontology in WikiCrimes.](image-url)
interval $[0,1]$, being 0 completely bad and 1 completely good. The Strength, represented by a real number belonging to the interval $[0,1]$, is a subjective measure set by the Source that indicates how reliable the evaluation is, 1 being the maximum reliability. This is an attempt to make the evaluations more accurate by including the element of reliability in the calculations made by the Source. In WikiCrimes, the Strength it is identified when the reputation of agents – who are confirming crimes and denouncing abuse – influences the calculation of the Target’s AR. The class Voice includes the necessary elements to represent the spreading of an evaluation; it is defined as a report on reputation. For example, “it was told that Agent X has reputation 0.81 in WikiCrimes.” Therefore the class has, as elements: the evaluation itself; the teller of the evaluation (i.e., the Gossiper); and the receiver of the evaluation (i.e., the Recipient).

4.3. General workflow

We have modeled the basic states of each one of the roles played by the agents, as well as the activities that represent the transition between two states by means of a colored Petri net [11]. Fig. 3 depicts the general flow. Basically, it describes the three roles that an agent can play: the Browser User, the Invited User and the Registered User. Both a Browser User and an
Invited User Agent, after registering in the system, can change to the role of Registered User that, in turn, can register and confirm crimes as well as denounce abuses. Another activity that brings about a change of state is the Registered User’s assuming the role of Certifier User.

4.4. Interacting with WikiCrimes

The WikiCrimes interface was based on the eXtreme Communication-Centered Design [3,4] of Semiotic Engineering [22]. The idea of semiotic engineering draws on concepts from semiotics and computer science to investigate the relationship between user and designer. Semiotics is the study of signs, and the essence of semiotic engineering is the communication between designers and users at interaction time; designers must tell the users, via the interface, how to use the signs that make up a system or program. This approach, which builds on – but goes further than – the currently dominant user-centered approach, allows designers to communicate their overall vision and therefore helps users understand designs, rather than simply which icon to click. Doing so, we had the goal of improving designer-user communication and therefore avoiding misunderstood communication and “mistaken” actions on the part of users.

The reflection of the application of this theory on the design of WikiCrimes was a clearer and more direct communication (designer-user), via interface, of how users could identify and meet their goals in the system (registering and commenting on crimes and denouncements, researching areas and locations, "knowing where it is not safe," and obtaining information about crime and about the system). This communication was carried out by conceptually highlighting and clustering the features necessary to meet the users’ main goals in a way as to permit the easy identification and execution thereof. Another important reflection of application of this theory is the communication of the designer to the user in the most direct manner, implemented in the system’s help function. In WikiCrimes, this help was developed on three levels:

- Contextual help, offered through “hints” that are displayed on the interface based on the position of the mouse pointer over certain items of the interface and areas of the system;
- Textual help, or system help, is offered based on a set of texts associated with images, providing information on “what it is,” “how to do it” and “where to find it” regarding the interface items, features, concepts and areas of the system. This help is indexed alphabetically and by topic, and is located in the upper right corner of the interface;
- Online help via chat. This help is aimed at serving users who have not found answers to their questions through the other forms of help offered, or who prefer direct interaction with the WikiCrimes team.

Fig. 4. Overview of WikiCrimes main interface with the map of homicides in Fortaleza, with pin points and texts describing its main areas.
Fig. 4 depicts the main screen of WikiCrimes, which offers a crime search function (right side of the Figure) that enables users to view the registers of crimes, filtered by crime type, date range and time range. The search can also be done by means of a direct search on the map by querying and viewing a location on the map. To post a criminal occurrence, the user must be registered and logged in to the system. Once logged in, the user must locate the exact address where the criminal fact occurred. Once the crime location is visible on the map, the selection of the type of crime can be done. The system currently permits the user to post the following types of crime:

- Robbery: robbery against the person, robbery against property, attempted robbery against the person and attempted robbery against property;
- Theft: theft against the person, theft against property, attempted theft against the person and attempted theft against property;
- Violent Crime: homicide, attempted homicide, robbery followed by murder (aggravated robbery involving death of the victim), brawls or fights, domestic violence, abuse of authority.

To obtain information about a crime, suffice it to click on the respective marker of the crime. Note that the complete description of the crime is accessible only for those who are logged in. In Fig. 4, the map of crimes is also shown by means of markers. There, the registers of homicides that occurred in Fortaleza from January to September, 2008 are depicted. Fig. 5 shows another way to visualize the same records, in which the map of crimes is depicted by means of hot spots. An algorithm for computing the map of Kernel [15] allows the identification of dangerous areas based on clusters represented by the intensity of a color in a given region.

5. Credibility of the information

In wiki-style open systems, there must be mechanisms for rewarding and/or punishing agents involved in the interoperation process. In these kinds of systems, there is a constant trade-off between diminishing the constraints imposed to the agents 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 report of false information. One of the main concerns in WikiCrimes is the number of hoaxes or, in our case, report of false criminal facts. How to prevent a person from spreading false information in the system? Why does that specific area have so many crimes? Is it really true or is someone trying to make a joke, to speculate about the real estate market, or to diminish the image of the local police?

Please cite this article in press as: V. Furtado et al., Collective intelligence in law enforcement – The WikiCrimes system, Inform. Sci. (2009), doi:10.1016/j.ins.2009.08.004
5.1. Reputation model

In WikiCrimes, there are not many requisites to become a member of the system. The only personal questions asked are a name and valid e-mail address; no document identification is required so that people will not be afraid to post information in the system. It is up to the user to provide the system with information that increases the credibility of the user reporting a crime. It is possible to add links to videos, newspapers, photos or any other document (such as a police report) that helps the users to believe in the informant making the posted information more credible. Moreover, for every criminal fact registered in the system, it is requested that there be an indication of at least one person who can confirm that the information posted is true; this is to increase the credibility of the information, possibly making the system as a whole more reliable. The more the information is confirmed, the more it is considered trustworthy. These indications for the confirmation of information generate a graph, where the vertices represent users of WikiCrimes and the edges represent the indication of others to confirm the criminal fact registered. The graph represents a Social Network formed by the WikiCrimes users. Doing so, we are able to create a reputation model that plays a strong role in the matter of identifying “bad agents” in the institutional environment, and the social net of WikiCrimes users is a fundamental tool for that. By "reputation," we mean a score that represents the community's view about a member of that community. Trust is a score that reflects the subjective view of an entity from another. Trust can be calculated upon the acknowledgement of a given reputation.

In WikiCrimes, reputation is built upon interaction of the agents, by the observation of breaking the commitments and consequently the norms of the system. We have thus defined the concept of reputation for the source of information; i.e., whoever registers the criminal fact. In other words, in WikiCrimes we have a social network formed by the users that are registered in the system, forming the Social Network Layer, and the information that is posted in the system – the criminal facts. The information posted in the system forms the Information Layer. The goal is to build a function that will calculate the reputation of the users and reflect that reputation in the trustworthiness of the information that is posted in the Information Layer. Some entities, such as the press and governmental agencies, are labeled as certifier entities, and therefore considered very well reputed. But that is not enough; the open characteristic of the system, in the sense that anyone could be a user, does not facilitate the task of knowing the reputation of all the users of the system. The attribution of reputation to the users who are not qualified as certifier entities is fundamental. We can assume that an agent delegates a commitment to another agent when the former indicates the latter to give an opinion about the a particular report. The act of acceptance of the commitment by the agent indicates a relationship of trust between the agents. If the commitment is broken by the second agent, it will be penalized in reputation points. This relationship suggests the adoption of mechanisms to propagate trust in the social network built in WikiCrimes. The Administrator and Certifier Entity roles have a very good reputation to start with. These agents serve as a starting point for the propagation of trust to the agents indicated by them, then to the ones indicated by those indicated, and so on, similar to the propagation of trust and distrust for demotion of web spam described in [24].

Basically the main steps of this process are the following:

1. Attribute initial reputation scores to the users (Propagation of Trust). A seed set of reputable users is selected and, from these users, the trust that the System has in them will be propagated to form a reputation of all in the system.
2. Once the initial reputation is calculated, the trustworthiness of the information posted by the users will be calculated, reflecting from the Social Layer to the Information Layer. This will give an idea of how credible the current posted information actually is.
3. In considering WikiCrimes as a Normative and Open MAS, the reputation and trust will be updated based on the interactions identified in the MAS, namely posting a crime, confirming a crime positively, confirming a crime negatively, and denouncing abuse. Observing these interactions, there is a social norm that is difficult to maintain: the user commits itself to register reliable information in the system. Agent interaction with the Information Layer will help to identify the agents that are breaking this social norm. These interactions will indicate how much trust the users of the system have in that information.

5.2. Acquired reputation

Two types of scores are considered in order to calculate the reputation based on the interaction activities performed by the user and the social evaluation of the peers. The interaction activities that count positively are those related to providing additional information in order to give credibility to report of crimes, such as posting videos, links to newspapers, official documents and photos. The rate of confirmations (or disconfirmations) made is also calculated, since it measures how responsive the user has been to others. The other type of score refers to the evaluation that others have of the crimes reported by the agent. The rate of confirmed and unconfirmed crimes and the rate of denouncement of abuses, as well as the administrator's evaluation, are seen as referrals from other users about the actions of a certain agent, which will reflect on the calculation of its reputation score. More precisely:

- the rate of confirmations made for each agent $a$, $R_c$ is the number of confirmations made by $a$ ($\text{NbConfirm}$) divided by the number of invitations to give opinion about reports received by $a$ (NIR). ($R_c = \frac{\text{NbConfirm}}{\text{NIR}}$)
• the rate of additional information provided \( R_{add} \) is measured from the number of crime reports in which an additional piece of information was provided for a crime \( (N_{add}) \) divided by the number of reports registered by the agent \( a \) \( (NbCrimes) \),
\[
R_{add} = (N_{add}/NbCrimes)
\]
• the total rate of positive confirmations \( (RtC\_+) \) is based on the number of positive confirmations made by the "guests" \( g \) invited by agent \( a \) pondered by their reputation. Let \( f_c(i,g) \) be a function that represents an invitation made by \( a \) to \( g \) asking for an opinion about the report of crime \( i \) and that assumes the following values,
\[
f_c(i,g) = \begin{cases} 
1 & \text{if crime } i \text{ is confirmed by } g \\
0 & \text{otherwise}
\end{cases}
\]
\[
RtC\_+ = \sum_{i=1}^{n} \sum_{j=1}^{m} (\rho(g)) \times f_c(i,g)/n \times m
\]
where, \( m \) is the number of invitations done by \( a \) and \( n \) is the number of crimes registered by \( a \)
• Conversely, the total rate of negative confirmations \( (RtC\_-) \) made by the "guests" \( g \) invited by agent \( a \) is \( RtC\_- = \sum_{i=1}^{m} (\rho(g)) \times f_d(i,g)/n \times m \)
where,
\[
f_d(i,g) = \begin{cases} 
1 & \text{if crime } i \text{ is unconfirmed by } g \\
0 & \text{otherwise}
\end{cases}
\]
• The total rate of qualified denouncements \( (RtQD) \) is based on the number of denouncements of abuse that other agents have made on crimes registered by \( a \) pondered by the reputation of the one who has denounced \( \rho(s) \)
\[
RtQD = 1 - \sum_{i=1}^{n} (\rho(s) \times f_d(i))/n \text{ where,}
\]
\[
f_d(i) = \begin{cases} 
1 & \text{if crime } i \text{ is denounced as abuse by source } s \\
0 & \text{otherwise}
\end{cases}
\]
The final acquired reputation, \( AR \), is computed from the aforementioned measures, as follows:
\[
AR = R_c + R_{add} + RtC\_+ - RtC\_- - RtQD
\]

5.3. Reputation from trust propagation

A complementary manner of computing the reputation is based on the propagation of the trust that the agents have in each other while considering the reputation of the ones involved. For the sake of simplicity, we are going to call this mechanism propagation of reputation, even though we are aware that the precise concept would refer to the calculus of a reputation from the trust value that an agent has in another agent. In other words, when a friend receives an invitation to confirm a crime from someone with a good reputation, we assume that there is a strong indication that friend trusts the other. This allows one to compute the reputation of the agent who has been invited to confirm the occurrence from the reputation of the friend who has made the invitation. The propagation of reputation between agents is done at several iterations. At each iteration, the agent’s reputation score is divided by the number of its outgoing links (representing the indicated friends) and each of them gets an equal share. So, a friend’s overall reputation score is the sum of the shares from all its parent friends. The score among the agent friends is divided depending on the number of friends and follows a logarithm function.

More formally, an agent \( a \) with a reputation score \( \rho(a) \) that has invited \( g \) friends will give \( d \times \rho(a) \times \log(1+g) \) to each friend \( (d \) is the decay factor, which determines how much of the parent friends’ score is propagated to its friends). This logarithm strategy takes into account the effect of the number of friends that an agent has indicated, giving more importance to crime reports in which the agent asks several confirmations from friends. For computing the overall scores, given the shares from all its parent friends, a strategy called maximum parent [23] is adopted for summing the reputation values in such a way as to never exceed the reputation score of the most reputed parent. In other words, the sum of reputation values from each parent is calculated and this sum is compared to the largest reputation score among each of its parents; the smaller of these two values is used as the child’s reputation score.

Finally, the score of reputation is then computed from the propagated reputation \( (PR) \) and the acquired reputation based on the participation and social evaluation \( (AR) \). More formally the reputation of agent \( a \), \( \rho(a) = \rho(a) = w \times PR + (1 - w) \times AR \), where, \( w \) is a weight to ponder the importance of one type of reputation over another.

The reputation scores have been used in WikiCrimes only in background analysis. There is a management module that is exclusively available to the system administrators who are responsible for moderating the registers of crime made in the system, taking into account the reputation of the denouncer. Further versions of the system will provide a visual functionality in which the user will be capable of visualizing the hot spots by varying the level of reputation of the informants.
6. Putting WikiCrimes to work

6.1. Articulation with organized sectors of society

Articulation with society is fundamental for the success of WikiCrimes. In Brazil we have designed a plan to constantly establish partnerships with non-governmental organizations, news corporations, business representatives, municipal governments, and other organized sectors. In order to do this, different strategies have been tried, such as a marketing campaign called “WikiCrimes – I use it,” which involves distribution of ads via folders, informational pamphlets and bumper stickers. One good example of partnership was the one established with car insurance brokers. Insurance brokers are an important focus of the project, since they have information about crimes and may register them in the system. Besides examples of bringing people to collaborate, we have defined three technological strategies to foster collaboration, which will be described as follows.
6.1.1. Providing alert services

Even though the rationale behind WikiCrimes is founded on the notion of solidarity (registering crime to help other people), providing services to the user is one of the main strategies to foster collaboration, because it gives a concrete product to the user. It is thus easier to realize the common wealth that an individual action can bring. In WikiCrimes, we have defined a service called “WikiCrimes Alert.” The user can receive alerts via e-mail of registers of crimes that have occurred within a user-defined geographic area. It is possible to draw a polygon representing the user’s residential area, for instance, and for a pre-defined time (every day, week or month) to receive alerts about the situation in that area.

6.1.2. Integration in social networks

Another important strategy to foster collaboration and advertise the project was to create means to integrate it into a social network application. In order to do that, we have defined a mini-application following Google’s proposal of an Open Social API to be used in Google’s Orkut (http://www.orkut.com) (the most popular social net application in Brazil). The idea is to provide a tool for users of Orkut to alert their friends about security problems faced in a particular region. The user can pinpoint a crime on the map, comment on it, and dispatch it to their friends to alert them about the fact. Their friends, in turn, can also comment on it and forward the message to their friends, and so on. Orkut’s mini-app was an opportunity to reach a wider audience and encourage more participation (mainly from youth). The use of WikiCrimes in Orkut is only in its early stages, and no precise stats can be extracted from such use yet.

6.1.3. Embedding in online newspapers

Even though web applications follow a very specific way of advertising, the use of traditional media is fundamental for advertising any public project. We have established cooperation with a local newspaper, Diário do Nordeste (DN) in Fortaleza, whereby we made it possible for the journalists themselves to generate a widget (an embedding of WikiCrimes) with a small map of crimes that can be inserted into online news displayed by the newspaper. Fig. 6 shows how this integration is perceived by the newspaper reader (see [6] for an example, in Portuguese). On average, the WikiCrimes widget is displayed in the DN newspaper once a week. We have measured the effectiveness of this proposition by computing how many new users subscribe to WikiCrimes when news is shown with the widget. That strategy was initially used in June, 2008 and two news reports using the widget were produced every month (the decision of when to create the widget and for what period it refers depends exclusively on the journalists). Overall there were nine insertions of the widget; on average, for these days, WikiCrimes had 150% more subscriptions of new users than on regular days (without that promotion).

6.2. Business model

The business model we have in mind for the project is based on the existence of a set of customized services that would be provided to the user from mobile devices. These services would be charged either from revenue share with a cell phone operator or via a monthly subscription fee; i.e., the user would subscribe to WikiCrimes in order to have the right to receive SMS alerts or other services. The idea is to provide information to citizens on their mobile phones; these services could be based both on contextual messaging (using GPS or other location service) as well as warning messages on demand of the citizen. These suggestions converge with two features that we have already described in this text: it will be possible to demarcate one or more areas of interest so any event that occurs within that area will generate an alert sent to the user by SMS. In parallel, we are finishing the WikiCrimes version for mobile phones (S60 platform, in priority) that will enable the register of crimes from cell phones. Similar services for use in browsers in the automobile industry can also be envisioned through car navigators.

7. Evaluation and implementation issues

7.1. Comparing WikiCrimes data with official data

The ultimate goal of WikiCrimes is not to compete with law enforcement authorities or with their data collection systems in terms of the number of crimes recorded. Evidently, the number of crimes in the official system tends to be higher than the number collected by WikiCrimes. However, WikiCrimes can be considered as an auxiliary system of data collection that adds more quality to this process. It can be an excellent indicator of the locations where under-reporting is high, for example. Even without the goal of having the same quantitative aspects, a qualitative analysis of WikiCrimes data may reflect a crime trend and dangerous areas. We performed a comparative analysis between the official data on homicides collected by the Police for the region of Fortaleza with those recorded in WikiCrimes. We chose homicide because it is the most important crime, is widely used around the world to classify the level of violence in cities, and due to the fact that we have several collaborators who participate actively in the registration of this type of crime. First, we obtained the surprising result that the number of crimes with the correct georeferencing in the official base was much lower than in WikiCrimes. The total number of homicides in the year 2008 in Fortaleza was over 800, but only 337 among those into the official database were georeferenced. The total number of georeferenced homicides for the same period in WikiCrimes was 572. This occurs because of deficiencies in the official data collection system, which is based on receiving the address of the occurrences over the telephone. These
addresses, in turn, are often not found by the Police dispatch system, which prevents georeferencing. Even with this difference, we performed a comparison of data by applying kernel methods for estimating clustering between crimes. The results of this comparison can be seen in Fig. 7, where one can clearly see the same shape that appears between the clusters. There is a slight difference in the lower right portion of the kernel maps, which is explained by the difference in the quantity of data. Statistically, the two datasets can be considered as coming from the same distribution as the Ripley [21] test plotted in Fig. 8 depicts. The Ripley technique is a distance-based global spatial statistics which can be adopted for evaluating the presence of crime hot spots. In Fig. 8 it is possible to see that both lines (official data and WikiCrimes data) share similar behavior being over the line of complete randomness, what indicates the concentration of crimes.

7.2. Technical information

The entire architecture of WikiCrimes has followed certain design patterns that are already quite common in Web application development. The Model-View-Control (MVC) standard was implemented through the Java Server Faces (JSF) framework [13]. This is a framework for Java-based Web applications that simplifies the development of user interfaces. It uses a components-based approach whereby it is possible – through a set of pre-defined interface components – to create addi-

![Fig. 7. Comparison between WikiCrimes and police records for homicides in Fortaleza during 2008.](image1)

![Fig. 8. Plot of the Ripley test for measuring existence of clustering data.](image2)
tional customized components for a particular application. Furthermore, the framework manages the statuses of these components and links the events in a similar manner. Another interesting feature of JSF is the ability to integrate it with other frameworks and thereby take advantage of some of the features of such other frameworks.

The Application was separated into three layers. The Presentation (or View) layer is responsible for the user's interaction with the system as well as the presentation of the results of such interactions. In addition to JSF components, it has several JavaScript classes that are executed on the client side. These vary from simple validations to the complete interaction with the maps, using the Google Maps API (http://maps.google.com). The Services layer has all of the WikiCrimes rules and logics. These range from services of user identify (registration, log in, log out) to the inclusion of crimes, comments, sending electronic messages, confirmations and statistics. The Integration and Data Access layer uses Hibernate and the database; MySql was configured with triggers and stored procedures for calculating statistics (counters of the number of crimes per city, for example). On the client side, we used the Google Maps API. It allows all the interaction and manipulation of the map through Javascript classes. Moreover, we used some services offered by the Google Maps API, such as Geocoding and some XMLHttpRequest requests. Geocoding is the process of "translating" addresses, avenues or streets into numbers that locate them (more specifically: latitude and longitude). These services were used to facilitate the location of areas or specific points on the map where the crimes are to be registered (or merely accessed). Viewing the push-pins (markers) that represent the crimes is done through XMLHttpRequest calls, responsible for loading the markers.

8. Conclusion and future work

WikiCrimes currently has around three thousand confirmed users. The project was initiated in January, 2008 and since then has had more than 100,000 visits from 156 countries. Most of the visits came from Brazil. The project was born in Fortaleza, a city of 2.5 million inhabitants in the Northeast of Brazil. Therefore this is the place that has the majority of users and registers of crime. More than half of the total number of crimes have been pinpointed there. Articulation with organized society, workshops, lectures and a diverse and consistent campaign of local advertising were particularly intense in the city. Expansion to other regions is gradually occurring, mainly by means of agreements with municipal governments of medium-sized cities and with collaborators who maintain blogs on a similar theme. In addition to the multifaceted technological component that characterizes WikiCrimes, certain scientific challenges are imposed and deserve special attention. Three of them are currently objects of study by the WikiCrimes team. First, it is worth mentioning the studies aimed at making WikiCrimes available on mobile devices. The goal of such actions is to bring WikiCrimes closer to its users, leading them to join the WikiCrimes community at different times and to expand on mere access to the site. A prototype for operating as a mobile version to run on Nokia’s S60 platform of cellular telephones is already being tested. The second topic of investigative research is in regard to the development of software capable of reading pages of online newspapers, describing news stories about crimes and registering such crimes automatically in WikiCrimes. This involves investigating semantic methods of exploring web content as well as methods of processing natural languages. Finally, but no less important, we are making efforts to develop an algorithm for rendering the geographic information around the place of crimes. By doing so, we can prepare a dataset to be mined regarding the patterns that can explain the reasons the crime occurred.

In parallel with the reputation assignment, in WikiCrimes it is important to identify malware by means of pattern identification. The major damage that malware can produce is to induce wrong interpretations in terms of geographic patterns identification; i.e., the tendency of crimes to be wrong. The basic strategy we are going to follow is to monitor the hot spots for a determined geographic area. On a daily basis, the reputation agent will compute the hot spots of the areas that have received crime registers. Those hot spot areas that had the number of crimes registered by the same person or group of persons above a pre-defined boundary are alerted to the institutional agent.

Acknowledgements

WikiCrimes is a work developed with the support of several people without whom its operation would be impossible. In particular, we would like to thank all of the members of the Knowledge Engineering group at UNIFOR and IVIA.

References


Please cite this article in press as: V. Furtado et al., Collective intelligence in law enforcement – The WikiCrimes system, Inform. Sci. (2009), doi:10.1016/j.ins.2009.08.004