CityWatch: The Personalized Crime Prevention Assistant

Cristina Kadar
MTEC, ETH Zurich, Switzerland
ckadar@ethz.ch

Irena Pletikosa Cvijikj
MTEC, ETH Zurich, Switzerland
ipletikosa@ethz.ch

ABSTRACT

Motivated by rising levels of crime against property and findings in criminology research, we are developing **CityWatch** - the first mobile application that supports crime prevention behavior at community level. **CityWatch** leverages data on past crime incidents, which are sourced both from trustworthy sources, like the national census and the insurance industry, and from its users through crowd-sourcing. It applies machine learning algorithms to analyze the past incidents together with further data characterizing the living areas and learns common patterns of crime. These patterns are then leveraged in a general forecasting component, as well as in generating personalized risk profiles and crime prevention tips for registered users based on their account information. The results are visualized in an interactive map, where users can analyze past crime in their neighborhood and view predictions of future crime. Users can report a new crime and opt to receive notifications about new incidents in their proximity or area of residence.

Categories and Subject Descriptors

H.3.5 [**Information Systems**]: Information Storage and Retrieval—*Online Information Services*

Keywords

crime prevention, crime prediction, mobile, crowd-sourcing, data mining, public good

1. MOTIVATION AND APPROACH

Every 8 minutes a burglary takes place in Switzerland. With 932 burglaries per year for every 100.000 inhabitants in 2012, Switzerland has become the top target for break-ins in Europe [3]. Furthermore, based on a survey [7] we have recently conducted, one in five Swiss inhabitants believe that they will be a victim of crime within the next 12 months.

These are alarming statistics, motivating the need for solutions that help individuals protect themselves against different types of attacks and increase their safety. Police departments across the country are undertaking steps to increase population awareness by providing tips on how to prevent burglaries. These are published as formal guidelines within a pile of different data sources, making it difficult for individuals to get hold of appropriate information in a targeted and timely manner.

Existing commercial solutions for crime prevention fall into four broad categories: visualization interfaces in form of individual points or heat maps; platforms for reporting or sharing incidents with other users or local authorities, applications listing a static set of prevention tips; and, finally, systems offering basic analytics on top of the data. So far, to the best of our knowledge, no studies have been conducted to analyze how such information systems influence significantly the safety perception of their users and if they motivate prevention behavior. HCI researchers aiming at providing crime prevention technologies have until this point designed solutions that provide single individuals with information to lessen their chances of being victimized [2, 9]. Yet criminology research suggests that collective action successfully decreases crime and anxiety. Lewis and Lewis[8], argue that technologies intended for crime prevention should be designed to support communication and group problem-solving, as opposed to simply providing information on victimization risk to the citizens.

Motivated by the aforementioned trends and findings, our goal is to address following research questions: **(1)** How to design an information system for crime prediction and prevention by means of big data analytics? **(2)** How to motivate individuals to contribute with their personal data and build together a crowd-sourced model of crime in their communities? **(3)** How to support individuals to undertake prevention measures and lower personal and communal crime risk levels?

2. SOLUTION AND FUTURE WORK

The system leverages incidents data in form of property insurance claims from a big Swiss insurance company that characterize the crime events rigorously with location, description, type, time, stolen goods. Next to the incidents reported by the application users, further external data sources including demographics data, local weather data, and other public data describing the neighborhoods (e.g. distance to highways, or presence of police stations) will be integrated in a future version.

Based on the input data, the system builds three internal models. The first one is a **spatio-temporal predic-**

MUM '14 November 25 - 28 2014, Melbourne, VIC, Australia
ACM 978-1-4503-3304-7/14/11 ...$15.00.
http://dx.doi.org/10.1145/2677972.2678008.

Figure 1: Visualizing crime levels.

tion model of the criminal incidents intended to predict the locations and times of future criminal events. The results are presented in the user interface in form of a hot spot map. Our current strategy is to model the temporal behavior of incidents by multivariate time series analysis [4] and the spatial behavior by means of a generalized additive model (GAM) [5], which can utilize a variety of data types (i.e. geographic, demographic and other statistical data) to make predictions. The textual descriptions of the incidents can give valuable insights into the way criminals operate and reveal patterns of crime. Towards this goal, we are employing Latent Dirichlet Allocation (LDA) [1], a **topic model** that can identify common themes (i.e. topics) that pervade the unstructured collection of event descriptions. The discovered topics can be then utilized to formulate prevention tips: general safety tips, tips mitigating crime patterns specific to a given city or neighborhood, or even profile-specific advice. Lastly, given the profiles of all past victims, we develop a **victimization risk model**, which estimates the victimization risk of any individual based on her location, demographic data, and housing details. We use Logistic Regression [6] to identify the risk factors that are associated with an increased risk of victimization.

The core functionality is covered by the following five use cases:

- **View crime map**: The user can view both historical data and future predictions of the crime levels in the country or in an area of interest, as the data is presented as a zoom-able heat map as shown in Figure 1. She can choose to filter by data source, that is choose to only visualize official data pulled from trustworthy sources such as the census or insurance companies, or browse through user-generated content. Presented incidentscan be narrowed by crime type – we currently support four types: burglary, theft, car theft, and robbery. Alternatively, the data can be viewed as charts presenting different statistics of the incidents per day or per type.

- **Report a crime**: Any user, if at some point in time becomes a crime victim, can report the incident in the application and chose to remain anonymous while doing so.The use case is kept simple and straightforward and requires filling in some meta-data: when (exact date and time), what (incident type and optional picture), where (location on the map) and how (short textual description) the incident took place.

- **Create profile**: In order to prevent misuse, every new user would need to create a user account providing basic information which is relevant for the application functionality: current address, date of birth, gender, type of housing, etc.The user can choose to import contacts from other accounts like e.g. Facebook and add them to her list of persons of interest. Initial settings are also required to define what types of notifications should be pushed to the user.

- **View risk profile and safety tips**: Based on the inputs in the user account, the application will compute a risk score expressing how likely the user is to be a crime victim within the next 12 months. Furthermore, the user will receive general as well as personalized safety tips to reduce her risk of victimization. The tips span different categories: tips on how to fit the doors and windows, on how to setup alarm systems, or on how to manage the relationship with the neighbors.

- **Receive notifications**: Based on the settings in their user profiles, users can receive notifications about incidents in their proximity, in their neighborhood, or any other area they have defined.

Future steps include improving the afore listed models and performing a series of experiments to identify means of increasing the motivation of individuals to use the application regularly and to contribute actively to it. Ultimately, we plan to deploy the system as a public free application and conduct "research in the large".

3. REFERENCES

[1] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. Latent Dirichlet allocation. *The Journal of Machine Learning Research*, 3:993–1022, March 2003.

[2] Jan Blom, Divya Viswanathan, Mirjana Spasojevic, Janet Go, Karthik Acharya, and Robert Ahonius. Fear and the city: Role of mobile services in harnessing safety and security in urban use contexts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10. ACM, 2010.

[3] Bundesamt für Statistik BFS. Polizeiliche kriminalstatistik (PKS). *Jahresbericht 2012*, 2013.

[4] James Douglas Hamilton. *Time series analysis*. Princeton Univ. Press, 1994.

[5] T. J. Hastie and R. J. Tibshirani. *Generalized additive models*. London: Chapman & Hall, 1990.

[6] Joseph M. Hilbe. *Logistic Regression Models*. Taylor & Francis, 2009.

[7] Bogdan Ivan. Evaluating the potential of crowdsourcing for risk estimation in the insurance industry. Master's thesis, Swiss Federal Institute of Technology Zurich, Switzerland.

[8] Sheena Lewis and Dan a. Lewis. Examining technology that supports community policing. *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*, 2012.

[9] Christine Satchell and Marcus Foth. Welcome to the jungle: Hci after dark. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11. ACM, 2011.