POLITECNICO DI TORINO
Repository ISTITUZIONALE

Monitoring the citizens' perception on urban security in Smart City environments

Original
Monitoring the citizens' perception on urban security
in Smart City environments / Cagliero, Luca; Cerquitelli, Tania; Chiusano, SILVIA ANNA; Garino, P.; Nardone, Marco; Pralio, B.; Venturini, Luca. - STAMPA. - 31st IEEE International Conference on Data Engineering Workshops (ICDEW):(2015), pp. 112-116. (Intervento presentato al convegno DAta Mining And Smart Cities Applications Workshop 2015 co-located with the 31st IEEE International Conference on Data Engineering (ICDE 2015) tenutosi a Seoul (South Korea) nel 13-17 April 2015) [10.1109/ICDEW.2015.7129559].

Availability:
This version is available at: 11583/2588158 since: 2016-04-15T17:12:59Z

Publisher:
IEEE

Published
DOI:10.1109/ICDEW.2015.7129559

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

12 October 2024

Monitoring the citizens' perception on urban security in Smart City environments

Luca Cagliero*, Tania Cerquitelli*, Silvia Chiusano*, Pierangelo Garino‡,
Marco Nardone*, Barbara Pralio†, Luca Venturini*
*Dipartimento di Automatica e Informatica Politecnico di Torino
Torino, Italy
Email: {name.surname}@polito.it
†Fondazione Torino Wireless
Torino, Italy
Email: {barbara.pralio}@torinowireless.it
‡Telecom Italia S.p.A.
Torino, Italy
Email: {pierangelo.garino}@telecomitalia.it

Abstract

Sensing the perception of citizens on urban security is a key point in Smart City management. To address non-emergency issues municipalities commonly acquire citizens' reports and then analyze them offline to perform targeted actions. However, since non-emergency data potentially scale towards Big Data there is a need for open standards and technologies to enable data mining and Business Intelligence analyses.

The paper presents an integrated data mining and Business Intelligence architecture, relying on open technologies, for the analysis of non-emergency open data acquired in a Smart City context. Non-emergency data are first enriched with additional information related to the context of the warning reports and then analyzed offline to generate (i) informative dashboards based on a selection of Key Performance Indicators (KPIs), and (iii) association rules representing implications between warning categories and contextual information (e.g., city areas, districts, time slots). KPIs and rules are exploited to selectively notify to municipality actors (assessors, area operators) potentially critical situations, according to their role and authority. The experiments demonstrate the effectiveness of the proposed approach in a real Smart City context.

I. Introduction

Smart cities are urban environments in which the municipality fosters the use of Information and Communication Technologies (ICTs) to engage citizens in city management and development [1]. A key aspect in Smart City governance is the participation of citizens in decision-making. The capability of ICTs to keep in touch citizens with municipality actors (e.g., assessors, area operators) is crucial for improving the effectiveness and efficiency of urban services.

Nowadays, *non-emergency data* on urban security issues are acquired by various Smart Cities all over the world. They consist of citizen warning reports that not require an emergency response. Smart Cities allow citizens to signal these potential warnings to the local administrators through Web portals, apps, emails, and contact centers. To improve the citizens' quality of life, non-emergency data are worth analyzing because they represent the citizens' perception on urban security from different viewpoints. Based on the signaled warnings and their level of severity, the municipality can plan targeted actions on the urban area, which may vary according to the temporal evolution of the citizens' perception on non-emergency issues.

The main research contributions to this scenario can be classified as follows: (i) definition of collaborative models and open standards (e.g., Open311 [2]), to facilitate the development of smart applications [3], [4] and the cooperation between cities, (ii) design and development of smart cities platforms by private vendors [5], researchers [6], [7] and public administrators [8], and (iii) characterization and analysis of the perception of urban security sensed by users [9], [10], [11], [12]. An overview of the key challenges in urban computing from the point of view of computer scientists is given in [13]. However, state-of the-art approaches are challenged by the increasing volume of analyzed data, which prompts the need for integrating data mining and Business Intelligence tools in non-emergency data analyses.

This paper presents Non-Emergency Data Analyzer (NED), a new integrated data mining and Business intelligence environment targeted to the analysis of open non-emergency data on urban security issues acquired in a Smart City context. The

The research leading to these results was funded partly by the EU project FI-WARE (Grant Agreement nr. 285248). We would like to thank Dr. G. Todesco (Chief Commissioner of the Scientific Investigation Unit of the Department of Criminal Police of Metropolitan Police of Turin) for his insightful comments and suggestions, Dr. G. Cazzin and Dr. A. Gachino (Engineering S.p.A.) for their technical support in using SpagoBI, and Dr. B. Moltchanov (Telecom Italia S.p.A.) for his technical support in using FIWARE technologies.

NED system relies on an ensemble of open components, which are easily portable in different Smart City contexts. To profitably analyze the citizens' perception on urban security, non-emergency data are acquired, enriched with additional information about the context of the warning reports (e.g., the related city area), and stored into a unique data repository. Next, two complementary analyses are performed: (i) Key Performance Indicator (KPI)-based analysis, and (ii) association rule discovery. In our scenario, KPIs are quantitative indicators measuring the level of warning of the citizens in a specific context, while association rules [14] are significant associations between warnings and contextual features (e.g., between a specific warning category and a city area). The aim of KPI generation is twofold: (i) to provide useful feedback to municipality users by generating informative dashboards on KPIs, and (ii) to automatically generate and notify alerting signals on critical situations. The aim of association rule extraction is to evaluate the strength of the correlations between warnings and contextual features and to automatically trigger alerts in case the extracted patterns highlight potentially critical situations [15]. For example, the city areas with maximal incidence of a specific warning category (e.g., *urban blight and renewal*) are those that may need maximal surveillance. On the other hand, a percentage decrease in the number of non-emergency warning signals from one year to the subsequent one may reveal a positive trend in the citizen perception on that specific urban security issue. The municipality can exploit such information, for example, to validate the effectiveness of the latest actions.

The effectiveness and usability of NED system have been evaluated on real non-emergency data acquired in a real Smart City context. We considered as analysis scenario the study of the non-emergency calls and emails received by the contact center of the Local Police Department of Turin, an important city located in the north-west of Italy, and we reported the results of a information and awareness campaign launched in Turin in September 2014 during a Europe-wide public event, namely the European Researchers' night.

This paper is organized as follows. Section II describes the NED system, while Section III summarizes the experiments and the dissemination activities. Finally, Section IV draws conclusions and presents future developments of this work.

II. The NED system

Non-Emergency Data Analyzer (NED) is a new data mining and Business intelligence environment aimed to analyze non-emergency open data acquired in a Smart City context. The main environment blocks are briefly introduced below. A more detailed description is given in the following sections.

Data preparation. To prepare non-emergency data to offline analyses, data is acquired, enriched through external open data sources with contextual data related to the warning reports, and then stored into a unique data repository (see Section II-A).
Data analysis. From the prepared data, informative dashboards are generated based on a selection of Key Performance Indicators (KPIs), which are quantitative indicators reflecting peculiar data features. In parallel, association rules are extracted using an established data mining algorithm [16]. Dashboards and association rules provide complementary information on the perception of citizens on urban security (see Section II-B).
Notification. Based on KPIs and association rules, alerts on critical situations are automatically generated and sent to the main municipality actors (e.g., city mayor, assessors, area operators). Notifications are selectively forwarded based on the role and authority of the municipality actors (see Section II-C).

The NED system implementation relies on the open FIWARE technologies (www.fi-ware.org) for Business Intelligence (SpagoBI v. 4.2, available at www.spagobi.org and soon becoming FIWARE Generic Enabler) and notification (CAP Context Broker, available at http://forge.fi-ware.org/plugins/mediawiki/wiki/fiware/index.php/), while it uses the opensource RapidMiner suite v.5.0 (https://rapidminer.com) for data mining analyses.

Although the NED system is general and it can be applied to data acquired in different Smart City environments, hereafter we will consider as use-case scenario the analysis of the warnings perceived by the citizens of Turin.

A. Data preparation

The contact center of the Local Police department of the Turin Smart City acquires non-emergency data and periodically integrates them into a unique data repository to allow offline data analyses. Warning categories and sub-categories are assigned according to the classification reported in Table I (http://aperto.comune.torino.it). For example, citizens may signal noise coming from the street adjacent to their house due to the presence of a group of young persons raising their voice night-time. This warning may be classified as *Youth gathering*, which belongs to the more general category *Civil tension*.

For each non-emergency report, the time and the address to which it refers to are also stored. In addition, the NED system enriches the initial data schema with extra contextual information acquired from external open data sources. Non-emergency data enriched with contextual information is stored in a data warehouse whose dimensional fact model [17] is depicted in Figure 1. More specifically,
(i) to analyze the *temporal distribution* of non-emergency data, the following time granularities are considered: *day*, *month*, *2-month*, *4-month*, *6-month*, and *yearly* time periods. Moreover, the day is classified as *working or high day*, and the warning report *time* is aggregated into the corresponding *daily time slot* (morning, afternoon, evening, or night).
(ii) to analyze the *spatial distribution* of non-emergency data, higher-level space granularities are also considered beyond the location addresses from which the warning report refers to. Specifically, each *address* is mapped to the corresponding *city area*

TABLE I. CATEGORIZATION OF NON-EMERGENCY REPORTS

Category	Sub-category
Social tension	Vandalism Other
Civil tension	Youth gathering Disturbing behaviors Disturbance from dogs Disturbance from public venues Disturbance from other animals Noise nuisance Improper use of common areas Other
Urban quality	Urban blight and renewal Abandoned vehicles Other

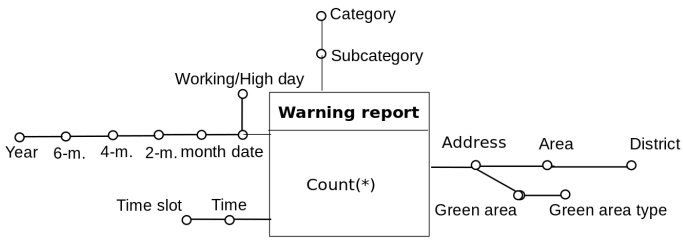

Fig. 1. Data warehouse dimensional fact model.

and to the *city district* including that area. While the address and district are recorded by the contact center employees, the city area corresponding to the address is added as additional contextual feature to the final repository to aggregate data at an intermediate granularity level. For warning reports referring to a green area, the *green area name* and *type* (e.g., park, garden) are stored.

Furthermore, topological and demographic information about city areas, districts, and green areas are integrated in the repository as well. Topologies are used to graphically analyze the most significant spatial trends in non-emergency data, while demographic information is exploited to characterize non-emergency data according to the distribution of citizens per area, district, and gender (e.g., the number of male citizens per city area). Demographic statistics and topologies about the city areas and districts of Turin were acquired from the official GeoPortal of the Turin municipality (www.comune.torino.it/geoportale). Topologies were encoded in GeoJSON, which is a standard format for encoding a variety of geographic data structures.

B. Data analysis

The prepared non-emergency data are analyzed to gain insights into the perceived citizen's warnings. The NED system performs two complementary (offline) analyses: (i) KPI analysis and (ii) data mining based on association rule analysis. The aim of data analysis is twofold: (i) to produce useful feedback to the municipality users by generating informative dashboards (using KPIs) and (ii) to automatically generate and notify alerting messages on critical situations (using KPIs and association rules). In the following we present KPI and association rule analyses.

KPI analysis. In Business Intelligence, the analysis of Key Performance Indicators (KPIs) is an established methodology [18]. KPIs help organizations define and measure progress toward organizational goals by monitoring most significant achievements. In our context, KPIs are quantitative indicators of the perception of citizens on urban security. To apply KPI analyses to data coming from a real scenario, we defined a set of KPIs related to the non-emergency data categories and sub-categories reported in Table I. These KPIs analyze non-emergency data from two main viewpoints: (i) the temporal dimension and (ii) the spatial dimension of warning reports. More specifically, to reveal potentially critical situations in a specific urban area, we analyzed the incidence per city area/district/green area of the perceived warnings related to a given category/sub-category. Furthermore, to keep track of the temporal evolution of the citizen perception on urban security, we also computed the percentage variation between the number of calls received from a given city area/district/green area in a given time period (e.g., in year 2013) and those received in the same area in a preceding time period (e.g., in year 2013). To avoid bias due to the unbalanced distribution of citizens per city area/district, we normalized both the aforesaid KPIs by the total number of citizens per city area/district. For example, a positive yearly differential between the number of calls related to sub-category *Abandoned vehicles* coming from a given city area may reveal a negative trend, which may be due to an insufficient surveillance of the area. Thus, for this area, the municipality may plan targeted actions.

Association rule analysis. This step aims at discovering interesting associations between warning categories/sub-categories and context features in the form of association rules. Association rule mining [14] is an established data mining technique to discover interesting yet hidden associations among data items. Association rules are implications that indicate the co-occurrence

of multiple items in the source data. They may represent underlying correlations among data items which are hardly inferable by performing KPI analyses. In our context, they represent sets of warning categories/sub-categories and contextual features (city areas, time periods) that are strongly correlated with each other. In the following, association rule definition is tailored to our context of analysis.

To introduce the concept of association rule, we first recall the notion of itemset. An *itemset* is a set of pairs (*feature*, *value*), called *items*, having distinct feature values. *Feature* can be a warning category, sub-category or a contextual feature (e.g., address, time, 6-month time period, green area type). *Value* is the value of the corresponding feature. For example, (Sub-category, Vandalism) is an item where the feature is Sub-category and the value is Vandalism. {(Sub-category, Vandalism),(Green Area, Pellerina Park)} is an example of an itemset consisting of 2 items belonging to features *Sub-category* and *Green area*, respectively.

An *association rule* is an implication $A \rightarrow B$, where A and B are disjoint itemsets (i.e., itemsets with no item in common). A and B are also denoted as the antecedent and consequent of the association rule $A \rightarrow B$, respectively. {(Sub-category, Vandalism)} $\Rightarrow$ {(Green Area, Pellerina)} is an example of association rule stating that the occurrence of item (Sub-category, Vandalism) "'implies"' those of item (Green Area, Pellerina) in the source data.

Many (statistics-based) quality measures can be exploited to select the most interesting association rules [19]. The NED system integrates three large used rule quality measures, i.e., the rule support, confidence, and lift. The *support* of rule $A \rightarrow B$ is the prior probability of A and B in the source dataset. The *confidence* is the conditional probability of occurrence of B given A. It measures the strength of the implication. The *lift* is a measure of the (symmetric) correlation between the antecedent and consequent of the extracted rules and it is computed as the ratio between the confidence of the rule and the prior probability of B in the source dataset. Lift values significantly above 1 indicate a positive correlation between rule antecedent and consequent, meaning that the implication between A and B holds more than expected in the source dataset.

In our context of analysis, co-occurrences between warnings and contextual features are deemed to be reliable if they are frequent and the corresponding items are strongly correlated with each other. Hence, in the NED system we selected all the association rules whose support value is equal to or above a minimum support threshold *minsup*, and the lift value is above or equal to the minimum positive correlation threshold $min^{+}lift$ ($min^{+}lift>1$) . The association rule extraction task is accomplished by using the Java-based RapidMiner implementation of the well-known FP-Growth algorithm [16] (http://rapid.com). However, different association rule mining algorithms can be easily handled by the NED system as well.

C. Notification

To allow the municipality to constantly monitor non-emergency data, the NED system automatically generates periodic notification messages when KPIs and rules may indicate potentially critical situations.

To receive notifications, users are asked to subscribe to the NED notification service. Since municipality actors have different roles and authorities, the system grants users to receive notification messages based on their area of expertise. Specifically, NED first defines a set of roles and then assigns a geographical scope (e.g., at the level of the municipality, district, or area) to each role based on its authority. During the subscription phase, users have to indicate their role in the Municipality (e.g., city mayor, area operator). According to the role and authority the NED system assigns them a scope (i.e., at level of Municipality, District, or Area).

For example, authoritative actors at the level of the municipality (e.g., the city mayor) receive all the notifications at the levels of city and district, but not those at the level of city area, because the latter provide information at a too fine granularity level. Conversely, each area operator is granted to receive only the notifications of the corresponding area.

To generate notification messages, a set of alerting rules is generated from the KPIs and association rules mined at the previous step (see Section II-B). Each alerting rule has a scope and may be characterized by a severity level of the warning. Each user receives by email the subset of alerting rules pertinent to his role scope, enriched with dashboards on the corresponding KPIs.

For example, alerting rules related to vandalism acts to street furnitures in specific city areas are sent to all users whose scope is at level of city area and whose area of expertise comprises street furnitures (e.g., the area operators).

Alerting rules from KPIs. To generate alerting rules from KPIs, the NED system considers the percentage differential between consecutive time periods per city area, district, and green area. The aim is to monitor the variation of the citizens' perception on urban security over time in specific urban areas, and notify it to the municipality actors. Based on KPI values, the severity level of the warnings is classified as *stable*, *substantial* increase/decrease or *critical* increase/decrease, respectively. *Stable* KPI variations indicate a relatively stable trend in the citizen perception on urban security. Conversely, KPI variations falling in the other levels indicate moderately/significantly decreasing/increasing trends. Usually, stable trends do not require triggering alerting signals. Instead, levels *substantial* and *critical* trigger pre-alerting and alerting signals, respectively. Notifications are selectively forwarded to the municipality actors in charge of the specific issue, who may decide to perform targeted actions (e.g., periodic surveillance of specific areas, redevelopment of green areas).

For each severity level of the warnings (stable, substantial, critical), the corresponding KPI value ranges are analyst-provided. In our experiments, we classified the KPI variations in the range [-2%; 2%] as *stable*, those in the ranges [-5%;-2%] and in

[2%;5%] as *substantial* decrease and increase, respectively, and those below -5% and above 5% as *critical* decrease and increase, respectively.

Alerting rules from Association Rules. Association rules represent potentially critical situations arising from non-emergency data when a specific city area/district/green area appears to be strongly correlated with a given combination of categories/sub-categories. To define alerting rules, the NED system first extracts the association rules from non-emergency data by enforcing *minsup* equal to 1% (i.e., the implications must hold for on at most 1% of the source data) and $min^{+}lift$ equal to 10 (i.e., the rule antecedent and consequent must be strongly correlated with each other). The top-K ranked rules in order of decreasing lift value (where K is an analyst-provided parameter) are exploited to automatically generate notifications to the municipality actors.

Rules represent implications between green areas/city areas/districts (one or more) and a specific warning category/sub-category. They can be exploited to trigger alerts related to a specific urban area. To trigger more detailed alerts, more complex rules can be extracted by combining the spatial information about the green area/city area/district with the day or the time slot at which the critical situation occurs.

For example, rule $\{$(Green Area, Pellerina Park)$\} \Rightarrow \{$(Sub-category, Vandalism)$\}$ may indicate a alerting situation related to a specific green area of Turin. Instead, rule $\{$(Green Area, Pellerina Park), (Time Slot, night)$\} \Rightarrow \{$(Sub-category, Vandalism)$\}$ specializes rule $\{$(Green Area, Pellerina Park)$\} \Rightarrow \{$(Sub-category, Vandalism)$\}$ by providing also the information about the time slot at which most of non-emergency data were received.

III. Analysis scenario

The NED system was validated on real non-emergency data acquired in the Turin Smart City environment. More specifically, we analyzed an open non-emergency call dataset consisting of 4,672 calls received by the contact center of the Local Police Department of Turin in years 2012 and 2013, which is available in open municipality portal AperTo (http://aperto.comune.torino.it). The main dataset attributes are described in Section II-A. The experiments were performed on a quad-core 3.30 GHz Intel Xeon workstation with 16 GB of RAM, running Ubuntu Linux 12.04 LTS.

Two representative examples of dashboards generated from the KPIs defined in the NED system are reported in Figures 2 and 3. They show the incidence of calls for disturbance from public venues per district in years 2012 and 2013, respectively. Districts are colored with a 4-level scale ranging from blue (low percentage of calls) to red (high percentage of calls). District 1 (*Centro Crocetta*) corresponds to the city center and it is characterized by an averagely high number of calls in both years. Since the level of warning perceived by citizens in this district remains critical over the two years, the municipality would need to undertake further actions. Oppositely, in district 8 (*Borgata Lesna*) the number of calls decreased from year 2012 to year 2013 thus the issue appears to be overcome.

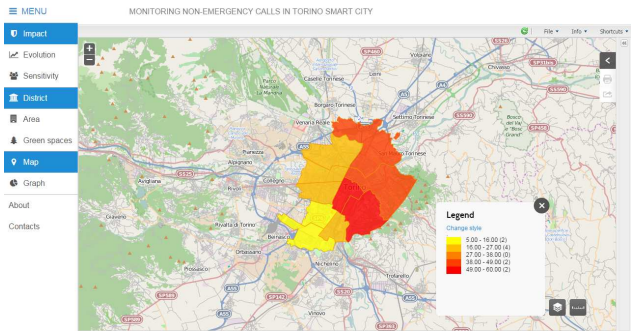

Fig. 2. Incidence of disturbance from public venues per district in year 2012.

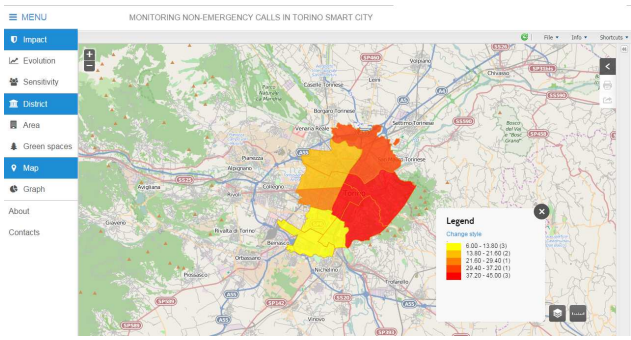

Fig. 3. Incidence of disturbance from public venues per district in year 2013.

Concerning district 1, the NED system extracted the following rules by enforcing *minsup*=1% and $min^{+}lift$=10: {(District, 1)} $\Rightarrow$ {(Category, Civil tension)} (sup=9%, conf=60%, lift$> 10^5$)
{(District, 1)} $\Rightarrow$ {(Sub-category, Disturbance from public venues)} (sup=3%, conf=19%, lift$> 10^5$)

These rules indicate that 60% of the warnings raised the citizens of district 1 belong to category *Civil tension*. Among these rules, *Disturbance from public venues* appear to be most correlated subcategory (i.e., the rule lift is maximal w.r.t. all the rules in the form {(District, 1)} $\Rightarrow$ {(Subcategory, *)}). Hence, the local police may increase night-time surveillance close to restaurants, pubs, and discos located within the city center.

To receive useful feedback from the Turin citizens in September 2014 we launched an information and awareness campaign during a Europe-wide public event, i.e., the European Researchers' night. Around 200 citizens browsed the dashboards, validated the signaled notifications, and provided fruitful comments to the municipality of Turin.

IV. Conclusions and future work

NED (Non-Emergency Data Analyzer) is a new data mining and Business Intelligence environment aimed to supporting the analysis of non-emergency data acquired in a Smart City context. The system integrates open tools for data analyses and mining to study the perception of citizens on urban security. The system has been validated in a real Smart City context and presented during a Europe-wide public event.

There is still room for improvements for our system. For example, the system may extract multiple-level rules with advanced data mining algorithms (e.g., [20], [21]) to analyze the correlations hidden in the non-emergency data at different abstraction levels.

References

[1] S. Pellicer, G. Santa, A. Bleda, R. Maestre, A. Jara, and A. Gomez Skarmeta, "A global perspective of smart cities: A survey," in *IMIS 2013*, July 2013, pp. 439–444.

[2] Open311, "A collaborative model and open standard for civic issue tracking. Available at http://www.open311.org/." 2014.

[3] Winnipeg311, "Winnipeg 311 Mobile App. Available at http://www.winnipeg.ca/Interhom/contact/app.stm/." 2014.

[4] Bloomington, "Open Source GeoReporter and uReport tools for Open311. Available at http://bloomington.in.gov/." 2014.

[5] IBM-IOC, "IBM Intelligent Operations Center. Available: http://www-03.ibm.com/software/products/it/intelligent-operations-center. Last access on November 2014," 2014.

[6] M. Hamilton, F. Salim, E. Cheng, and S. L. Choy, "Transafe: A crowdsourced mobile platform for crime and safety perception management," *SIGCAS Comput. Soc.*, vol. 41, no. 2, pp. 32–37, Dec. 2011.

[7] M. Behrens, N. Valkanova, A. F. Schieck, and D. Brumby, "Smart citizen sentiment dashboard: A case study into media architectural interfaces," in *Inter. Symp. on Pervasive Displays*, 2014.

[8] Smartdatanet, "Smart data platform. Available at http://www.smartdatanet.it/." 2014.

[9] M. Blythe, P. C. Wright, and A. F. Monk, "Little brother: could and should wearable computing technologies be applied to reducing older people's fear of crime?" *Personal and Ubiquitous Computing*, vol. 8, no. 6, pp. 402–415, 2004.

[10] M. Williams, O. Jones, C. Fleuriot, and L. Wood, "Children and emerging wireless technologies: investigating the potential for spatial practice," in *Conf. on Human Factors in Computing Systems 2005*, pp. 819–828.

[11] F. Naceur, "Impact of urban upgrading on perceptions of safety in informal settlements: Case study of bouakal, batna," *Frontiers of Architectural Research*, vol. 2, no. 4, pp. 400–408, Dec. 2013.

[12] C. Bach, R. Bernhaupt, C. S. D'Agostini, and M. Winckler, "Mobile applications for incident reporting systems in urban contexts: lessons learned from an empirical study," in *European Conference on Cognitive Ergonomics 2013, ECCE '13, Toulouse, France, August 26 - 28, 2013*, 2013, p. 29.

[13] Y. Zheng, L. Capra, O. Wolfson, and H. Yang, "Urban computing: Concepts, methodologies, and applications," *ACM Trans. Intell. Syst. Technol.*, vol. 5, no. 3, pp. 38:1–38:55, Sep. 2014.

[14] R. Agrawal, T. Imielinski, and Swami, "Mining association rules between sets of items in large databases," in *ACM SIGMOD 1993*, 1993, pp. 207–216.

[15] R. Agrawal and G. Psaila, "Active data mining," in *KDD 1995*, pp. 3–8.

[16] J. Han, J. Pei, and Y. Yin, "Mining frequent patterns without candidate generation," *In SIGMOD'00, Dallas, TX*, May 2000.

[17] M. Golfarelli, D. Maio, and S. Rizzi, "The dimensional fact model: A conceptual model for data warehouses," *International Journal of Cooperative Information Systems*, vol. 7, pp. 215–247, 1998.

[18] R. Kimball and M. Ross, *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*, 2nd ed. New York, NY, USA: John Wiley & Sons, Inc., 2002.

[19] P.-N. Tan, V. Kumar, and J. Srivastava, "Selecting the right interestingness measure for association patterns," in *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'02)*, July 2002, pp. 32–41.

[20] E. Baralis, L. Cagliero, T. Cerquitelli, V. D'Elia, and P. Garza, "Expressive generalized itemsets," *Inf. Sci.*, vol. 278, pp. 327–343, 2014.

[21] E. Baralis, L. Cagliero, T. Cerquitelli, S. Chiusano, P. Garza, L. Grimaudo, and F. Pulvirenti, "Misleading generalized itemset mining in the cloud," in *IEEE International Symposium on Parallel and Distributed Processing with Applications, ISPA 2014, Milan, Italy, August 26-28, 2014*, 2014, pp. 211–216. [Online]. Available: http://dx.doi.org/10.1109/ISPA.2014.36