CitymisVis: a Tool for the Visual Analysis and Exploration of Citizen Requests and Complaints

Rocío B. Hubert, Ana G. Maguitman, Carlos I. Chesñevar
ICIC CONICET-UNS, Dep. of Computer Science and Eng., National University of the South
San Andrés 800 - B8000FTN B. Blanca, Argentina
Tel.: +54-291-459-5135
{rocio.hubert, agm, cic}@cs.uns.edu.ar

Marcos A. Malamud
Mismatica Management
Martín y Omar 129. Oficina 206,
CP 1642, San Isidro,
Buenos Aires, Argentina
Tel.: +54-11-60939386
malamud@mismatica.com

ABSTRACT
Citymis Community is an EGOV app through which municipalities listen and respond to requests and complaints from the citizens. As a result of the successful adoption of Citymis Community in several municipalities, large amounts of data are daily generated, which require analysis by the corresponding service departments. The intelligent analysis of citizen requests and complaints can lead to improved levels of service coordination and can help in the decision-making process by relevant authorities. Advances in dynamic visualization technologies and data mining offer the opportunity to develop novel ways of analyzing these data. In this paper we present CitymisVis, a tool that extends Citymis Community by applying specifically developed methods for clustering and visually exploring geolocated reports. CitymisVis allows to perform a visual analysis of the geographic areas more affected by different kinds of issues and generates navigable representations offering statistics based on the collected data. In this way, CitymisVis helps to understand what issues require urgent intervention and allows to adequately reflect the proportions of different problems.

CCS Concepts
• Information systems~Decision support systems • Computing methodologies~Artificial intelligence • Applied computing~E-government

Keywords
Government 2.0; e-Governance; Citizen Requests and Complaints; Visualization; Data mining

1. INTRODUCTION AND MOTIVATIONS
Citizen requests and complaints provide a valuable source of insight into infrastructure and service-related problems within municipalities. The application of information and communication technology within public administration to collect and process citizens' requests and complains provides the involved stakeholders (citizens, government officials, etc.) with suitable channels for efficient electronic communication and coordination [1,2,3]. In the particular context of intelligent city management, a number of platforms have been developed to listen and respond to requests and complaints from the citizens [4,5,6]. Nevertheless, existing platforms only provide a partial solution to the problem of efficiently and effectively analyzing the large amounts of data that are daily generated by their users. In some cases, citizen requests and complains correspond to isolated issues, while in other cases they are associated with bigger problems. Issue-related trends could correspond to persistent problems impacting specific geographic areas or particular kinds of services, which relevant authorities need to identify and solve by prioritizing the most urgent demands.

Advances in dynamic visualization technologies and data mining make it possible to develop tools that facilitate the analysis of geolocated requests and complaints. The clustering-based visual analysis of data from citizen complaints and requests provides useful insights into the proportion of the existing infrastructure and service-related problems.

Citymis Community is an EGOV application through which municipalities listen and respond in real time in a collaborative network, resulting in a significant reduction of response time and improvements in the quality of the interaction with citizens. This paper presents CitymisVis, a tool that provides a number of interactive functionalities for the visual analysis and exploration of the requests and complaints submitted through the Citymis Community application. In this way, CitymisVis helps recognize a range of problems in a community infrastructure and service delivery, providing insight into the proportions of the identified problems. Therefore, it is a valuable aid in the decision-making process for relevant authorities. The tool is also useful to produce surveys to keep citizens informed by visual data and statistics. In this paper we illustrate different functionalities of CitymisVis using reports submitted by users from San Isidro, a municipality with a population of approximately 350,000 inhabitants located near Buenos Aires City.

2. CITYMIS - OVERVIEW
Citymis Community (https://citymis.co/ - Citizen Centric EGOV) is an ecosystem of Web and mobile applications expert in the holistic management of municipal services, designed for both citizens and municipal public agents, developed by Mismatica Management (http://www.mismatica.management), which allows users to report issues directly from their computers or smartphones. This section reviews the main features of Citymis Community and is based on the material presented in [7].

To report an issue, the user has to specify the category of the issue that is being reported. The possible categories include street conditions, transit and traffic signals, public lighting system, urban hygiene, running water, urban trees, parks and squares, lost pets and pets for adoption. In the meantime, each of these categories has a

DOI: http://dx.doi.org/10.1145/3047273.3047320

list of associated issues. For instance, the public lighting system category includes issues such as streetlight off, blinking streetlight, new light request, and streetlight repair request, among others. The user has to type in the street and number where the problem has been noticed. Alternatively, a map can be used to enter the location of the issue, providing a mechanism that not only allows to specify the exact geographic position but also to find and examine other issues reported in the same or surrounding area. In addition, the user can fill in a comment field and provide a photo to better describe the reported issue. The application also requests the contact information of the user.

From an operational perspective, a municipality is divided into geographic sub-regions, usually referred to as "areas of assignment''. This allows to allocate reports to specific employees (inspector, maintenance team, etc.) to solve or verify the reported issues. Each report is firstly associated with a "management unit", which is basically a geographic location (latitude, longitude, street and number), and it initially enters the "pending moderation" state. At this point, the report can be accepted, in which case it is assigned the "in hand" status, or it can be dismissed, receiving the "rejected" status. Once the report has been accepted, the remaining step is to solve the reported problem in order to attain the "fulfilled" status and complete the process. Each time there is a change in the status of the report, the user that submitted the report receives an email notification and is invited to add new comments, to provide additional data that can help solve the problem, or simply to give an opinion of the service.

Currently, Citymis Community is actively being used by 43 municipalities and organization in Argentina, Bolivia, Mexico, Venezuela and Chile. In the meantime, approximately 98 municipalities are testing the demo version of the application, covering in total more than 28.000.000 citizens.

3. THE CITYMISVIS TOOL

With the purpose of obtaining useful insight from the data collected through the Citymis Community application, we have developed CitymisVis, a web-based tool with various functionalities that allow citizens and municipality employees to interactively analyze maps and statistics. To illustrate the use of the tool we provide examples created using reports from the San Isidro municipality. The dataset employed by the tool is in Spanish and contains 68,368 reports coming from citizens that used a mobile app, the Citymis Community web system or a toll-free phone number to report a problem. CitymisVis has access to various attributes associated with each report, such as the report date, the type of service, the type of problem and the location of the problem (latitude, longitude, street name and number). In addition, the tool also has access to the user's comments (description of the claim), photo (if available) and the solution to the problem. CitymisVis offers three main mechanisms to explore and analyze the data, namely heatmaps, clusters and statistics, which are described next.

3.1 Heatmaps

Heatmaps allow to depict the intensity of requests and complaints at geographical points. Colors in a heatmap range from green (slightly affected) over yellow (moderately affected) to red (highly affected). CitymisVis uses Mapbox [8] to visualize the maps and to generate the corresponding heatmaps. The developed tool offers *zoom in* and *zoom out* options, making it possible to visualize an entire city or to focus on specific areas. In addition, it is possible to visualize a heatmap for the entire set of services or to select a specific service. Visualizing reports about certain types of services through a heatmap can help to identify prevailing problems and patterns related to those services. Figure 1 presents a heatmap for the public lighting system service of San Isidro. This map allows to distinguish different concetrations of reports across different geographic locations, offering a valuable aid to recognize the most affected areas.

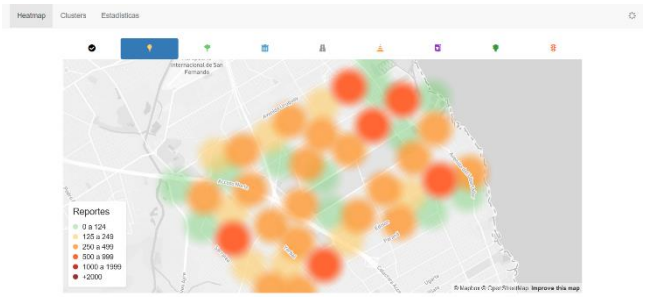

Figure 1: Heatmap showing different concentrations of reports associated with the public lighting system.

3.2 Clusters

Clusters allow to visualize the groupings of requests and complaints. The clusters in CitymisVis are generated using the *density-based spatial clustering of applications with noise* algorithm, also known as DBSCAN [9]. According to the DBSCAN algorithm, clusters are areas with higher density of items than other areas. Items within a cluster are mutually density-connected. In addition, if a point is density-reachable from any point of the cluster, it is part of the cluster as well. The algorithm marks as outliers those points that lie alone in low-density regions.

In the same way as for the heatmap representations, the cluster-based maps allow to visualize clusters generated over the entire set of services, or to select a specific service. The tool also allows to visualize clusters of reports at different granularity levels. These granularities depend on different parameter values associated with the DBSCAN algorithm, such as the distance to compute a neighborhood and the number of points required for a region to be considered dense. Clusters are color-coded to indicate if they include a large (red), medium (yellow) or small (green) number of points. CitymisVis uses the implementation of the DBSCAN algorithm available through the ELKI environment [10] while the Mapbox tool is used to display the generated clusters on the corresponding map. Figure 2 shows different clusters obtained for the analyzed dataset, with colors representing different concentrations of reports.

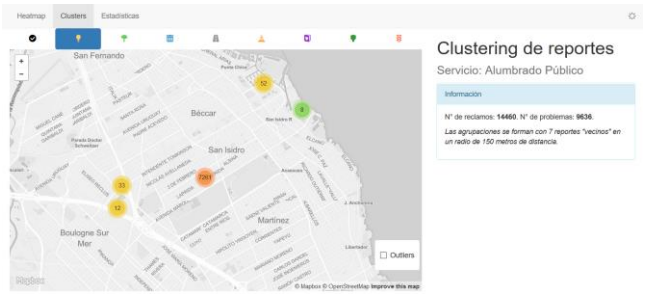

Figure 2: Clustering of reports associated with the public lighting system.

CitymisVis provides a mechanism through which a user is able to visualize sub-clusters of a selected cluster. The *hierarchical greedy clustering algorithm* [11] is used to identify clusters and its sub-clusters at different zoom levels in a map. Given a point in the dataset, this algorithm finds all the neighbor points within a specific

radius. These points are used to form a new cluster. A new point that is not part of the existing clusters is selected in a greedy manner and the steps are then repeated until all points in the dataset are considered. The DBSCAN algorithm is used to identify the main clusters, and then combined with the hierarchical greedy clustering method offered by Mapbox to dynamically and efficiently identify sub-clusters at different granularity levels. This results in a set of nested polygons associated with the cluster selected by the user and its sub-clusters. The left-hand side of Figure 3 shows the polygons resulting from applying hierarchical greedy clustering to the reports associated with the public lighting system service of the San Isidro municipality. On the right-hand side of Figure 3, there is a pie chart showing the proportions of reports belonging to different subcategories of the service under analysis. For instance, the public lighting system service category includes 16 subcategories, such as streetlight off, blinking streetlight, new light request, streetlight repair request, etc. Hovering the mouse over individual slices of the pie chart allows to identify the corresponding subcategory. In the presented example, the streetlight off subcategory (light blue) is the most common one, accounting for 61.98% of the reports corresponding to the public lighting system category.

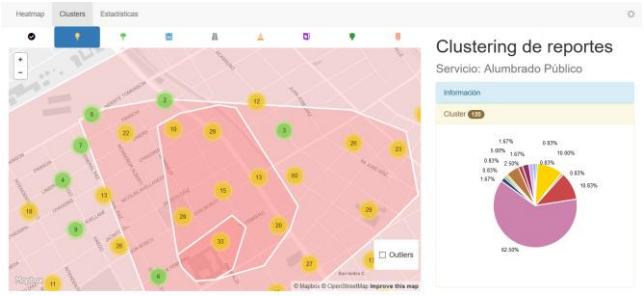

Figure 3: Hierarchical greedy clustering of reports (left) and pie chart for subcategories of the public lighting system (right)

If needed, the user can inspect an individual report by clicking on the marker representing the report. This gives access to additional information of the report, including its location and the comment submitted by the user. This functionality is illustrated in Figure 4.

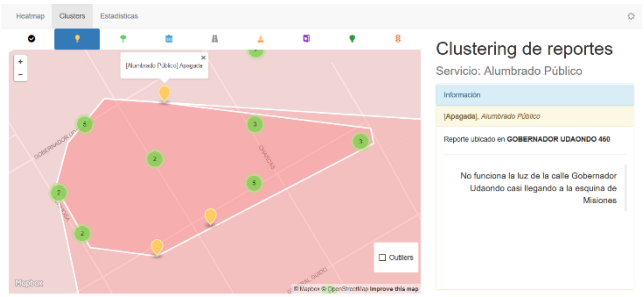

Figure 4: Information associated with a report within a cluster.

The clustering functionality in CitymisVis offers a valuable aid to help define strategies for improved levels of service coordination. In particular, it allows to design better logistics for attending citizen demands by considering the type of service and their geographical vicinity.

3.3 Statistics

Another feature of CitymisVis is the possibility of accessing to statistics computed over the dataset and to visually represent these statistics by means of zoomable sunburst charts. A sunburst is a tree-like (hierarchical) representation that uses a radial layout. The root node of the tree is at the center while the leaves lay on the circumference. The area of each slice corresponds to its value. Figure 5 presents a sunburst summarizing the proportions of reports related to all the service categories and subcategories in the San Isidro municipality.

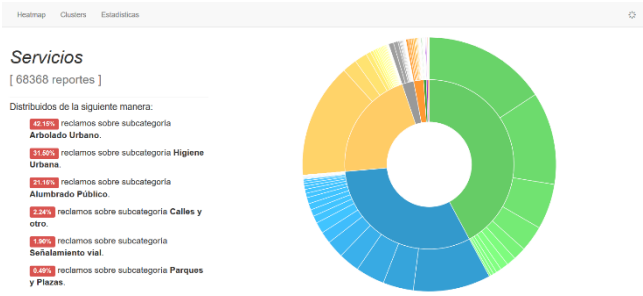

Figure 5: Sunburst chart for the reports associated with all the service categories and subcategories.

Categories and subcategories in the sunburst can be identified by hovering the mouse over them. A special feature of these sunbursts is that they are "zoomable". In other words, they offer the possibility of clicking on a slice at any level to produce a new sunburst for the selected category or subcategory. Figure 6 illustrates this feature by presenting the statistical information and the sunburst corresponding to the public lighting system category. The *zoomable sunburst with updating data* implemented with *d3* [12] is used to generate these charts.

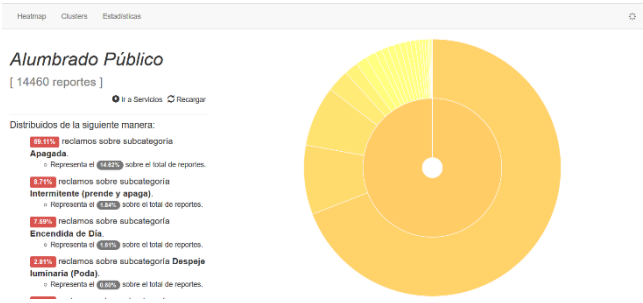

Figure 6: Sunburst resulting from zooming in on the public lighting system category.

The use of zoomable sunbursts results in an intuitive way to present statistics in a visual hierarchical structure that is easy to navigate. As a consequence, it offers a suitable tool to get a clear vision of the dimensions of the problems in a municipality based on their corresponding service categories and subcategories.

4. RELATED WORK

Several platforms have been developed to offer an access point for a two-way communication process between governments and citizens. These platforms allow citizens to report issues related to service delivery, healthcare, education, public transportation, safety, and security, among others (e.g. [4,5,6,13]). The major benefit of these platforms is the reduction of response time and improvements in the quality of service delivery [14,15]. A number of studies have attempted to gain insights into prevailing problems and trends based on the analysis of data collected from citizen reports (e.g. [16,17]). The analyses carried out in many of these studies apply traditional methodologies, such as geospatial analysis and regression. The application of data mining techniques, such as clustering, association rule analysis and metaheuristics to identify citizen priorities and monitor citizen perception has been addressed by a number of proposals (e.g. [18,19]). The use of data

mining techniques is sometimes complemented by the application of business intelligence methodologies, such as the investigation of key performance indicators, which can be used as quantitative indicators of the perception of citizens on different urban-related issues [20]. However, different from CitymisVis, the proposals identified in the literature carry out the analysis offline rather than through specially developed interactive and visual tools.

5. CONCLUSION

In this paper we have presented CitymisVis, a prototype that helps to identify prevailing problems in a municipality. CitymisVis supports the analysis of citizen complaints and reports through the application of specifically developed methods based on clustering and visualization techniques. These methods include the use of heatmaps to represent the intensity of reports at geographical points, the application of density-based and hierarchical clustering to visualize groupings of reports at different granularity levels, and the use of zoomable sunbursts to produce navigable representations of statistics obtained from the collected data. Through these methods, CitymisVis offers a valuable aid to help recognize the most affected areas, plan logistics for attending citizen demands in an efficient manner, design strategies to timely and proactively improve the service infrastructure in a community, quantitatively infer cause-effect relations resulting from municipality's public policies by evaluating concrete actions in delimited geographic areas, and produce surveys to keep citizens informed by visual data and statistics. Accordingly, the use of CitymisVis leads to a more efficient internal organization by lowering maintenance costs. It also results in a better use of public resources by leading to smarter decisions and better investments with measurable quantitative effects. Finally, it helps to improve the political communication, and hence to develop understanding between citizens and government officials.

As part of our future work we plan to include a temporal dimension to the clustering algorithms to allow analyzing reports in a specific window of time. We also plan to extend the analysis offered by the tool by considering additional sources of data such as citizen opinions reflected in social media or comments from readers of digital newspapers. The incorporation of these new data opens new research challenges, such as developing tailored methods for topic-based search, named-entity recognition, sentiment analysis, etc. Finally, we plan to conduct experiments and user studies to assess the effectiveness of CitymisVis.

6. ACKNOWLEDGMENTS

We thank the municipality of San Isidro for permission to use their data and Elsa Estevez for her fruitful suggestions. This research was partially supported by LACCIR (Microsoft Research Latin America) through Project R1211lAC004, PICT-ANPCyT 2014-0624, PIP-CONICET 112-2012010-0487, and PGI-UNS 24/N039.

7. REFERENCES

[1] Oliveira, G. H. M. and Welch, E. W., 2013. Social media use in local government: Linkage of technology, task, and organizational context. Government Information Quarterly 30.4. 2013: 397-405.

[2] Hong, H., 2013. Government websites and social media's influence on government-public relationships. Public Relations Review 39.4. 2013: 346-356.

[3] Steibel, F. and Estevez E., 2015. Designing Web 2.0 tools for online public consultation. Impact of Information Society Research in the Global South, 2015, pp 243-263, Springer.

[4] Fedotova, O., Teixeira, L., and Alvelos, H., 2012. E-participation in Portugal: evaluation of government electronic platforms, Procedia Technology, Volume 5, 2012, 152-161.

[5] Desouza, K. C. and Bhagwatwar, A, 2012. Citizen apps to solve complex urban problems. Journal of Urban Technology 19.3. 2012: 107-136.

[6] Sun, P.-L., Ku, C.-Y., and Shih, D.-H., 2015. An implementation framework for E-Government 2.0, Telematics and Informatics, Volume 32, Issue 3, August 2015, 504-520.

[7] Gamboa, F., Malamud, M., Maguitman, A., and Chesñevar, C., 2016. Citymis OpTree: Intelligent Citizen Management Using Sentiment Analysis and Opinion Trees. ICEGOV 2016, Montevideo, Uruguay, March 1-3, 2016

[8] Mapbox Developers Tools. Retrieved 20 August 2016, from https://www.mapbox.com/developers/

[9] Ester, M, Kriegel, H.-P., Sander, J., and Xu, X., 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. KDD. Vol. 96. No. 34. 1996.

[10] ELKI Data Mining Framework. Elki.dbs.ifi.lmu.de. Retrieved 20 August 2016, from http://elki.dbs.ifi.lmu.de/

[11] Agafonkin, V. Clustering millions of points on a map with Supercluster. Mapbox. Retrieved 20 August 2016, from https://www.mapbox.com/blog/supercluster/

[12] Zoomable sunburst with updating data. Bl.ocks.org. Retrieved 20 August 2016, from http://bl.ocks.org/kerryrodden/477c1bfb081b783f80ad

[13] Chakravarti, B and Venugopal, M., 2008. Citizen Centric Service Delivery through e-Governance Portal. A White Paper published by National Institute for Smart Government Hyderabad, India.

[14] Dirks, S., Gurdgiev, C., and Keeling, M., 2010. Smarter cities for smarter growth: How cities can optimize their systems for the talent-based economy. IBM Institute for Business Value.

[15] Dahiya, D. and Mathew, S. K., 2013. Review of Strategic Alignment, ICT Infrastructure Capability and E-Governance Impact.

[16] Clark, B. Y., Brudney, J. L., and Jang S.-G., 2013. Coproduction of government services and the new information technology: Investigating the distributional biases. Public Administration Review 73.5. 2013: 687-701.

[17] Offenhuber, D., 2014. Infrastructure legibility: a comparative analysis of open311-based citizen feedback systems. Cambridge Journal of Regions, Economy and Society.

[18] Mostafa, M, Alesheikh, A. A., and Saeidian, B., 2016. Analyzing public participant data to evaluate citizen satisfaction and to prioritize their needs via K-means, FCM and ICA. Cities 55. 2016:70-81.

[19] Panagiotou, N., Zygouras, N., Katakis, I., Gunopulos, D., Zacheilas, N., Boutsis, I., Kalogeraki,V., Lynch, S., and O'Brien., B., 2016. Intelligent Urban Data Monitoring for Smart Cities. In Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pp. 177-192. Springer International Publishing, 2016.

[20] Cagliero, L., Cerquitelli, T., Chiusano, S., Garino, P., Nardone, M., Pralio, B., and Venturini, L., 2015. Monitoring the citizens' perception on urban security in Smart City environments. ICDEW 2015. IEEE, 2015.