Towards Understanding Real-Estate Ownership in New York City: Opportunities and Challenges

[Extended Abstract]

Tuan-Anh Hoang-Vu
New York University
tuananh@nyu.edu

Vicki Been
Furman Center for Real Estate & Urban Policy
vicki.been@nyu.edu

Ingrid Gould Ellen
Furman Center for Real Estate & Urban Policy
ingrid.ellen@nyu.edu

Max Weselcouch
Furman Center for Real Estate & Urban Policy
max.weselcouch@nyu.edu

Juliana Freire
New York University
juliana.freire@nyu.edu

1. MOTIVATION AND OPPORTUNITIES

Understanding who is investing in real estate, and the patterns of their investments, is critical both for assessing the need for, and the effects of, policy interventions by governments, lenders, and non-profit community development organizations. If we knew more about patterns of ownership, for example, we could target buildings that seem to "produce" disproportionate numbers of homeless families seeking housing in the City's shelter system. Many other policies could also be made more effective, including policies related the city's property tax assessment, water and sewer lien practices, outreach to small property owners for energy upgrades, enforcement of rent regulation rules and policies to encourage investment through tax subsidies and zoning changes. Not only is understanding these patterns critical for deciding how to target interventions, but they may have significant implications for the effectiveness of interventions ranging from policing to the opening or closing of schools to land use policies such as inclusionary zoning.

Researchers have been stymied in their attempts to understand patterns of ownership for several reasons. Understanding the patterns of investments requires attention to a large number of variables – from the timing and location of the purchases, to types of housing, the financing used, and evidence of deterioration in the housing or neighborhood. To identify patterns, it is necessary to inspect a large number of data slices, over time and space, and perform many different analysis over the different facets of the data. This problem is compounded due to the fact that the ownership of properties is opaque. Generally, properties are owned by limited liability companies created specifically for the purpose of holding one particular building. While common individuals behind multiple LLCs will reveal overlapping ownership, information about the officers of the LLCs is not easily ascertainable. Furthermore, even though property purchases and mortgage deeds are available on-line, it is challenging to identify names that are likely the same person.

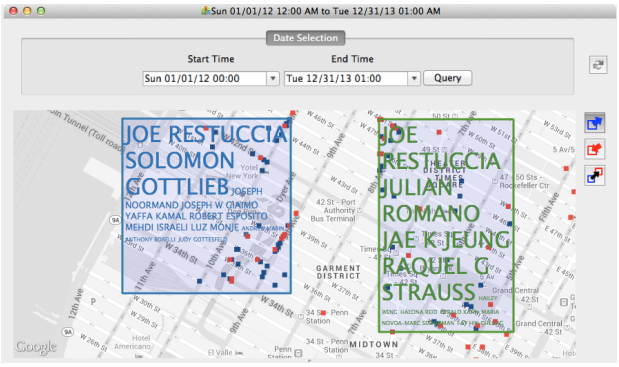

Figure 1: Exploring property owners through tag-cloud lenses

In what follows, we describe preliminary work we have done on providing support for exploration of property data through visualization. We also discuss some open problems and directions for future research.

2. A FIRST STEP: VISUALLY EXPLORING PROPERTY DATA

With the goal of helping social scientists and policy experts to identify patterns of ownership in rental properties in New York City, we are developing a framework that allows social scientists and policy experts to interactively explore multiple data sets that contain information related to ownership. For our initial prototype, we used several data sets collected by The Furman Center [4], including data on building characteristics, property sales and mortgages, records of building code violations and evictions, and information about the unused zoning capacity on the lots.

Building upon our previous work on the exploration of spatio-temporal data [3], we have developed a series of visual interfaces through which users can explore these data and answer questions that help understand ownership patterns,

DSMM'14, June 27, 2014, Snowbird, Utah, USA
ACM 978-1-4503-3012-1/14/06.

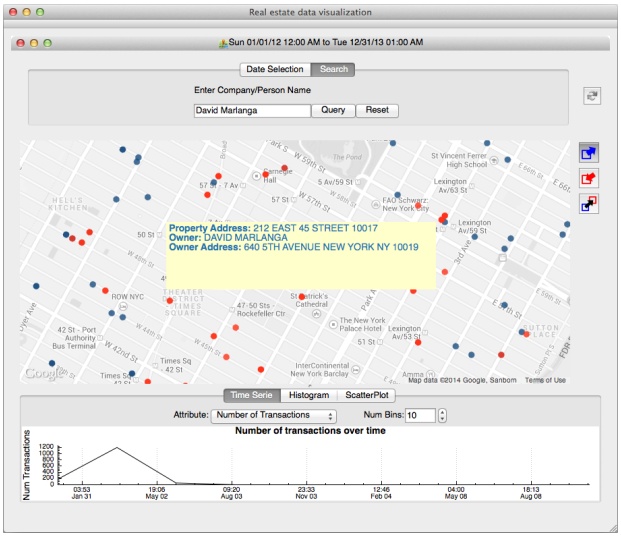

Figure 2: Analyzing buying and selling patterns for an owner. This interface allows users to select data slices by visually specify spatio-temporal queries, and analyze their results in a single step.

including exploring the portfolio of different owners, the set of owners in a region, and ownership changes over time.

To support the visual specification of spatial queries, we have adapted tag-cloud lenses (TCLs) [2]. These lenses provide an easy-to-use, interactive mechanism for researchers to explore multiple variables in a geographical context. As illustrated in Figure 1, users can place a lens on a region (or multiple regions) as well as interactively move them over a map. Each lens-region combination corresponds to a spatial query that retrieves the names of the owners in the selected region. The results are displayed in a tag-cloud where the scaling corresponds to the number of buildings owned by a given entity.

The interface also supports analyses that look for trends in the data over time. Figure 2 shows one such analysis, where we examine the sales and purchases for one owner in a period of one year (1/1/12–12/31/13). Sales and purchases are represented as blue and red dots, respectively. Users can also restrict their analysis to particular regions (e.g., zip codes and neighborhoods). One important feature of this interface is that it combines data selection and analytics, allowing users to analyze the data without losing the spatio-temporal context. When the user selects a data slice (represented by a region and time interval), the contents of the slice are displayed using the visual representation of their choice. In this figure, we show a line plot that summarizes the number of transactions over time for the selected owner.

3. CHALLENGES AND OPEN PROBLEMS

Identifying Owners.

The information the users visualize shows the owners as they are listed in the different data sets. However, a number of these actually correspond to the same person, company or holding. To provide a more accurate view of the ownership patterns, we are investigating techniques to match properties and owners across multiple datasets. This is an instance of the entity matching problem. We have implemented a simple, rule-based entity matching approach which attained promising results. However, automated methods such as the one we used are not fool-proof. We are currently investigating mechanisms whereby users can help in the reconciliation of owner information as they visually explore the data.

On-the-Fly Data Integration.

As researchers analyze data and find interesting patterns, they often have to obtain additional data sets to try and explain their findings. For example, to understand the cause of a big spike in the number purchase transactions, it may be necessary to correlate that against mortgage rates. However, the level of effort required to gather and integrate data is staggering. While there has been substantial work on data integration [5], we still lack techniques and tools that support on-the-fly integration.

Automatic Event Detection.

Given the large number of data slices present in many urban data sets, manual exploration can be impractical. Aggregation of the data could overcome this problem, but might result in the occlusion of small or local patterns of the data [1]. Hence, it is important to help *guide* users towards interesting patterns and/or data slices to enable efficient analysis of the data. Additionally, once a pattern is identified, searching the data for occurrences of this pattern is also a non-trivial task. We will explore techniques for automatically identifying events, as well as interfaces that guide users towards interesting data slices and support querying over these events, for example, to find similar events in the data.

4. ACKNOWLEDGMENTS

This project is funded by an NYU Provost Seed Grant.

5. REFERENCES

[1] G. Andrienko, N. Andrienko, C. Hurter, S. Rinzivillo, and S. Wrobel. From movement tracks through events to places: Extracting and characterizing significant places from mobility data. In *Visual Analytics Science and Technology (VAST), 2011 IEEE Conference on*, pages 161–170. IEEE, 2011.

[2] N. Ferreira, L. Lins, D. Fink, S. Kelling, C. Wood, J. Freire, and C. Silva. Birdvis: Visualizing and understanding bird populations. *IEEE Transactions on Visualization and Computer Graphics*, 17:2374–2383, 2011.

[3] N. Ferreira, J. Poco, H. T. Vo, J. Freire, and C. T. Silva. Visual exploration of big spatio-temporal urban data: A study of new york city taxi trips. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2149–2158, 2013.

[4] Furman center for real estate & urban policy. http://furmancenter.org.

[5] A. Halevy, A. Rajaraman, and J. Ordille. Data integration: The teenage years. In *VLDB*, pages 9–16, 2006.