International Journal of E-Planning Research
Volume 8 • Issue 1 • January-March 2019

Mining Social Media to Measure Neighborhood Quality in the City of Atlanta

Subhrajit Guhathakurta, Center for Spatial Planning Analytics and Visualization, Georgia Institute of Technology, Atlanta, USA
Ge Zhang, Georgia Institute of Technology, Atlanta, USA
Guangxu Chen, Georgia Institute of Technology, Atlanta, USA
Caroline Burnette, Georgia Institute of Technology, Atlanta, USA
Isabel Sepkowitz, Georgia Institute of Technology, Atlanta, USA

ABSTRACT

This article presents a model to classify perceptions of various Atlanta neighborhoods based on social media. Tweets were extracted using Twitter's API and categorized to determine 1) whether they are neighborhood related; 2) whether a positive or negative sentiment could be assigned, and 3) whether they belong to one of eight categories of neighborhood quality assessments. These eight categories are public safety, transportation, density, walkability, maintenance, aesthetics, open space, and quality of dining and entertainment venues. Tweets that were related to neighborhood quality and geo-tagged accounted for 4% of all filtered Tweets. Overall 49% of neighborhood perception related Tweets were extracted to create an indicator of perceived neighborhood quality. The study then compared the perception of neighborhoods from social media analysis with quantitative indicators of neighborhood quality.

KEYWORDS

Amenities, Machine Learning, Neighborhood, Perception, Quality of Life, Social Media

1. INTRODUCTION

Since neighborhood quality is an important attribute of residents' quality of life, choosing the right neighborhood is a critical task undertaken by households at one or more points during their lifecycle (Sirgy and Cornwell, 2002). Given that neighborhood quality is closely related to housing satisfaction, moving to a new area requires substantial research about the potential neighborhoods where a household might choose to live (Lee et al., 2008; Lovejoy et al., 2010; Lu, 1999; Oakley et al., 2013). During 2015-2016, around 1 out of 9 people in the U.S. moved to a new residence, and this statistic has been consistent in the recent past (U.S. Census 2016). The perception of a neighborhood is also closely tied to housing values (McCluskey and Rausser, 2001; Poor et al., 2001). Housing in desirable neighborhoods tends to maintain high resale values compared to similar housing in less desirable areas. Also, a household's social status is often partly derived from the perceived quality of the neighborhood where the household is located. While neighborhood quality matters for households'

DOI: 10.4018/IJEPR.2019010101

financial and social status, changes in neighborhood quality as perceived may precede quantitative measures of neighborhood quality. Therefore, it is important to track the perception of neighborhood quality as an early indicator of neighborhood transition.

There is a sizable literature on the development and application of quantitative indicators of neighborhood quality (Frank et al., 2006, 2010; Gupta et al., 2012; Lalloué et al., 2013; Maly, 2000). Many such neighborhood indicators have been operationalized through web-based tools, such as RentLingo (http://www.rentlingo.com/), Zillow (http://www.zillow.com), Atlanta's Neighborhood Quality of Life and Health Portal (www.cgis.gatech.edu/NQOLH) and the Charlotte-Mecklenburg Quality of Life Explorer (https://mcmap.org/qol/), among others. However, these objective data-based indicators fail to capture people's perceptions about their own and other neighborhoods, which is often the basis for their behavior and decisions (De Jesus et al., 2010; Inoue et al., 2010; Mota et al., 2005; Rhodes et al., 2006). While the perceived neighborhood quality has been examined with the help of surveys (Dawson et al., 2007; Hale et al., 2013), these can be expensive and lengthy processes. Also, the results of the survey can be biased if that survey is not conducted in close to ideal conditions with statistically determined minimal sample size and comprehensive coverage (Olson, 2006). In this study, we demonstrate the potential of crowdsourcing as a complementary solution for obtaining information about neighborhood perceptions which can reduce the cost of the traditional survey and its potential for biased results. By augmenting traditional surveys with crowdsourcing, we can easily validate the survey results with limited effort and resources. Both the promise and the pitfalls of this approach are discussed in this paper.

Social media sites, such as Instagram, Twitter or Facebook, have become a popular data source for extracting information on people's attitudes and perceptions (Chou et al., 2009; Endarnoto et al., 2011; Olsen and Christensen, 2015; Yates and Paquette, 2011; Zeitzoff, 2011). Of the various social media platforms, Twitter is especially useful for this study given that it has many more users than most other social media sites. Currently, Twitter produces 500 million Tweets per day from 100 million active users (Protalinski, 2013). The data from Twitter are freely available for downloading with the help of their Application Processing Interface (API). In the following sections, the methodology used to collect, process, and classify the neighborhood quality related Tweets is discussed. The perceived neighborhood characteristics are then extracted from those Tweets. The methodology was applied to the City of Atlanta. The Twitter-based qualitative assessment is then compared to the quantitative measurement of the Neighborhood Quality of Life and Health Portal (http://www.cgis.gatech.edu/NQOLH/) to observe whether neighborhood perception aligns with objective data.

2. A REVIEW OF RELEVANT LITERATURE

Prior studies have found that neighborhood perceptions are often instrumental in understanding the behavior and health of residents (Ferreira, César, Camargos, Lima-Costa, & Proietti, 2010; Gary et al., 2008; Weden, Carpiano, & Robert, 2008). For example, Hume et al. (2005) investigated children's perceptions of their environments and found significant associations between these perceptions and objectively measured physical activity. Several studies assessing the walkability of places also concluded that positive perceptions of the built environment contributed to increased walking (Nagel et al., 2008). The relationship between perception of environmental factors and physical as well as mental health seem to cut across different demographic groups, characterized by attributes such as age and gender (Ball, Bauman, Leslie, & Owen, 2001).

Typically, neighborhood perception analysis is based on data collected through surveys, which are administered via telephone, mail, and the Internet, or via face to face interviews (Burdette & Whitaker, 2005; Hartnagel, 1979; Subramanian & Kennedy, 2009). Examples of such studies include: 1) research by Timperio et al. (2004) which surveyed children aged 5–6 years and 10–12 years from 19 Australian primary schools to obtain their perceptions about the local neighborhood (Timperio, Crawford, Telford, & Salmon, 2004); 2) a study by Boehmer et al. (2007) in which telephone surveys