Technological Forecasting & Social Change xxx (2013) xxx–xxx

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

Predictive analytics can facilitate proactive property vacancy policies for cities☆

Sheila U. Appel, Derek Botti, James Jamison, Leslie Plant, Jing Y. Shyr, Lav R. Varshney*

IBM Thomas J. Watson Research Center, 1101 Kitchawan Road, Yorktown Heights, NY 10598, United States

ARTICLE INFO

Article history:
Received 4 September 2012
Received in revised form 6 July 2013
Accepted 28 August 2013
Available online xxxx

Keywords:
Property vacancy
Systems of systems
Predictive analytics
Urban planning

ABSTRACT

Is it possible for a city to understand, analyze, predict, and therefore prevent vacant properties? In this paper, we demonstrate the feasibility of using techniques from machine learning and data mining to determine the future vacancy risks for individual properties and for neighborhoods using a variety of structural, demographic, socioeconomic, and city activity features with high accuracy. Within a larger systems-of-systems framework that we develop, these predictive analytics will allow a city to move from decision-making based on 'educated anecdotes' and reactive strategies aimed at the most urgent need, to policy development based on informed, holistic insight and proactive interventions that prevent and reverse decline. A demonstration of the use of predictive analytics within the sociotechnical system is provided using data from Syracuse, New York.

1. Introduction

Many cities in the industrial Northeast and Midwest regions of the United States, the so-called rust belt, have seen the proliferation of rusting factories, declining home prices, population losses, high unemployment, and economic malaise [2]. As there has been an outmigration of jobs and people from city centers, one of the major consequences has been a rise in vacant residential properties [3]. Elected and appointed officials in many of these cities perceive property vacancy as a major problem that affects all citizens [4].

Although abandoned homes are symptomatic of other problems, they also contribute to neighborhood decline and frustrate revitalization, e.g. in Baltimore, Maryland [5]. Indeed, housing abandonment can attract criminal activity, lead to an increased risk of residential fire, and lead to unwelcome public health trends, independently of the socioeconomic status of the area [6]. There are a variety of policy actions that can be taken to address property vacancy [7], but require an understanding of underlying causes.

Broadly speaking, we have found that property vacancy can be linked to three hierarchical levels of cause. An overarching factor is regional population dynamics: when people leave a region it results in a mismatch between supply of housing stock and demand for housing. Local spatial factors such as nearness to bus routes, schools, and grocery stores, as well as blight and crime in local neighborhoods are also contributing factors. Finally, there are property-specific factors such as the floor area, the number of bathrooms, and the owner's residency status.

Several previous studies have also attempted to understand causes for property vacancy. Bassett et al. found that housing abandonment in Flint, Michigan is not due to any single cause but is significantly related to a variety of economic, spatial, and demographic factors [8]. In Buffalo, New York, Silverman et al. found that the vacant residential property rate of a census tract increases with the poverty rate, the rate of renters receiving rental assistance, and higher percentages of business addresses [9]. In Philadelphia, Pennsylvania, Hillier et al. found that outstanding housing code violations, and tax arrearages, as well as characteristics of nearby properties were predictive of abandoned properties [10]. They also developed a basic predictive algorithm.

☆ This paper is based in part on a technical report [1], prepared in Nov. 2011.

* Corresponding author at: IBM Thomas J. Watson Research Center, 1101 Kitchawan Road, Yorktown Heights, NY 10598, United States.
E-mail address: lrvarshn@us.ibm.com (L.R. Varshney).

0040-1625/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.techfore.2013.08.028

As part of the IBM Smarter Cities Challenge (www.smartercitieschallenge.org), we worked with the government of the City of Syracuse, New York to help understand, analyze, predict, and therefore prevent vacant residential properties. The city's goal is to move from decision-making based on 'educated anecdotes' and reactive strategies aimed at the most urgent need, to policy development based on informed, holistic insight, and proactive interventions that prevent and reverse decline. Specifically, Syracuse asked us how to:

1. identify indicators for factors contributing to the causes of property vacancy,
2. integrate and analyze relevant data from disparate sources across a broad ecosystem of stakeholders, and
3. develop a predictive, flexible model to show the impact various events or actions could have on a neighborhood's stability.

In this paper, we report our results.

Since Syracuse does not have tens of thousands of abandoned houses like Detroit, Michigan or Philadelphia, the problem seems more manageable [11] and a proactive approach based on data-driven risk forecasting more amenable to affecting social change.

In developing a proactive approach to the residential vacancy problem, we developed a systems-of-systems framework drawing on the field of engineering systems [12]. Within this framework, we defined a specific information technology architecture that would support the data gathering, data analysis, and knowledge dissemination necessary for the various city departments, regional agencies, not-for-profit organizations, and citizen groups to work together. Such an information technology system would also integrate into the several city subsystems, leading to coordinated preventative actions.

Government databases and data warehouses in Syracuse and elsewhere are siloed, and so it was a non-trivial task to bring different data sources together. Furthermore, many useful data were held by not-for-profit organizations rather than by government departments. Data, such as those maintained in codes enforcement systems, in housing partner data systems, and in police intelligence systems, are meant for specific tasks within the realm of each specific system. However, these data reveal a comprehensive view of the city if combined. Our research and development focused on combining and exploiting these data for the development of a predictive analytics solution.

The keystone of the system is a predictive analytics subsystem with algorithms drawn from machine learning and data mining [13–15]. Indeed, many of the algorithmic techniques developed for business analytics and service delivery in the private sector can be adapted almost directly to municipal government service delivery [16]. Our predictive algorithms operate both to identify neighborhoods that are on the bubble with respect to vacancy and to identify individual vacant properties that should by all rights be occupied. These neighborhoods and individual properties are where most effort should be devoted. Consistent and unbiased estimates of predictive accuracy demonstrate that our algorithms will be very effective. Important factors include male unemployment rates and nuisance crime rates in neighborhoods, and housing code violations in individual properties.

In closing this introductory section, let us note that we believe that the analytical frameworks and techniques described herein may be applicable throughout the rust belt and beyond. The reason for this belief is because there are several universal laws which describe how cities are structured, how they behave, and how they evolve [17–19]. By deriving insight from the vast oceans of data that are now starting to be collected and collated with digital devices, cities can positively affect issues that fundamentally the quality of their citizen's lives and become smarter [20,21].

2. City of Syracuse

The City of Syracuse had evolved from a crossroads settlement in the early nineteenth century to a bustling industrial and transportation hub by 1900. Since the 1950s Syracuse has grappled with balancing the ebb and flow of its population and aligning that to its housing stock. At its peak, Syracuse was home to 250,000. While the population has stabilized in the past 10 years, the distribution of those residents has followed a common trend among many cities, particularly those in the rust belt: the outmigration of jobs and people from the city center to suburbs, as shown in Figs. 1 and 2.

The City of Syracuse is located in the geographic center of New York State within Onondaga County. More than 85% of the 42,000 parcels in the City of Syracuse are residential in nature. There are roughly 25,000 single family homes in the City and an additional 10,000 multi-unit residential structures housing more than 60,000 households. The nature, type, and condition of these residential uses vary widely but all fit together to form a patchwork of neighborhoods that provide a variety of living experiences. Of the total housing units, about 75% were built before 1960 and 47% were constructed in 1939 or earlier. Houses built after 1980 make up only 6% of the total. By contrast, only about 53% of housing units in the county were built before 1960, and three times as many houses were built in the county after 1980 than in the city, reflecting a continuation of residential suburban sprawl [22].

Of the approximately 35,000 residential parcels in the city about 1500 are vacant today and the mayor, members of the Common Council, and other civic leaders have identified vacant properties as one issue that unites all Syracusans, regardless of ethnicity, age, income, or education. Their byproducts—blight, crime and declining property values and tax revenues—impact the quality of life of a diverse population which includes refugees, academics, artists, and blue collar workers, among others.

Global economic factors mean Syracuse shares a housing dynamic common among many cities. Declining property values and neighborhood degradation have removed the impetus for many homeowners to upgrade or maintain these properties; under- or unemployment has made it impossible for others. Declining property values have also led to an influx of speculators who, unfamiliar with local market dynamics, purchase rental properties as investments. In fact nearly 50% of Syracuse's housing stock is occupied by renters, compared to 33% occupied by owners. Absentee landlords or poor landlord management has led to rentals becoming untenanted and abandoned, exacerbating the problem.

2.1. Making Syracuse a smarter city

Syracuse offers an ideal opportunity to demonstrate the tenets of a smarter city: its relatively small size and population

Please cite this article as: S.U. Appel, et al., Predictive analytics can facilitate proactive property vacancy policies for cities, Technol. Forecast. Soc. Change (2013), http://dx.doi.org/10.1016/j.techfore.2013.08.028