

CSCI-548 Project Proposal

Lu Zhen, Yilei Qian, Chao Wang

University of Southern California

Abstract. We propose a mashup system that provides useful housing information in Los Angeles. In our proposed system, we not only provides traditional mashup of the housing address and online maps, but also integrates crime reports, public traffic and local business information to help users evaluate the living environment.

1 Background

When it comes to looking for housing, the famous online classified advertisements website, CraigslistTM, may be one of the best source to find information. Due to the simplicity and inconvenience of Craigslist, many mashup websites, including HousingMaps, MapsKrieg and MapCraigs, have been established to mash up Craigslist with other online maps, which facilitates users to find housing. Although these works is impressing, price and location is not the only factor people will care. The safety and convenience is also important for most of housing finders.

2 Goal

For the above sake, we propose a mashup system that provides more useful housing information in Los Angeles. In our proposed system, we not only provides traditional mashup of the housing address and online map, but also integrates other information such as crime reports, public traffic and local business (eg. groceries, restaurants) to help users evaluate the living environment around. Further, we will use artificial intelligence techniques to build an intelligent filter to filtrate housing by safety and convenience.

3 Data Sources

The data sources we planned to mash into our system:

Housing Craigslist¹

Maps Google Maps²

¹ <http://losangeles.craigslist.org/apa/>

² <http://code.google.com/apis/maps/>

Crime Reports Crime Mapping³
Public Traffic Metro⁴, ...
Local Business Yelp⁵, Google Place⁶, ...

4 Methods

Our basic work flow is described as follows:

1. Extract related information from the above data sources
2. Clean and normalize the extracted data
3. Use Geocoding to convert street address to geographic coordinates
4. Mapping the geographic coordinates onto Google Maps

5 More Magic

Rather than simply display all information on the map, we planned to apply more sophisticated unsupervised learning techniques to rank the safety and convenience index based on the crime reports, public traffic and local business. Then we associate the index with the house in the area in order to facilitate users to evaluate the living environment beyond the information obtained from the original house advertisement. We can also try to utilize some data mining methods to dig the underlying relation between certain region and price or any other interesting relation.

³ <http://www.crimemapping.com/>

⁴ <http://developer.metro.net/>

⁵ <http://www.yelp.com/developers/>

⁶ <http://places.google.com/>