Tian-ran Hu, Jie-bo Luo, Henry Kautz, Adam Sadilek, 2016. Home location inference from sparse and noisy data: models and applications. *Frontiers of Information Technology & Electronic Engineering*, **17**(5):389-402
http://dx.doi.org/10.1631/FITEE.1500385

Home location inference from sparse and noisy data: models and applications

Front Inform Technol Electron Eng

Key words: Home location, Mobility patterns, Healthcare

Contact: Tian-ran Hu
E-mail: thu@cs.rochester.edu
ORCID: Tian-ran Hu, http://orcid.org/0000-0003-0086-2447

Motivation

- Home location is key to a better understanding of human lives.
- Many urban computing applications could benefit from knowing precise home location.
- Studing the possibility of predicting home location helps preventing privacy leakage of social media users.

Front Inform Technol Electron Eng

Contribution

- Precise home location prediction (100 m×100 m).
- Experimental results validate that our model outperforms baseline methods on two standard metrics.
- Based on the prediction results, we show several possible applications based on home locations. The results suggest the effectiveness of our model.

From Inform Technol Electron Eng

Method: temporal and spatial features

- Check-in rate
- Check-in rate during midnight
- Last destination of a day
- Last destination with inactive midnight
- PageRank and reverse PageRank scores
- Etc.

Front Inform Technol Electron Eng

Evaluations on two large datasets

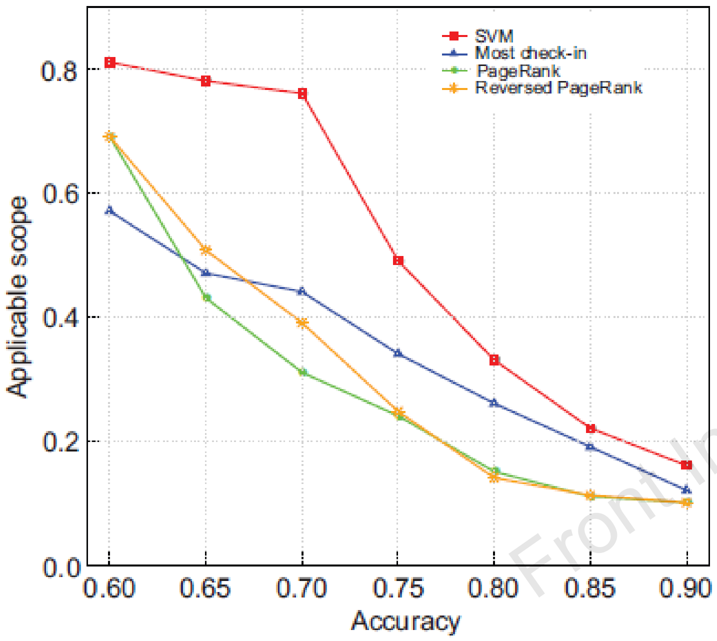

Fig. 4 The applicable scope and accuracies of different methods on the NYC dataset

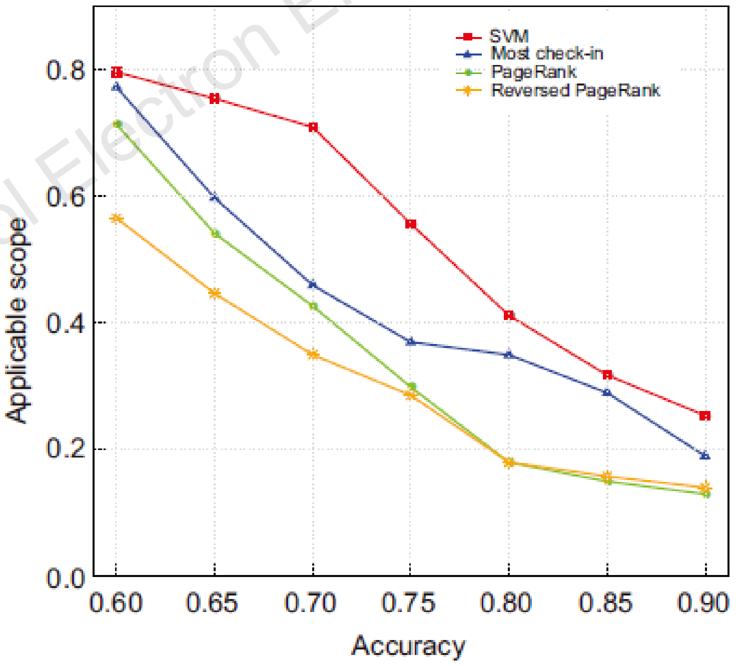

Fig. 5 The applicable scope and accuracies of different methods on the Bay Area dataset

Front Inform Technol Electron Eng

Applications: health condition of a district

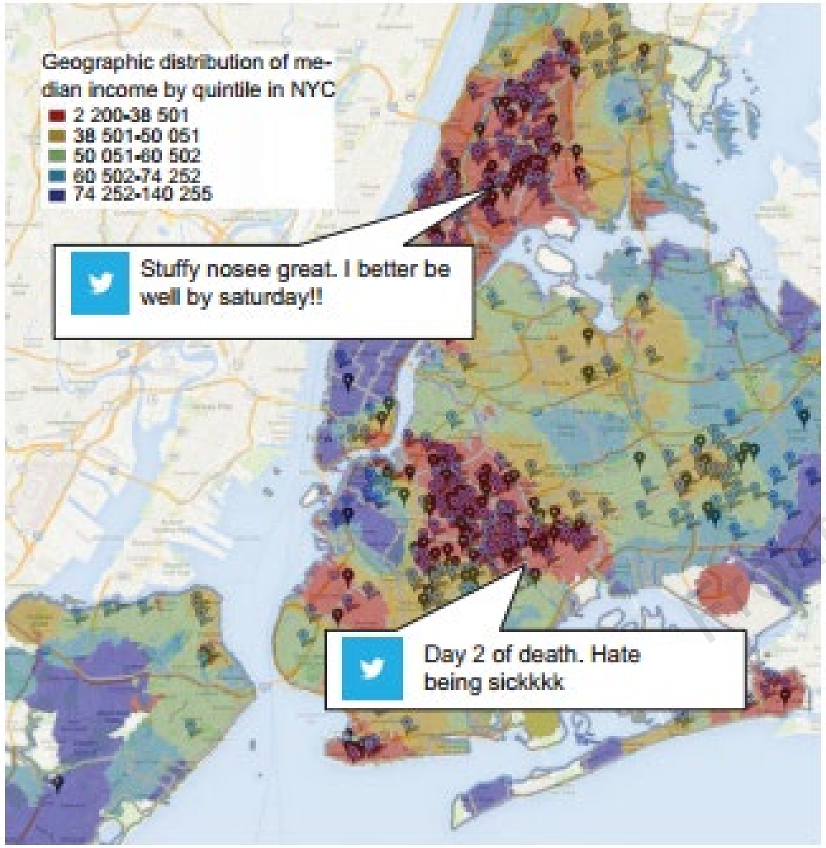

Table 5 The sickness score, percentages of people whose health are in 'excellent' or 'good' condition

Area name	Sickness score	Percentage (%) Excellent	Percentage (%) Good
Upper West	0.046	30.7	26.3
Chelsea	0.018	29.3	20.0
Gramercy	0.029	26.4	19.8
Flatbush	0.100	23.6	30.5
Central Harlem	0.062	23.2	23.9
Lower Man.	0.019	23.2	22.0
Southeast Q.	0.043	22.1	27.6
Astoria	0.042	21.0	35.1
Crown Heights	0.066	20.8	34.7
Heights/Slope	0.061	20.5	27.1
Inwood	0.049	20.2	38.8
Bushwick	0.070	20.0	30.5
Southwest Q.	0.084	16.9	33.5
South Bronx	0.050	13.2	38.6
Fordham	0.083	13.0	46.1
Pelham	0.085	10.5	38.7
Correlation	NA	−0.569	**0.601**

Inform Technol Electron Eng

Applications

Factors that affect health

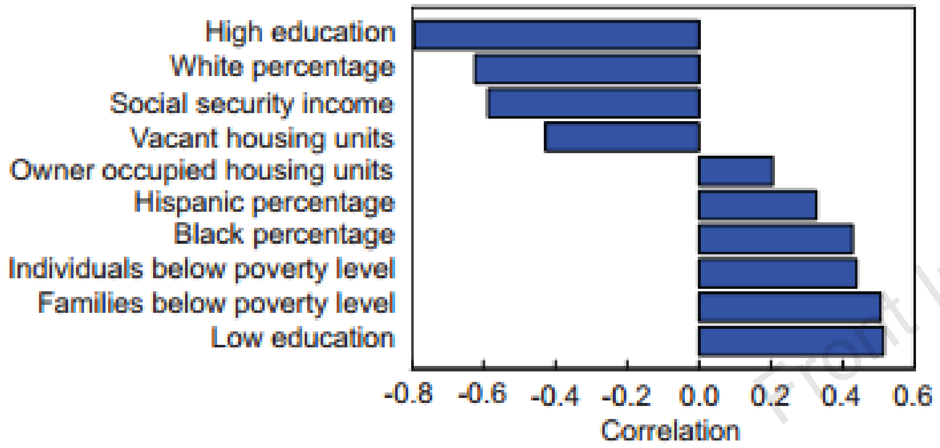

Fig. 10 Correlations between health and factors including poverty level, education, race, and income

City buzz

Table 3 Buzz by home and buzz by location in different cities of the Bay Area

City name	Home buzz	Location buzz
Mountain View	#qa	#mountainview #followmeskip #skipfollowme
Fremont	#hospitality	#fremont
San Mateo	#webdesign	#sanmateo
Emeryville	#backbenoit #emeryville	#emeryville
Sunnyvale	#engineering #healthcare #apocono	#sunnyvale #apocono
Santa Clara	#marketing	#santaclara
San Leandro	#love #fashion	#sanleandro
Redwood	#manufacturing #geekcamp #tryhard	#redwoodcity #tryhard #scifi
Hayward	#healthcare #accounting #letsgooakland	#hayward #green

Front Inform Technol Electron Eng