北京城市实验室
Beijing City Lab

ID of the slides

45

Slides of BCL

www.beijingcitylab.com

How to cite

Author(s), Year, Title, Slides at Beijing City Lab, http://www.beijingcitylab.com

E.g. Long Y, 2014, Automated identification and characterization of parcels (AICP) with OpenStreetMap and Points of Interest, Slides at Beijing City Lab, http://www.beijingcitylab.com

LifeSpec.

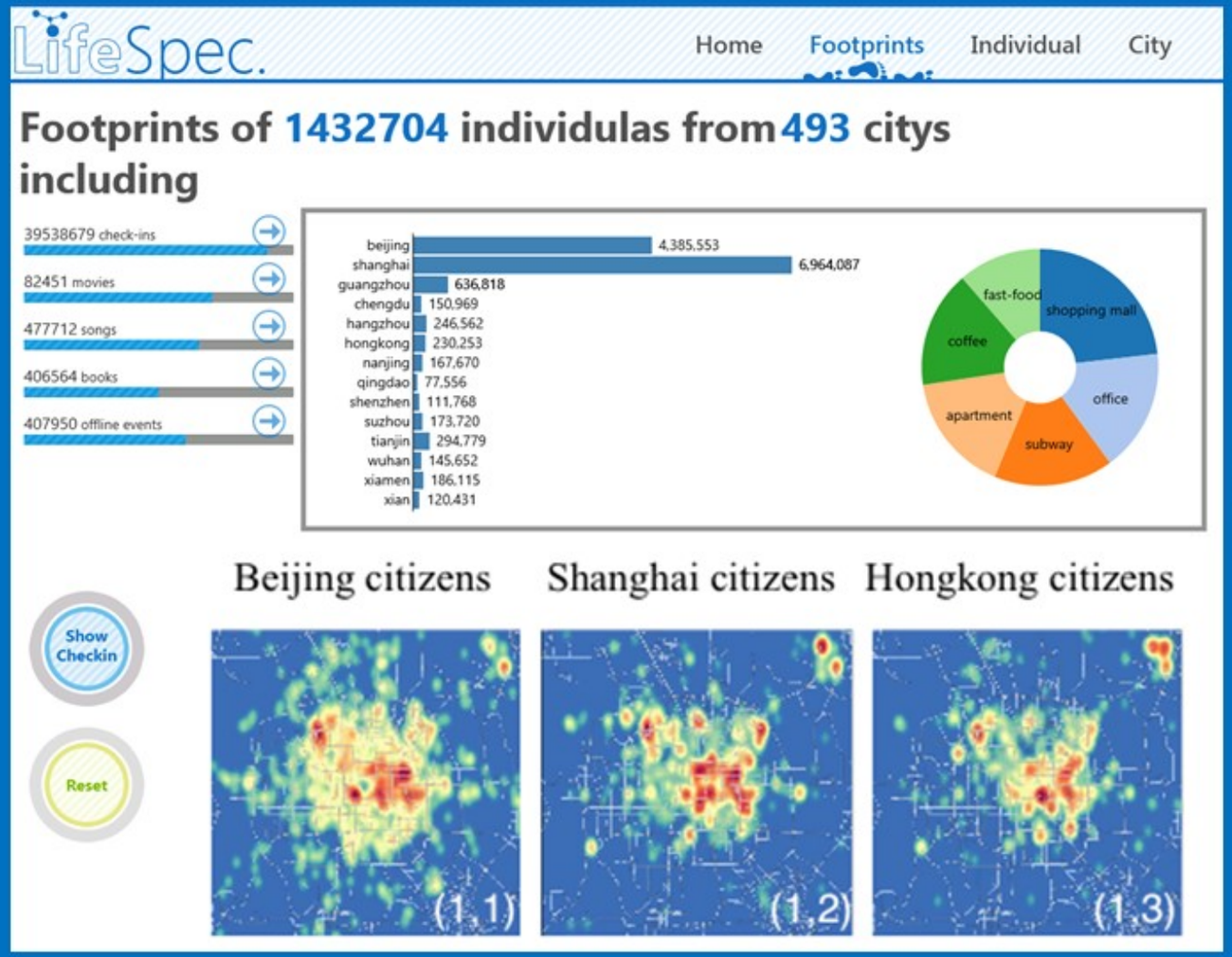
Exploring the Spectrum of Urban Lifestyles

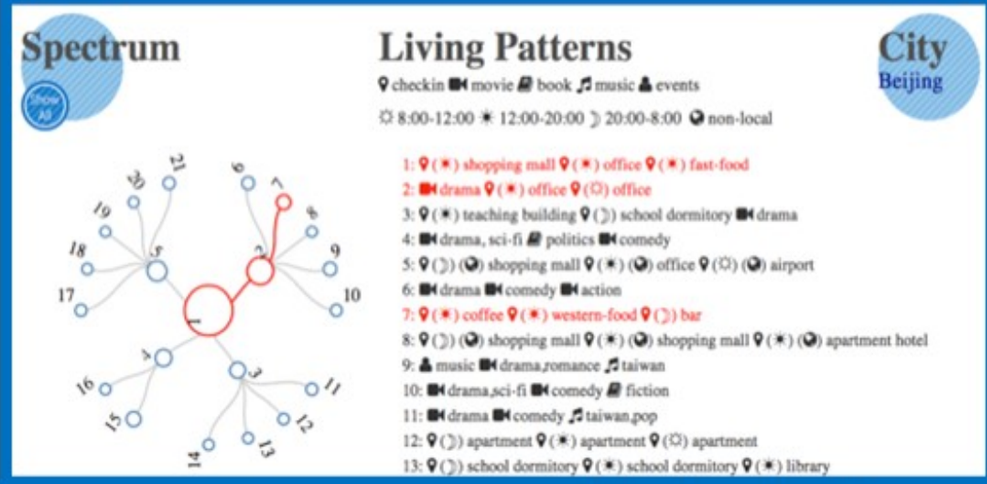
Nicholas Yuan[1], Fuzheng Zhang[1], Defu Lian[1], Kevin Zheng[2], Siyu Yu[3], and Xing Xie[1]

[1]Microsoft Research

[2]School of Information Technology & Electrical Engineering, The University of Queensland

[3]Department of Sociology, University of California, Berkeley



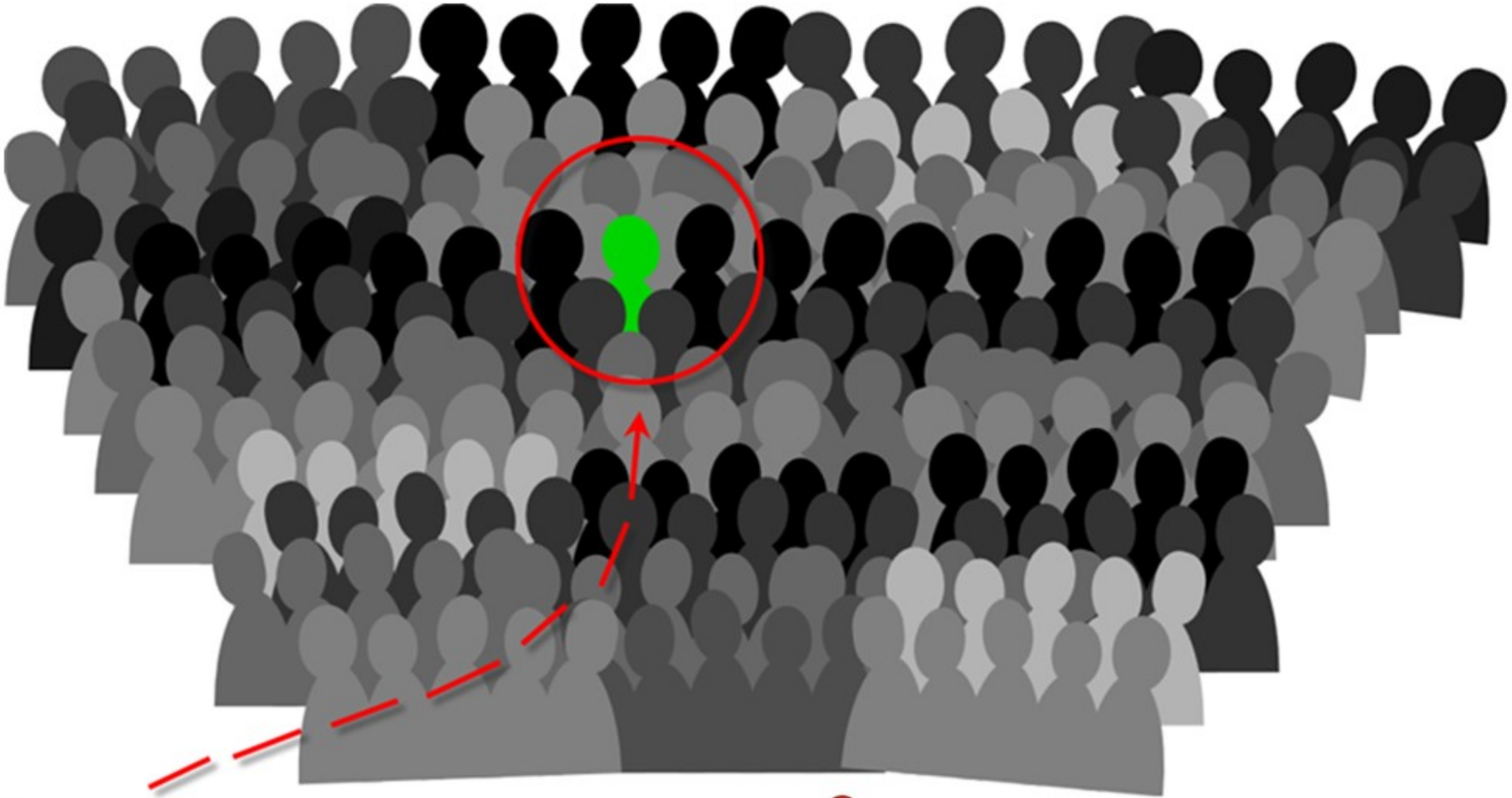


http://mobius/lifespec

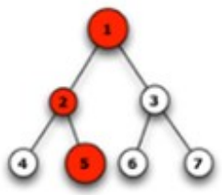
- Lifestyle specification of an individual
- Lifestyle spectrum of a group of individuals

LifeSpec.

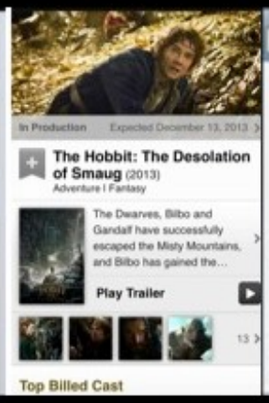
"Consumer lifestyles are strong indicators of their buying behaviors"
—Lifestyle marketing: reaching the new American consumer.

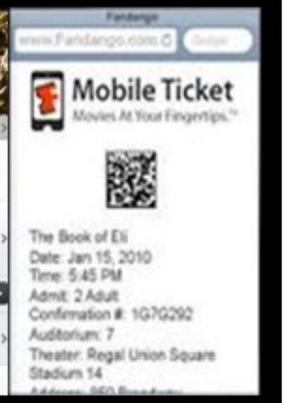


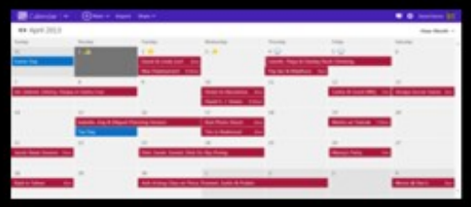
Targeting Computationally

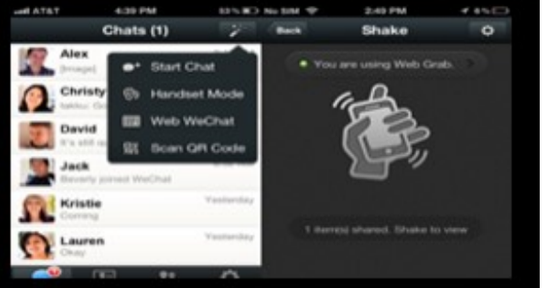


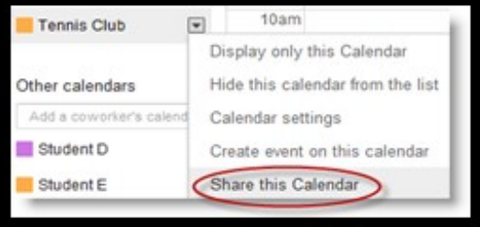
FOOTPRINT





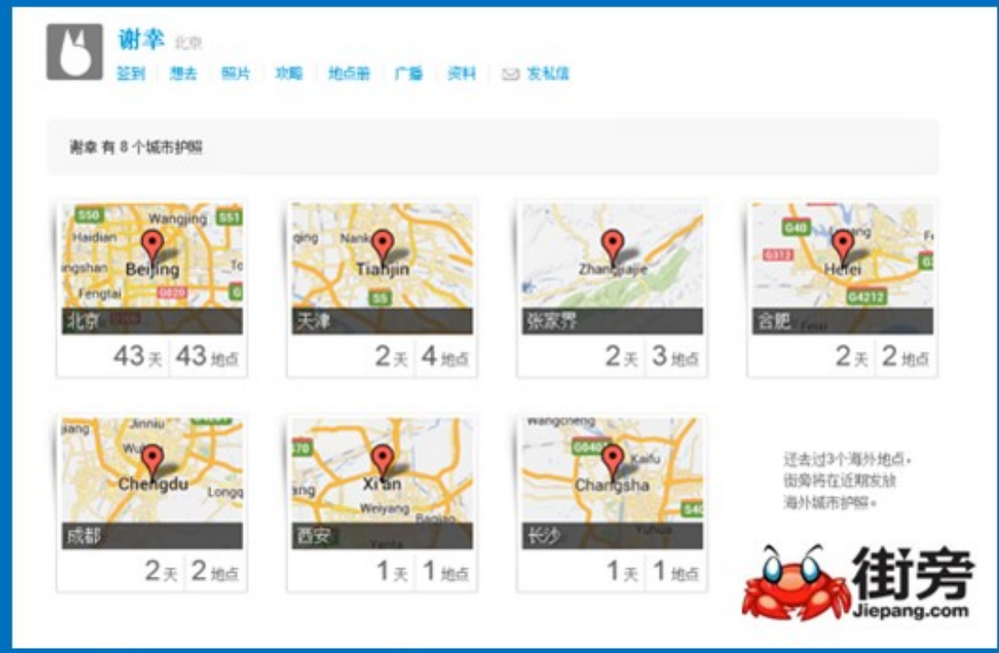






A Day in Your Life

FOOTPRINT



Heterogeneous Online Networks

LifeSpec.

Self-Disclosure

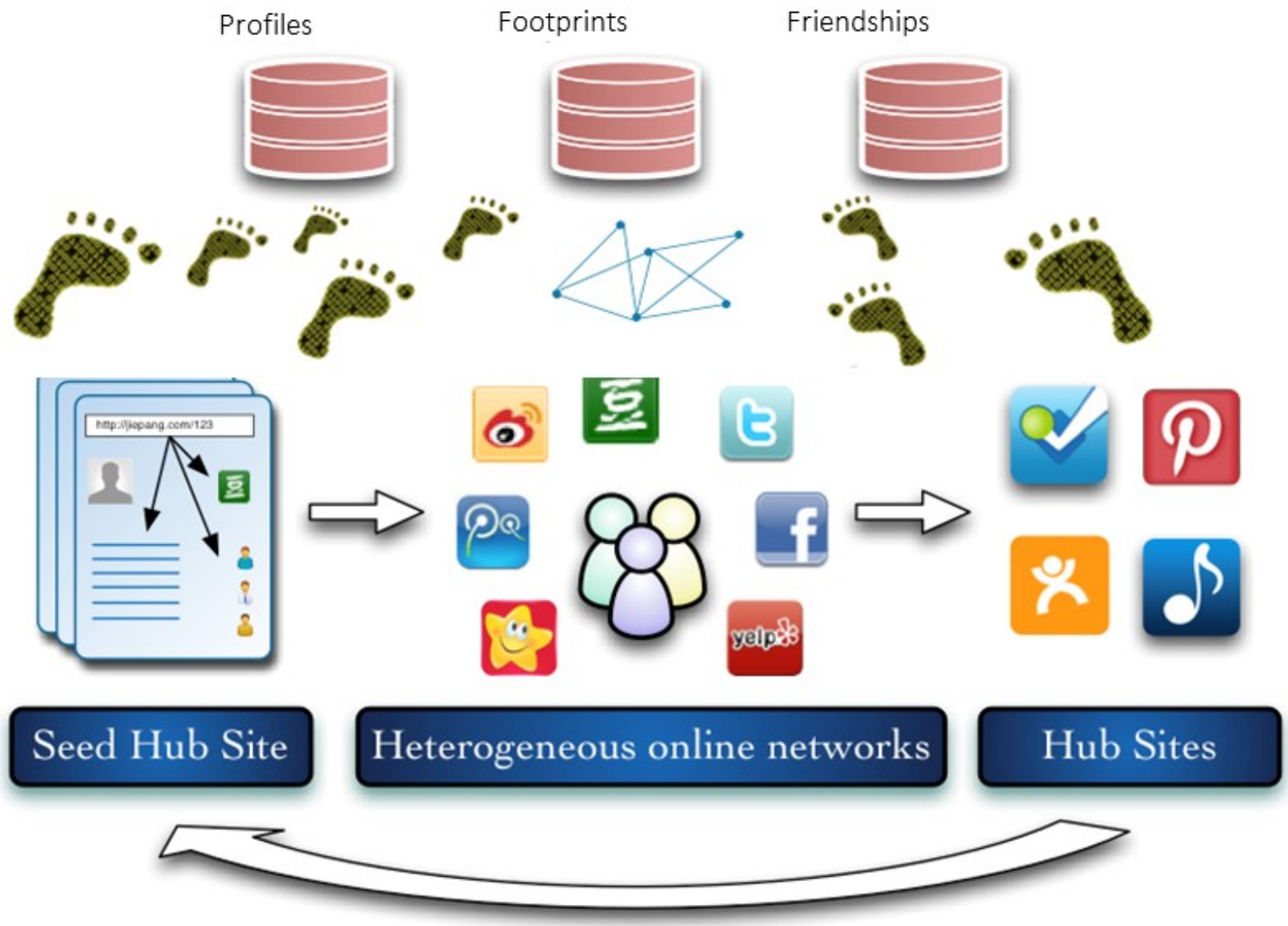
Index users with entities

Cross-Domain Posting

Hub Sites

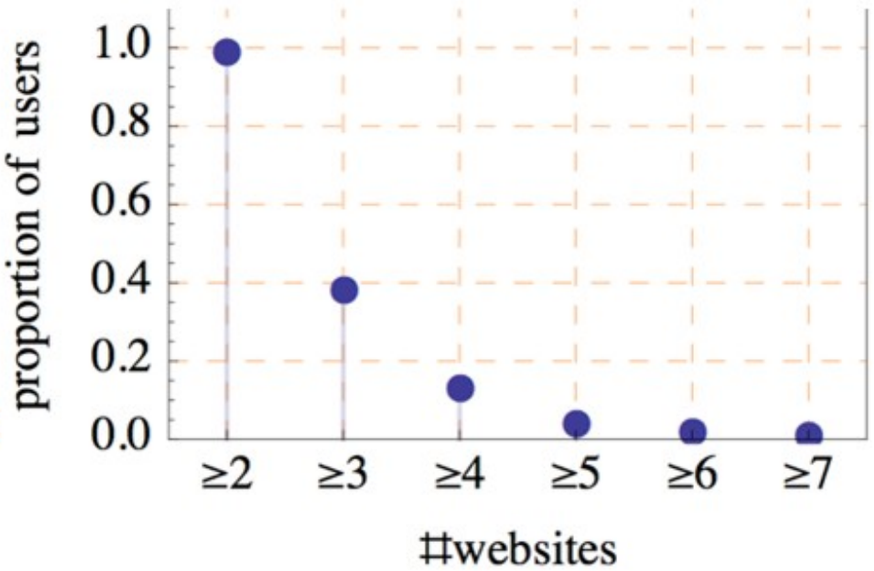
```
<a href="http://
www.facebook.com/profile.php?
id=...
alt="Facebook" width="16"
height="16" /></a>
<a href="http://twitter.com/
..." alt="Twitter"
width="16" height="16" /></a>
```

ICONNECT Algorithm



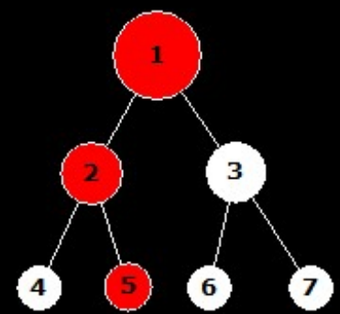
Data

- 4 (major) networks: Jiepang, Weibo, Douban, Dianping
- 1.4M+ unique (deterministically identified) users accounts
- Heterogeneous footprints: tweets, photos, check-ins, movies, books, music, offline events, online purchase history, etc.
- Rich user profiles integrated from different sites (publicly available)



Age	Gender	Residence	Relationship	Occupation	College	High School	Self description	...

SPECTRUM



Relational Hierarchical LDA (RH-LDA)

checkin movie book music events

8:00-12:00 12:00-20:00 20:00-8:00 non-local

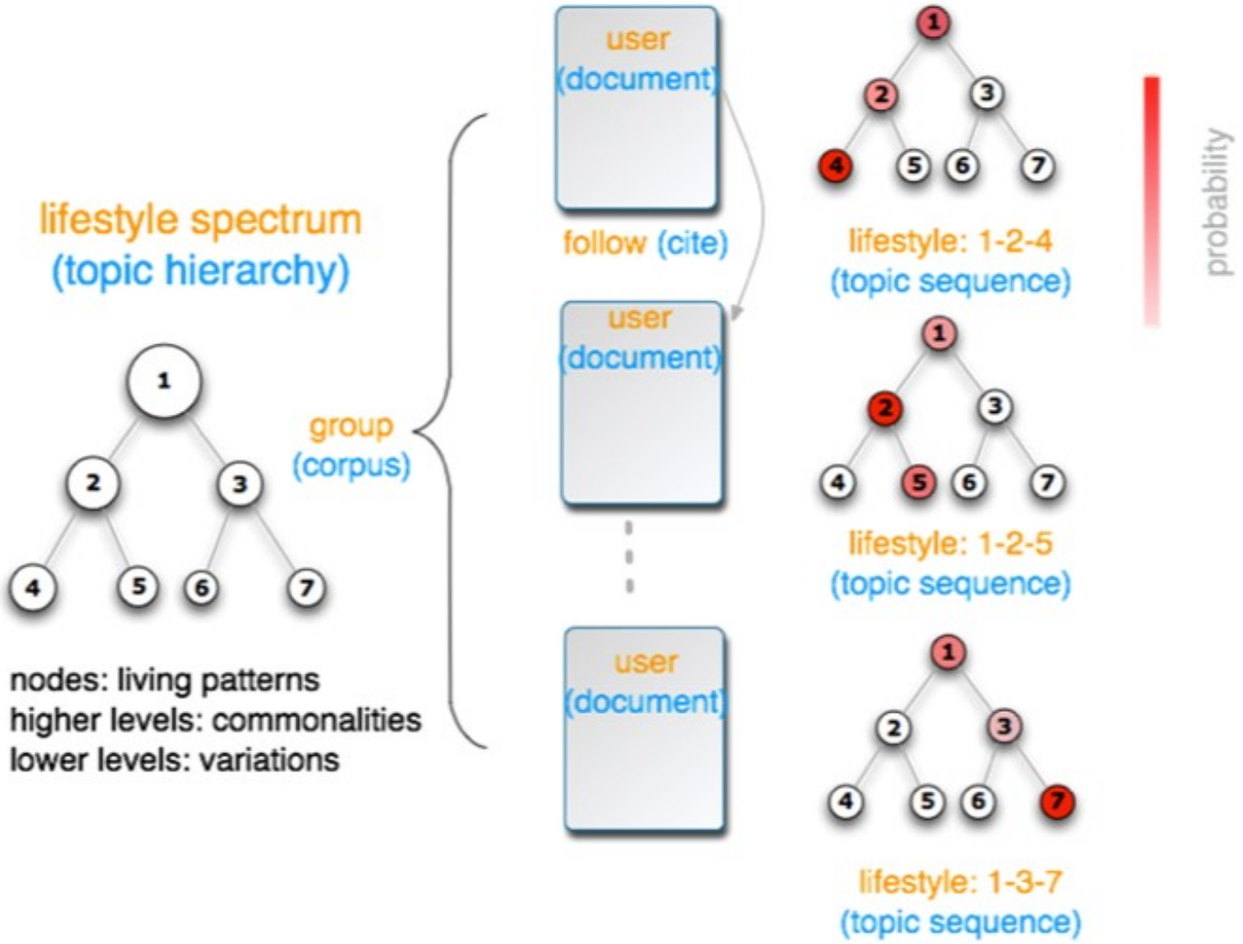
footprint (word): combination of domain specific tags (category)

(☀) shopping mall drama, sci-fi taiwan,pop lecture

living pattern (topic): frequently co-occurring footprints

(☀) shopping mall + taiwan,pop + ( ) bar

lifestyle spectrum: tree-structured topic hierarchy



[1] N. Yuan, F. Zhang, D. Lian, K. Zheng, S. Yu, X. Xie, "We Know How You Live: Exploring the Spectrum of Urban Lifestyles", *ACM Conference on Online Social Networks*, 2013.

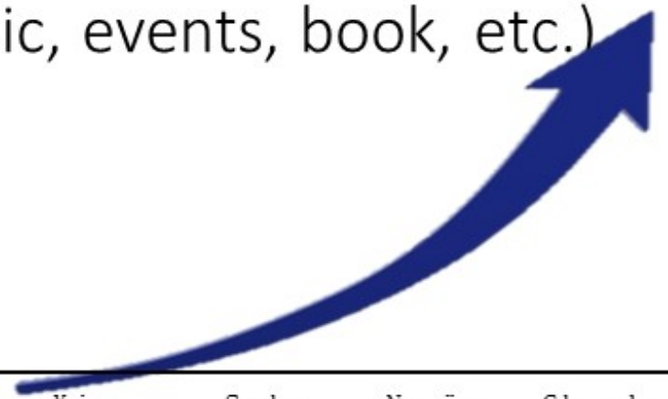
[2] D. Blei, T. Griffiths, M. Jordan, J. Tenenbaum, "Hierarchical topic models and the nested Chinese restaurant process." *Advances in neural information processing systems*, 2004.

Data

53 million footprints (check-in, movie, music, events, book, etc.)

3 million social links

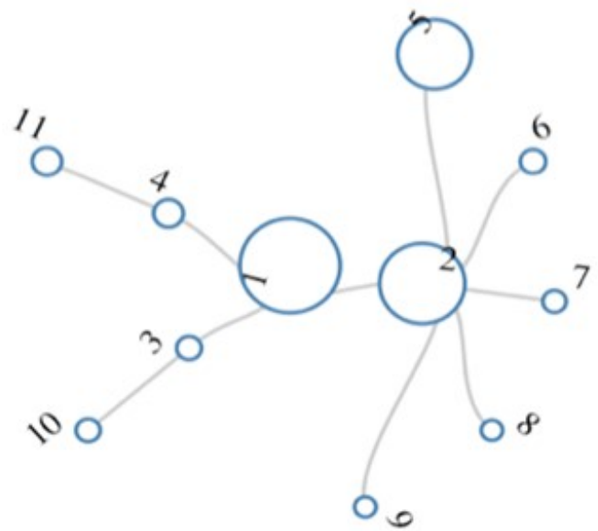
39 million check-ins



	city	Shanghai	Beijing	Guangzhou	Tianjin	Hangzhou	Hongkong	Xiamen	Suzhou	Nanjing	Chengdu	Wuhan	Xian
	users	417,681	162,764	53,089	15,490	34,322	12,599	10,123	19,673	21,558	23,372	20,975	15,261
Footprints	check-in	25,178,189	5,898,447	1,092,138	392,943	619,219	424,650	369,231	560,274	414,202	327,634	321,646	229,678
	movie	1,661,214	1,466,479	171,789	118,775	238,721	57,003	70,172	89,706	174,664	191,042	166,337	123,223
	music	766,165	737,254	85,953	60,658	103,936	30,313	29,716	39,701	82,513	88,426	76,316	62,876
	book	402,318	387,138	51,913	28,188	57,835	18,117	18,516	19,521	44,345	42,241	44,804	28,435
	event	609,076	803,158	101,246	52,133	78,587	18,277	20,889	27,400	46,788	66,640	44,764	72,902
	total	28,616,962	9,292,476	1,503,039	652,697	1,098,298	548,360	508,524	736,602	762,512	715,983	653,867	517,114

checkin movie book music events

8:00-12:00 12:00-20:00 20:00-8:00 non-local



a) financial practitioners

1: economics ( ) ( ) apartment hotel ( ) ( ) shopping mall

2: ( ) japanese cuisine ( ) fast-food lecture

3: ( ) hot-pot ( ) bar ( ) snack

4: ( ) snack ( ) fast-food ( ) japanese cuisine

5: drama,romance drama,comedy drama,action

6: ( ) bank ( ) bank ( ) subway

7: fiction,hongkong fiction,love mystery,japan

8: folk,indie indie,folk drama

9: ( ) car-4s fiction,society cartoon,philosophy

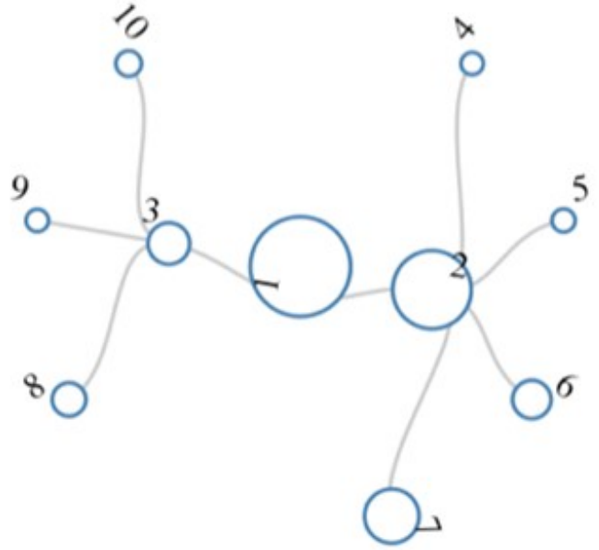
10: music drama,romance drama,comedy

11: ( ) ( ) scenic ( ) ( ) airport ( ) ( ) office

a') financial practitioners

checkin movie book music events

8:00-12:00 12:00-20:00 20:00-8:00 non-local

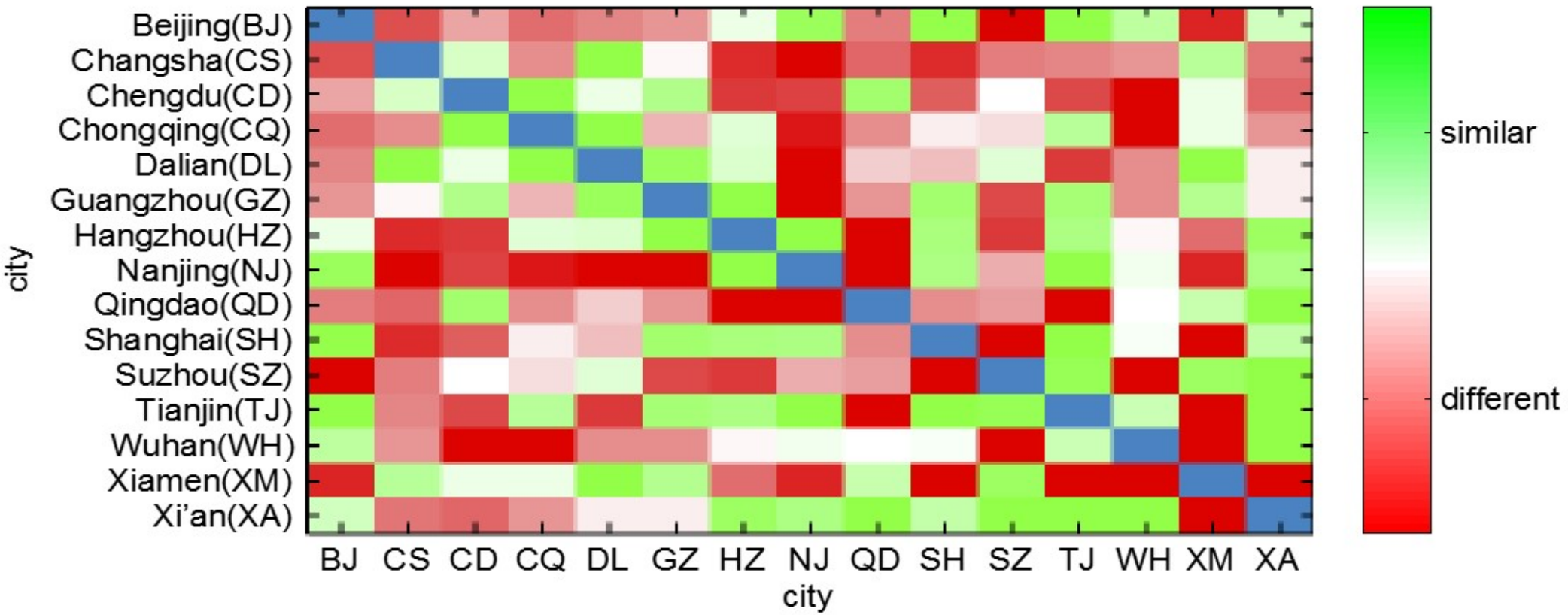


b) software practitioners

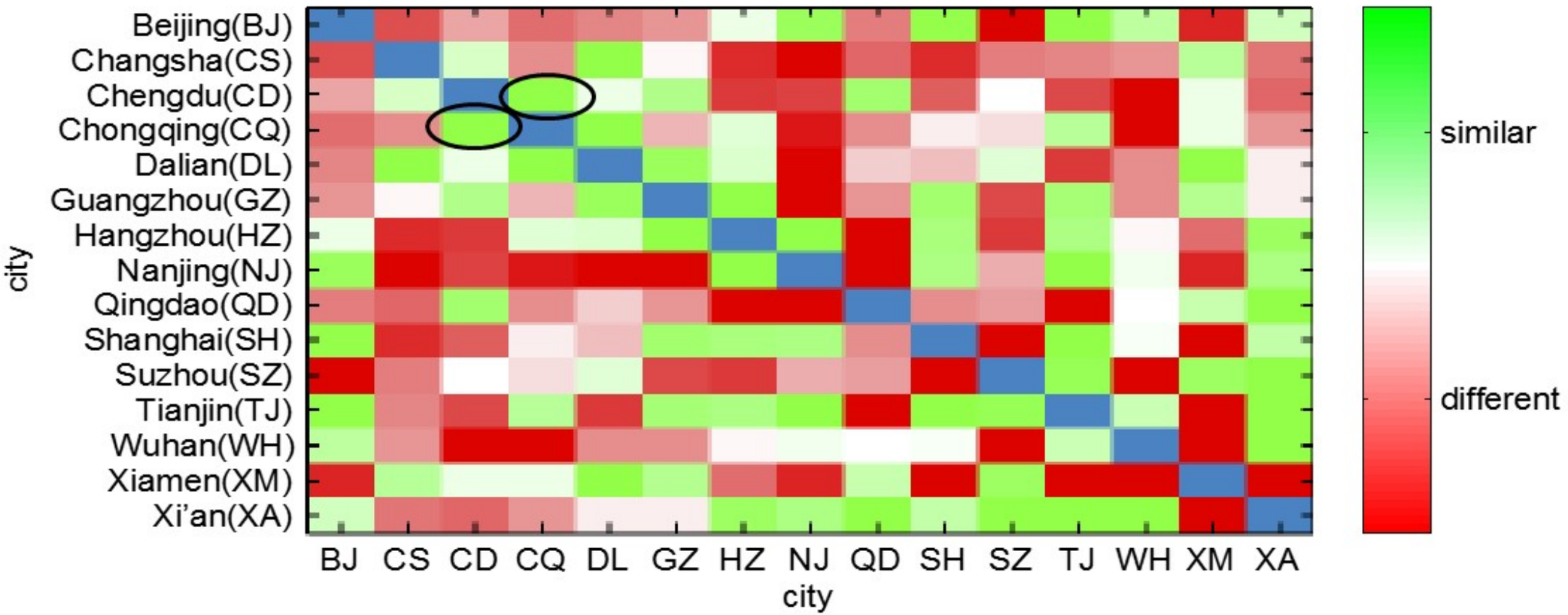
1: computer programing,computer movie
2: drama,romance drama,comedy taiwan,pop
3: ux,design fiction,foreignliterature fiction,chineseliterature
4: mystery,japan comedy,action cartoon,mystery
5: taiwan,pop music chineserock,rock
6: ( ) apartment ( ) office ( ) ( ) apartment
7: drama,romance drama,action drama,comedy
8: lecture music get-together
9: programing,computer algorithm,computer drama,suspense
10: drama,romance music taiwan,pop

b') software practitioners

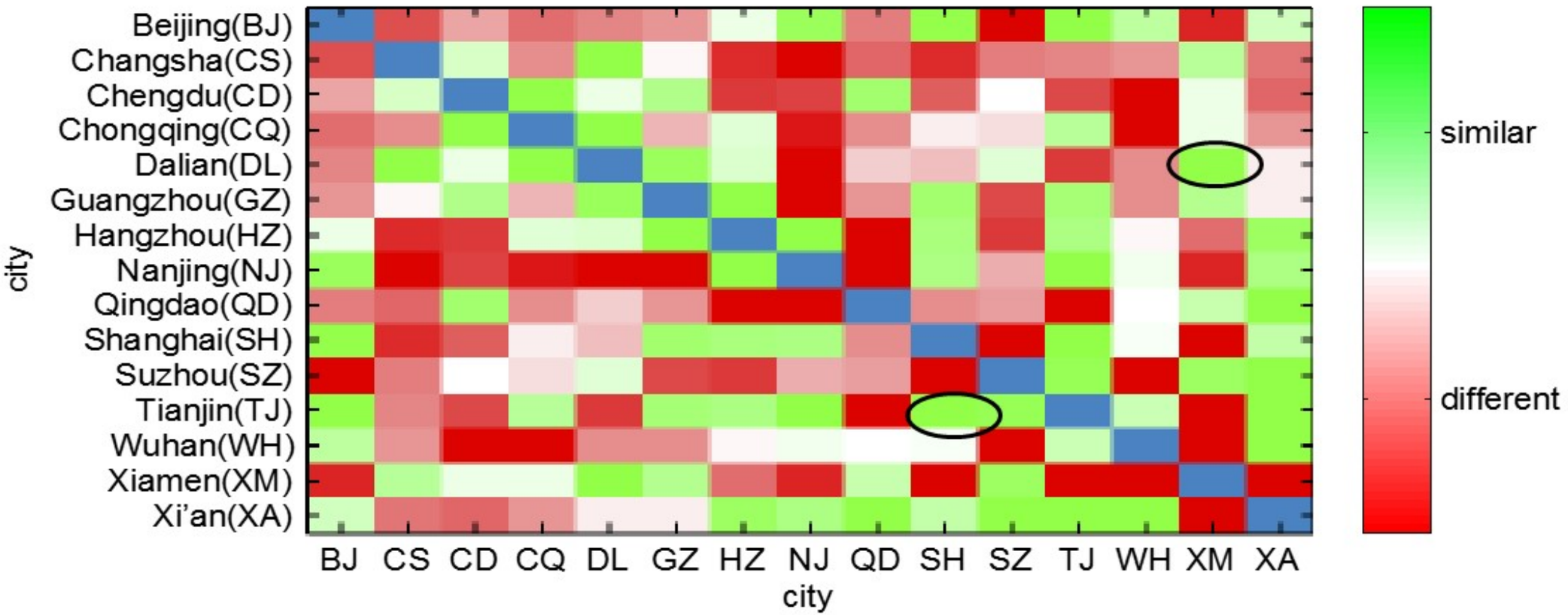
Similarity Matrix of Lifestyle Spectrum



Similarity Matrix of Lifestyle Spectrum



Similarity Matrix of Lifestyle Spectrum



Intergroup difference and intragroup variations (place of residence)

Table 2: Average recognition ratio

Method	RTM	hLDA	RH-LDA
Check-in	0.361	0.500	0.667
+Movie	0.389	0.556	0.694
++Music	0.444	0.583	0.722
+++Book	0.472	0.639	0.806
++++Events	0.472	0.667	0.833

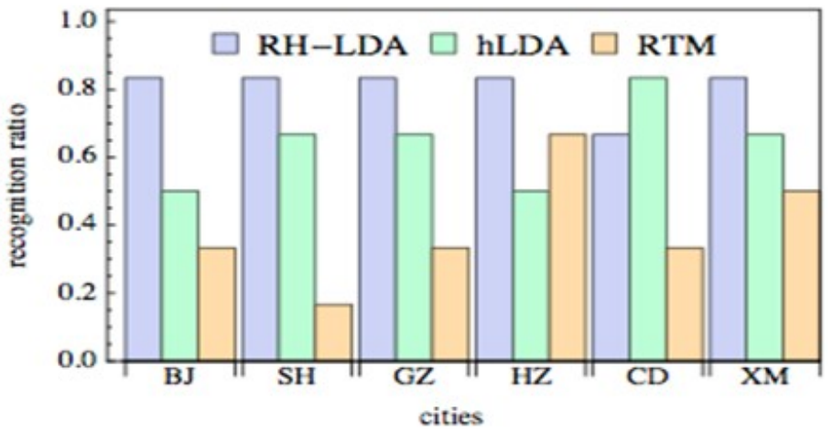


Figure 12: Recognition ratio.

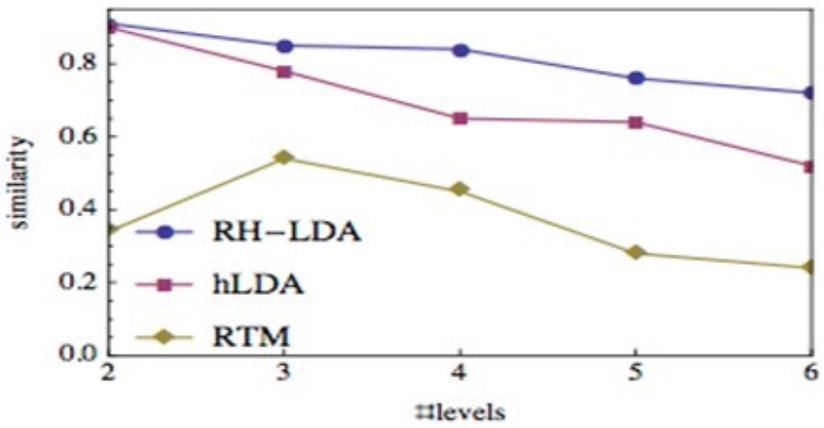


Figure 13: Jaccard Similarity.

36 participants
aged 21-45
20 males 16females
6 for each city
8+years

Train models for 6 cities
Select based on lifestyles
RR: correctly inferred/all

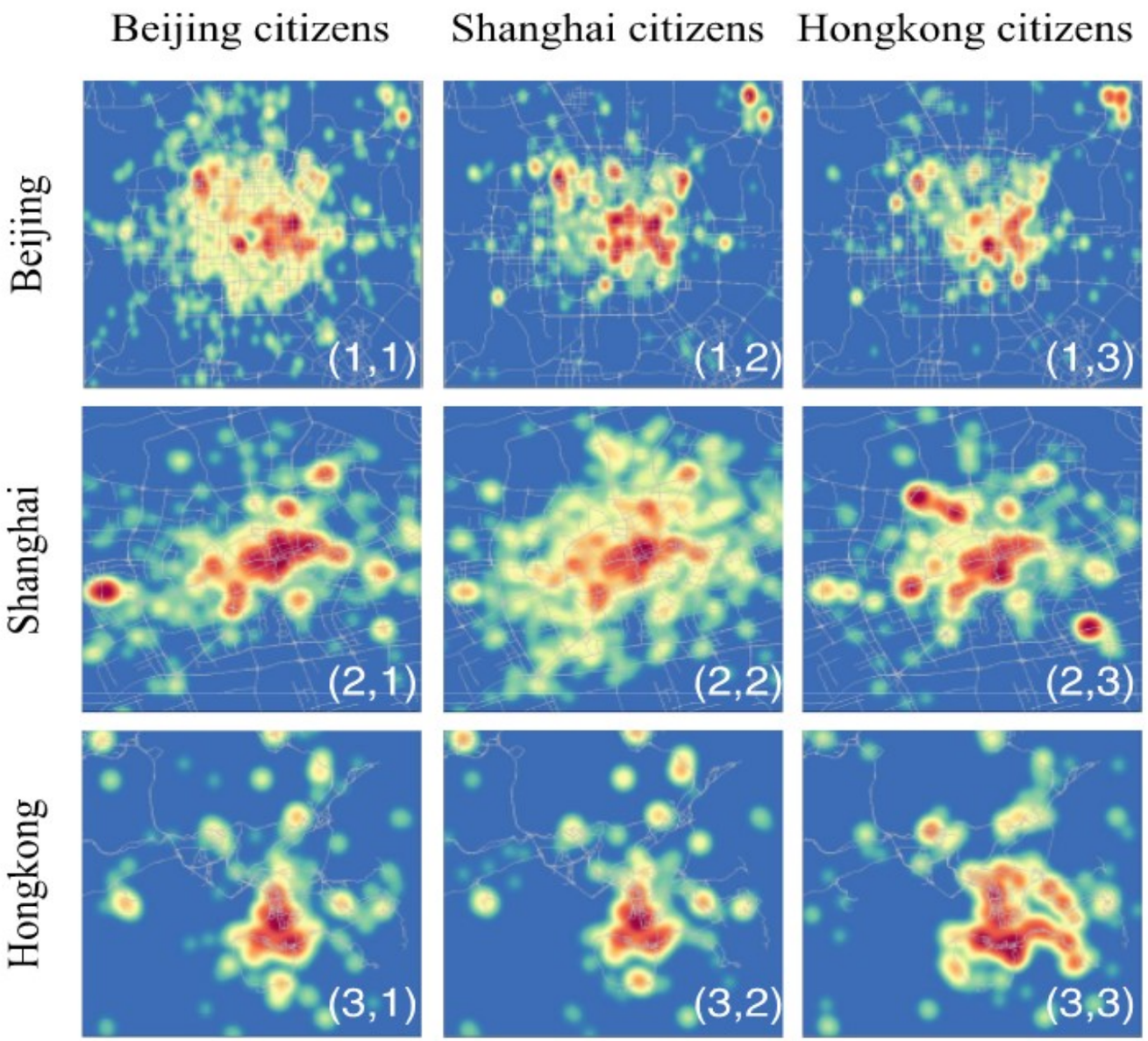
Collect social data
Pick a path in the spectrum
Measured by JS

LifeSpec as A Data Repository

- Industry Applications:
  - Recommendation
  - Advertising

- Scientific Research:
  - sociology
  - economics
  - public health

Mobility



Density distributions showing where people check-in in each others' cities

LifeSpec.

Summary

- Computational framework for exploring urban lifestyles
- IConnect: identifying connected user accounts based on self-disclosure
- RH-LDA: relational-hierarchical model for summarizing lifestyles
- LifeSpec as a data platform for scientific and industry applications

Nicholas Yuan
nichy@microsoft.com
Microsoft Research