# Learning Networks of Places and People Using Location Data

**Tony Jebara** 

# Sense Networks

COLUMBIA UNIVERSITY





#### From online to real networks?



a network of real places

a network of real people

Online data is easy to get, what about the real world?

#### From online to real networks?



Online data is easy to get, what about the real world?



#### GPS and location data





### CitySense: where is everyone

Citysense: real-time density of users at every street corner
Poisson models find most active bars/restaurants







#### Next: where's everyone like me

#### Need to have a network of people

Each dot is a user

Dot's color is user's social cluster



### **Network of People**

who is like whom? who colocates with whom?



#### **Network of People** Hard to say if User A is like User B... User A User B Emeryville Emeryville Moraga 1) Moraga 80 880 San San 880 Francisco Alameda Francisc Alameda Oakland Oakland ò ò San San 0 Leandro Ca Leandro. Cas San San Daly City Va Francisco Daly City o South San San Francisco South San San Bay Lorenzo Bay Francisco Francisco Lorenzo Haywardo Haywardo San Bruno (92) San Bruno 92 Pacifica Pacifica Millbrae Millbrae Burlingame Burlingame San Mateo San Mateo Foster City Foster City (84) New s Beach 84 News Moss Beach Redwood Redwood Half Menio Park Menlo Park Half City 280 City Moon Bay Moon Bay East East 0 alo Alto Palo Alto herton Atherton o Alto Mour Palo Alto Moun

... don't just see if they colocate physically... do they overlap semantically (network of places)

### **Network of Places**

#### Is place A like place B? Look at each place's Flow, Commerce & Demographics



#### Network of Places: Flow

Look at flow A to B

Markov transition

Minimum Volume Embedding (MVE)

Color code clusters



### Network of Places: Commerce

Get each block's SIC (standard industrial categorization) code & cluster



# Network of Places: Demographics

#### Get each block's census demographic data & cluster



### Encoding people

For each user, convert GPS trail into matrix of probabilities for week hour probability of being in

flow cluster
 sic cluster
 demographic cluster

Week Hour	FLO 1	FLO 2	 FLO 20	SIC 1	SIC 2	 SIC 97	DEM 1	DEM 2	 DEM 78
1	.03	.31	.14	.03	.05	.41	.11	.04	.01
2	.14	.34	.02	.04	.05	.52	.01	.01	.00
168	.07	.34	.51	.02	.06	.48	.02	.01	.00





# **Encoding people**

#### 9 example users



compute pair-wise overlap from weekly exposure matrices







	. 101	
10.000		
		and the second second
at tarica		
		日本 日本
	1 11 11 11 11 11 11 11 11 11 11 11 11 1	

real friends should colocate!



	the second
and the second s	A DESCRIPTION OF A REAL PROPERTY OF A
	A REAL PROPERTY OF A REAL PROPERTY.
	a state of the second sec
	Contraction of the first state of the second state of the secon
	Charles and the second second
	AND THE REPORT OF A DAMAGE OF
	HE ROLL & LOUIS CONTRACTOR
	the second s
	a state the second s
	the particular of the particular of the
	- Contraction of the lower of a second
	The second se
The second second second second	
	120 State 1 1 1 1 1 1 1 1 1 1 1
<ul> <li>A second s</li></ul>	and the state of the state of the second state of the second
	· ···································
	- The state of the
	a second s
	All the second s



### **Connecting Similar People**



#### **Network of People**



#### Network of People $\rightarrow$ Tribes



\*Weekend Mole" • Out occasionally on weeknights, typically middle-aged, Latino, middleincome neighborhoods

#### Mature Homebody" Rarely goes out, typically spends nights in mature, white, higher income neighborhoods

### Network of People Predictions



# The Next Net

#### (Stephen Baker, BusinessWeek)

