

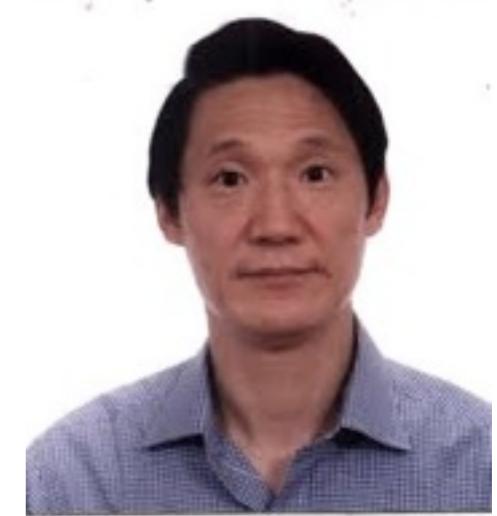
Transferring Heterogeneous Links across Location-Based Social Networks



Jiawei Zhang



Xiangnan Kong

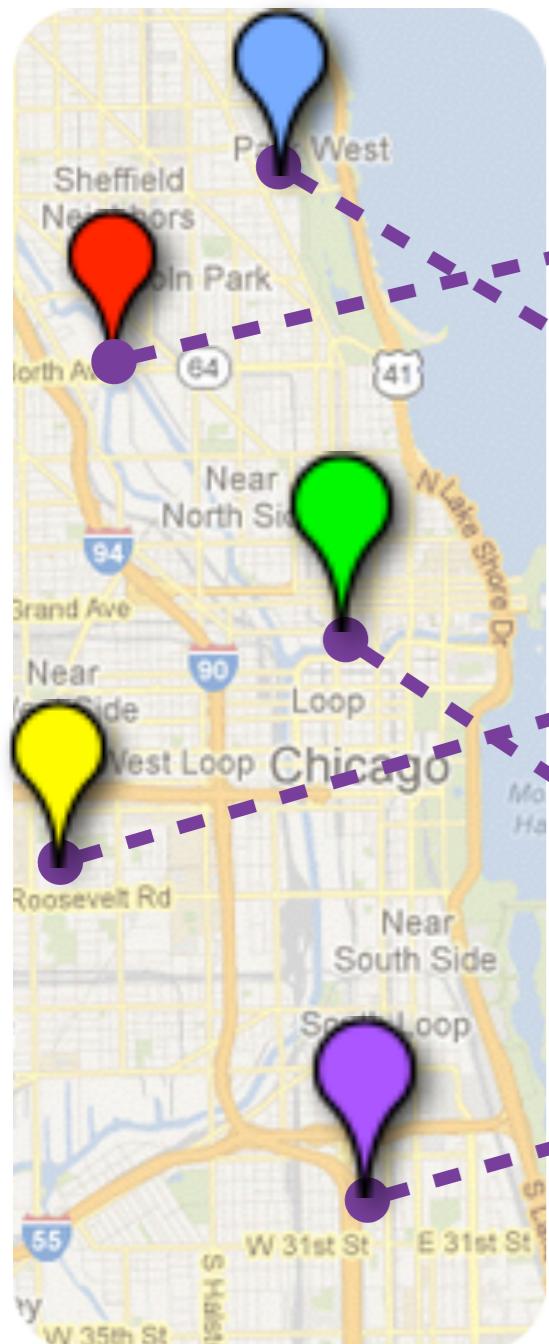


Philip S. Yu

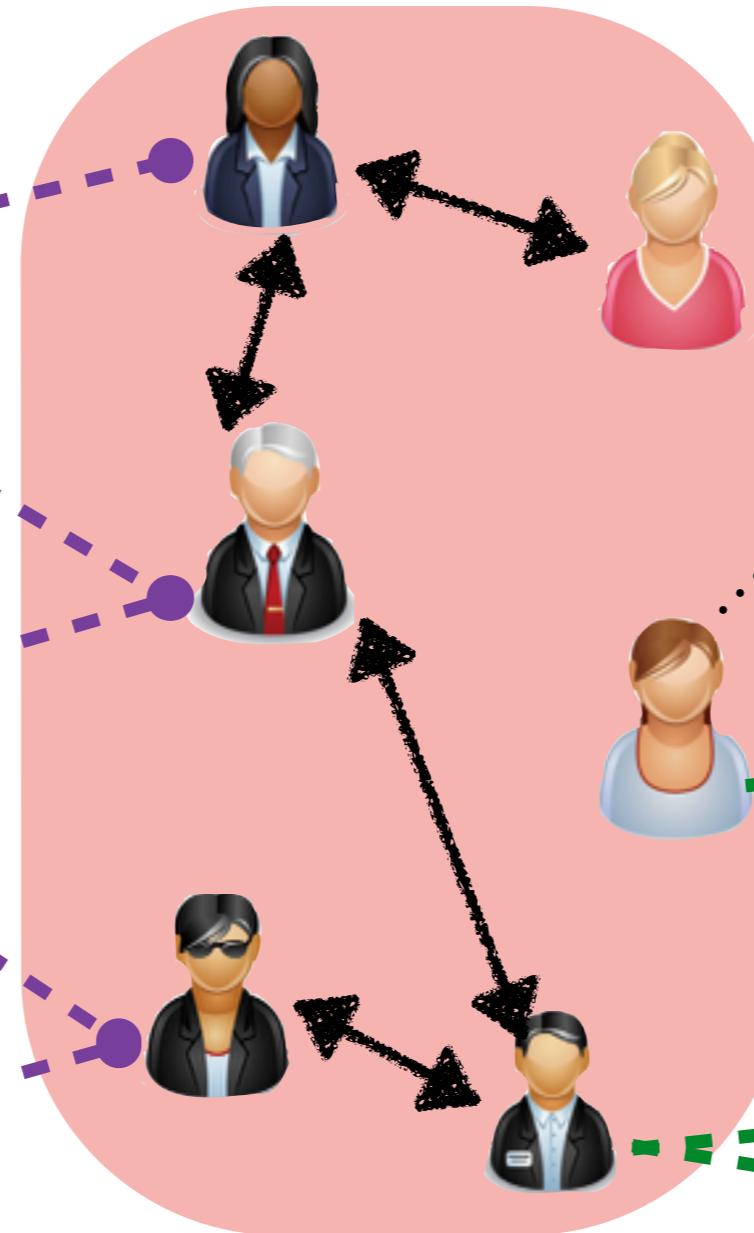
University of Illinois at Chicago

location Links

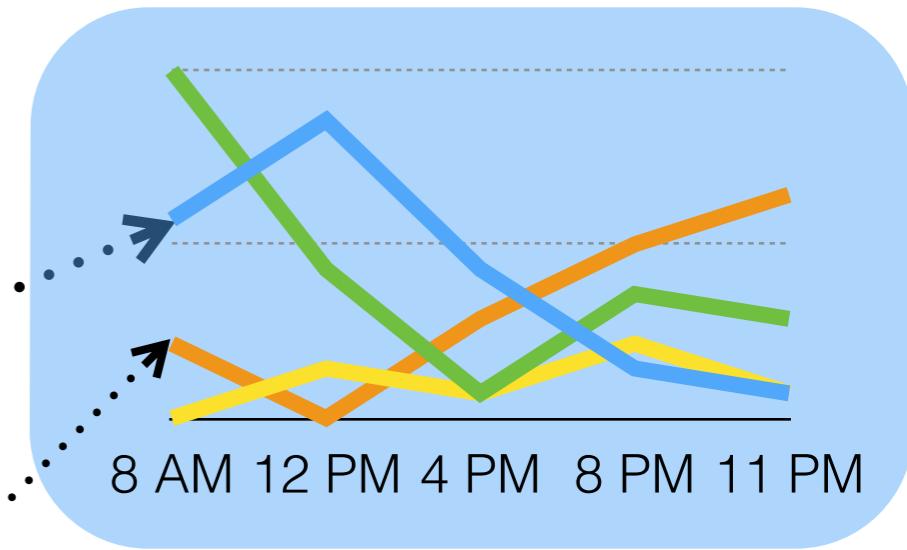
Locations



Social Links



Temporal Activities



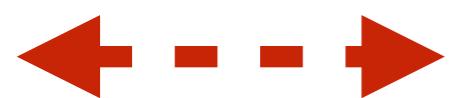
Contents: Tweets



Social Network:

Who Where What When

Problem Description: Collective Link Prediction

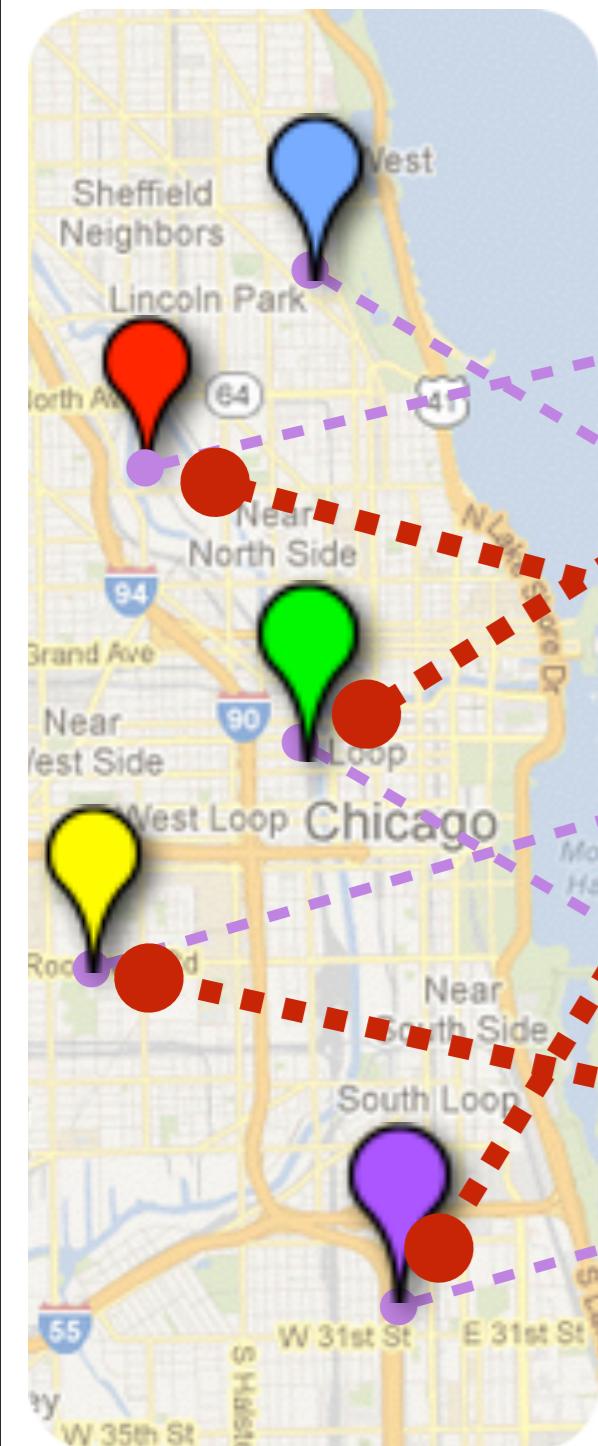


social links
to be predicted

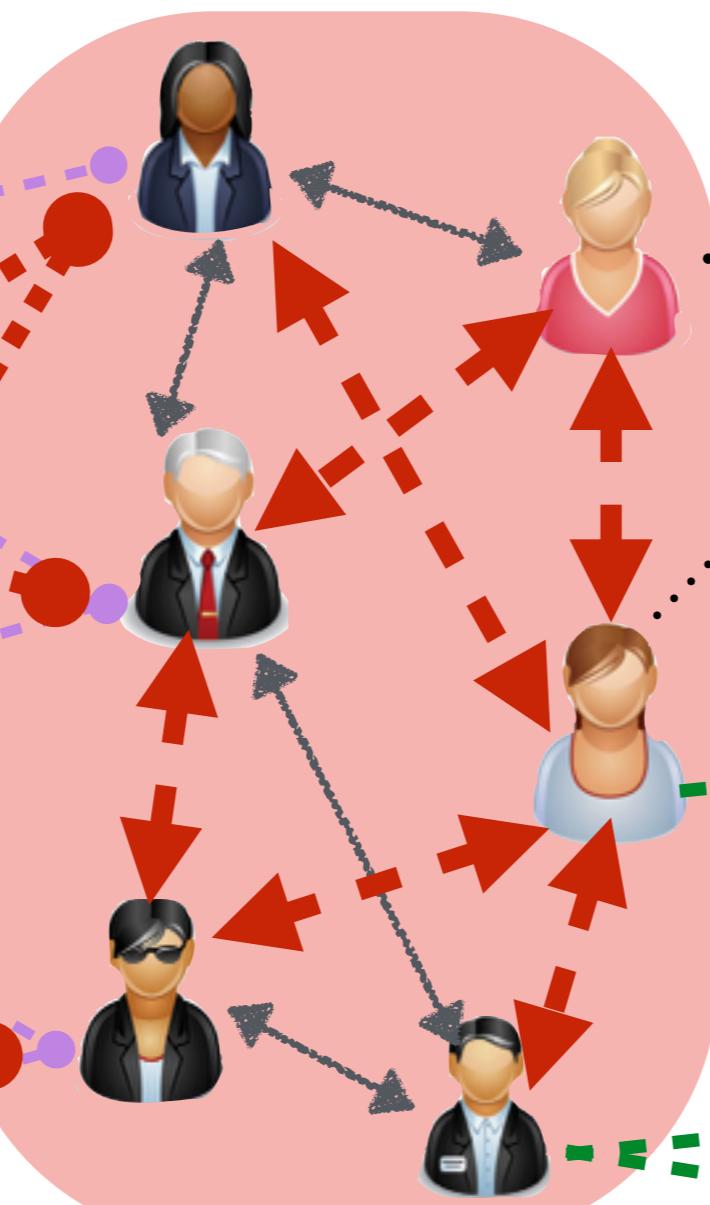


location links
to be predicted

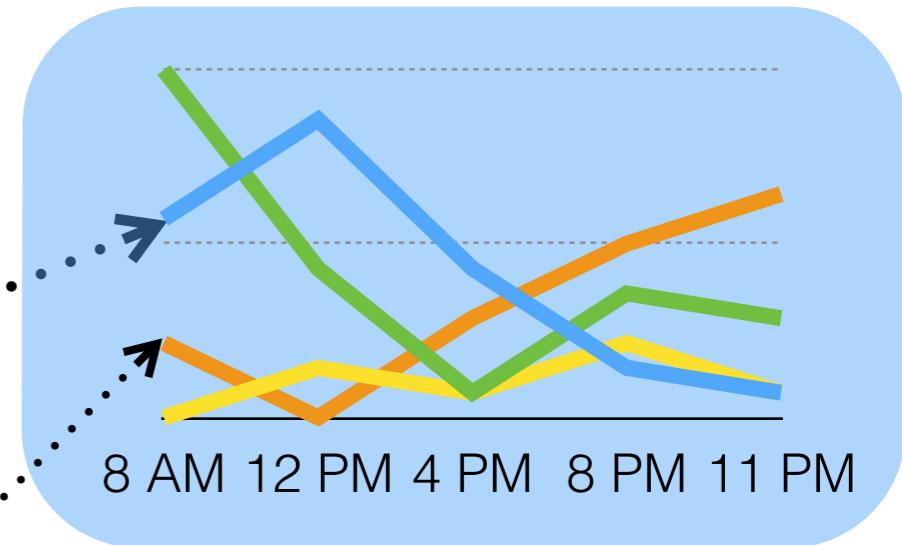
Locations



Social Links



Temporal Activities

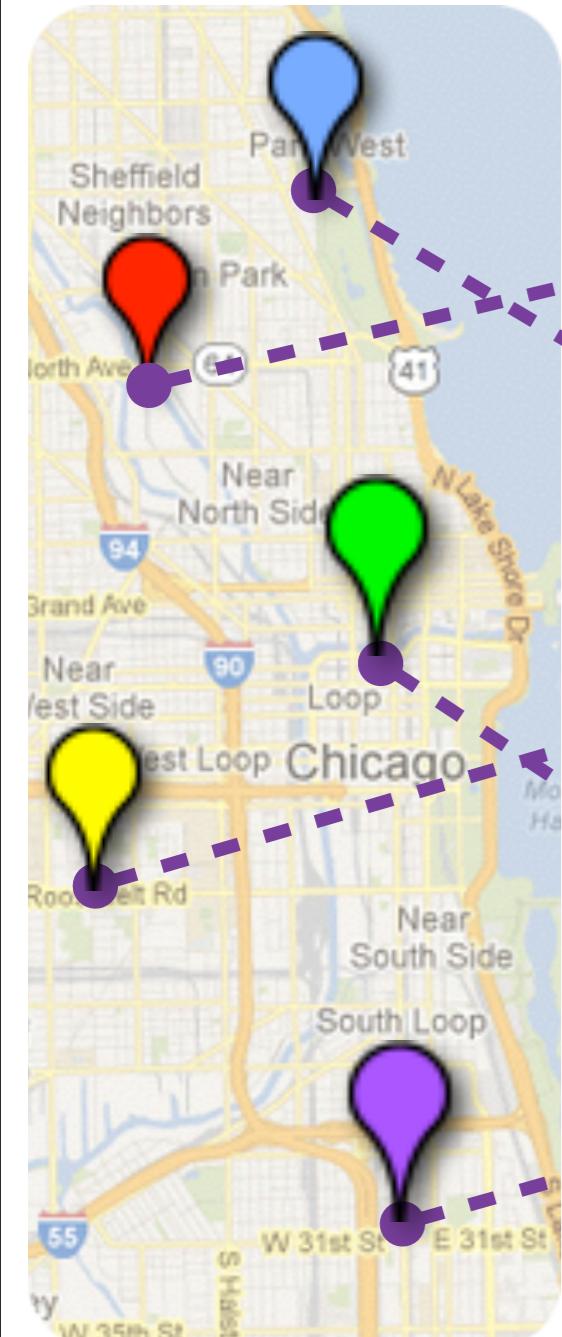


Contents: Tweets

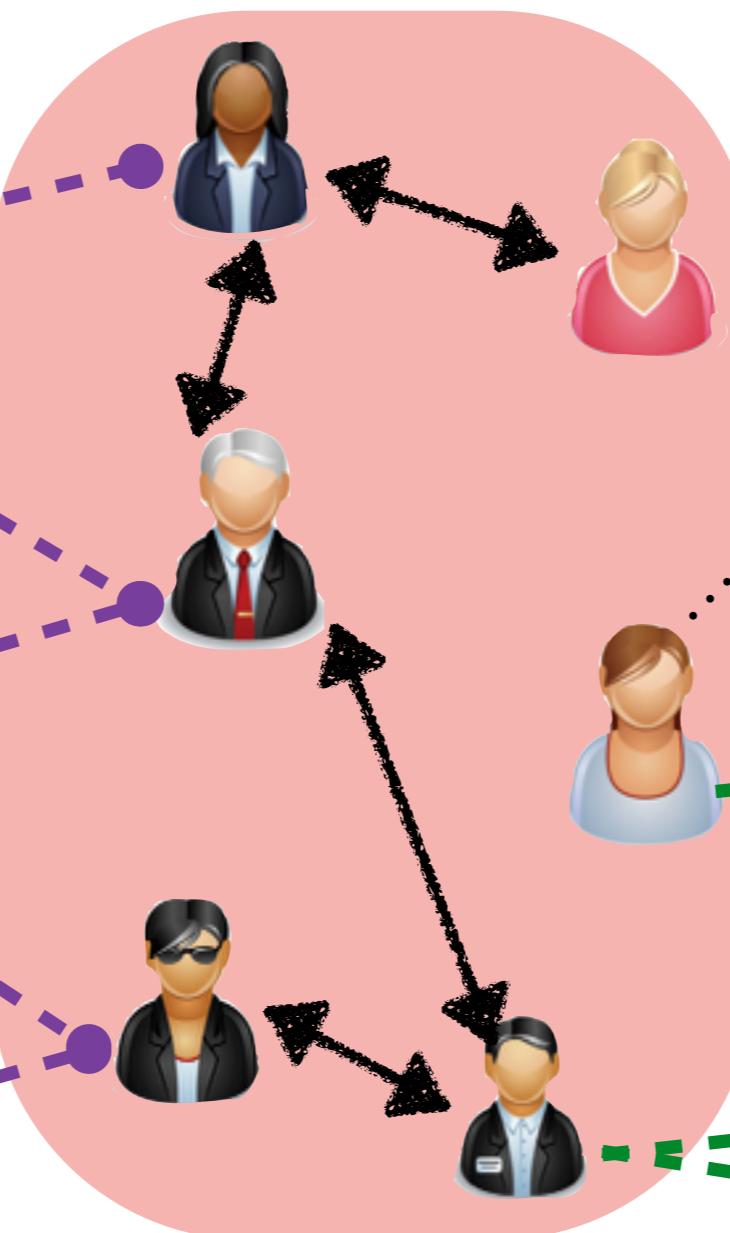


Solve Challenge 1: Heterogeneous Features

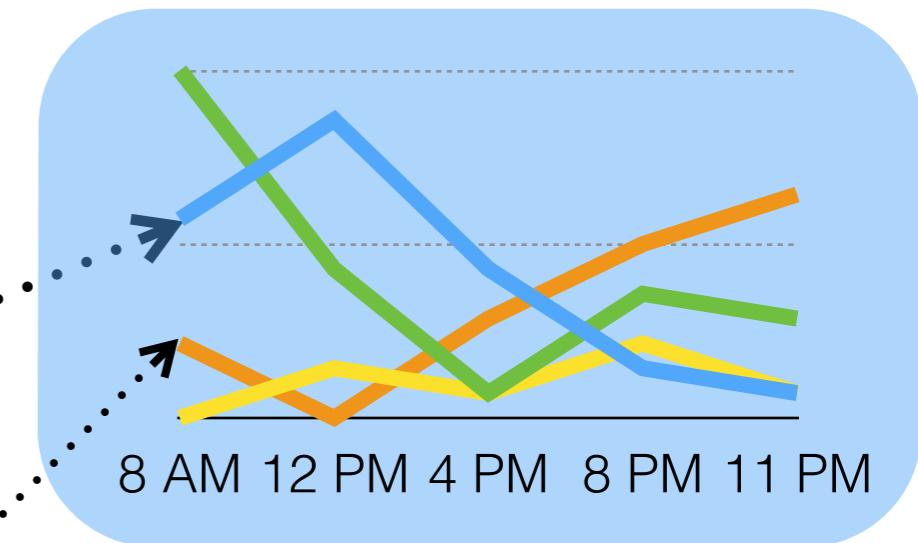
Locations



Social Links



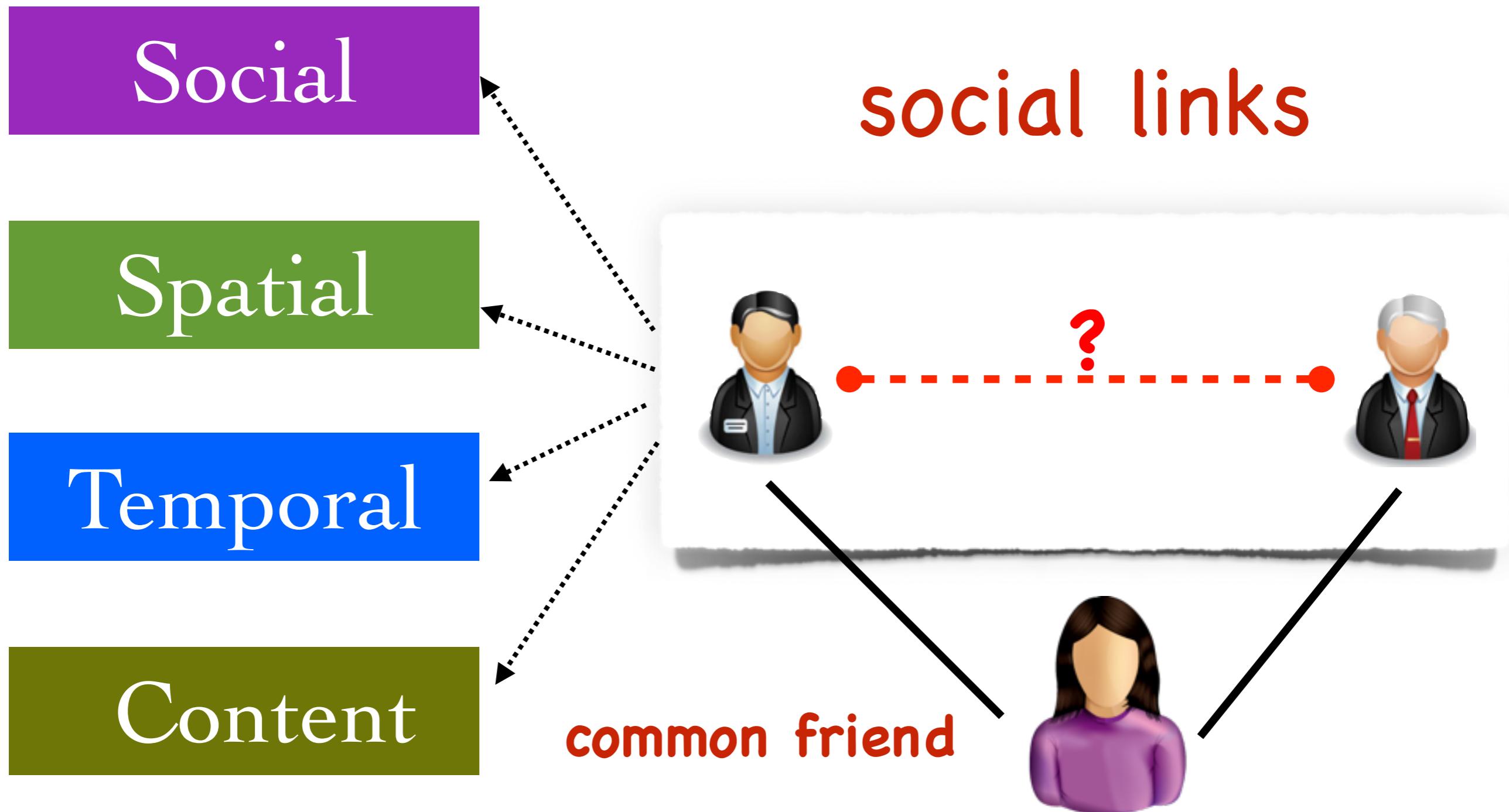
Temporal Activities



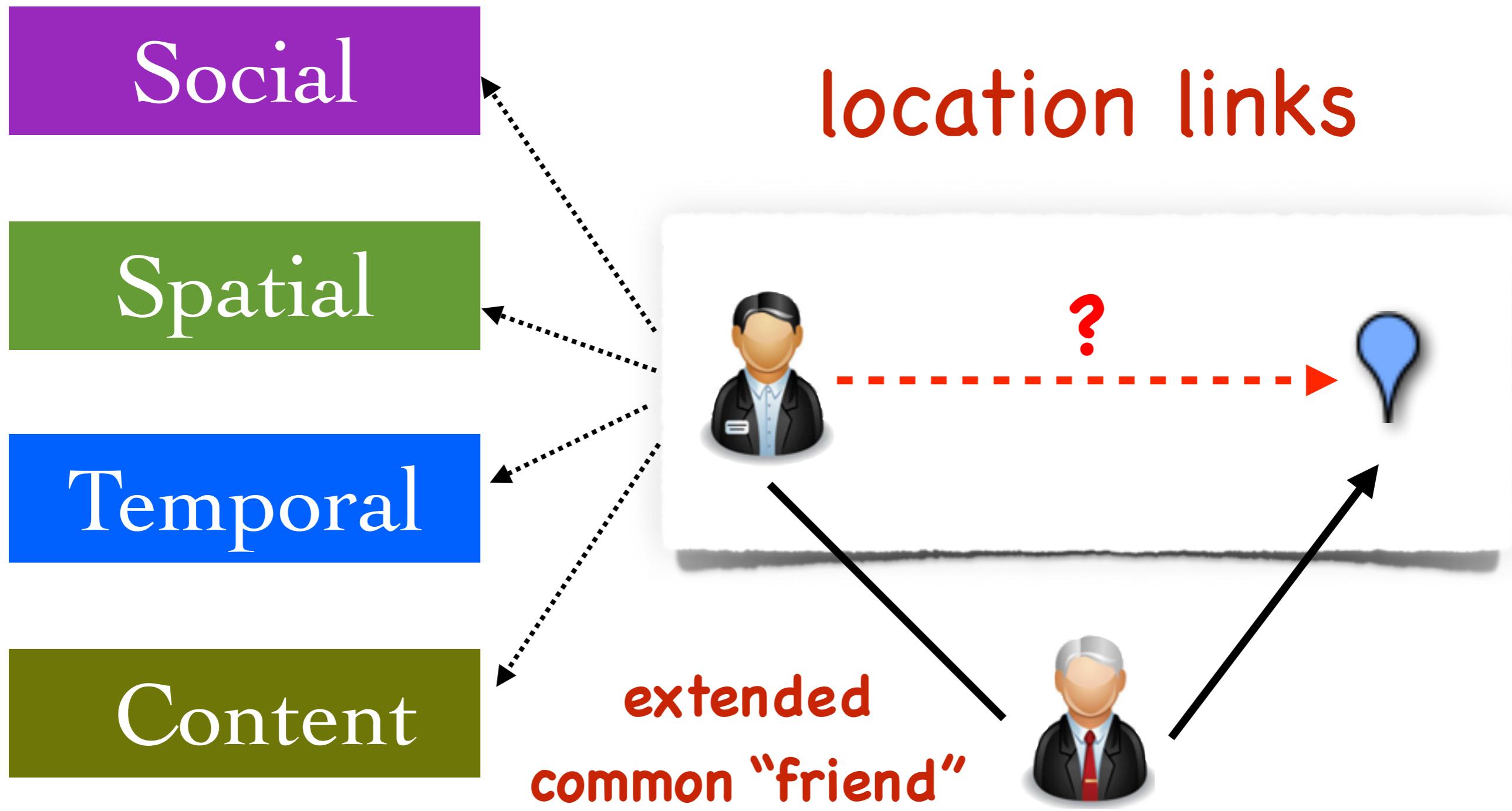
Contents: Tweets



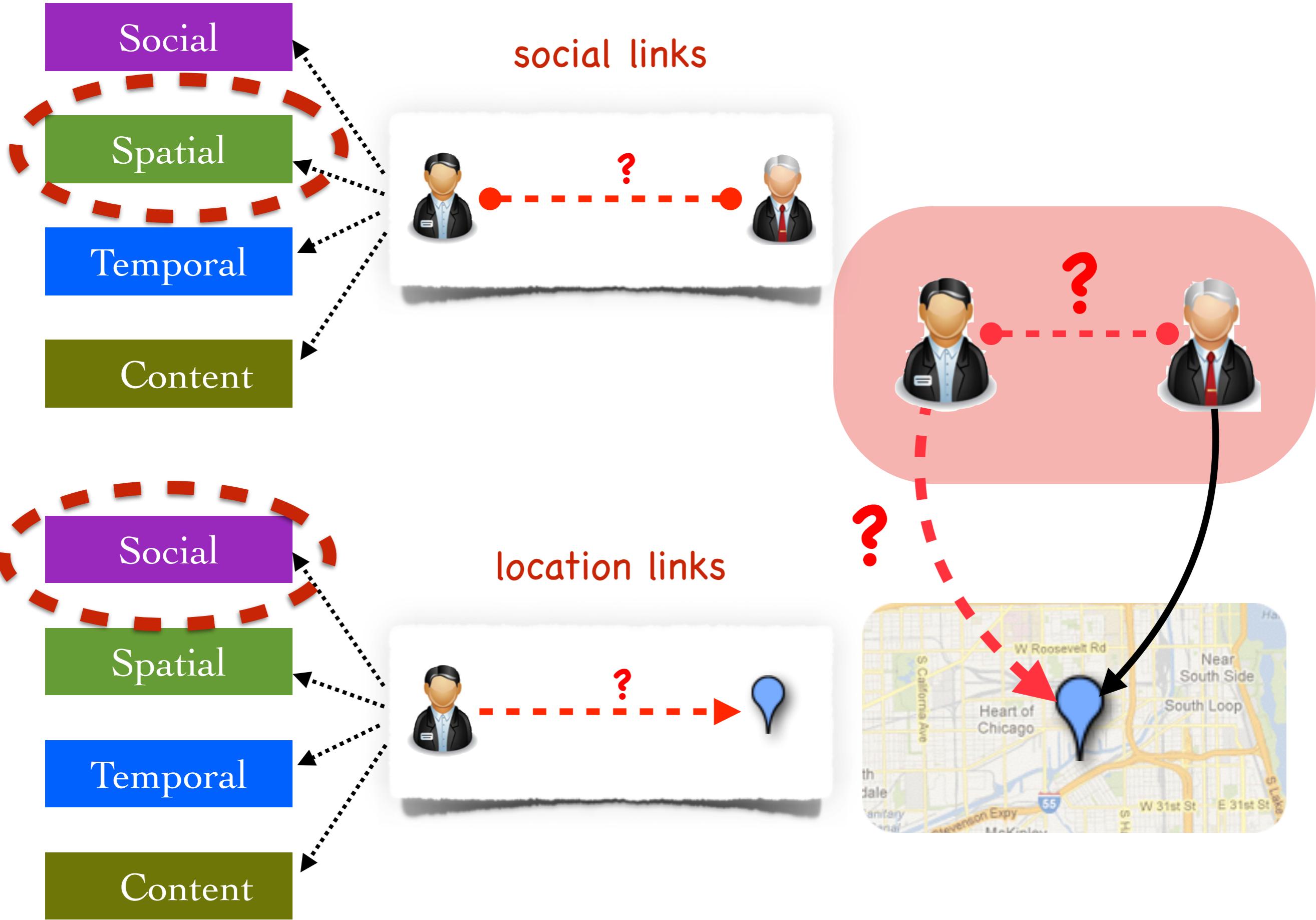
Extract Heterogeneous Features (1)

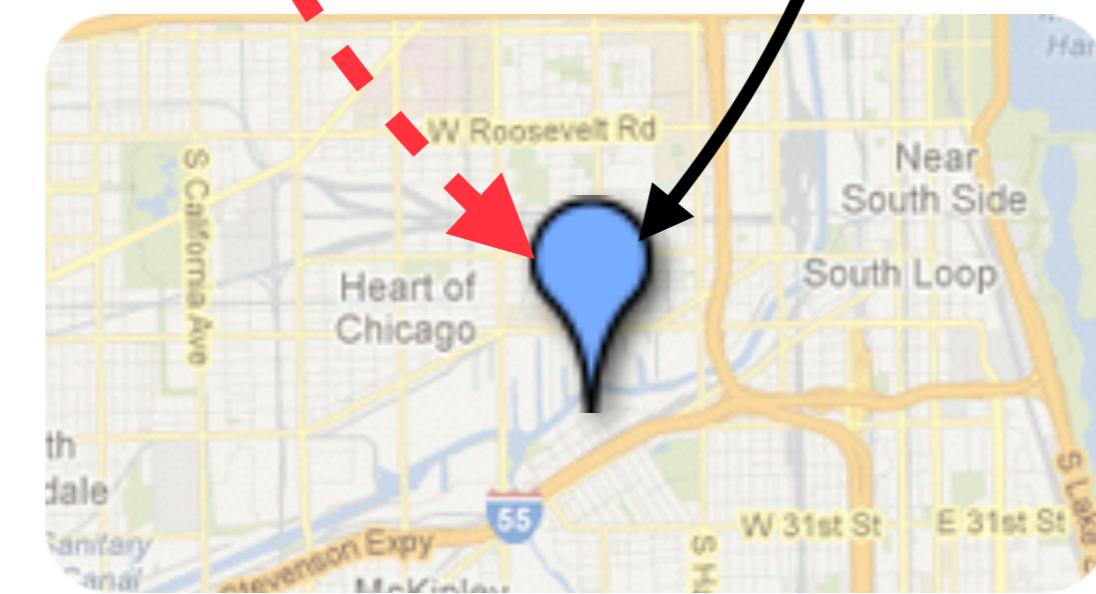
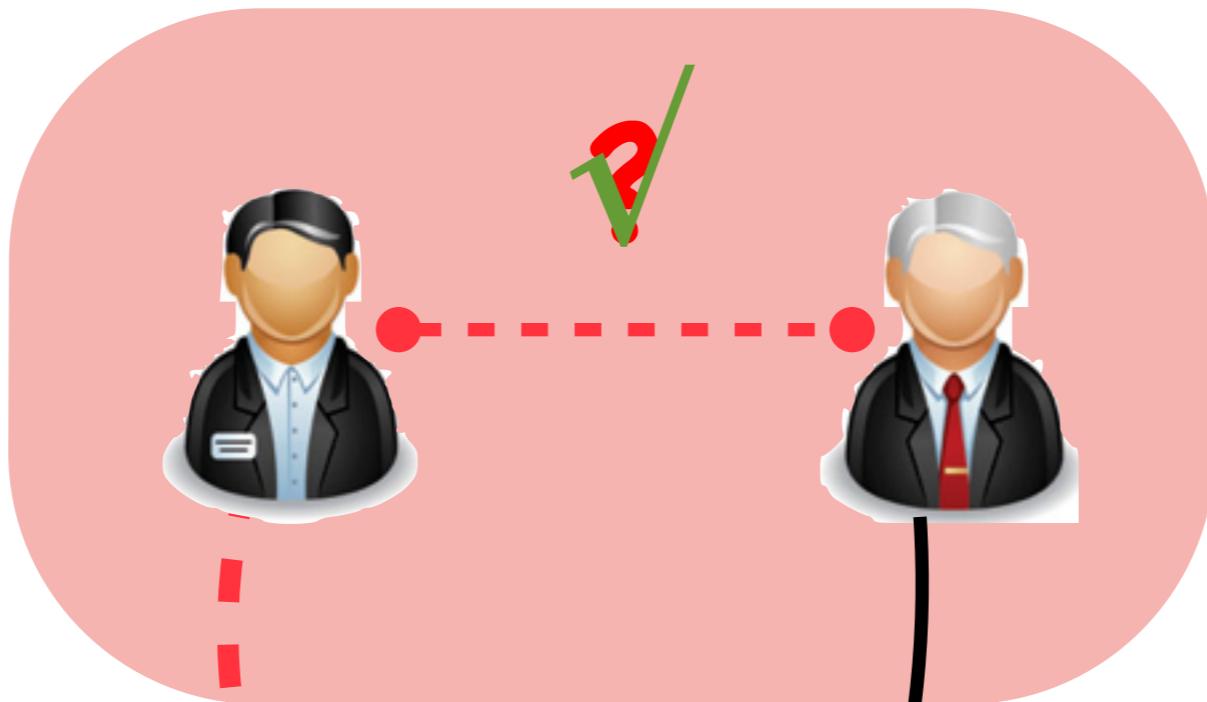


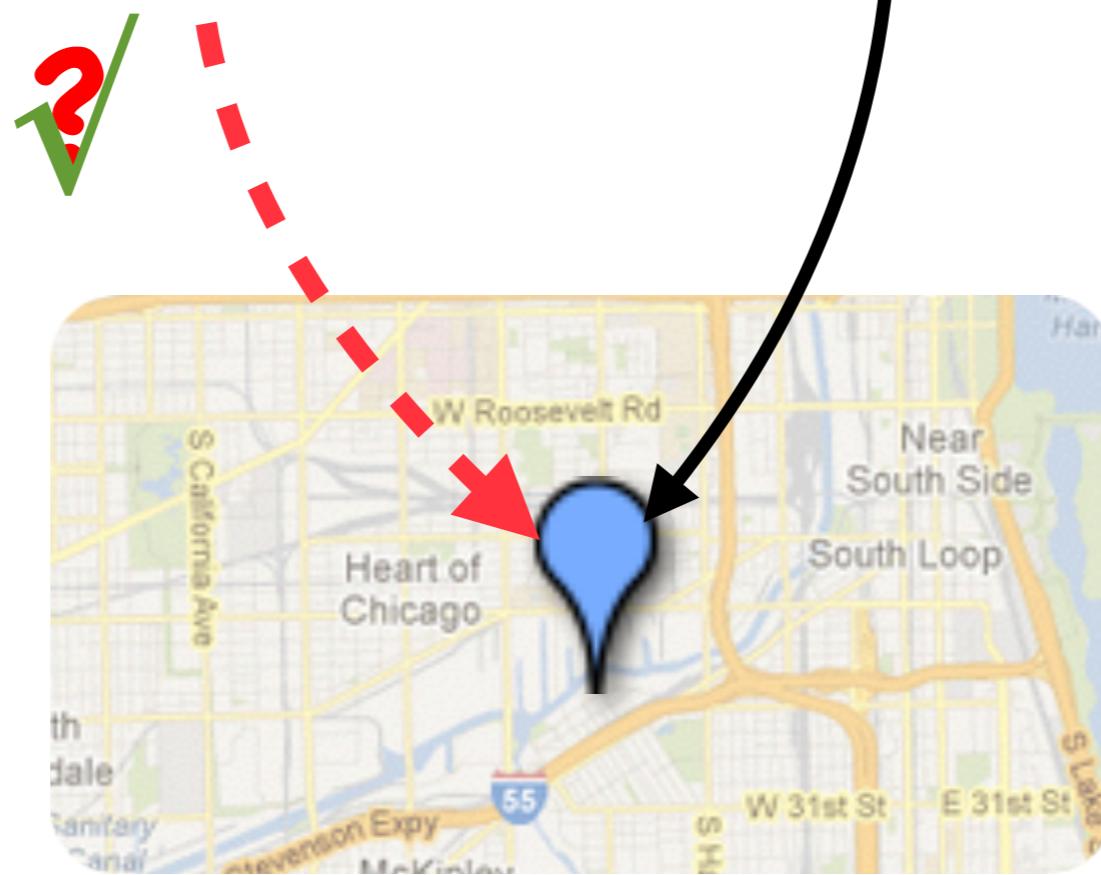
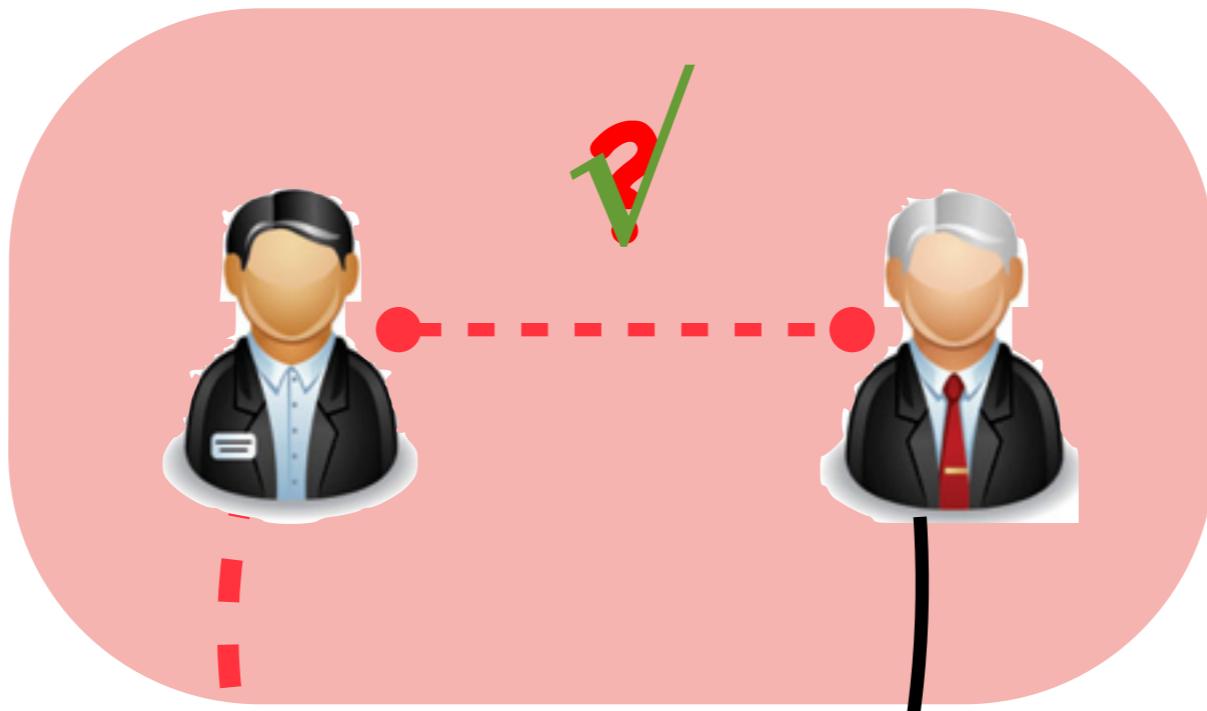
Extract Heterogeneous Features (2)



Solve Challenge 2: Collective Link Prediction



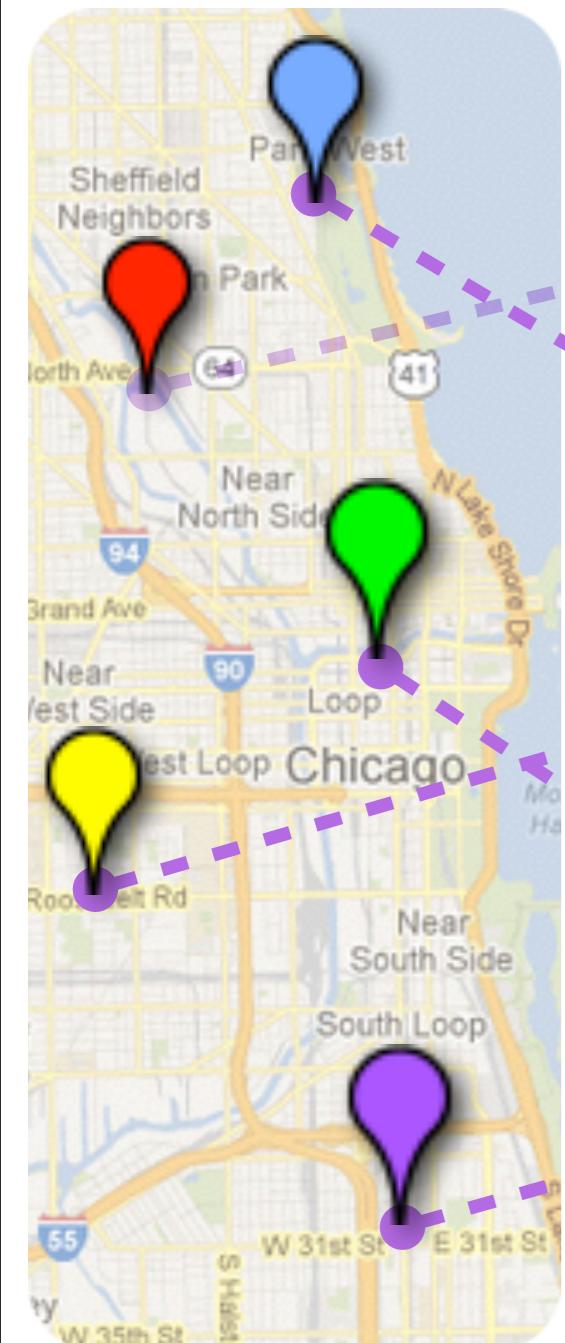




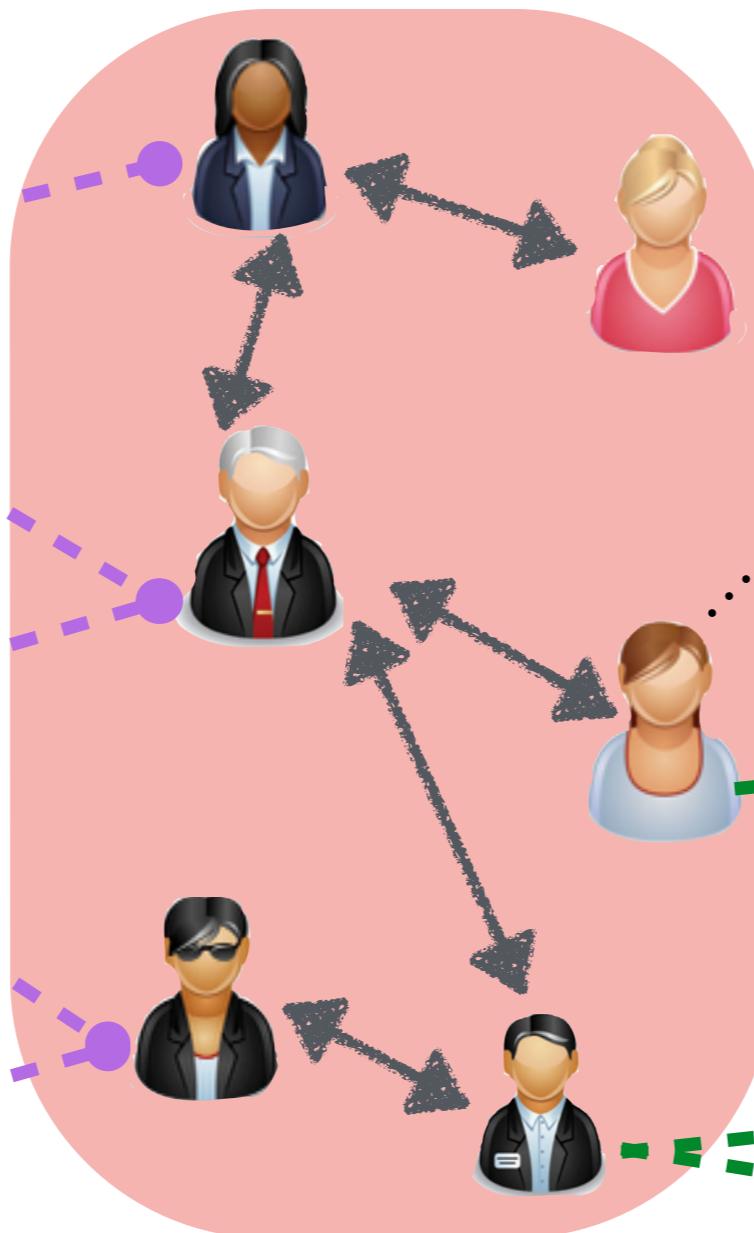
Solve Challenge 3: Cold Start Problem

Brand New Network

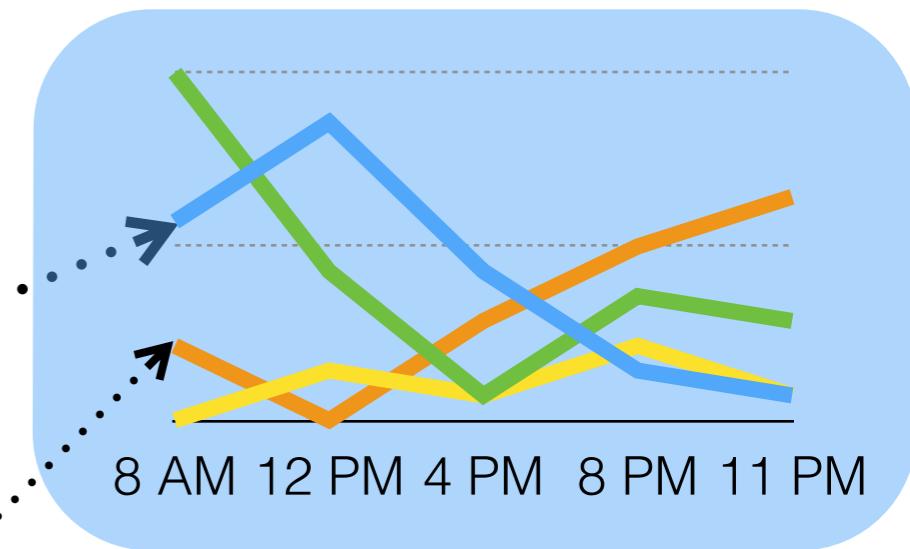
Locations



Social Links



Temporal Activities



Contents: Tweets







Add Friends

Foursquare is better with your friends!

Find friends already using Foursquare via other networks around the web, or invite your friends using their email address

facebook

Gmail™
by Google



YAHOO!

Friends not on Foursquare? Invite them!

Invite your friends to Foursquare [via Email](#).

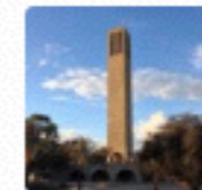
People you may know



ADD FRIEND



ADD FRIEND



ADD FRIEND



New



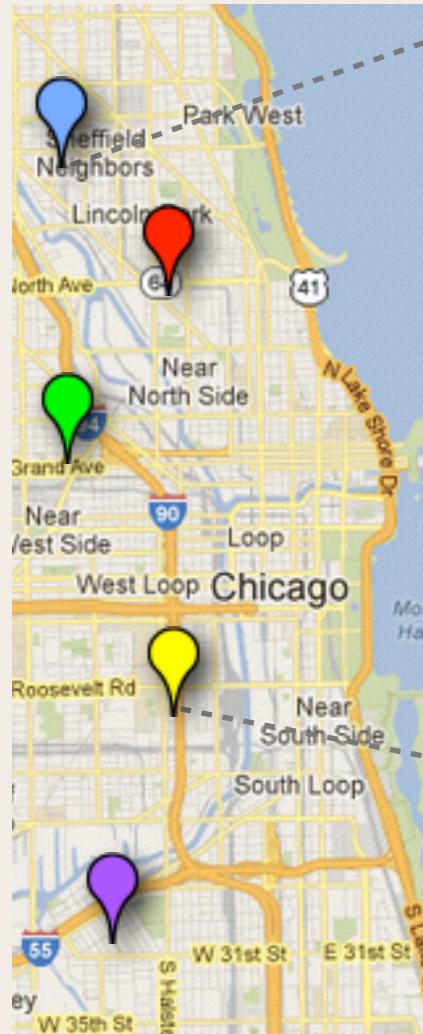
Temporal Activities

target network



User Accounts

Locations



Tips



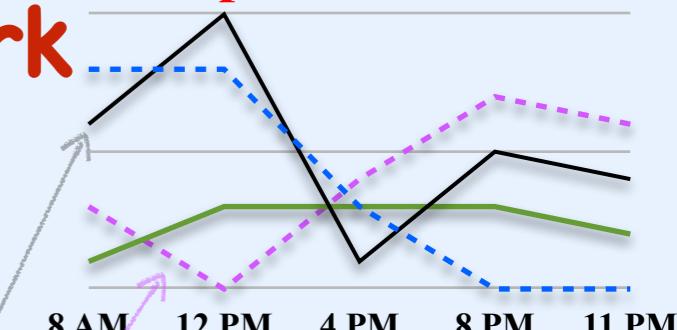
Old



source network

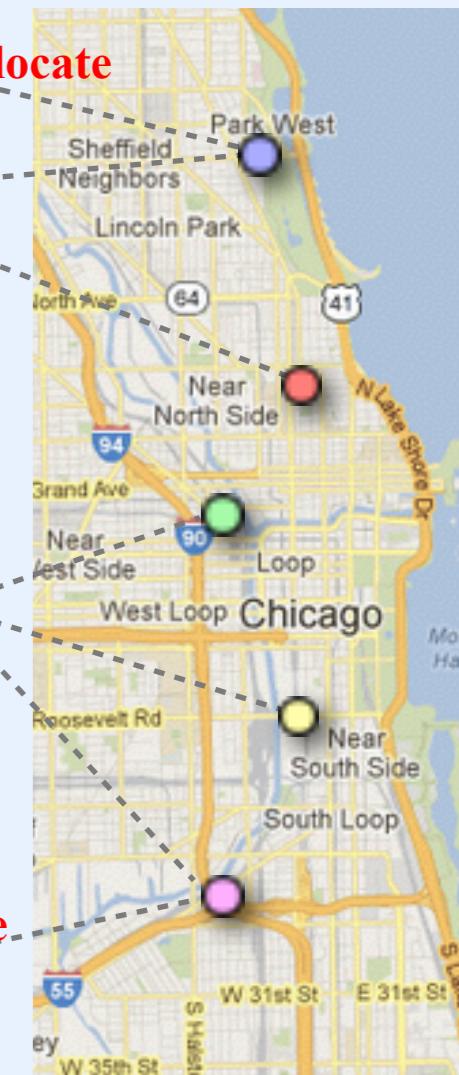
User Accounts

Temporal Activities



Locations

locate



Tweets

locate



Anchor Links across Aligned Networks



New



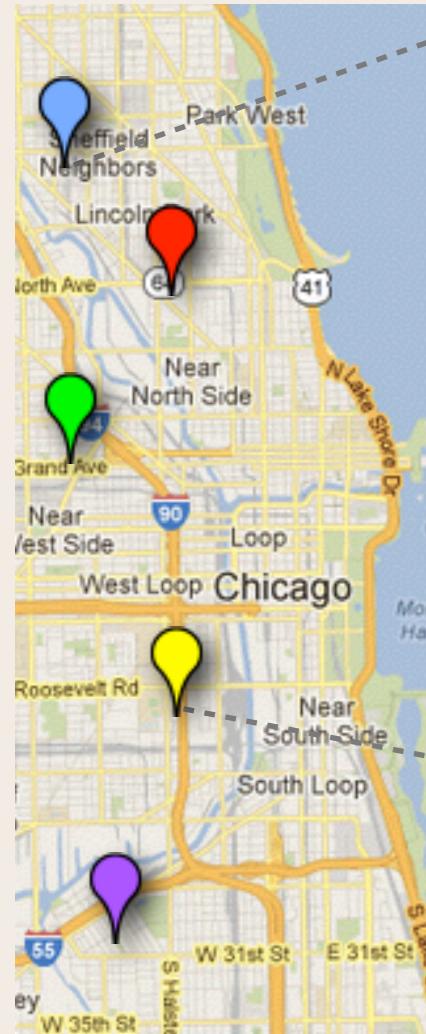
Temporal Activities

target network



User Accounts

Locations



Tips



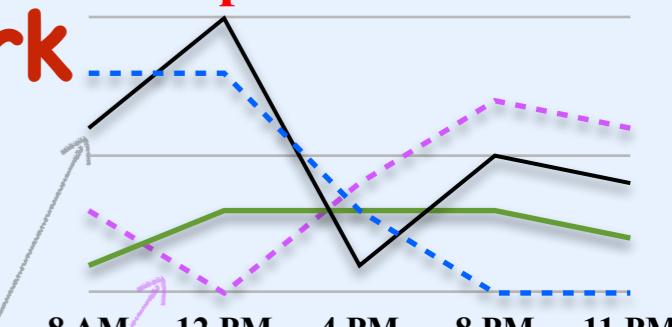
Old



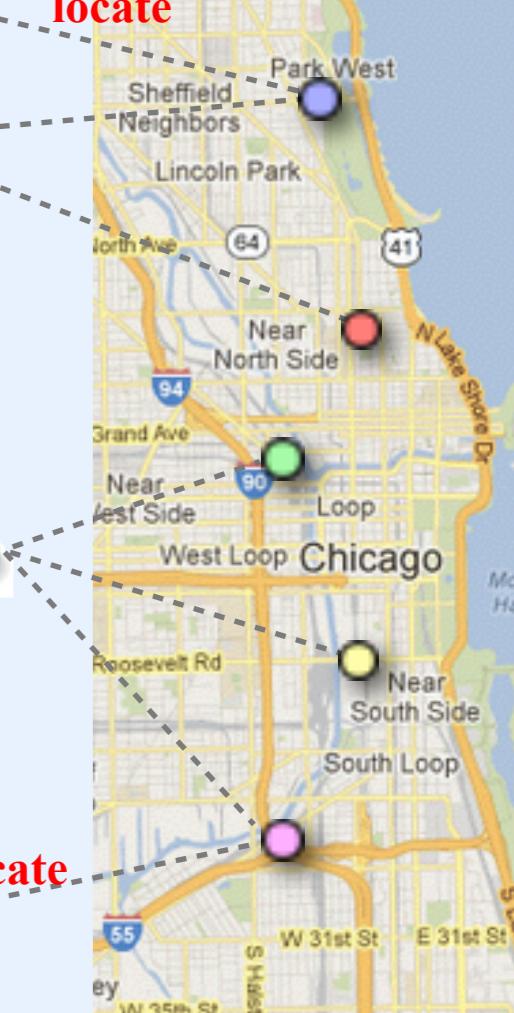
source network

User Accounts

Temporal Activities



Locations



locate

Tweets



Experiments

Data Sets

	property	Twitter	Foursquare
# node	user	5,223	5,392
	tweet/tip	9,490,707	48,756
	location	297,182	38,921
# link	friend/follow	164,920	31,312
	write	9,490,707	48,756
	locate	615,515	48,756



Evaluation Metric

1. Ground Truth

existing social and location links

2. Evaluation Metric

(1) Accuracy

(2) AUC

Experiment Results

collective link prediction

independent link prediction

link	measure	methods	remaining information rates σ				
			0.1	0.2	0.3	0.4	0.5
social	Accuracy	TRAIL	0.810±0.012	0.824±0.009	0.837±0.008	0.844±0.009	0.832±0.003
		TRAILT	0.691±0.040	0.684±0.039	0.704±0.033	0.729±0.006	0.718±0.020
		TRAILS	0.572±0.007	0.578±0.007	0.580±0.004	0.575±0.012	0.580±0.011
		SCAN	0.772±0.050	0.788±0.004	0.811±0.009	0.830±0.005	0.809±0.004
		SCANT	0.524±0.023	0.559±0.008	0.559±0.017	0.554±0.044	0.630±0.008
		SCANS	0.583±0.005	0.579±0.003	0.583±0.010	0.562±0.005	0.579±0.004
	AUC	CN	0.494±0.002	0.500±0.015	0.504±0.006	0.496±0.012	0.495±0.018
		JC	0.497±0.003	0.503±0.004	0.501±0.002	0.502±0.010	0.496±0.008
		AA	0.494±0.002	0.499±0.014	0.501±0.006	0.494±0.012	0.492±0.018
		TRAIL	0.855±0.002	0.849±0.004	0.850±0.008	0.854±0.005	0.850±0.003
		TRAILT	0.622±0.046	0.627±0.036	0.655±0.022	0.676±0.009	0.674±0.019
		TRAILS	0.548±0.004	0.551±0.006	0.552±0.004	0.549±0.000	0.551±0.002
location	Accuracy	SCAN	0.747±0.003	0.752±0.007	0.748±0.000	0.754±0.008	0.746±0.005
		SCANT	0.512±0.009	0.522±0.002	0.520±0.001	0.537±0.006	0.554±0.008
		SCANS	0.557±0.002	0.547±0.006	0.553±0.002	0.545±0.006	0.552±0.007
		NAIVE	0.525±0.014	0.526±0.006	0.525±0.008	0.526±0.007	0.525±0.013
		TRAIL	0.848±0.005	0.856±0.010	0.870±0.010	0.878±0.007	0.899±0.007
		TRAILT	0.839±0.006	0.850±0.003	0.857±0.009	0.866±0.008	0.862±0.005
	AUC	TRAILS	0.631±0.003	0.632±0.002	0.631±0.001	0.634±0.001	0.634±0.002
		SCAN	0.712±0.010	0.757±0.002	0.758±0.009	0.770±0.005	0.775±0.005
		SCANT	0.676±0.009	0.711±0.005	0.730±0.005	0.749±0.003	0.756±0.001
		SCANS	0.633±0.003	0.633±0.003	0.633±0.001	0.636±0.001	0.637±0.000
		FCF	0.598±0.008	0.638±0.015	0.638±0.005	0.654±0.012	0.664±0.007
		TRAIL	0.719±0.004	0.736±0.001	0.749±0.006	0.754±0.003	0.753±0.002
geographical	Accuracy	TRAILT	0.674±0.009	0.697±0.004	0.706±0.005	0.709±0.001	0.717±0.006
		TRAILS	0.536±0.003	0.527±0.001	0.537±0.005	0.553±0.003	0.560±0.002
		SCAN	0.658±0.000	0.670±0.002	0.682±0.001	0.697±0.003	0.699±0.003
	AUC	SCANT	0.610±0.001	0.623±0.001	0.631±0.001	0.647±0.001	0.653±0.002
		SCANS	0.536±0.025	0.531±0.008	0.535±0.002	0.547±0.004	0.557±0.004
		NAIVE	0.536±0.014	0.536±0.002	0.536±0.001	0.537±0.008	0.536±0.012

Parameter Analysis

link	measure	methods	anchor link sample rates ρ			
			0.0	0.2	0.4	0.6
social	AUC	TRAIL	0.712±0.004	0.733±0.019	0.761±0.017	0.782±0.007
		TRAILT	0.712±0.012	0.711±0.007	0.711±0.012	0.711±0.010
		TRAILS	0.500±0.000	0.507±0.005	0.524±0.005	0.555±0.036
		SCAN	0.603±0.020	0.621±0.036	0.539±0.022	0.664±0.026
		SCANT	0.603±0.009	0.603±0.014	0.603±0.016	0.603±0.027
		SCANS	0.500±0.000	0.496±0.001	0.513±0.013	0.515±0.015
	CN	CN	0.525±0.000	0.525±0.008	0.524±0.013	0.525±0.005
		JC	0.527±0.008	0.527±0.011	0.527±0.010	0.528±0.002
		AA	0.493±0.006	0.490±0.006	0.490±0.012	0.490±0.009
	Accuracy	TRAIL	0.654±0.014	0.746±0.009	0.756±0.009	0.764±0.008
		TRAILT	0.655±0.004	0.653±0.008	0.655±0.014	0.655±0.008
		TRAILS	0.500±0.000	0.501±0.003	0.535±0.009	0.529±0.006
		SCAN	0.554±0.028	0.567±0.009	0.563±0.007	0.605±0.014
		SCANT	0.553±0.002	0.553±0.004	0.553±0.003	0.554±0.002
		SCANS	0.500±0.000	0.498±0.003	0.515±0.008	0.529±0.003
location	AUC	NAIVE	0.500±0.000	0.508±0.001	0.514±0.006	0.517±0.002
		TRAIL	0.871±0.020	0.876±0.011	0.891±0.006	0.881±0.028
		TRAILT	0.871±0.015	0.872±0.004	0.872±0.013	0.872±0.003
		TRAILS	0.500±0.000	0.492±0.002	0.479±0.004	0.504±0.002
		SCAN	0.745±0.005	0.746±0.011	0.773±0.010	0.788±0.012
		SCANT	0.745±0.021	0.744±0.011	0.745±0.025	0.744±0.020
	FCF	SCANS	0.500±0.000	0.490±0.002	0.481±0.002	0.504±0.001
		FCF	0.682±0.006	0.683±0.002	0.682±0.007	0.683±0.002
		TRAIL	0.734±0.008	0.754±0.005	0.765±0.006	0.775±0.003
	Accuracy	TRAILT	0.735±0.002	0.734±0.007	0.734±0.007	0.734±0.006
		TRAILS	0.500±0.000	0.509±0.003	0.514±0.006	0.511±0.001
		SCAN	0.731±0.002	0.753±0.001	0.754±0.002	0.755±0.002
		SCANT	0.732±0.013	0.732±0.010	0.732±0.016	0.732±0.009
		SCANS	0.500±0.000	0.511±0.002	0.516±0.006	0.517±0.005
		NAIVE	0.500±0.000	0.509±0.001	0.517±0.001	0.517±0.005

Summary

1. we study the **collective link prediction problem** simultaneously: **social links & location links**
2. we use information from **multiple aligned networks** simultaneously: **new network & aligned old network.**
3. we propose a tentative method to solve the **cold start problem!**

Q & A