

A Data Mashup for Understanding Fine Scale Human Dynamics

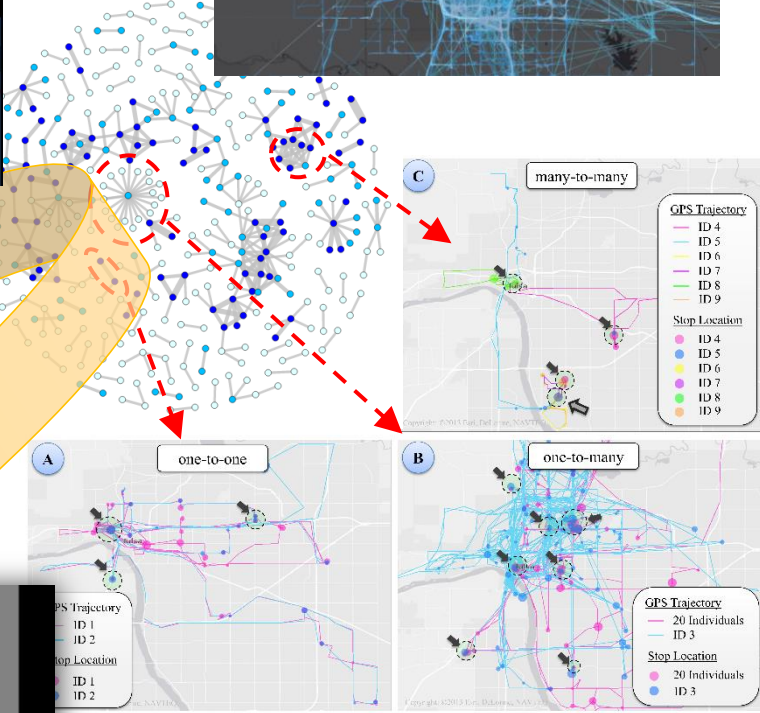
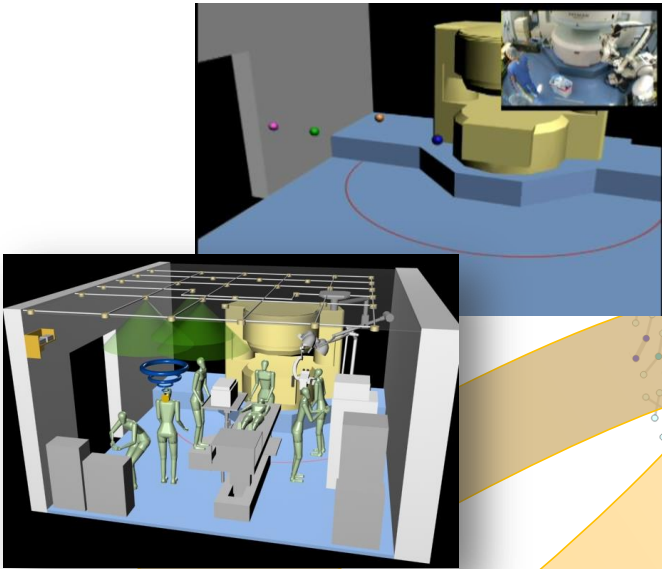
Atsushi Nara

Assistant Professor, Department of Geography

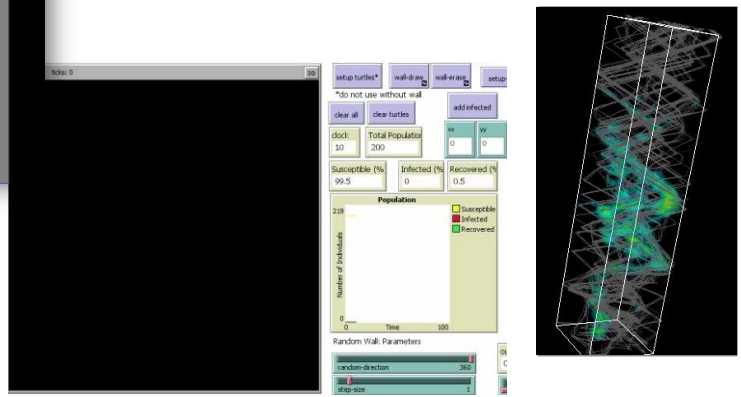
Associate Director, HDMA

San Diego State University

Introduction



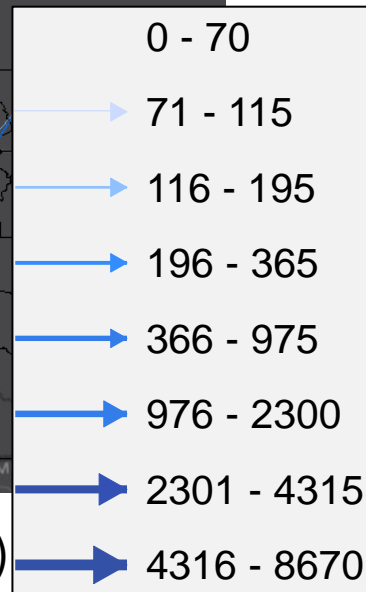
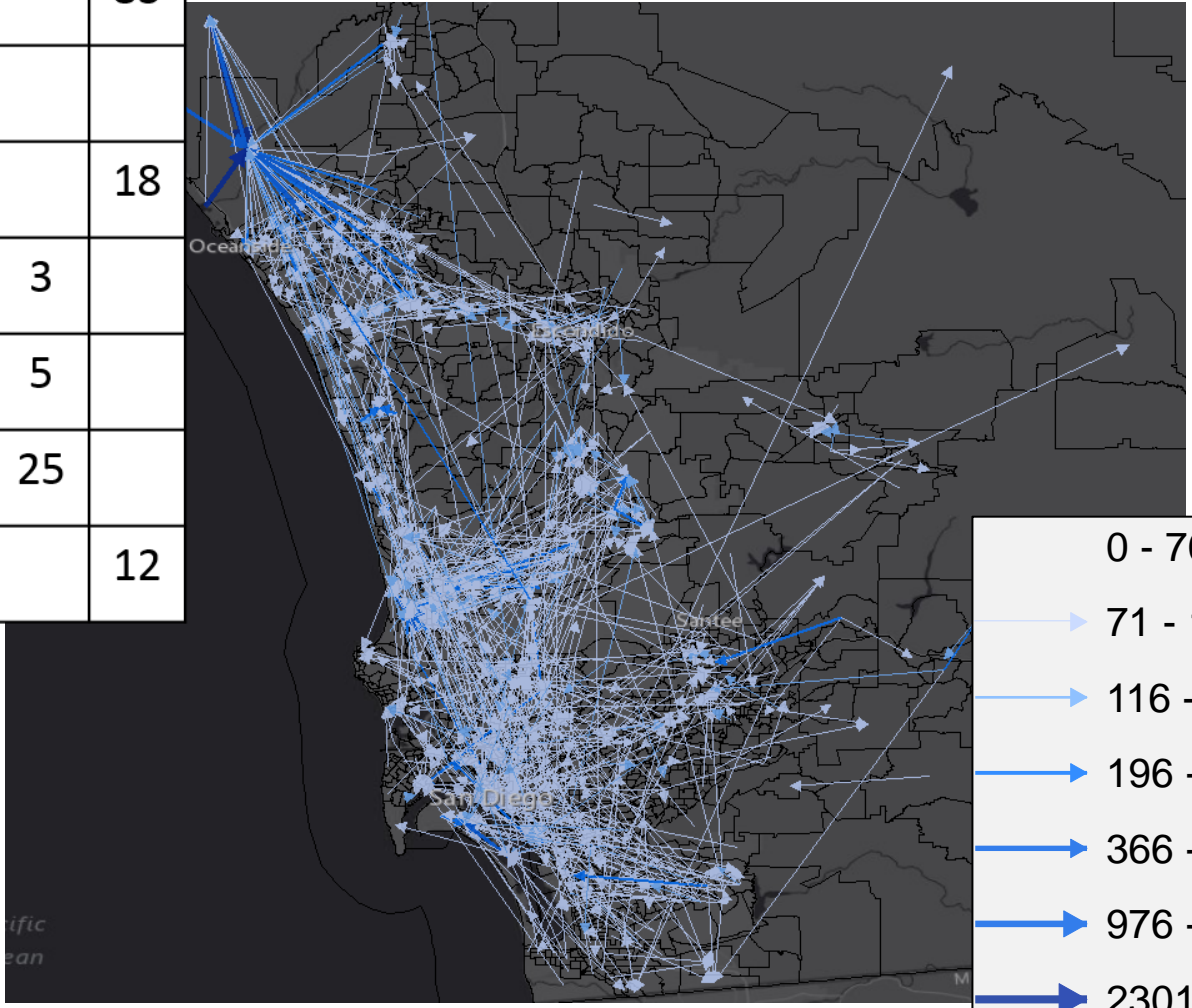
Evacuation
Evacuation2
Evacuation3
Evacuation4
Evacuation5
Evacuation6
Exit



Introduction

OD-Matrix

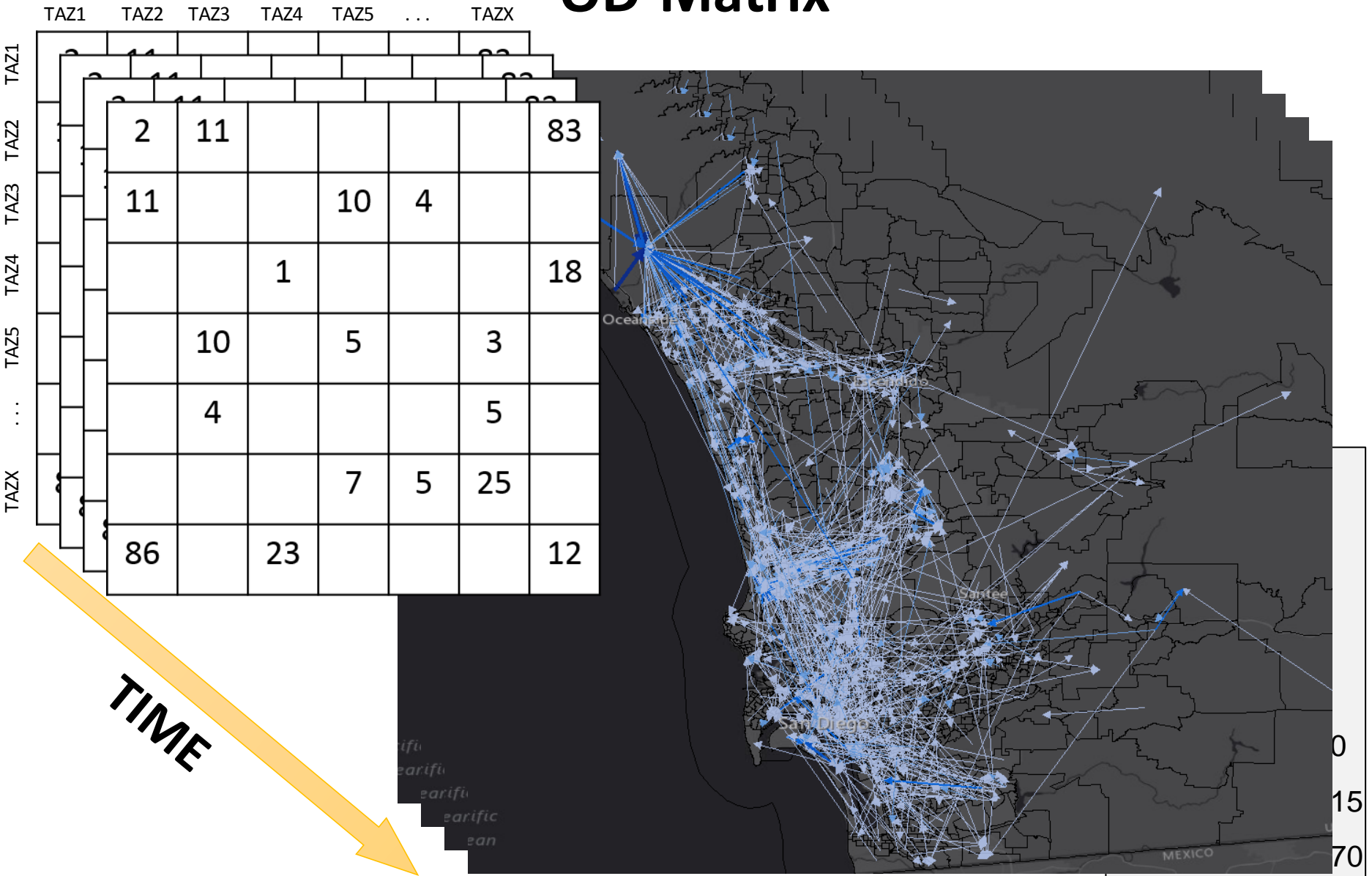
	TAZ1	TAZ2	TAZ3	TAZ4	TAZ5	...	TAZX
TAZ1	2	11					83
TAZ2	11			10	4		
TAZ3			1				18
TAZ4		10		5		3	
TAZ5		4				5	
...				7	5	25	
TAZX	86		23				12



Flow based on the number of workers (TAZ)

Introduction

OD-Matrix



OD-Matrix Data Sources

- Census
- Department of Transportation
- Person Trip Survey Data
(e.g., PFLOW project at Univ. of Tokyo)
- Commercial Company
 - Car navigation
 - Mobile Phone
 - IT

- **Micro Human Dynamics**
 - Finer Spatial Scale
 - Finer Temporal Scale
- **Multi Human Dynamics**
 - More variations
- **More Accessible Data**



Data Mashups using Social Media Data ?

Approach

Data Mashup is the integration of heterogeneous data from multiple sources as a way of producing new insights and services, a mashup is often done in a web platform.

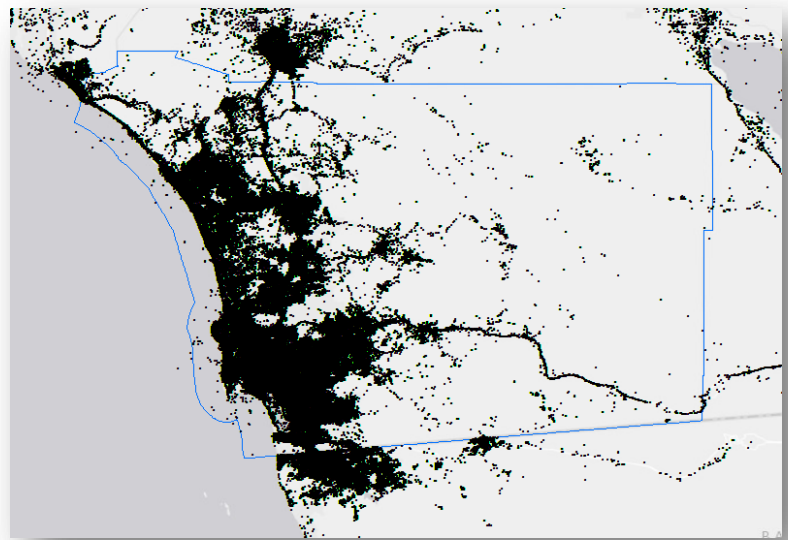


Data Collection

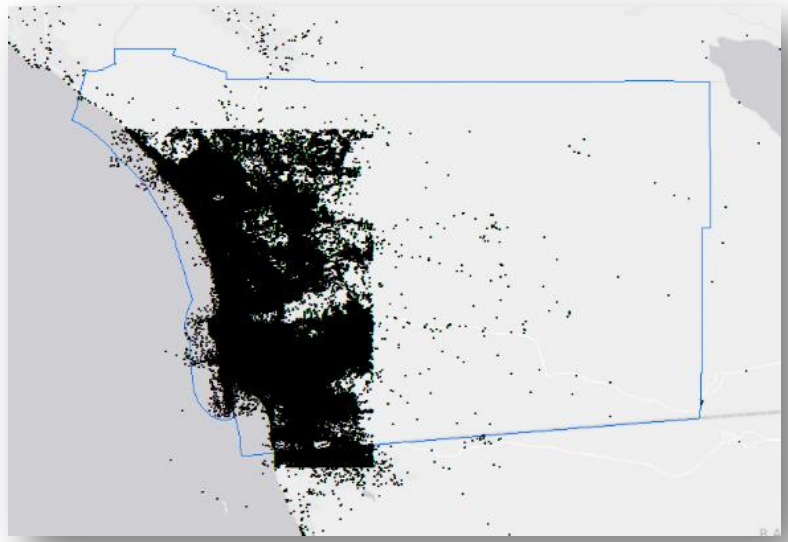
Social Media	# of Geotagging Users	# of Geotagged Posts
Instagram	410,528	4,362,176
Twitter	56,836	1,916,580
Foursquare-Tw	3,789	40,161

From 2014/12/07 To 2015/05/17 (161 Days)

Twitter



Instagram



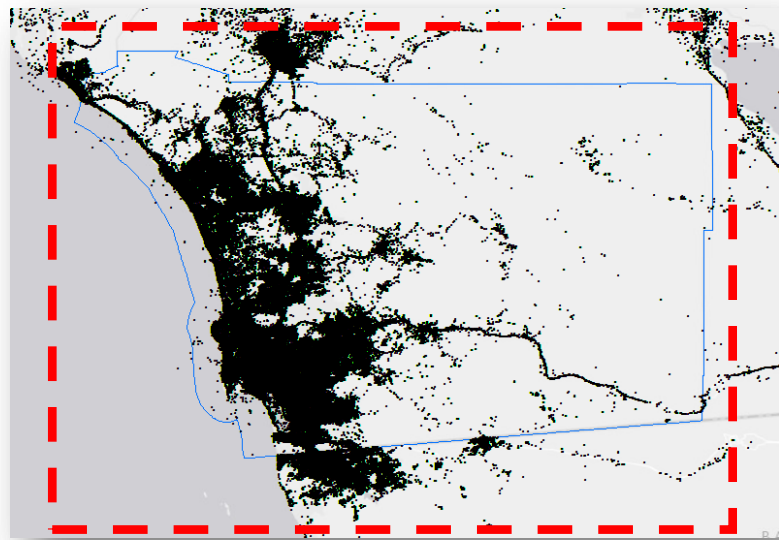
Foursquare cross-posted to Twitter



Data Collection



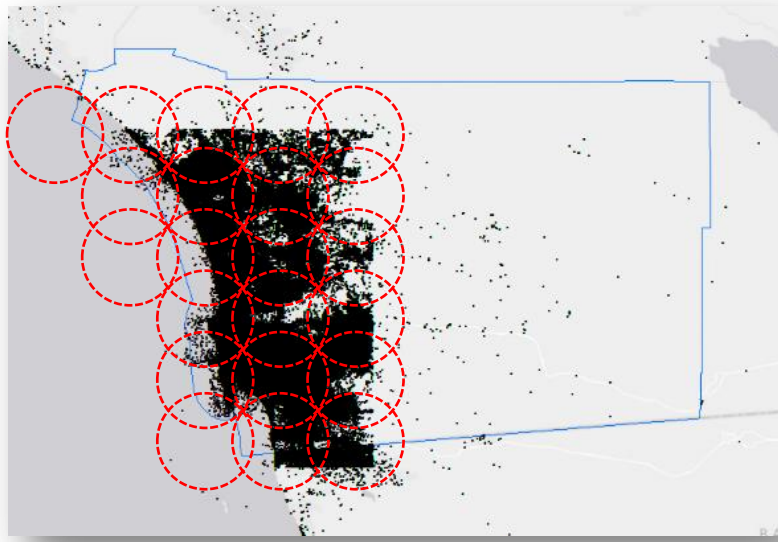
Twitter



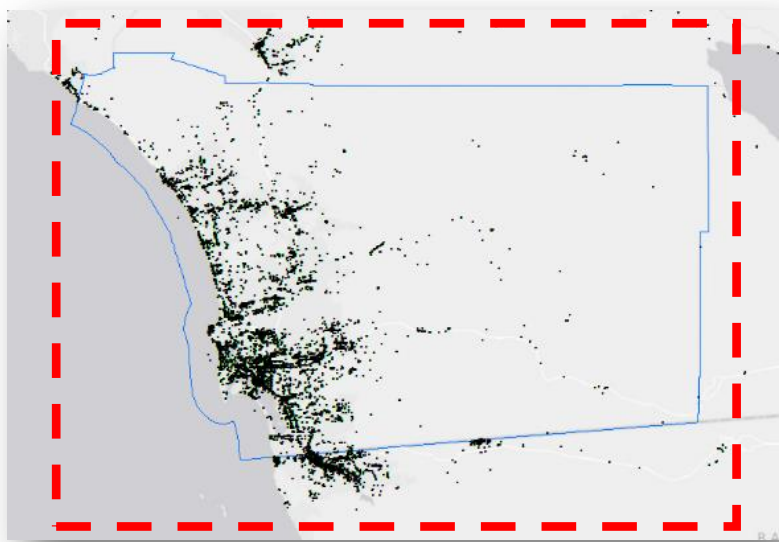
Social Media	# of Geotagging Users	# of Geotagged Posts
Instagram	410,528	4,362,176
Twitter	56,836	1,916,580
Foursquare-Tw	3,789	40,161

From 2014/12/07 To 2015/05/17 (161 Days)

Instagram



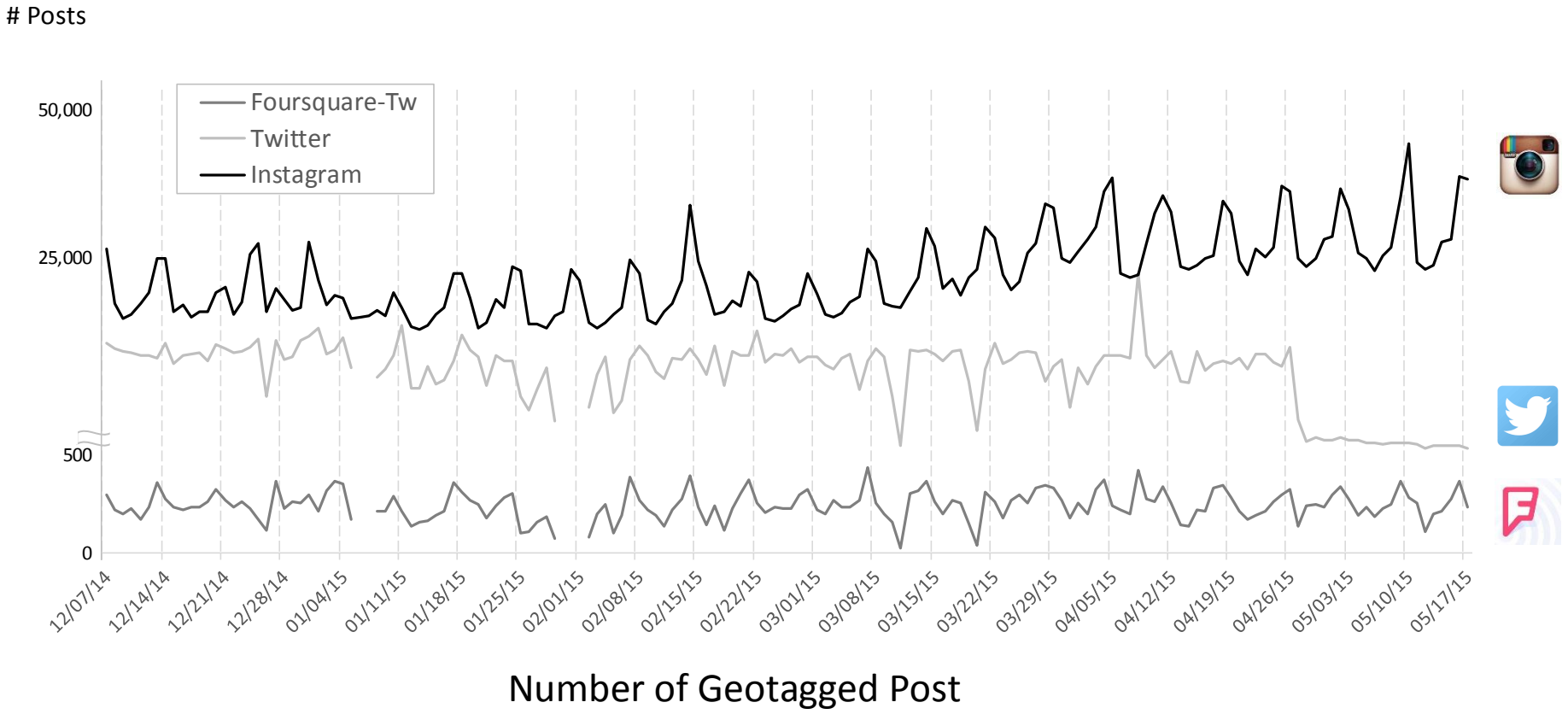
Foursquare cross-posted to Twitter



Data Collection

Social Media	# of Geotagging Users	# of Geotagged Posts
Instagram	410,528	4,362,176
Twitter	56,836	1,916,580
Foursquare-Tw	3,789	40,161

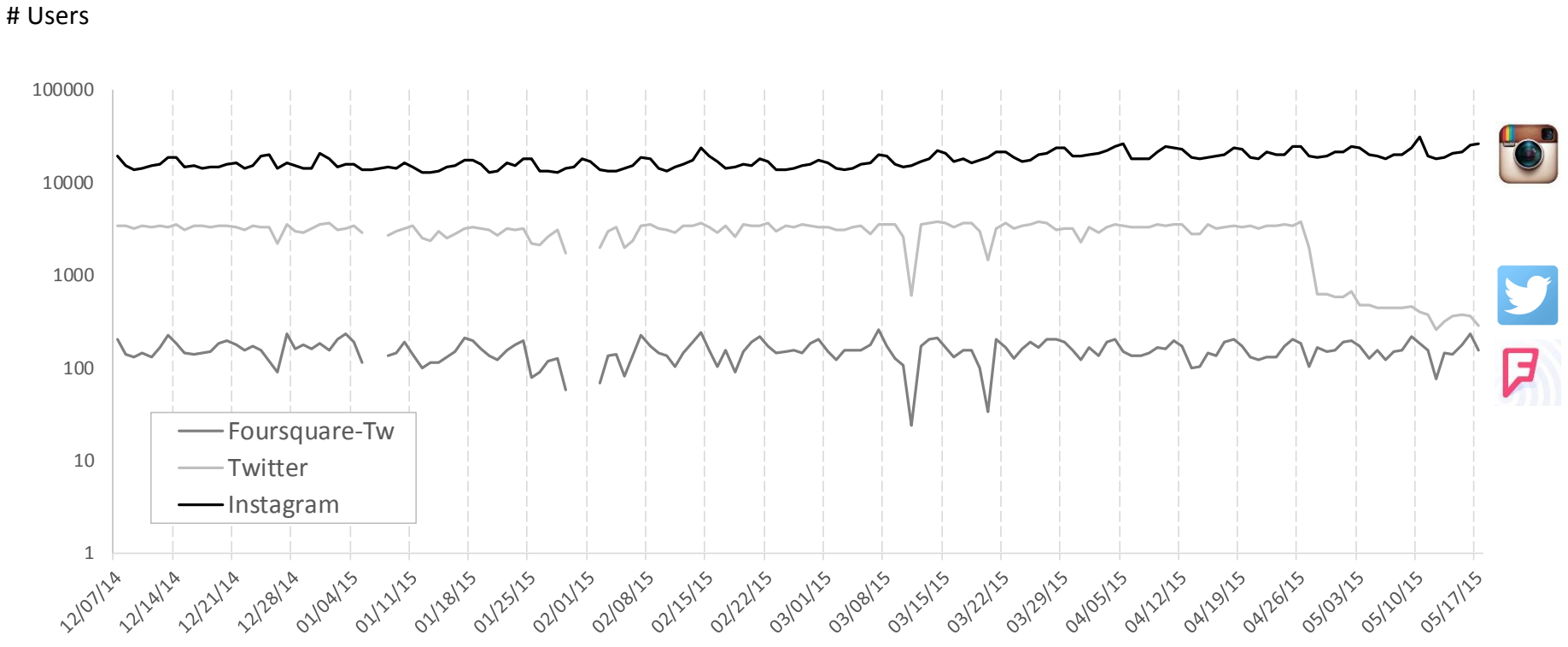
From 2014/12/07 To 2015/05/17 (161 Days)



Data Collection

Social Media	# of Geotagging Users per Day (Avg)	# of Geotagging Users per Day (Min)	# of Geotagging Users per Day (Max)
Instagram	17,857.48	13,049	31,039
Twitter	2,827.27	260	3,823
Foursquare-Tw	155.48	24	257

From 2014/12/07 To 2015/05/17 (161 Days)

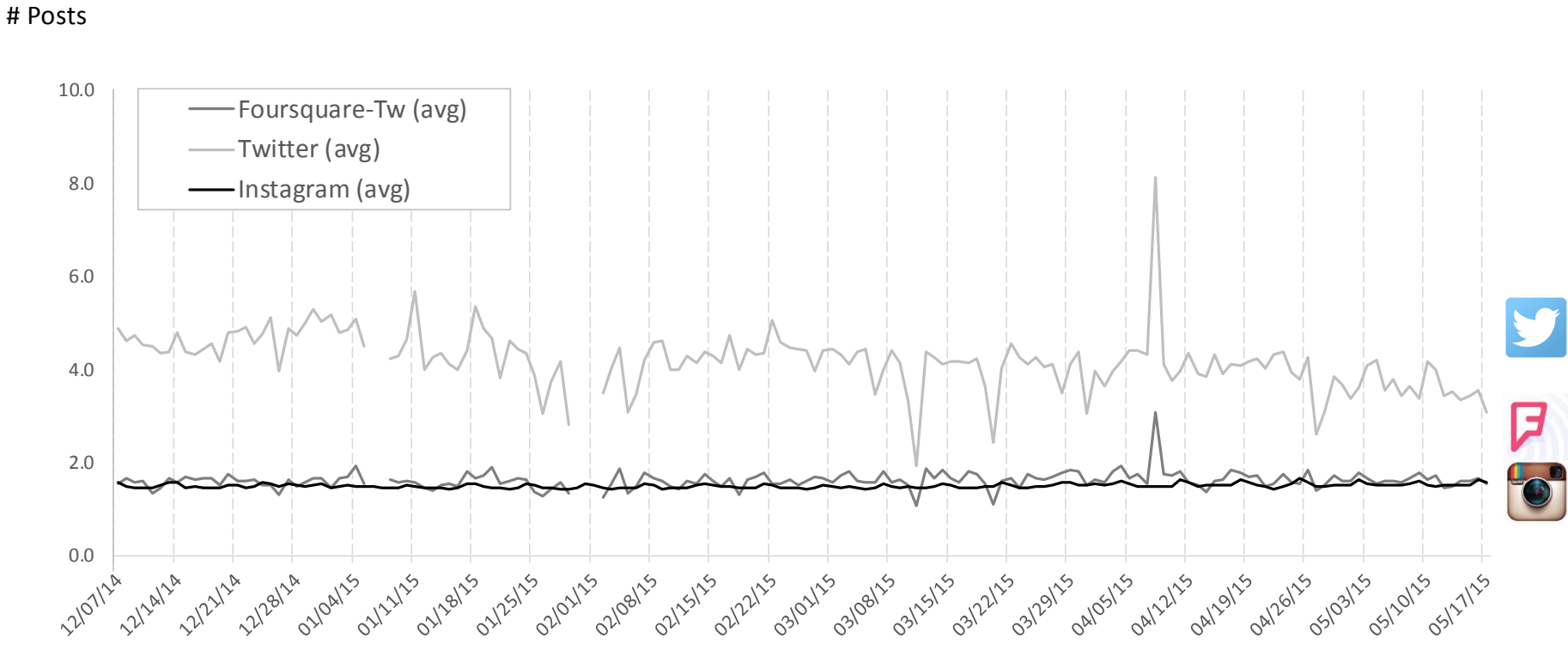


Number of Users who Post Geotagged Messages per Day

Data Collection

Social Media	Avg of Avg Geotagging Posts per Users per Day
Instagram	1.50
Twitter	4.19
Foursquare-Tw	1.62

From 2014/12/07 To 2015/05/17 (161 Days)



Daily Average of Geotagged Posts per Users

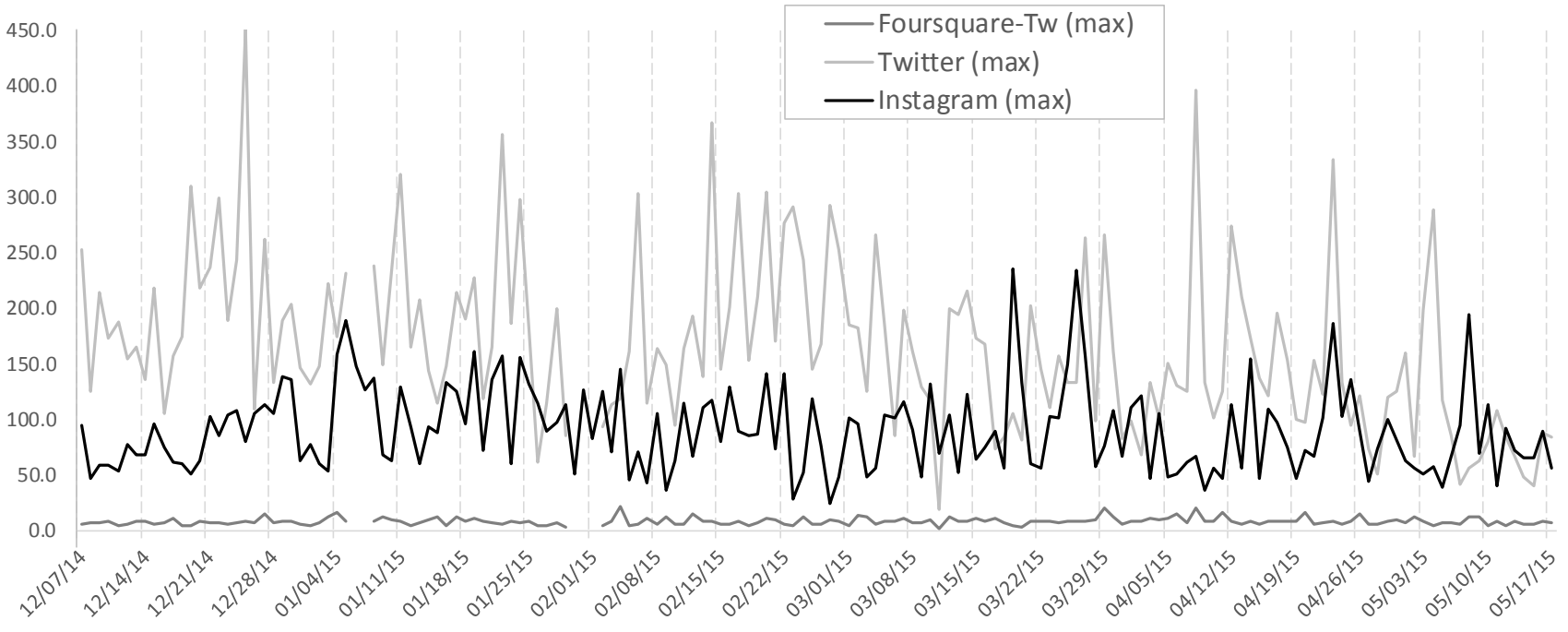


Data Collection

Social Media	Avg of Avg Geotagged Posts per Users per Day	Avg of Min Geotagged Posts among Users per Day	Avg of Max Geotagged Posts among Users per Day
Instagram	1.50	1.00	90.74
Twitter	4.19	1.00	166.86
Foursquare-Tw	1.62	1.00	8.46

From 2014/12/07 To 2015/05/17 (161 Days)

Posts



Daily Highest of Geotagged Posts by a Single User

Data Collection

- Instagram: Largest number of users, Largest amount of posts, Light users
- Twitter: Large amount of posts, Existence of heavy twitter users
- Foursquare: Less number of users and posts
- Larger amounts of posts during weekends and holidays for all social media
- # of Geotagging users are stable over time
- Technical issues affected data collection

Data Filtering

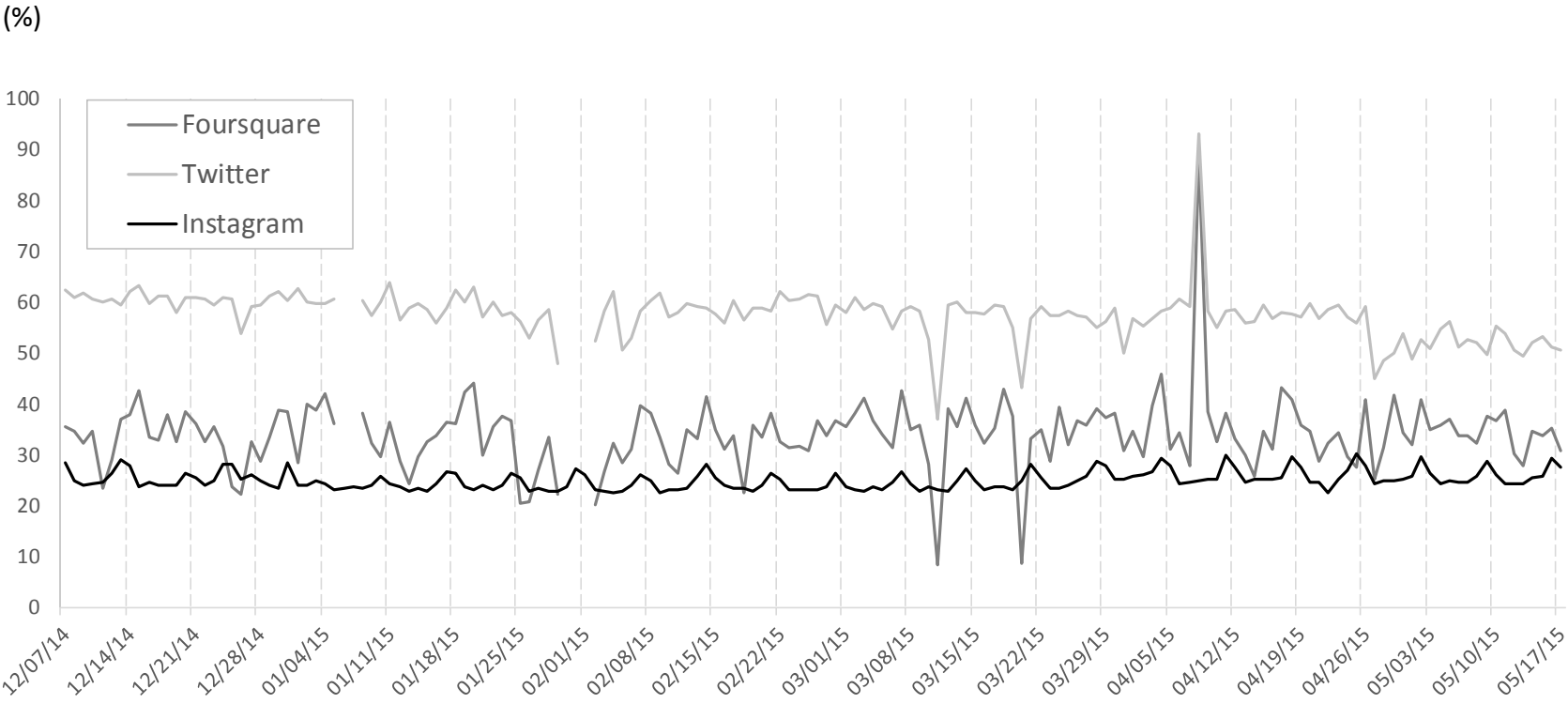


Only movements within San Diego county.
Excluding movements coming from/heading to outside of San Diego county

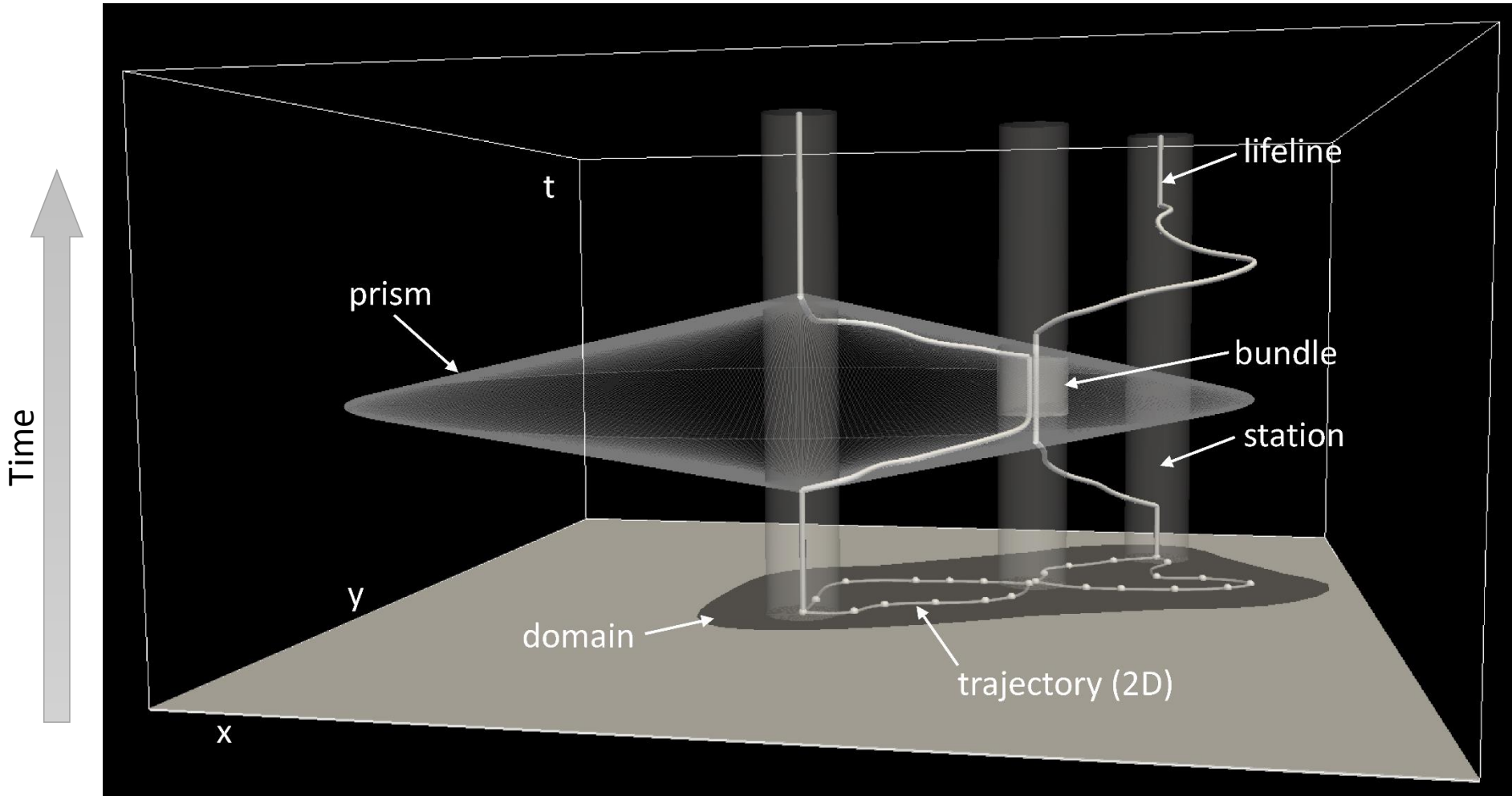
Data Filtering

Social Media	% of Users with 2 or more geotagged posts per day (Avg)
Instagram	25.03
Twitter	57.63
Foursquare-Tw	34.00

From 2014/12/07 To 2015/05/17 (161 Days)

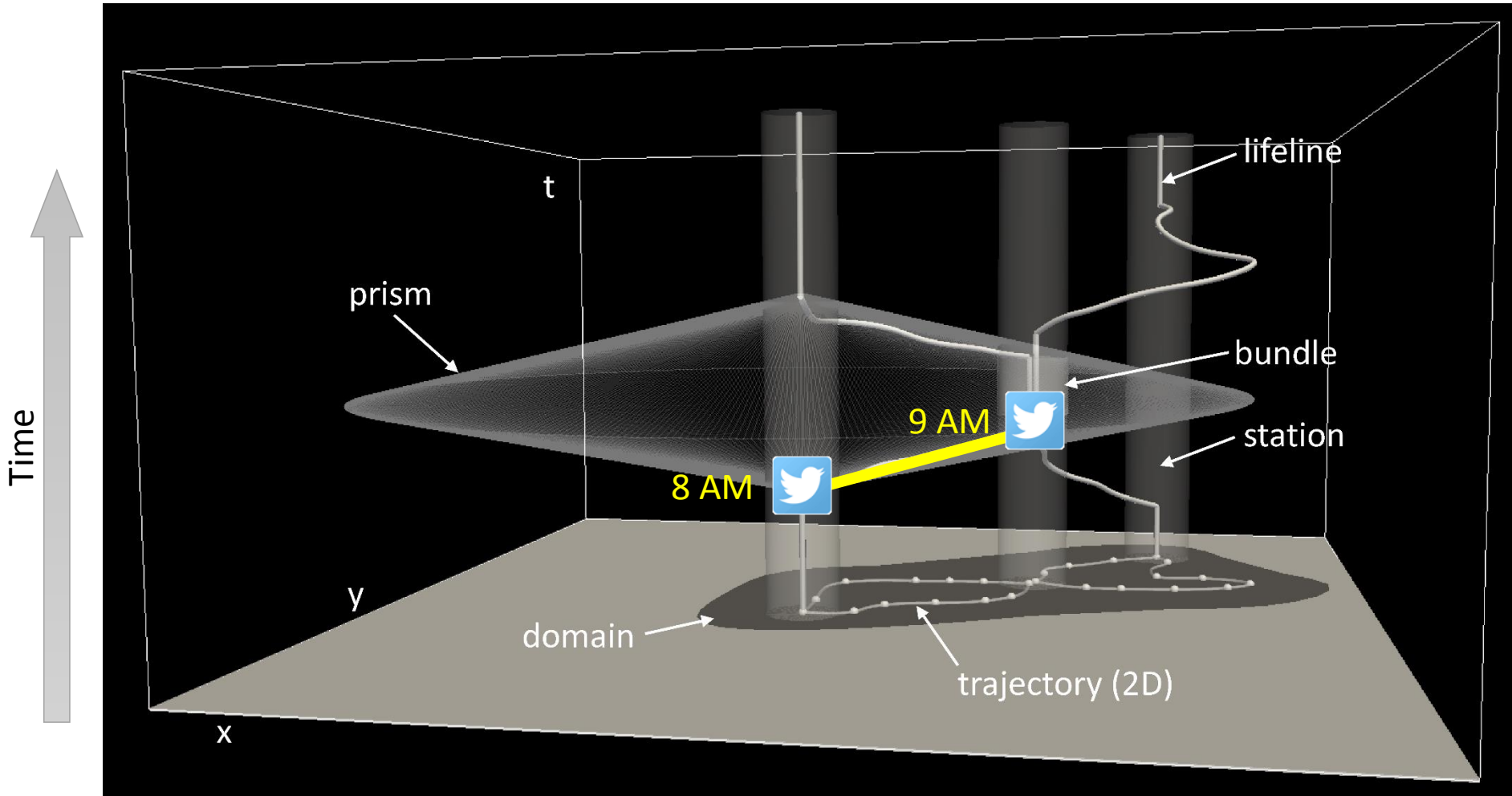


Data Filtering



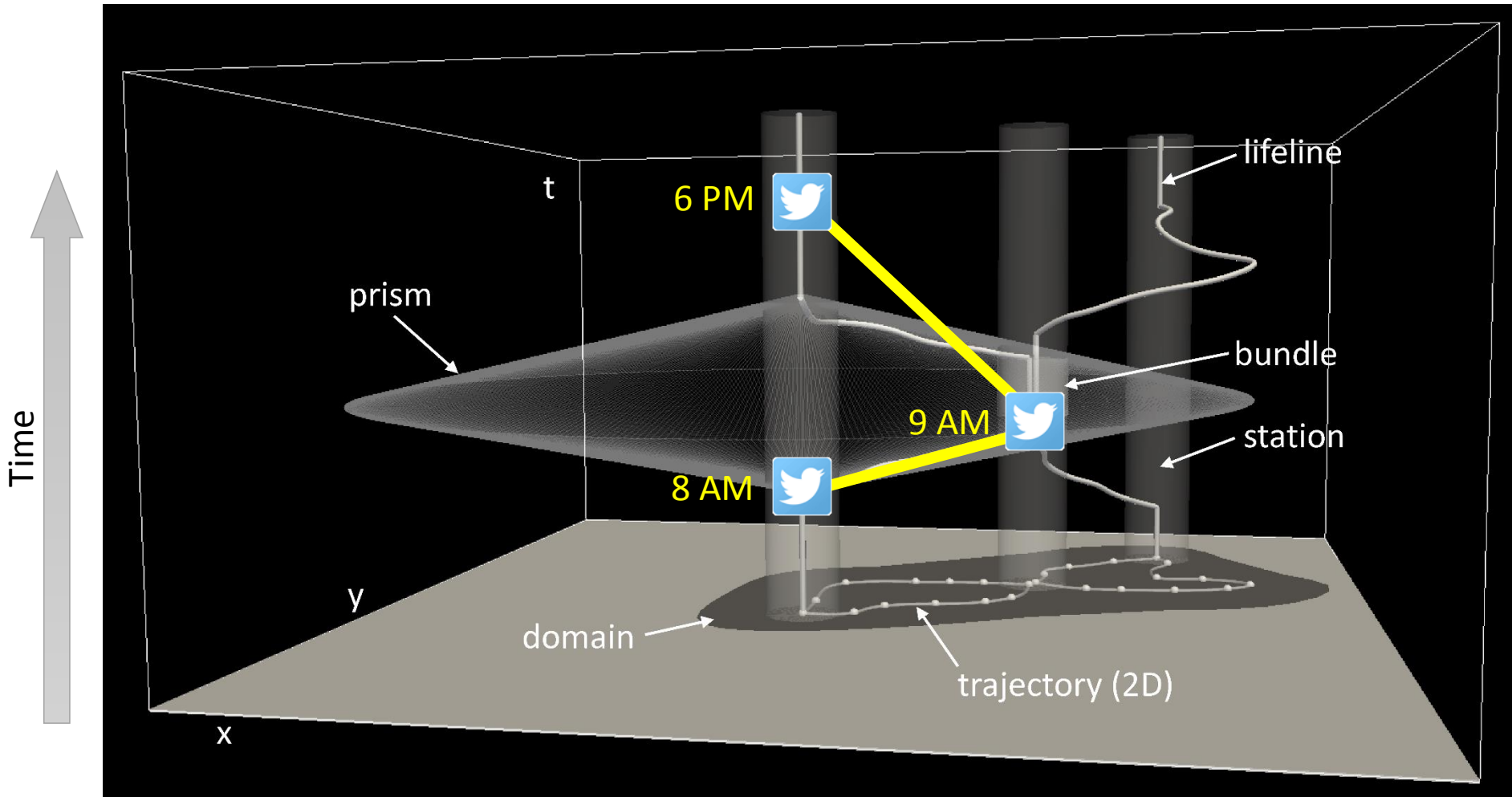
(Yuan & Nara, 2015)

Data Filtering



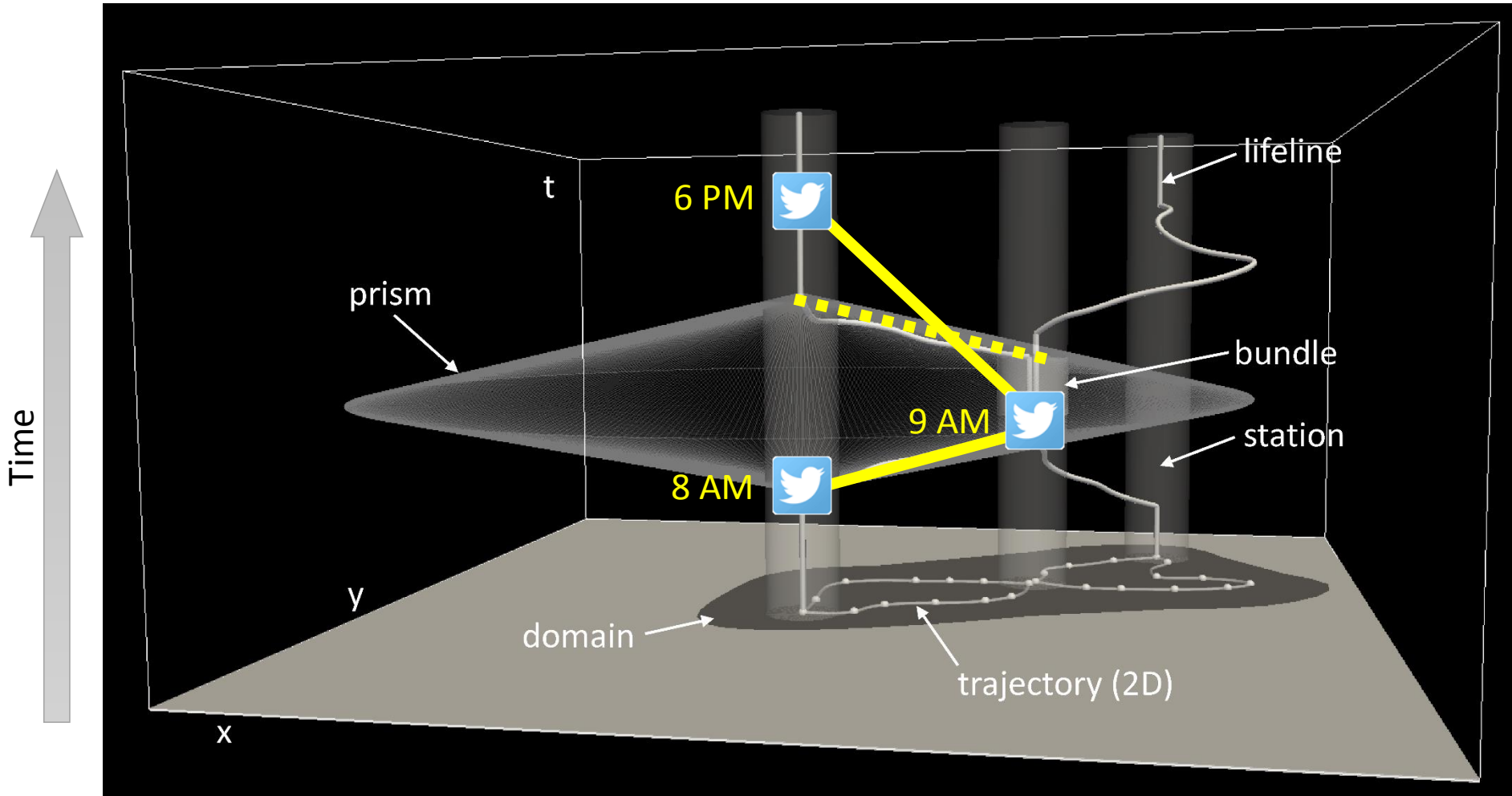
(Yuan & Nara, 2015)

Data Filtering



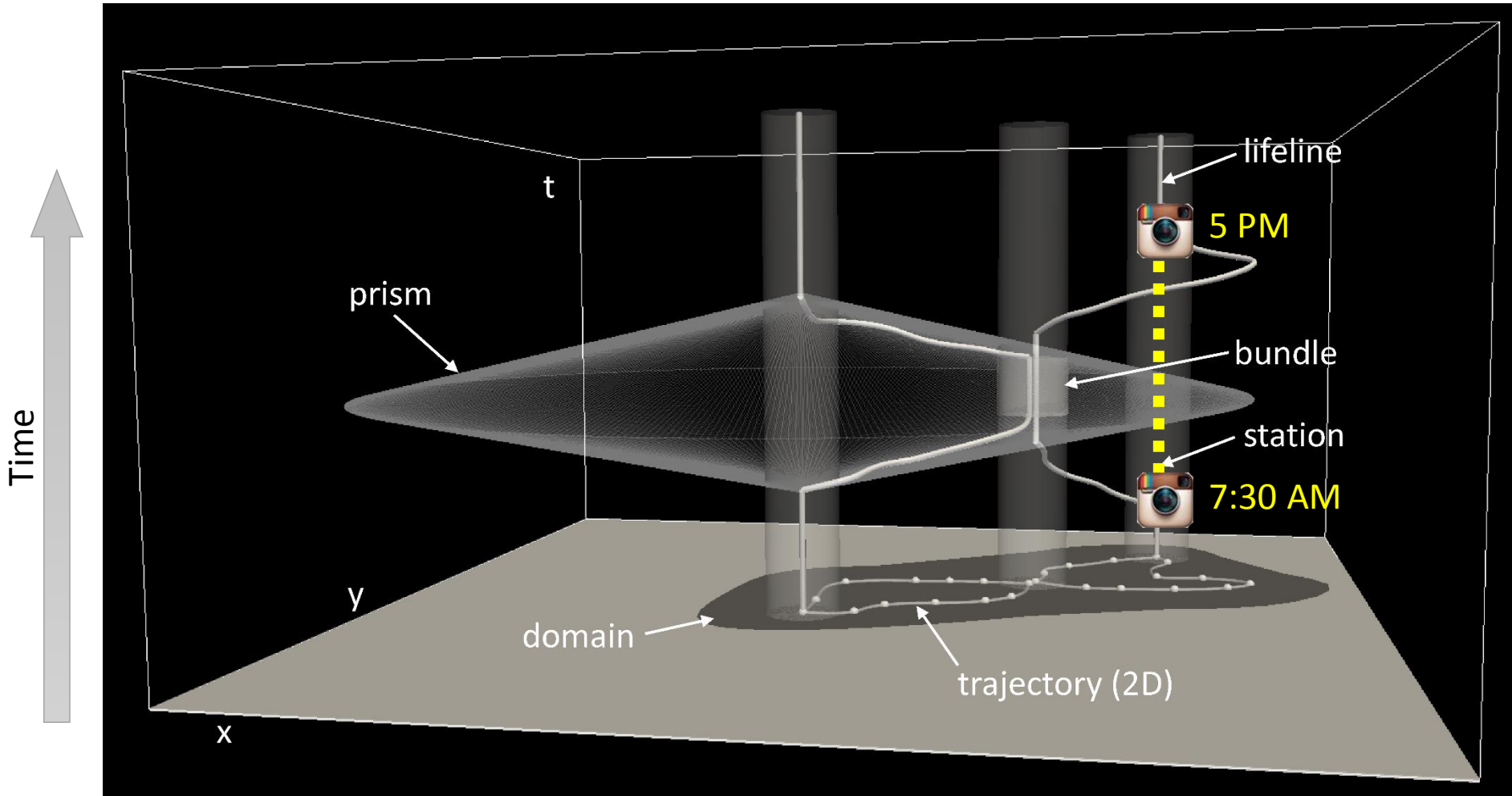
(Yuan & Nara, 2015)

Data Filtering



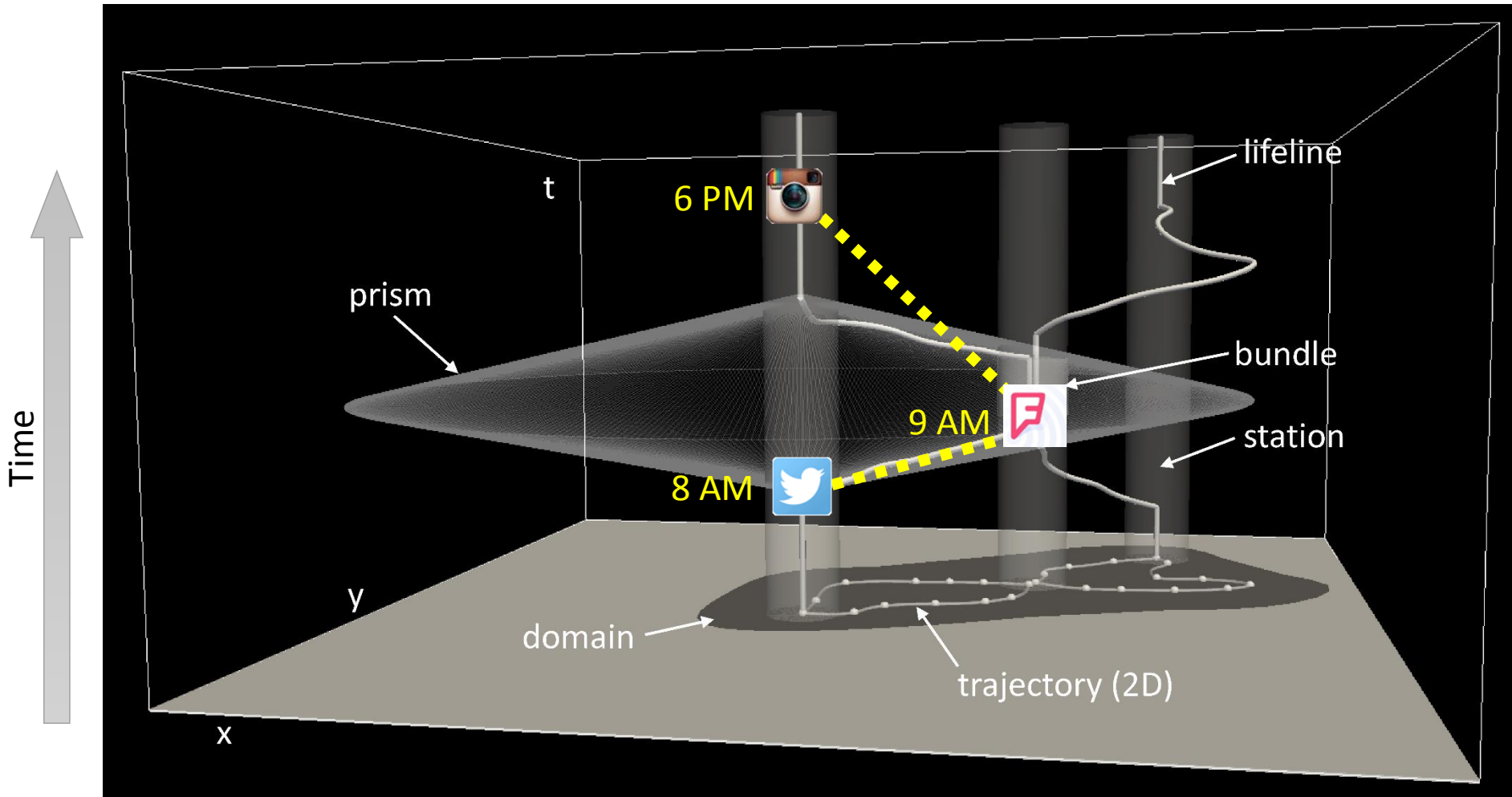
(Yuan & Nara, 2015)

Data Filtering



(Yuan & Nara, 2015)

Data Filtering



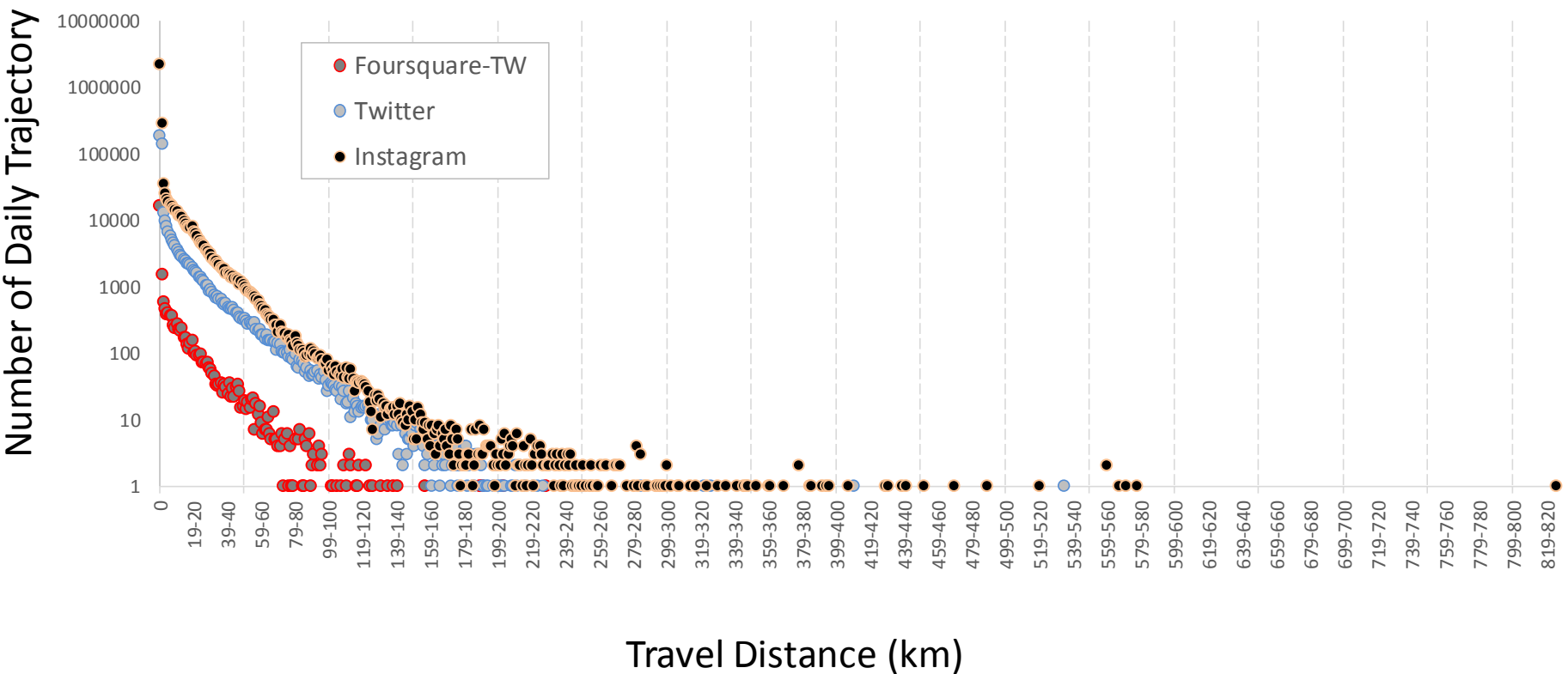
(Yuan & Nara, 2015)

Data Filtering

of Daily Trajectory

Social Media	Total	Distance = 0	D > 0
Instagram	2,892,912	2,238,847 (77.4%)	654,065 (33.6%)
Twitter	443,881	187,831 (42.3%)	256,050 (57.7%)
Foursquare-Tw	24,410	16,219 (66.4%)	8,191 (33.6%)
Total	3,361,203	2,442,897 (72.7%)	918,306 (27.3%)

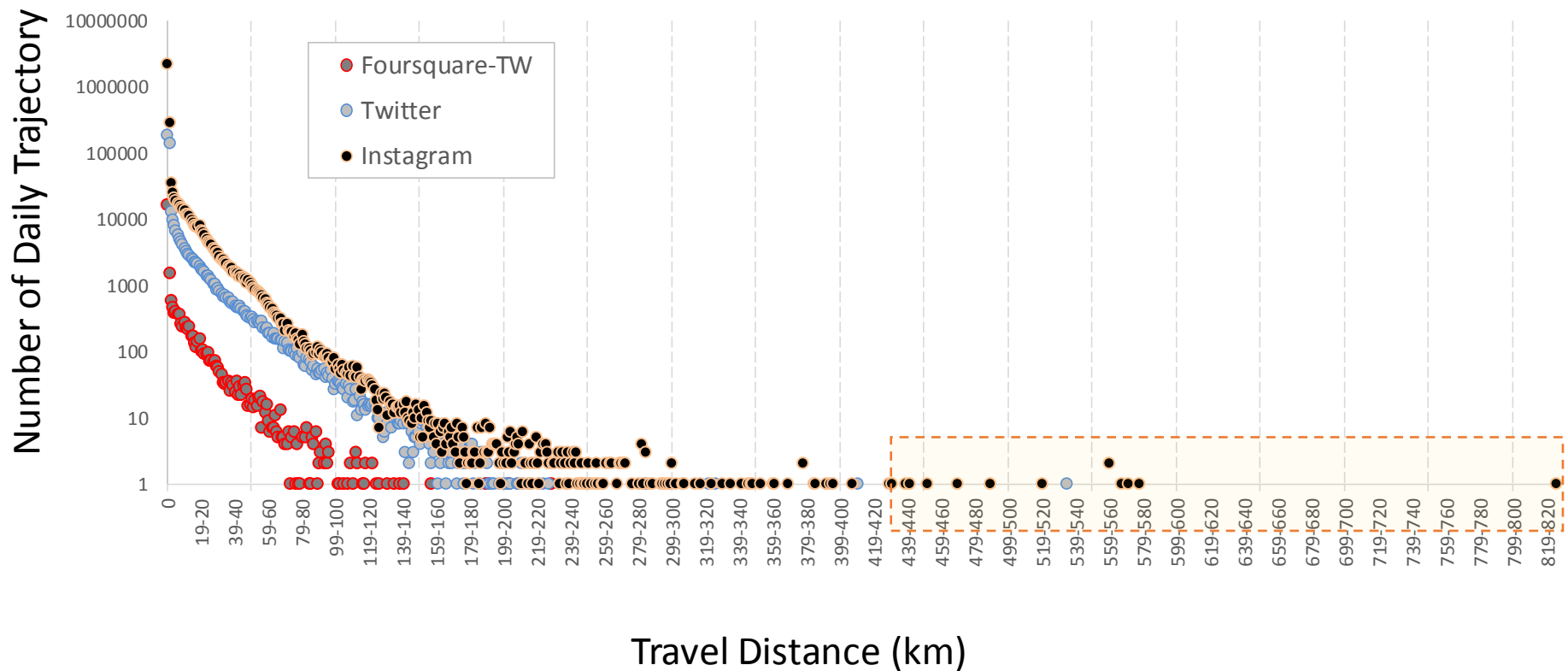
Frequency Distribution of Daily Trajectory Travel Distance (Bin Size: 1km)



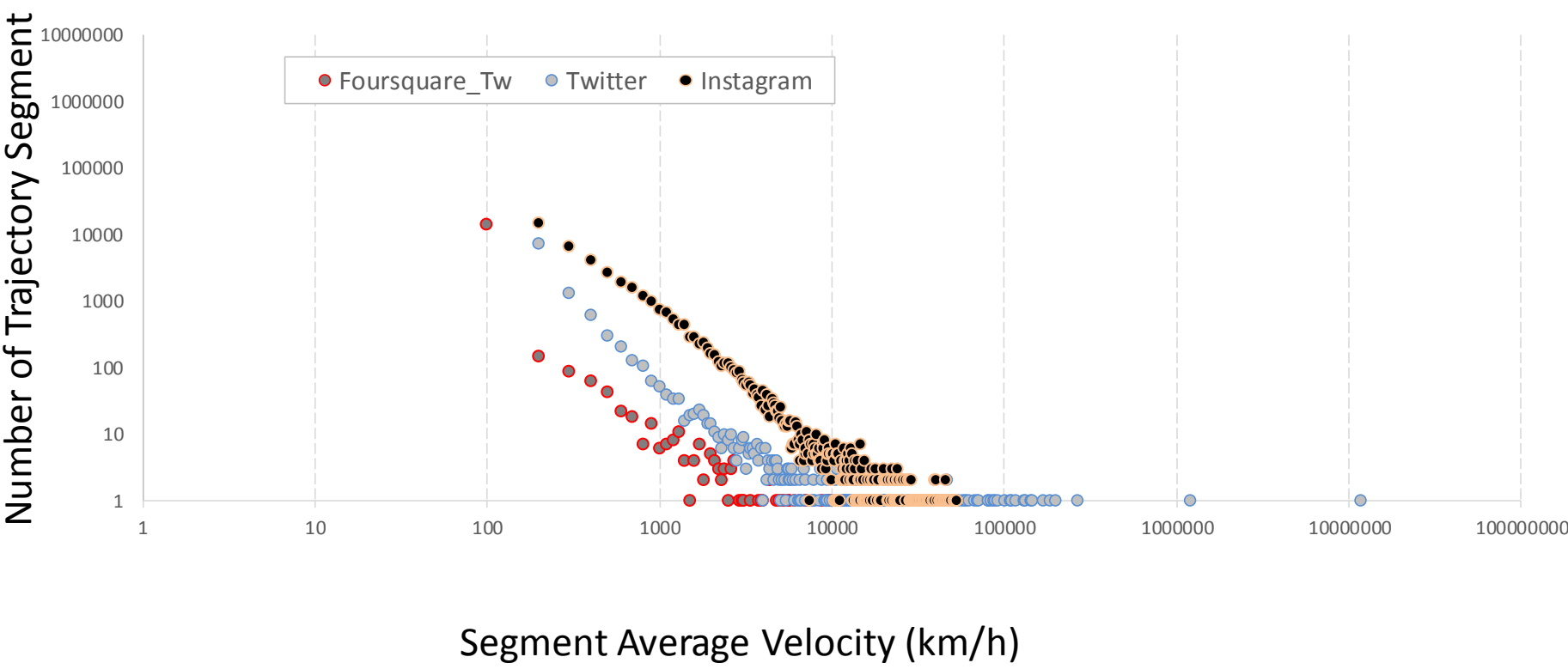
Data Filtering

Social Media	# of Daily Trajectory		
	Total	Distance = 0	D > 0
Instagram	2,892,912	2,238,847 (77.4%)	654,065 (33.6%)
Twitter	443,881	187,831 (42.3%)	256,050 (57.7%)
Foursquare-Tw	24,410	16,219 (66.4%)	8,191 (33.6%)
Total	3,361,203	2,442,897 (72.7%)	918,306 (27.3%)

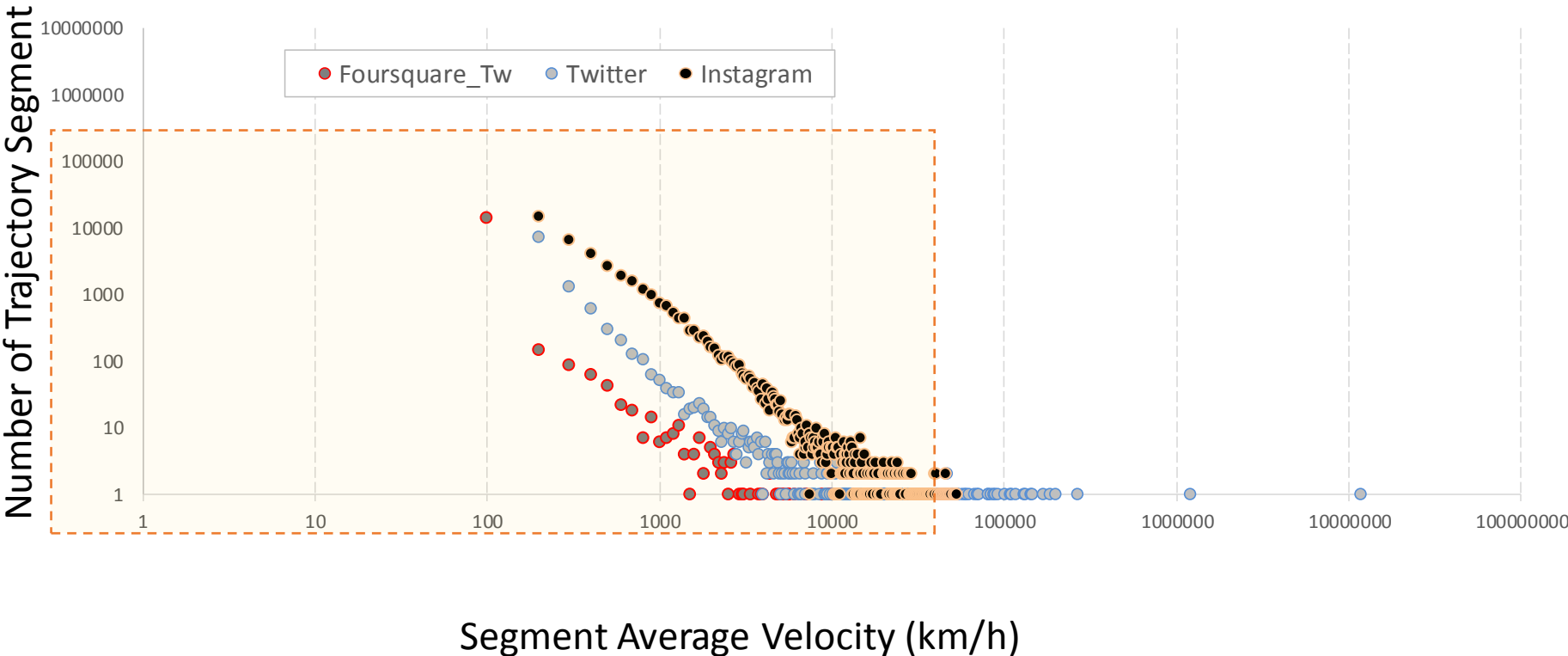
Frequency Distribution of Daily Trajectory Travel Distance (Bin Size: 1km)



Data Filtering

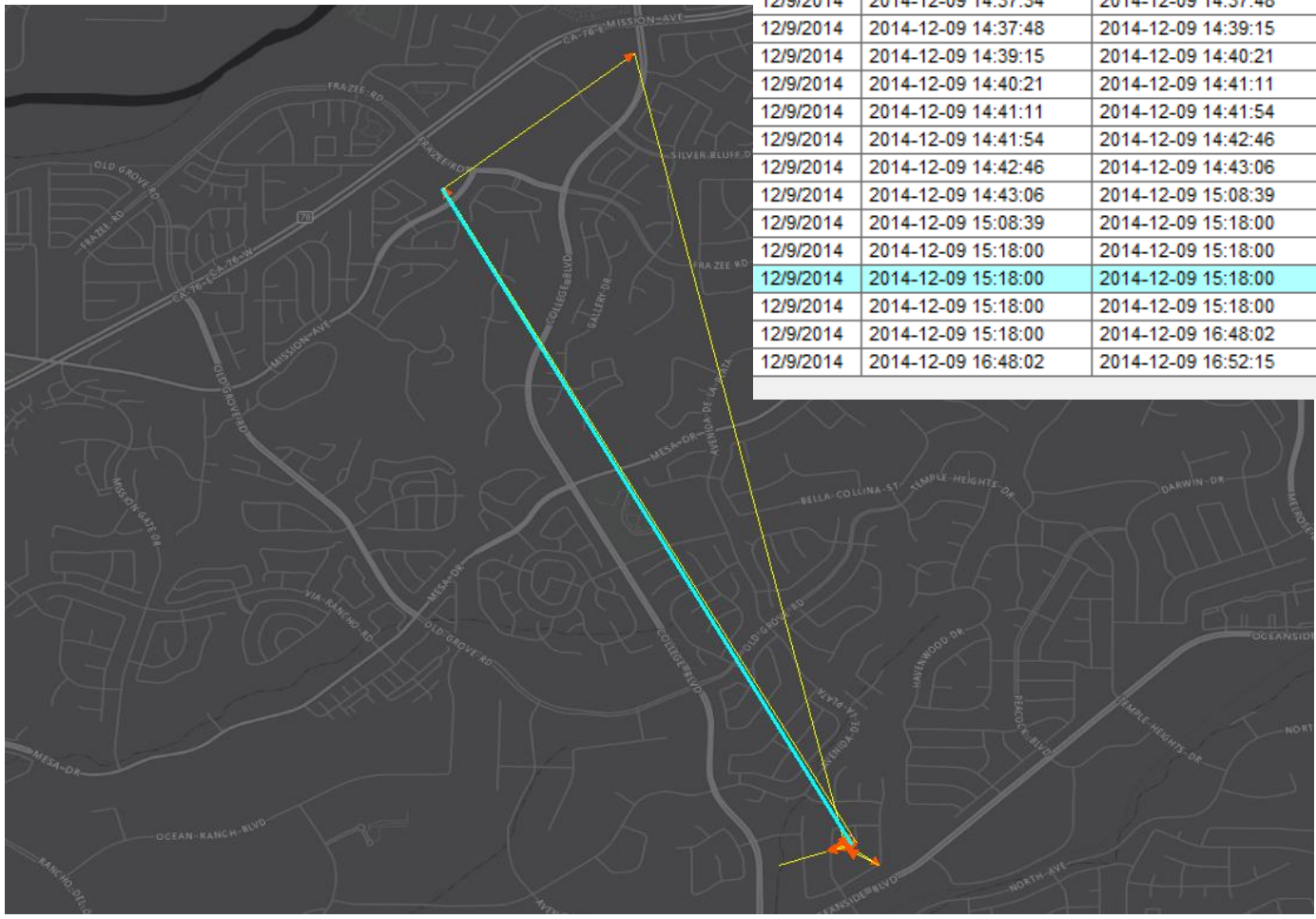


Data Filtering



Data Filtering

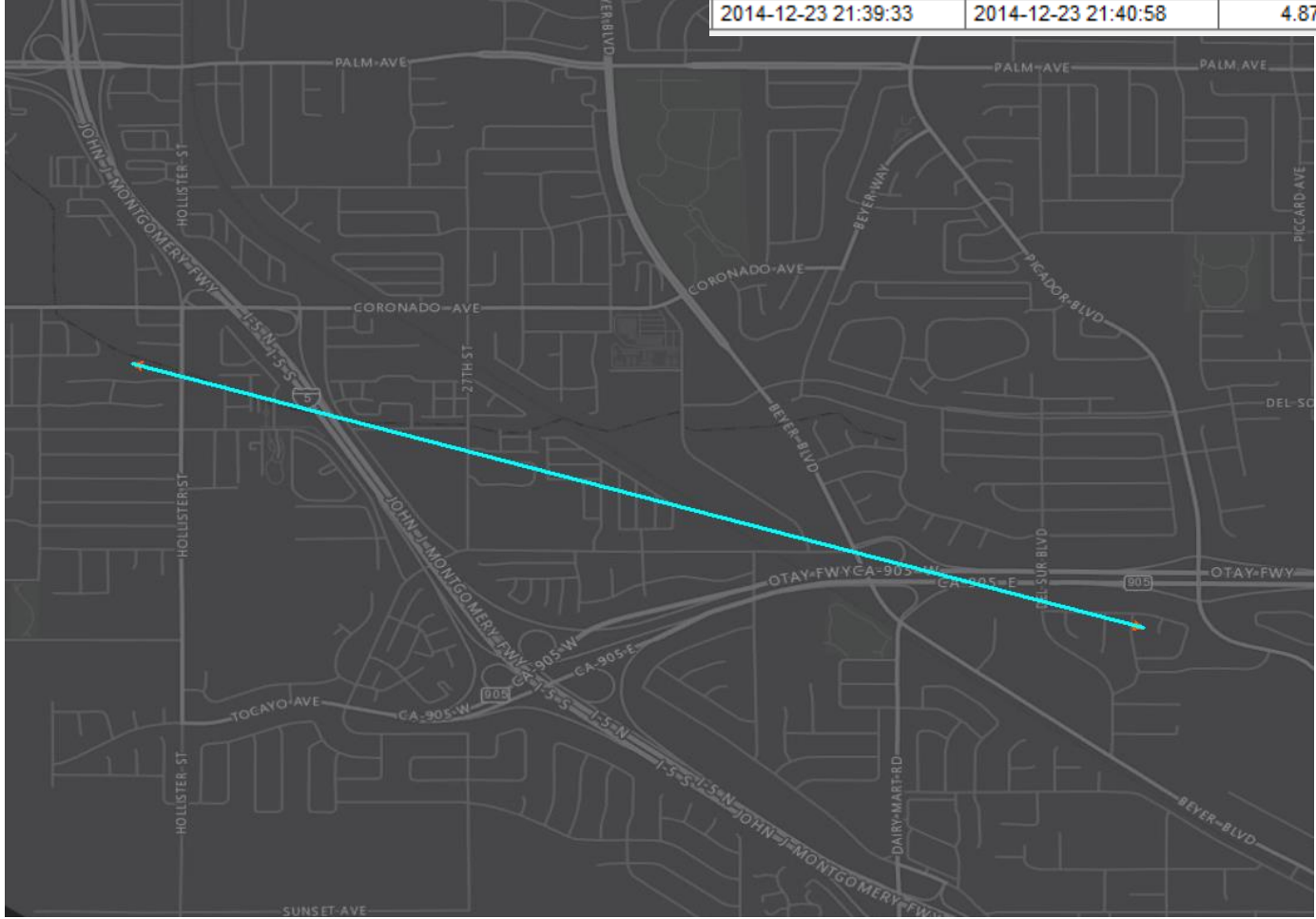
DATE	FTIME_LOCA	TTIME_LOCA	LENGTH_M	DURATION	V_KM_PER_H
12/9/2014	2014-12-09 09:30:36	2014-12-09 09:31:03	288.16198	00:00:26.649	38.927657
12/9/2014	2014-12-09 09:31:03	2014-12-09 14:33:09	18.17244	05:02:06.066	0.003609
12/9/2014	2014-12-09 14:33:09	2014-12-09 14:36:23	22.298074	00:03:14.26	0.413225
12/9/2014	2014-12-09 14:36:23	2014-12-09 14:37:14	136.964501	00:00:50.913	9.684603
12/9/2014	2014-12-09 14:37:14	2014-12-09 14:37:25	138.357374	00:00:11.194	44.49585
12/9/2014	2014-12-09 14:37:25	2014-12-09 14:37:34	45.668322	00:00:08.978	18.312092
12/9/2014	2014-12-09 14:37:34	2014-12-09 14:37:48	0	00:00:13.722	0
12/9/2014	2014-12-09 14:37:48	2014-12-09 14:39:15	84.828717	00:01:26.907	3.51391
12/9/2014	2014-12-09 14:39:15	2014-12-09 14:40:21	35.640019	00:01:06.351	1.933717
12/9/2014	2014-12-09 14:40:21	2014-12-09 14:41:11	53.89638	00:00:49.907	3.887771
12/9/2014	2014-12-09 14:41:11	2014-12-09 14:41:54	36.69954	00:00:43.142	3.062407
12/9/2014	2014-12-09 14:41:54	2014-12-09 14:42:46	37.529349	00:00:51.427	2.627135
12/9/2014	2014-12-09 14:42:46	2014-12-09 14:43:06	0	00:00:20.418	0
12/9/2014	2014-12-09 14:43:06	2014-12-09 15:08:39	10.981874	00:25:32.979	0.025789
12/9/2014	2014-12-09 15:08:39	2014-12-09 15:18:00	3270.450259	00:09:21.009	20.98651
12/9/2014	2014-12-09 15:18:00	2014-12-09 15:18:00	3273.718376	00:00:00.236	49938.076916
12/9/2014	2014-12-09 15:18:00	2014-12-09 15:18:00	3271.563842	00:00:00.001	11777629.832533
12/9/2014	2014-12-09 15:18:00	2014-12-09 15:18:00	934.569822	00:00:00.047	71584.071456
12/9/2014	2014-12-09 15:18:00	2014-12-09 16:48:02	3601.412764	01:30:01.538	2.400258
12/9/2014	2014-12-09 16:48:02	2014-12-09 16:52:15	72.351747	00:04:13.233	1.028564



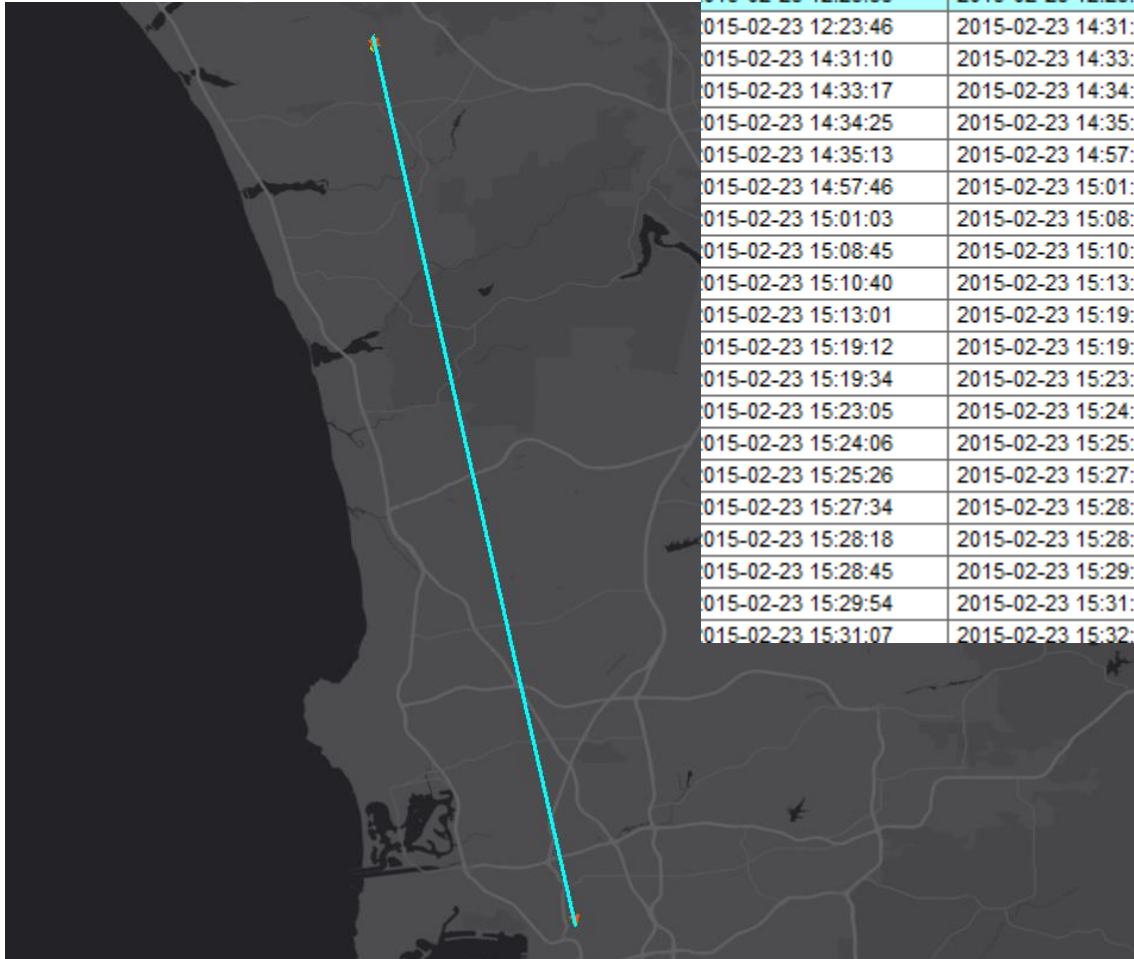
Travel Distance = 3.3km
Time 0.001 second

Data Filtering

FTIME_LOCA	TTIME_LOCA	LENGTH_M	DURATION	V_KM_PER_H
2014-12-23 21:38:39	2014-12-23 21:38:39	3008.595146	00:00:00.009	1203438.058592
2014-12-23 21:38:39	2014-12-23 21:39:33	3008.105539	00:00:53.406	202.770849
2014-12-23 21:39:33	2014-12-23 21:40:58	4.87853	00:01:25.644	0.205066



Data Filtering



FTIME_LOCA	TTIME_LOCA	LENGTH_M	DURATION	V_KM_PER_H
015-02-23 10:24:15	2015-02-23 12:20:34	106.597806	01:56:18.723	0.054989
015-02-23 12:20:34	2015-02-23 12:23:33	48678.973747	00:02:59.684	975.291654
015-02-23 12:23:33	2015-02-23 12:23:46	48683.992808	00:00:12.805	13687.026483
015-02-23 12:23:46	2015-02-23 14:31:10	669.590409	02:07:23.429	0.315372
015-02-23 14:31:10	2015-02-23 14:33:17	159.525809	00:02:07.836	4.492419
015-02-23 14:33:17	2015-02-23 14:34:25	133.655767	00:01:07.704	7.106829
015-02-23 14:34:25	2015-02-23 14:35:13	58.232663	00:00:48.004	4.367086
015-02-23 14:35:13	2015-02-23 14:57:46	90.551079	00:22:32.408	0.24104
015-02-23 14:57:46	2015-02-23 15:01:03	102.487449	00:03:17.639	1.866812
015-02-23 15:01:03	2015-02-23 15:08:45	34.913252	00:07:41.756	0.272195
015-02-23 15:08:45	2015-02-23 15:10:40	33.04001	00:01:54.978	1.034494
015-02-23 15:10:40	2015-02-23 15:13:01	40.792375	00:02:20.998	1.041522
015-02-23 15:13:01	2015-02-23 15:19:12	54.410378	00:06:11.36	0.52746
015-02-23 15:19:12	2015-02-23 15:19:34	25.275389	00:00:21.496	4.232946
015-02-23 15:19:34	2015-02-23 15:23:05	83.739117	00:03:30.761	1.430344
015-02-23 15:23:05	2015-02-23 15:24:06	123.374865	00:01:01.185	7.259124
015-02-23 15:24:06	2015-02-23 15:25:26	14.866256	00:01:20.514	0.664711
015-02-23 15:25:26	2015-02-23 15:27:34	91.455208	00:02:07.842	2.575357
015-02-23 15:27:34	2015-02-23 15:28:18	105.9989	00:00:44.149	8.643368
015-02-23 15:28:18	2015-02-23 15:28:45	0	00:00:26.693	0
015-02-23 15:28:45	2015-02-23 15:29:54	46.857628	00:01:08.717	2.454814
015-02-23 15:29:54	2015-02-23 15:31:07	16.854622	00:01:13.134	0.829664
015-02-23 15:31:07	2015-02-23 15:32:33	1.787182	00:01:26.519	0.074364

Data Filtering

Helicopter
Westland Lynx
401 km/h (249 mph)



Wheel-driven
Vesco Turbinator
775 km/h (763 mph)



Propeller-driven aircraft
Tupolev Tu-114
870 km/h (540 mph)



Land vehicles
ThrustSSC
1,228 km/h (763 mph)



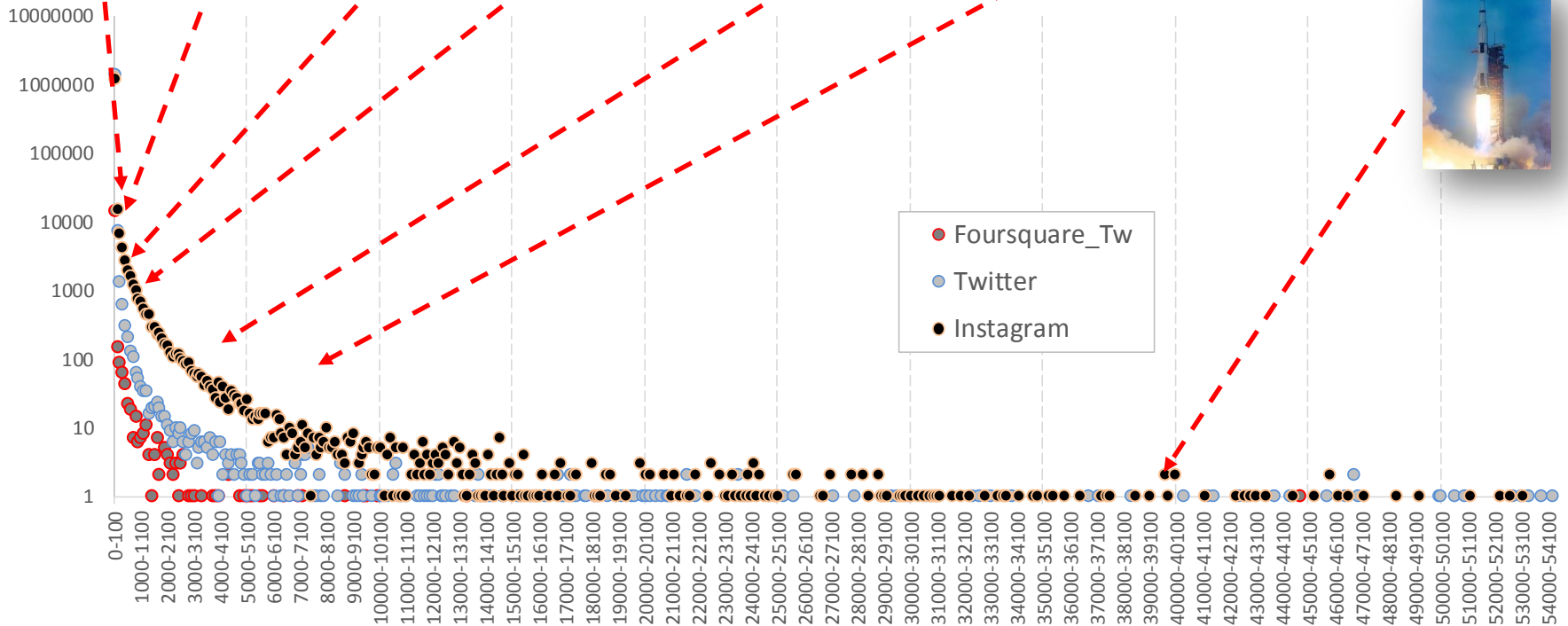
Manned air-breathing craft
SR-71 "Blackbird"
3,529 km/h (1,193 mph)



Manned rocket-powered aircraft
North American X-15
7,274 km/h (4,520 mph)



Manned spacecraft
Apollo 10
39,897 km/h (24,791 mph)



Segment Average Velocity (km/h)

Sources: <https://en.wikipedia.org/>
<http://www.teamvesco.com/>

Data Filtering

Helicopter
Westland Lynx
401 km/h (249 mph)



Wheel-driven
Vesco Turbinator
775 km/h (763 mph)



Propeller-driven aircraft
Tupolev Tu-114
870 km/h (540 mph)



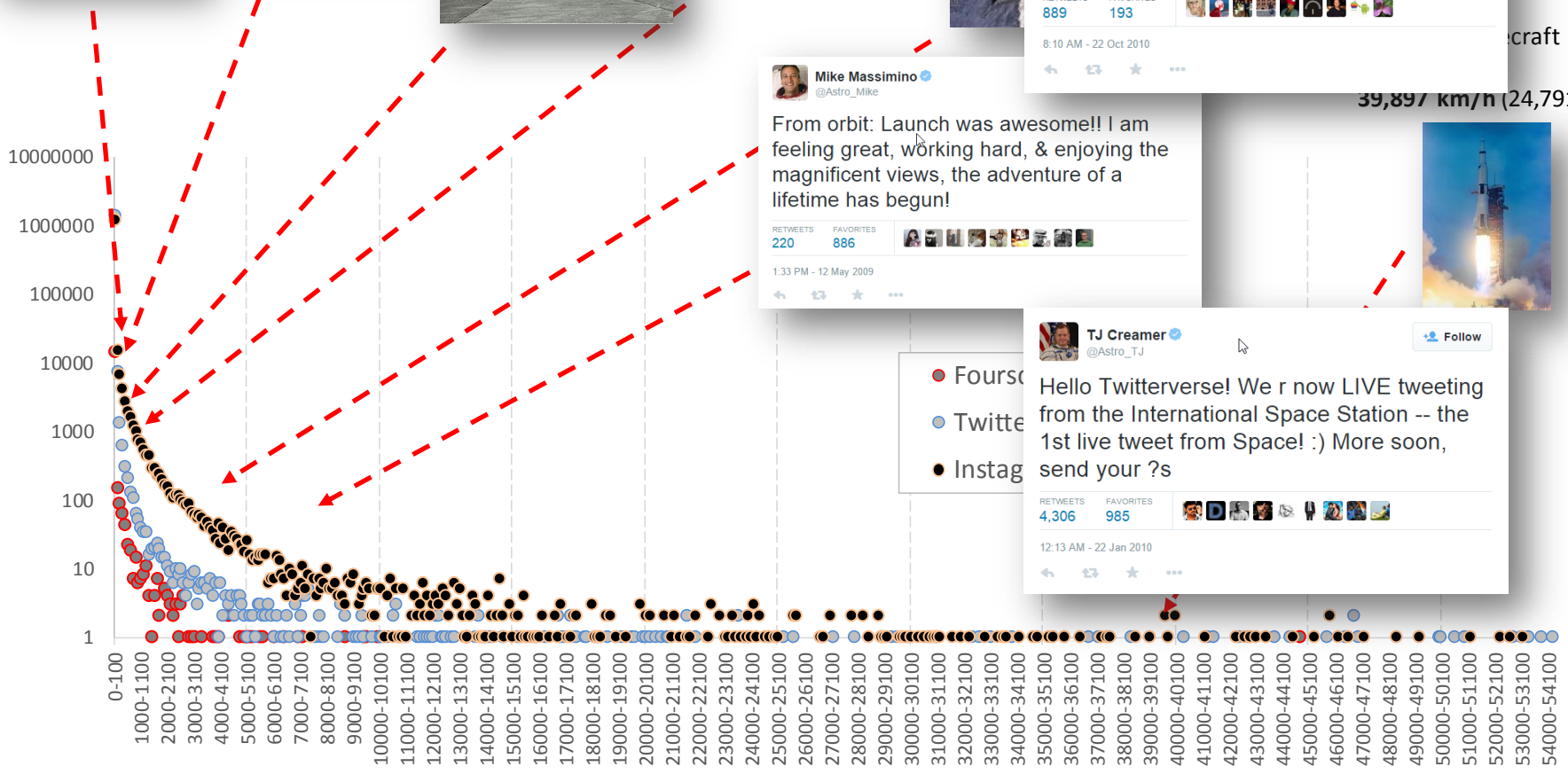
Land vehicles
ThrustSSC
1,228 km/h (763 mph)



Manned air-breathing craft
SR-71 "Blackbird"
3,529 km/h (1,193 mph)



Manned rocket-powered aircraft
North American X-15
7,274 km/h (4,520 mph)



Douglas H. Wheelock @Astro_Wheels
I just unlocked the "NASA Explorer" badge on @foursquare! <http://4sq.com/dseNyl>

RETWEETS: 889 FAVORITES: 193

8:10 AM - 22 Oct 2010

Mike Massimino @Astro_Mike
From orbit: Launch was awesome!! I am feeling great, working hard, & enjoying the magnificent views, the adventure of a lifetime has begun!

RETWEETS: 220 FAVORITES: 886

1:33 PM - 12 May 2009

TJ Creamer @Astro_TJ
Hello Twitterverse! We r now LIVE tweeting from the International Space Station -- the 1st live tweet from Space! :) More soon, send your ?s

RETWEETS: 4,306 FAVORITES: 985

12:13 AM - 22 Jan 2010

39,897 km/h (24,791 mph)

Segment Average Velocity (km/h)

Sources: <https://en.wikipedia.org/>
<http://www.teamvesco.com/>

Data Filtering

401 km/h (249 mph)



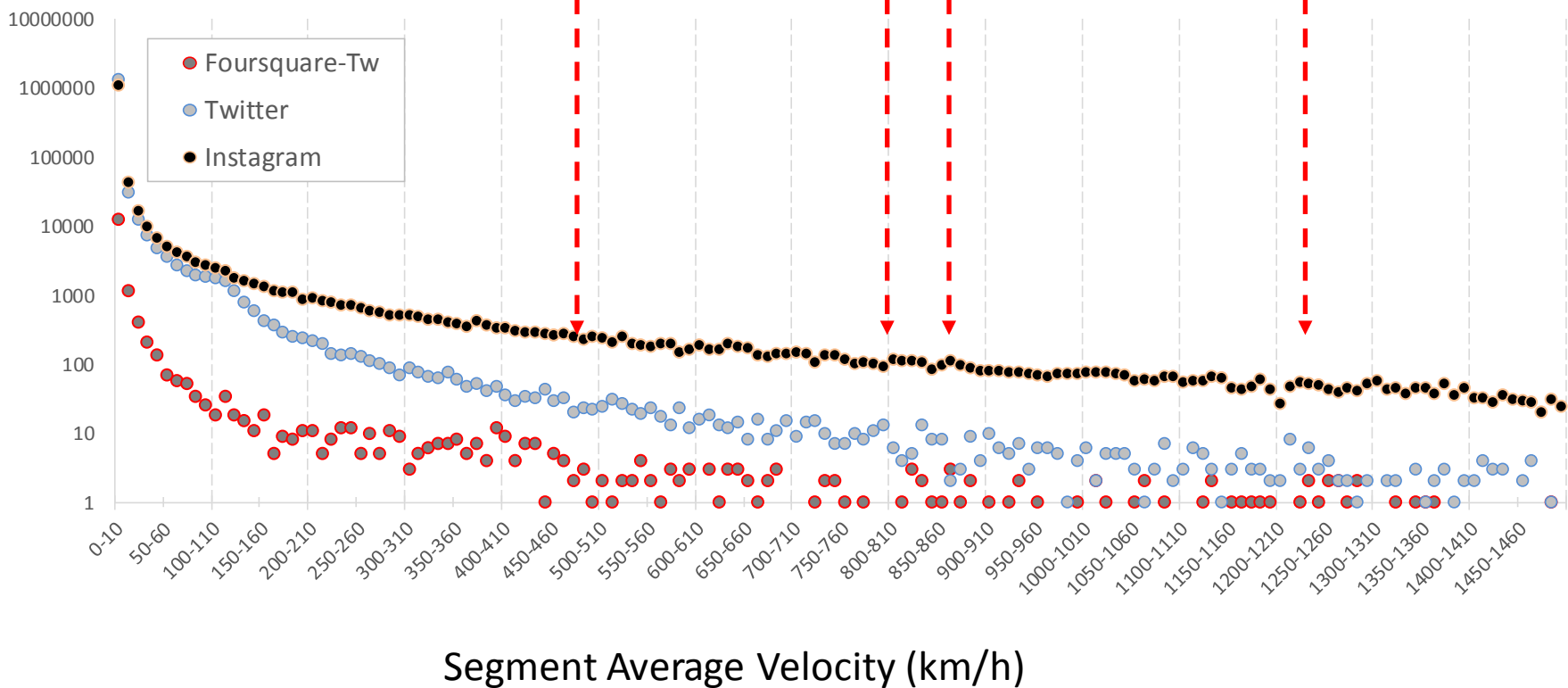
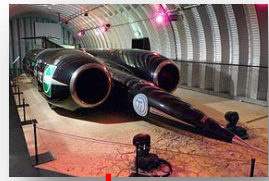
775 km/h (763 mph)



870 km/h (540 mph)



1,228 km/h (763 mph)



Data Filtering

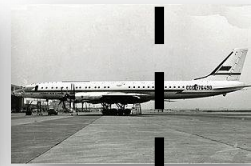
401 km/h (249 mph)



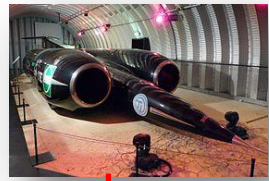
775 km/h (763 mph)



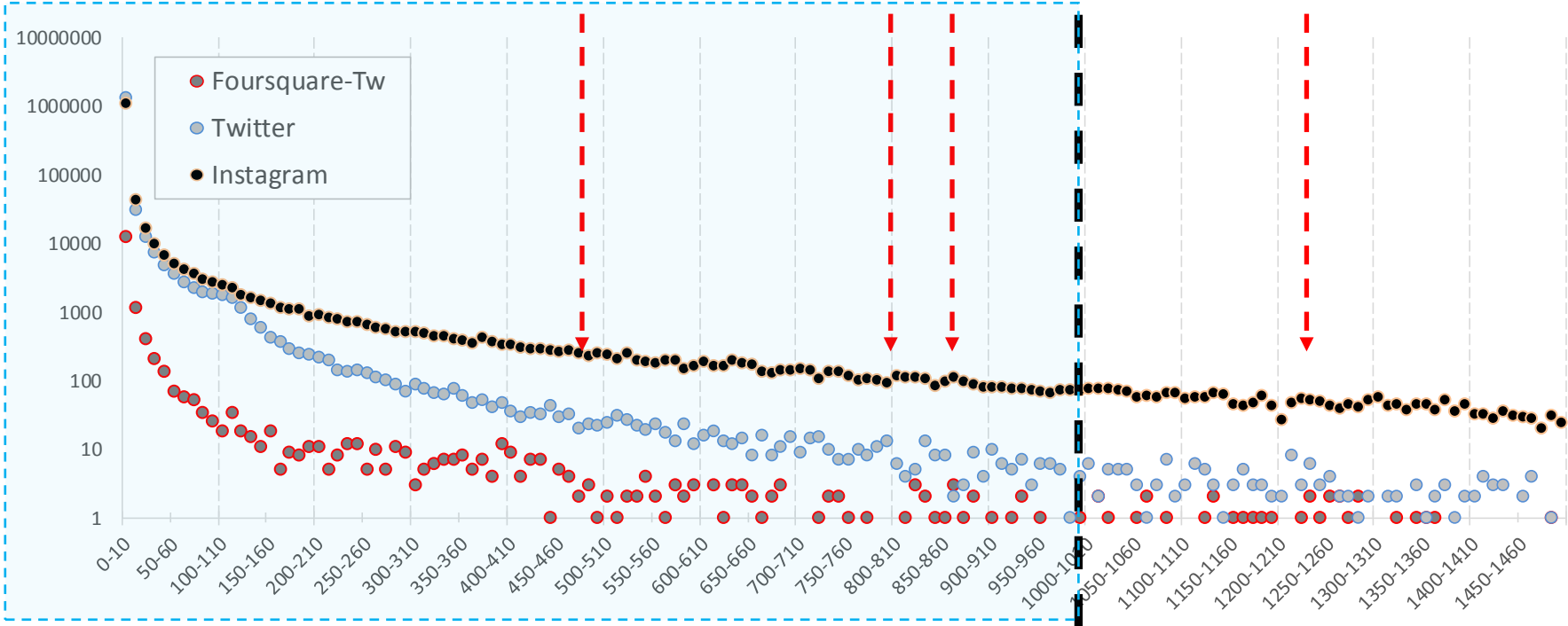
870 km/h (540 mph)



1,228 km/h (763 mph)

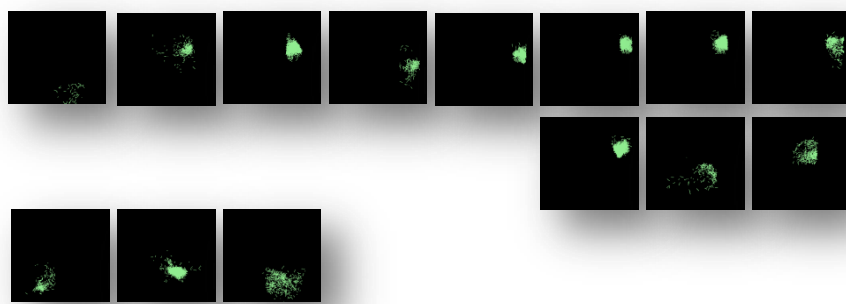
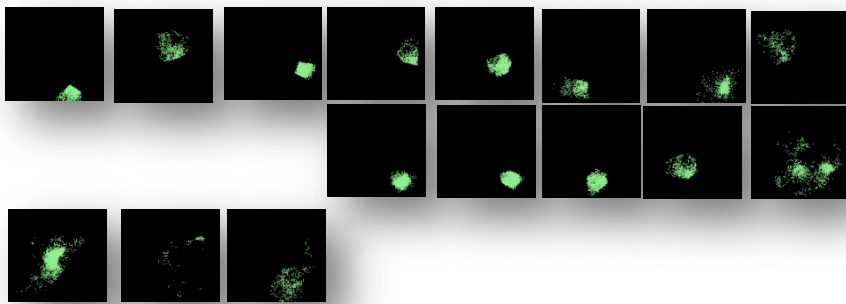
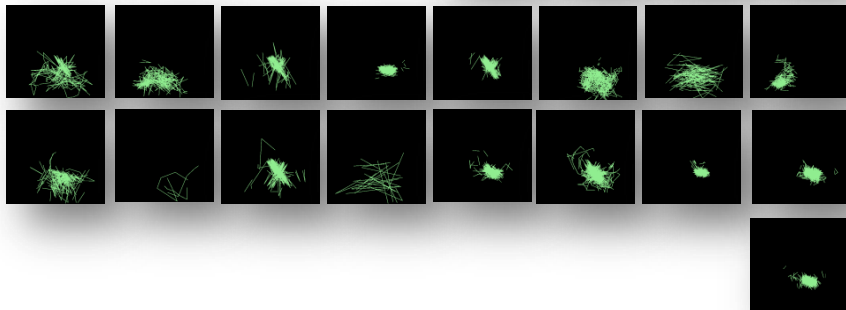
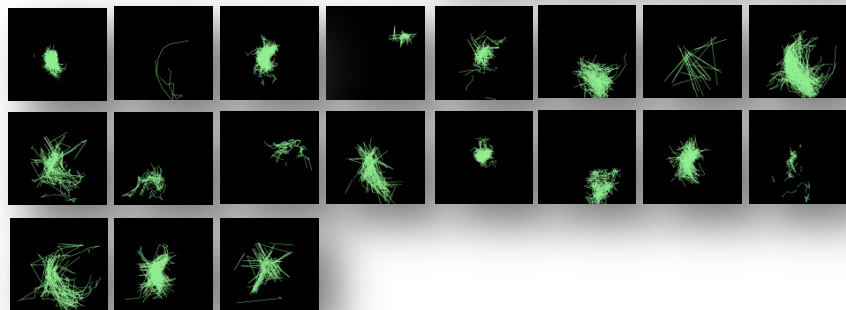
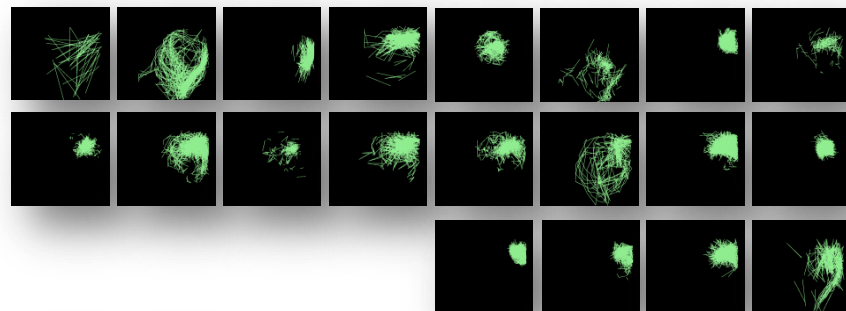
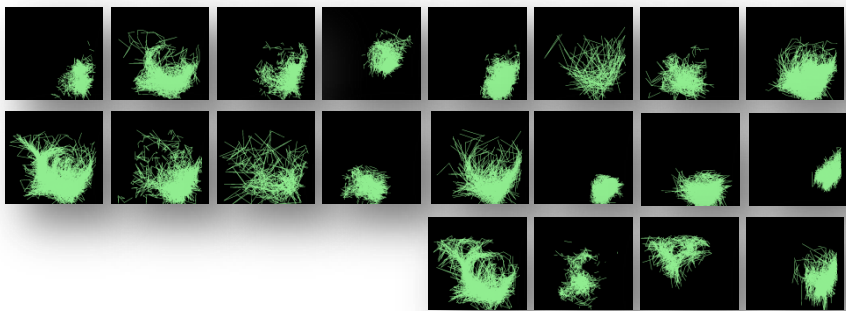


1,000 km/h

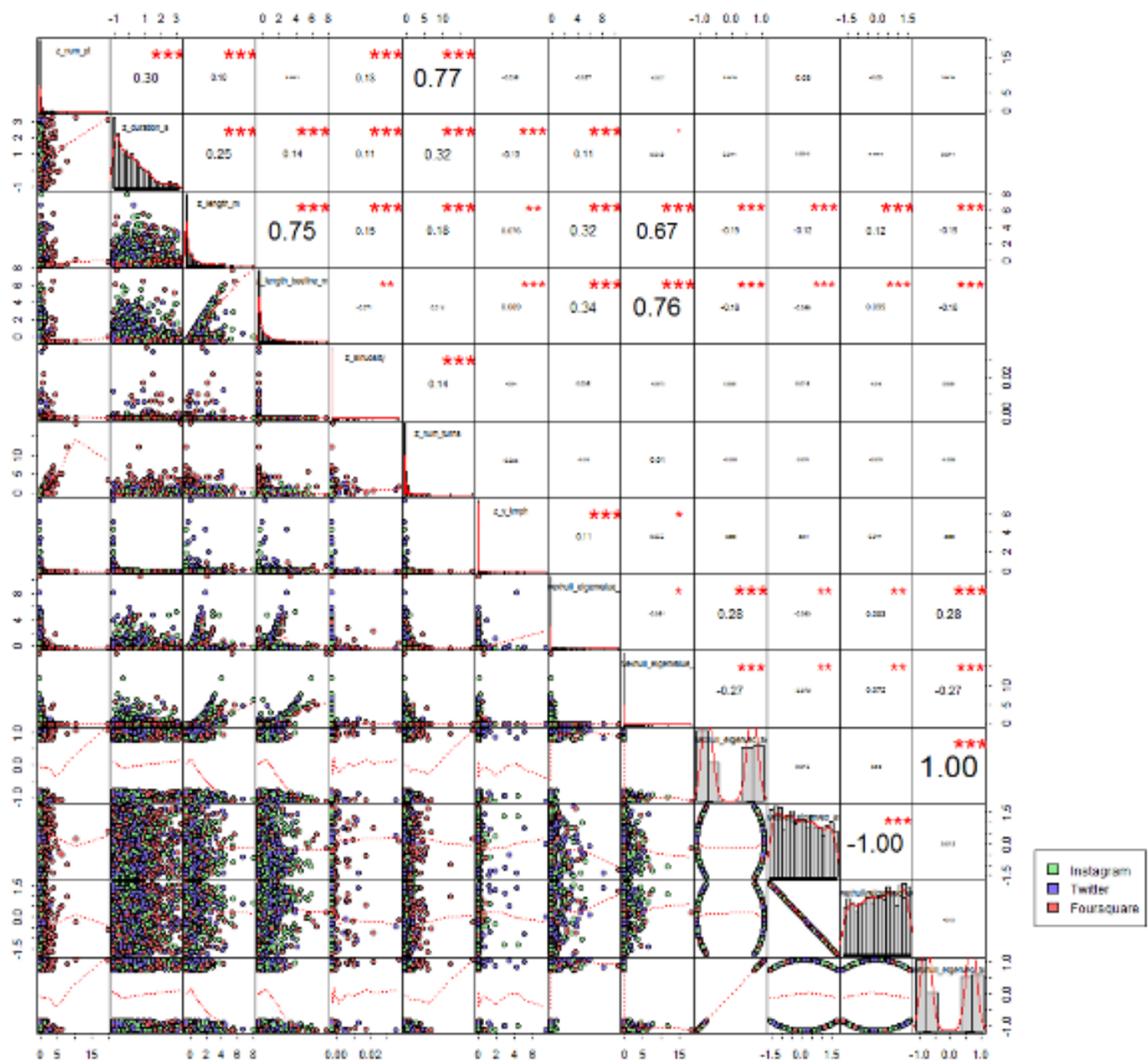


Segment Average Velocity (km/h)

Data Filtering: Trajectory Clustering

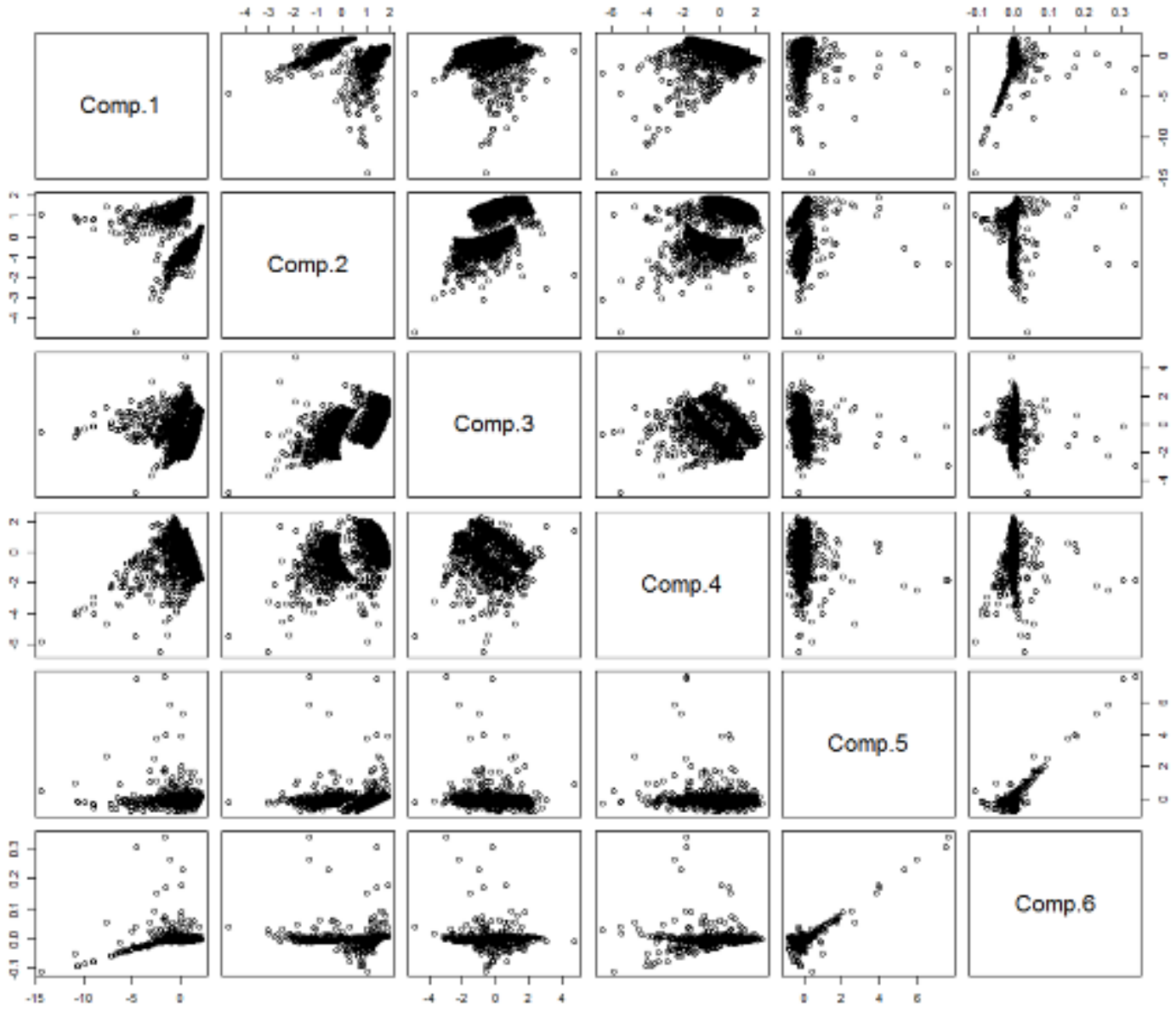


Data Filtering: Trajectory Clustering



Plot is based on 500 samples from each social media data

Data Filtering: Trajectory Clustering

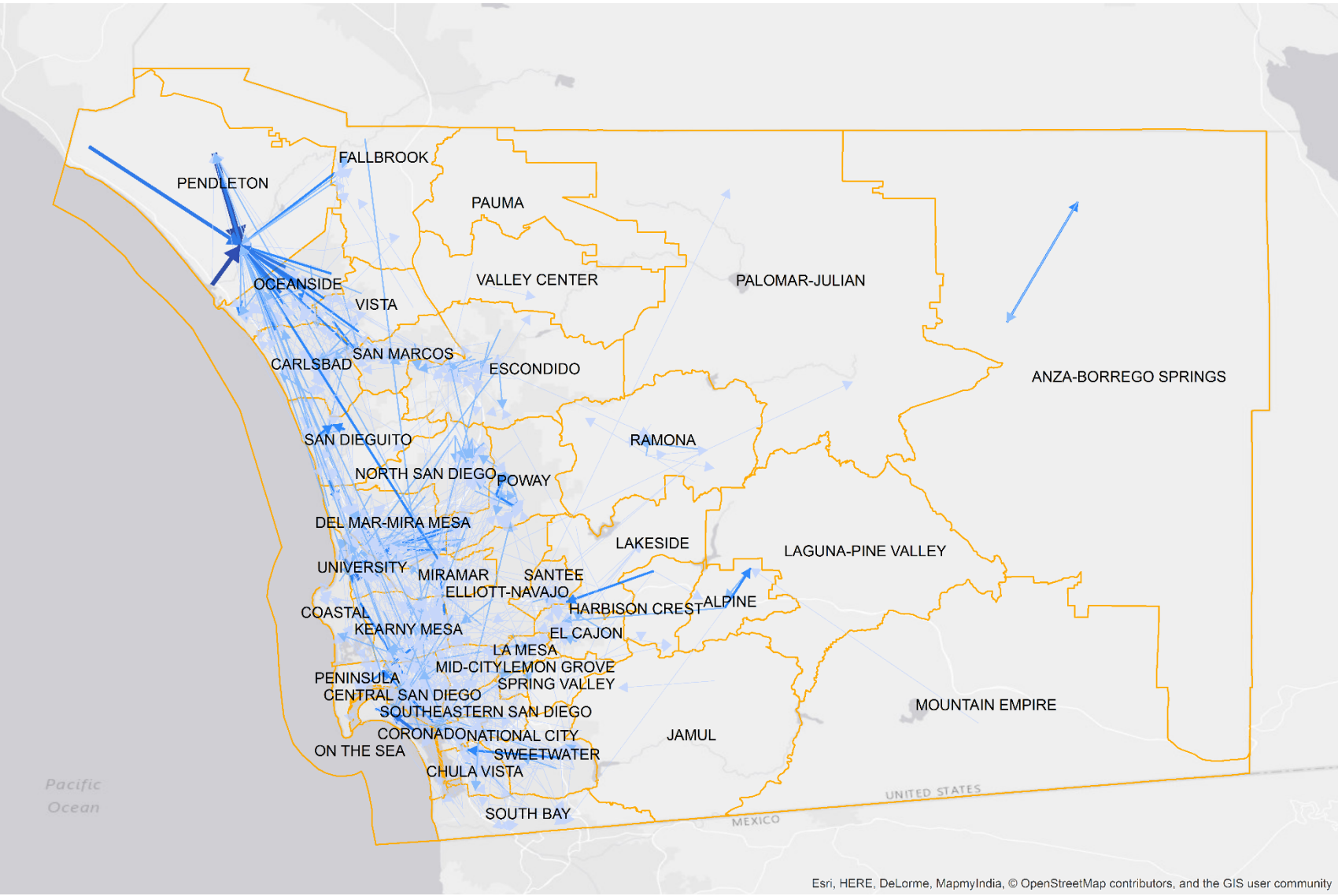


Plot is based on 1,500 samples of data points

Data Filtering: Trajectory Clustering

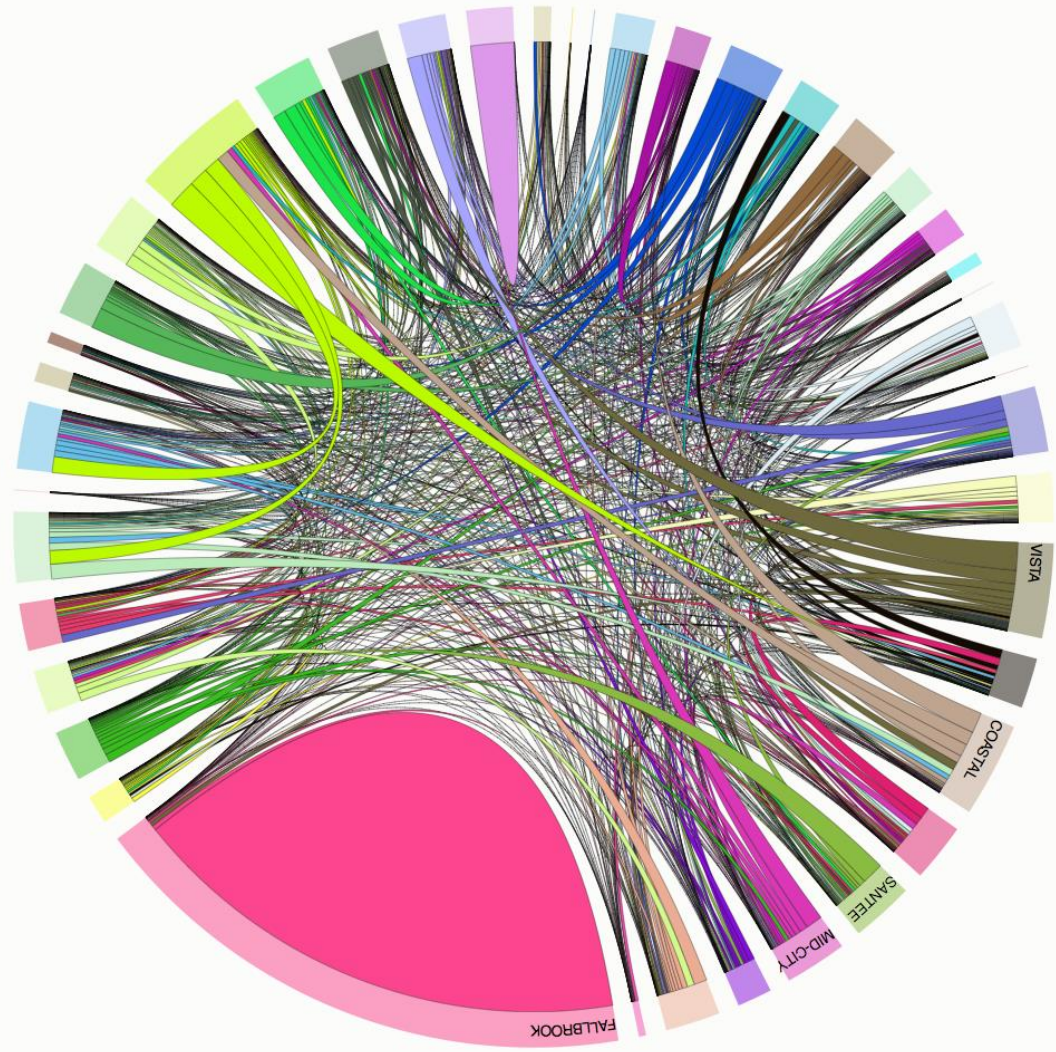
Results to be computed sometime soon...

Data Mashup



OD Matrix (All)

Built with [d3.js](#).



Data Mashup

