Probe Data Analysis for Road Slope

Boyu Wang , Jingya Xun

Problem Statement

- GPS Probe Data can be used as a source of real-time traffic information. Probe data is generated from tracking the position, velocity, heading and other information of moving vehicles.
- There are many factors that can bring down the precision of the probe data collected. Some common factors include low quality of GPS hardware and vehicle location, such urban canyons and tunnels.
- The goal of this project is to explore ways to increase the accuracy of collected probe data by matching them to the correct, pre-defined link data.

Approach - Map Matching

- We compare the latitude and longitude of the probe data and link data to find the pairs with minimum distance between them. This way we match every probe data entry to a link data entry. We save the matched data pairs to a csv file for future use.

Approach - Road Slope Derivation

- Using the matched data pairs from the last step, we calculate the slope between two consecutive points on the map for all points in the file.
- The slope is calculated as

slope = (p2.longitude - p1. longitude) / (p2.latitude - p1.latitude)

Approach - Evaluation

- In this step we evaluate the derived road slope from the last step with the surveyed road slope, which can be seen as ground truth.
- To calculate the average slope of both the probe data and link data, we take the average the slope data calculated from each point.
- Finally we calculate the error as follows

$$error = | probe\_avg - surveyed\_avg | / surveyed\_avg$$

Formula

Haversine Formula: the haversine formula determines the great-circle distance between two points on a sphere given their longitudes and latitudes. Important in navigation, it is a special case of a more general formula in spherical trigonometry, the law of haversines, that relates the sides and angles of spherical triangles.

Haversine Formula

$a = \sin^2(\Delta\varphi/2) + \cos\varphi_1 \cdot \cos\varphi_2 \cdot \sin^2(\Delta\lambda/2)$

$c = 2 \cdot \arctan(\sqrt{a}/(1-a))$

$R = 6371*1000$

$Dist = R * c$

Here, R = radius of Earth in km, φ = latitude in radians, λ = longitude in radians

Intermediate result

sampleID	dateTime	sourceCode	Latitude	Longitude	Altitude	Speed	Heading	linkPVID	direction	distFromNoc	distFromLinkLine
3496	##########	13	51.4968682	9.38602223	200	23	339	567329767	B	35.1249716	N/A
3496	##########	13	51.4966822	9.38615726	200	10	129	567329767	B	12.4292141	N/A
3496	##########	13	51.4967049	9.3864223	201	21	60	567329767	B	19.2388976	N/A
3496	##########	13	51.4967493	9.38683997	201	0	360	567329767	B	46.2371349	N/A
3496	##########	13	51.4968639	9.38729402	199	0	360	567329767	B	80.1450679	N/A
3496	##########	13	51.4969302	9.38771596	198	5	89	567329767	B	109.998341	N/A
3496	##########	13	51.4969575	9.38779383	198	1	288	567329767	B	116.115717	N/A
3496	##########	13	51.4969521	9.38780523	197	0	310	567329767	B	116.638896	N/A
3496	##########	13	51.496949	9.38781814	196	0	274	567329767	B	117.353759	N/A
3496	##########	13	51.4969443	9.38784018	196	0	226	567329767	B	118.605385	N/A
3496	##########	13	51.4969411	9.38785175	197	0	201	567329767	B	119.238535	N/A
3496	##########	13	51.4969385	9.38785544	197	0	182	567329767	B	119.382162	N/A
3496	##########	13	51.4969375	9.38785669	197	0	232	567329767	B	119.423826	N/A
3496	##########	13	51.4969363	9.38785946	197	0	202	567329767	B	119.561108	N/A
3496	##########	13	51.4969358	9.3878603	197	0	199	567329767	B	119.597347	N/A
3496	##########	13	51.4969345	9.38786323	197	0	179	567329767	B	119.742922	N/A
3496	##########	13	51.4969339	9.38786533	197	0	184	567329767	B	119.858469	N/A
3496	##########	13	51.4969324	9.38786742	197	0	199	567329767	B	119.94058	N/A
3496	##########	13	51.4969318	9.38786885	197	0	178	567329767	B	120.012528	N/A
3496	##########	13	51.496931	9.38786952	197	0	183	567329767	B	120.026075	N/A
3496	##########	13	51.4969306	9.38787002	197	0	194	567329767	B	120.043857	N/A
3496	##########	13	51.4969303	9.38787103	197	0	176	567329767	B	120.100643	N/A
3496	##########	13	51.4969307	9.38787245	197	0	172	567329767	B	120.206042	N/A
3496	##########	13	51.4969308	9.38787337	197	0	177	567329767	B	120.272464	N/A

Intermediate result

4553	##########	13	53.0514388	8.80883859	33	15	303	799517966	B	42.4258057	13.15
4553	##########	13	53.0516051	8.80850097	34	26	303	799517966	B	71.5061684	12.15
4553	##########	13	53.0517575	8.80794894	34	31	290	799517966	B	111.629303	12.15
4553	##########	13	53.0519227	8.80731535	34	40	309	799517966	B	157.451921	12.15
4553	##########	13	53.0521656	8.80639242	33	53	291	799517966	B	224.564474	13.15
4553	##########	13	53.0524122	8.80545205	33	0	360	799517966	B	293.007465	13.15
4553	##########	13	53.0527942	8.80453834	37	42	306	799517966	B	366.793095	9.15
4553	##########	13	53.0530851	8.80381666	37	43	307	762732452	B	24.1814883	9.21
4553	##########	13	53.0533996	8.80307176	37	44	305	762732452	B	84.9919459	9.21
4553	##########	13	53.0537149	8.80227925	39	44	309	762732452	B	148.472596	7.21
4553	##########	13	53.0540206	8.80160342	41	41	310	762732452	B	204.983025	5.21
4553	##########	13	53.0543768	8.80106715	41	31	317	762732452	B	257.414198	5.21
4553	##########	13	53.0546384	8.80060254	41	32	315	762732452	B	299.819002	5.21
4553	##########	13	53.0549228	8.80007582	39	34	312	762732452	B	347.077689	7.21
4553	##########	13	53.0552369	8.79942446	38	43	308	762732452	B	402.891388	8.21
4553	##########	13	53.0555756	8.79874477	39	40	313	762732452	B	461.897617	7.21
4553	##########	13	53.0558974	8.79814387	39	37	313	762732452	B	515.638735	7.21
4553	##########	13	53.0562273	8.79758262	39	37	315	716671988	B	219.190037	7.02
4553	##########	13	53.0565878	8.79698935	39	43	315	716671988	B	162.820627	7.02
4553	##########	13	53.056993	8.79633422	40	46	316	716671988	B	100.021457	6.02
4553	##########	13	53.0574452	8.79564615	38	48	314	716671988	B	31.8858372	8.02
4553	##########	13	53.0578707	8.7950533	38	38	319	716671988	B	29.8469351	8.02
4553	##########	13	53.0582288	8.79451459	38	39	315	716671988	B	83.4825812	8.02
4553	##########	13	53.0585791	8.79393658	39	37	320	716671988	B	138.264052	7.02
4553	##########	13	53.0588625	8.79355537	39	26	320	716671988	B	178.730773	7.02
4553	##########	13	53.0591252	8.79320894	41	30	319	716671988	B	215.954864	5.02
4553	##########	13	53.0594878	8.79270226	41	40	318	716671988	B	268.599791	5.02
4553	##########	13	53.0598768	8.79215945	42	42	319	762772849	B	29.2655309	4.18
4553	##########	13	53.0602976	8.79156073	42	43	318	762772849	B	90.8242303	4.18
4553	##########	13	53.0606761	8.7910237	41	37	319	762772849	B	146.13572	5.18
4553	##########	13	53.0610546	8.79049103	41	40	320	762772849	B	201.257925	5.18
4553	##########	13	53.0614467	8.78993313	42	40	319	762772849	B	258.616694	4.18

Final result

linkPVID	groundTruth	calculatedSlope	error
51881767	0.10433333333333300	-0.003963719657807880	1.0379909232377800
51881768	-0.0165	-0.009129268148705050	0.44671102129060300
51883238	0.10875000000000000	0.008854941386381020	0.918575251619485
51883890	-0.35825	-0.0001475476365570800	0.9995881433731830
51881767	0.10433333333333300	0.0001317559955217500	0.9987371629822200
51883238	0.10875000000000000	0.013299336998072600	0.8777072459947340
51881767	0.10433333333333300	-0.003964894284980810	1.0380021816451800
51883238	0.10875000000000000	0.01787931013111860	0.8355925505184500
51883890	-0.35825	-0.006905950442665140	0.9807230971593440
554728228	0.9421428571428570	-0.027374633238274800	1.0290557138237900
799517966	-0.05566666666666670	-0.0039835326021523500	0.9284395340930720
762732452	-0.0055	4.52688275386554E-05	1.0082306959161200
51882112	-0.07166666666666670	0.01365818699515650	1.190579353420790
716671988	-0.013666666666666700	-0.005511784016293140	0.596698730515136
51888926	-0.023500000000000000	0.01338227841242700	1.569458655847960
554728228	0.9421428571428570	-0.01819212367679940	1.0193093048881900
799517966	-0.05566666666666670	0.002543230693557700	1.0456867789261900
51882112	-0.07166666666666670	-0.002015523790470760	0.9718764122259890
716671988	-0.013666666666666700	-3.39842033100293E-05	0.997513350977315
51888926	-0.023500000000000000	0.013289745338137100	1.5655210782186000
799517966	-0.05566666666666670	0.006809454077161130	1.122325522344210
762732452	-0.0055	6.84734116240439E-05	1.0124497112043700
716671988	-0.013666666666666700	0.005345070753048160	1.391102738027910
762772849	0.0115	0.013262032839539200	0.15322024691645500
51888926	-0.023500000000000000	-0.011681921350556700	0.5028969638060970

Problems and Findings

- One problem we faced was long data preprocessing time due to the large data size and limitation in our computation power.
- We noticed in the link data file many data entries do not have recorded slope information. We filtered out those data points in the final results.
- Some link data did not include altitude in shapeinfo, which makes it not possible to compute distfromlink.

Discussion

- When evaluation slope we observed some calculated slopes differ a lot from the surveyed slopes. This may be caused by complex urban transportation structures such as tunnels and roadway bridges. For example, if there are multiple roadway bridges aligned close together with different height, then we would mistakenly match our probe data to the wrong link. The mismatched data pair would have very close longitude and latitude, but very different altitude, which may cause the drastic different in slope.

Reference

- https://infolab.usc.edu/csci587/Fall2016/papers/Map-Matching%20for%20Low-Sampling-Rate%20GPS%20Trajectories.pdf
- https://en.wikipedia.org/wiki/Haversine_formula