Wenyong ZHANG, Dawen XIA, Guoyan CHANG, Yang HU, Yujia HUO, Fujian FENG, Yantao LI, Huaqing LI, 2022. APFD: an effective approach to taxi route recommendation with mobile trajectory big data. *Frontiers of Information Technology & Electronic Engineering*, 23(10):1494-1510.
https://doi.org/10.1631/FITEE.2100530

APFD: an effective approach to taxi route recommendation with mobile trajectory big data

Key words: Big data analytics; Region extraction; Artificial potential field; Dijkstra; Route recommendation; GPS trajectories of taxis

Corresponding authors: Dawen XIA; Huaqing LI
E-mail: dwxia@gzmu.edu.cn; huaqingli@swu.edu.cn
ORCID: https://orcid.org/0000-0002-0151-9643;
https://orcid.org/0000-0001-6310-8965

Motivation

Urbanization has led to the growth of traffic congestion. The complex road network and traffic congestion have made taxis less efficient in carrying passengers. On one hand, this increases waiting time and creates a poor ride experience for passengers, and on the other hand, it results in issues such as high taxi operating cost, high energy consumption, and environmental pollution. To improve passenger experience, increase the efficiency of taxi operation, and reduce energy consumption, we propose an effective method for recommending the fastest taxi routes.

Main idea

1. The origin and destination coordinates are used to extract the area that is likely to include the optimal route. The area is taken as the effective area for method retrieval to improve the execution efficiency.

2. A potential energy field is constructed in the road network area. A critical equipotential line is established according to the current coordinate position, and redundant points are removed by comparing the potential energy of each connection point with the critical potential.

3. A distance matrix, a marker matrix, and a visited matrix are constructed for calculating the optimal route. These matrices are calculated using the node coordinates and side length data to obtain the optimal route.

Method

We present an effective taxi route recommendation approach (called APFD) based on the artificial potential field (APF) method and Dijkstra method with mobile trajectory big data. First, we propose a region extraction method that searches for a region including the optimal route through the origin and destination coordinates. Then, based on the APF method, we put forward an effective approach for removing redundant nodes. Finally, we employ the Dijkstra method to determine the optimal route recommendation.

Method

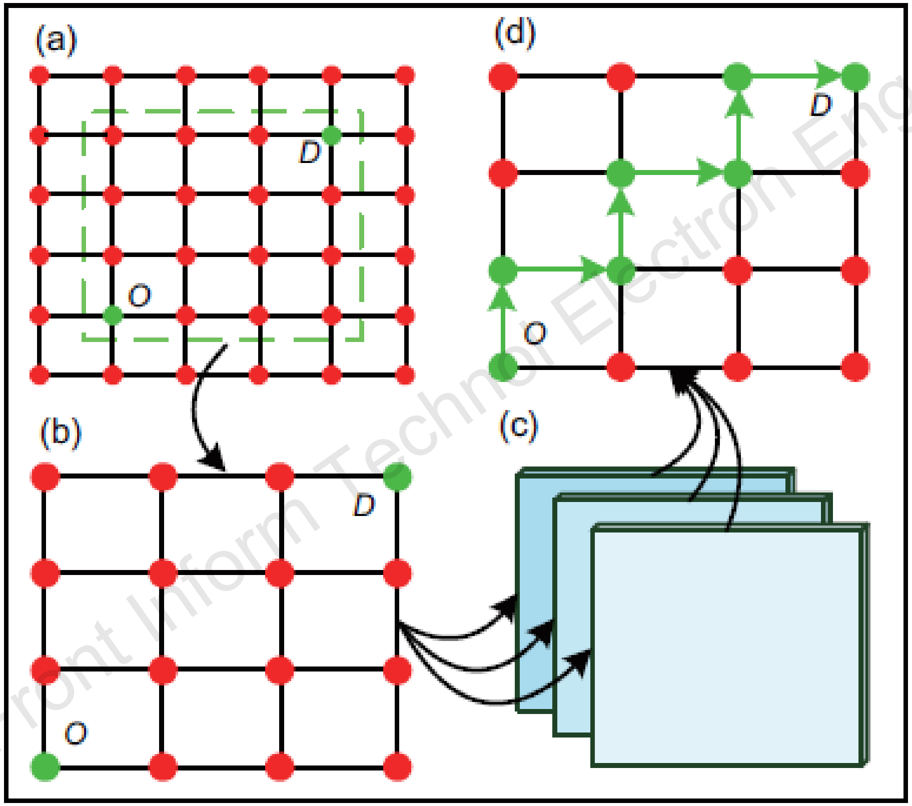

Fig. 3 Process of the APFD method: (a) simulating the road network; (b) extracting the effective area; (c) establishing the calculation matrices; (d) searching for the optimal route

Major results

Route recommendation evaluation

Table 2 Total distance of each method in the simulation map

Origin and destination points	Coordinate pair	Total distance						
		APFD	AC	A*	APF	RRT	NSGA-II	PSO
(12, 03) (08, 10)	A19, B19	610	753	944	703	708	650	689
(03, 15) (11, 10)	A17, B17	1044	1457	1610	1343	1141	1257	1444
(15, 13) (05, 28)	A20, B20	2216	2617	2962	2598	2667	2379	2596
(16, 03) (08, 21)	A10, B10	2368	2551	3258	2767	2765	2442	2618
(00, 05) (10, 28)	A1, B1	2519	2605	2636	3065	3661	2650	2665
(01, 04) (19, 15)	A12, B12	2531	2872	2981	3079	2922	2789	2917
(12, 35) (05, 12)	A14, B14	2701	2807	3108	3121	2906	2753	2847
(09, 03) (15, 27)	A7, B7	2714	3149	3814	3144	2921	2957	2814
(10, 15) (02, 38)	A3, B3	2774	3188	3406	3035	3123	2889	2854
(00, 05) (10, 28)	A18, B18	2843	3457	3813	3379	3367	3108	2912
(07, 08) (00, 36)	A16, B16	3237	3501	3720	3515	3746	3488	3574
(09, 09) (19, 36)	A11, B11	3074	3421	3423	4219	3727	3303	3511
(18, 30) (00, 07)	A6, B6	3088	3509	3768	4134	4085	3117	3447
(18, 28) (01, 02)	A5, B5	3659	3712	4307	4742	4410	3899	3997
(01, 34) (11, 01)	A2, B2	3897	4244	5370	4471	4094	4007	4314
(14, 36) (01, 03)	A13, B13	3972	4317	4718	4789	4384	4212	4551
(00, 30) (17, 01)	A8, B8	3986	4227	4490	4636	4255	4336	4219
(17, 03) (03, 38)	A15, B15	4314	5199	5563	5274	4934	4779	5012
(00, 00) (19, 39)	A4, B4	5030	5227	5468	6010	5579	5578	5315
(00, 39) (19, 00)	A9, B9	5106	5486	5823	6084	6061	5418	5656

Major results (Cont'd)

Route recommendation evaluation

Table 3 Performance improvement of APFD compared with other methods in the simulation map

Origin and destination points	Coordinate pair	Performance improvement of APFD (%)					
		AC	A*	APF	RRT	NSGA-II	PSO
(12, 03) (08, 10)	A19, B19	23.44	54.75	15.25	16.07	6.56	12.95
(03, 15) (11, 10)	A17, B17	39.56	54.21	28.64	9.29	20.40	38.31
(15, 13) (05, 28)	A20, B20	18.10	33.66	17.24	20.35	7.36	17.15
(16, 03) (08, 21)	A10, B10	7.73	37.58	16.85	16.77	3.13	10.56
(00, 05) (10, 28)	A01, B01	3.41	4.64	21.68	45.34	5.20	5.80
(01, 04) (19, 15)	A12, B12	13.47	17.78	21.65	15.45	10.19	15.25
(12, 35) (05, 12)	A14, B14	3.92	15.07	15.55	7.59	1.93	5.41
(09, 03) (15, 27)	A07, B07	16.03	40.53	15.84	7.63	8.95	3.68
(10, 15) (02, 38)	A03, B03	14.92	22.78	9.41	12.58	4.15	2.88
(00, 05) (10, 28)	A18, B18	21.60	34.12	18.85	18.43	9.32	2.43
(07, 08) (00, 36)	A16, B16	8.16	14.92	8.59	15.72	7.75	10.41
(09, 09) (19, 36)	A11, B11	11.29	11.35	37.25	21.24	7.45	14.22
(18, 30) (00, 07)	A06, B06	13.63	22.02	33.87	32.29	0.94	11.63
(18, 28) (01, 02)	A05, B05	1.45	17.71	29.60	20.52	6.56	9.24
(01, 34) (11, 01)	A02, B02	8.90	37.80	14.73	5.06	2.82	10.70
(14, 36) (01, 03)	A13, B13	8.69	18.78	20.57	10.37	6.04	14.58
(00, 30) (17, 01)	A08, B08	6.05	12.64	16.31	6.75	8.78	5.85
(17, 03) (03, 38)	A15, B15	20.51	28.95	22.25	14.37	10.78	16.18
(00, 00) (19, 39)	A04, B04	3.92	8.71	19.48	10.91	10.89	5.67
(00, 39) (19, 00)	A09, B09	7.44	14.04	19.15	18.70	6.11	10.77

Major results

Route recommendation evaluation

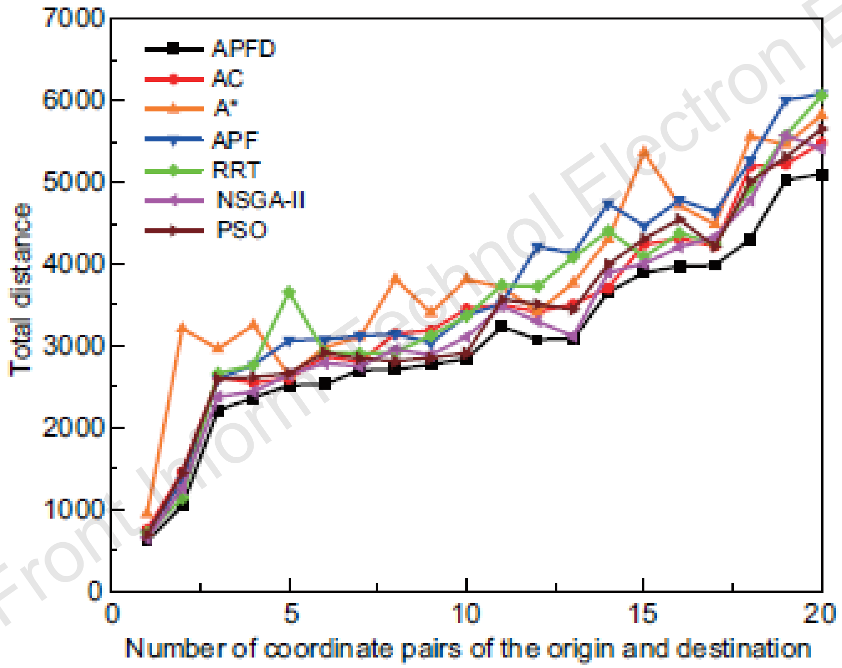

Fig. 11 Comparisons of the shortest route of each method with random coordinates in the simulation map

Major results (Cont’d)

Route recommendation evaluation

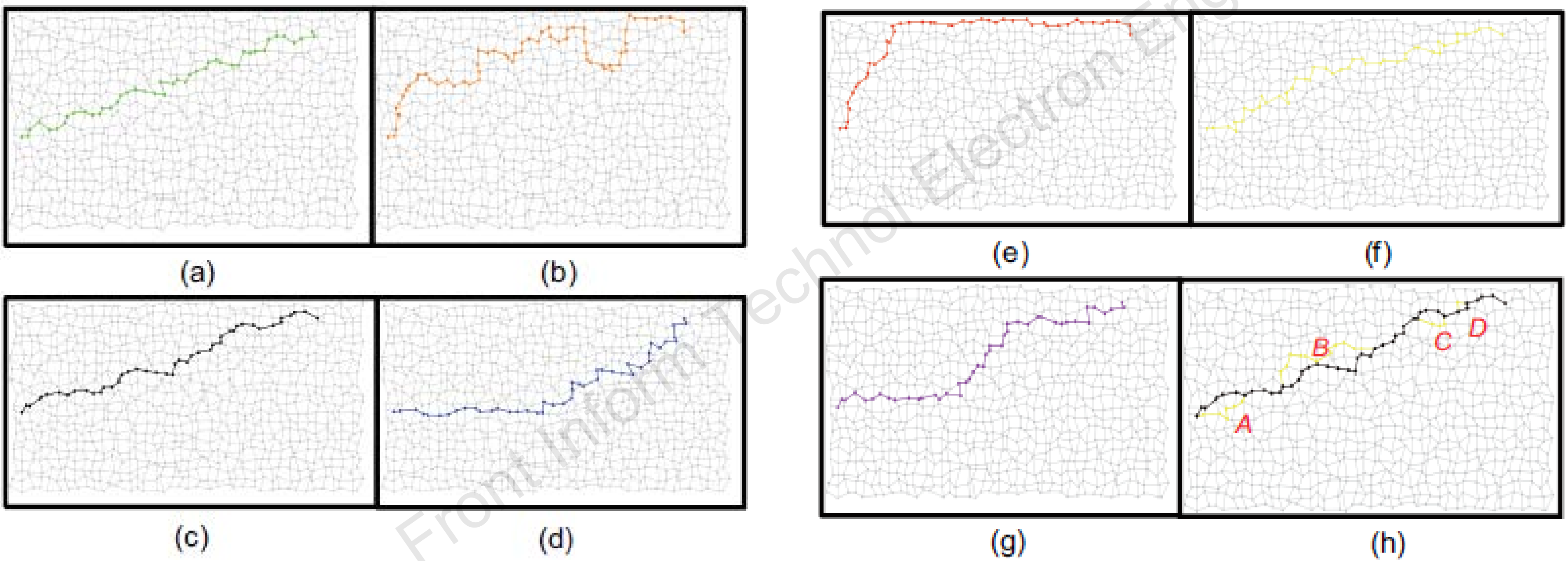

Fig. 12 Comparison of the shortest route of each method with random coordinates: (a) RRT; (b) A*; (c) APFD; (d) APF; (e) AC; (f) NSGA-II; (g) PSO; (h) APFD with NSGA-II

Major results (Cont'd)

Route recommendation evaluation

Table 4 Efficiency comparison between APFD and Dijkstra in the simulation map

Origin and destination points	Coordinate pair	Execution time (ms)		η
		APFD	Dijkstra	
(12, 03) (08, 10)	A19, B19	4	111	27.75
(03, 15) (11, 10)	A17, B17	20	283	14.15
(15, 13) (05, 28)	A20, B20	85	822	9.67
(16, 03) (08, 21)	A10, B10	81	616	7.60
(00, 05) (10, 28)	A1, B1	150	732	4.88
(01, 04) (19, 15)	A12, B12	173	684	3.95
(12, 35) (05, 12)	A14, B14	94	785	8.35
(09, 03) (15, 27)	A7, B7	79	779	9.86
(10, 15) (02, 38)	A3, B3	111	1071	9.65
(00, 05) (10, 28)	A18, B18	153	793	5.18
(07, 08) (00, 36)	A16, B16	129	1020	7.91
(09, 09) (19, 36)	A11, B11	211	1086	5.15
(18, 30) (00, 07)	A6, B6	363	1058	2.91
(18, 28) (01, 02)	A5, B5	426	1079	2.53
(01, 34) (11, 01)	A2, B2	284	1013	3.57
(14, 36) (01, 03)	A13, B13	43	1083	25.19
(00, 30) (17, 01)	A8, B8	535	1085	2.03
(17, 03) (03, 38)	A15, B15	537	1117	2.08
(00, 00) (19, 39)	A4, B4	1053	1116	1.06
(00, 39) (19, 00)	A9, B9	1093	1123	1.03

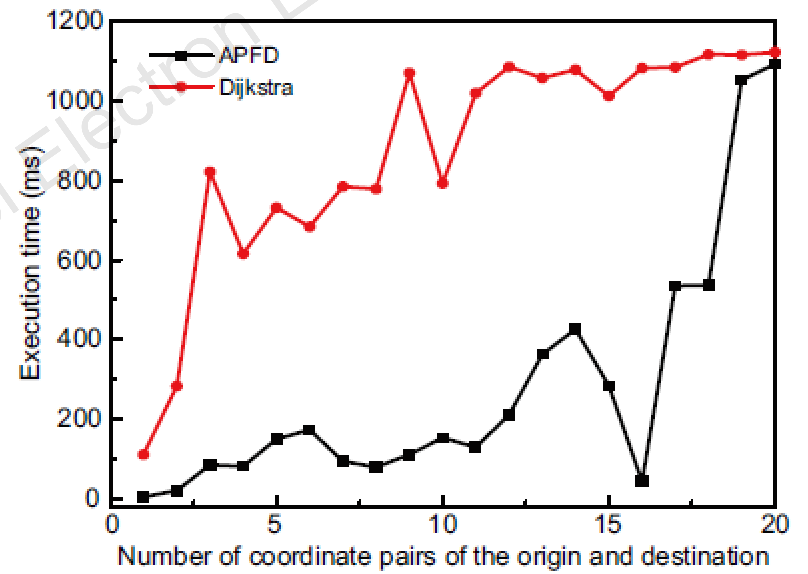

Fig. 13 Execution time comparison between APFD and Dijkstra in the simulation map

Major results (Case study)

Table 5 Total distance of each method in the real-word map

Coordinate pair	Total distance (m)						
	APFD	AC	A*	APF	RRT	NSGA-II	PSO
A15, B15	5697	9415	10 525	8879	7530	6540	5914
A6, B6	5962	6555	7055	7148	8636	6418	7214
A2, B2	8271	9293	10 098	9911	9519	8596	9317
A20, B20	10 604	11 321	12 029	11 998	14 031	12 550	11 178
A12, B12	10 650	11 650	13 151	12 244	13 002	11 473	12 487
A3, B3	12 439	12 947	14 920	13 259	13 512	13 369	13 889
A5, B5	12 444	13 640	15 894	14 889	14 412	12 967	14 187
A17, B17	12 682	13 278	15 451	16 589	17 771	13 131	14 789
A10, B10	13 055	13 920	14 569	15 598	18 879	13 996	14 501
A11, B11	13 514	16 026	18 286	16 519	16 747	14 140	15 547
A4, B4	14 959	16 752	17 002	15 994	16 687	15 887	16 667
A16, B16	15 342	16 231	17 309	16 991	19 002	15 998	17 417
A19, B19	15 542	16 011	16 911	17 017	17 809	16 646	16 468
A7, B7	17 012	17 977	18 388	18 014	19 998	17 579	18 789
A9, B9	19 660	24 265	25 507	24 488	25 916	21 121	23 896
A8, B8	22 401	24 753	33 086	28 146	27 115	23 540	24 789
A1, B1	24 159	24 336	28 391	26 889	26 110	24 878	25 501
A13, B13	26 644	30 036	31 724	28 985	29 936	27 187	29 879
A18, B18	27 321	30 745	35 268	33 132	35 957	28 137	30 024
A14, B14	28 088	32 674	33 733	29 994	32 110	29 898	33 217

Major results (Case study)

Table 6 Performance improvement of APFD compared with other methods in the real-word map

Coordinate pair	Performance improvement of APFD (%)					
	AC	A*	APF	RRT	NSGA-II	PSO
A15, B15	39.49	84.75	55.85	32.17	14.80	3.81
A6, B6	9.05	18.33	19.89	44.85	7.65	21.00
A2, B2	11.00	22.09	19.83	15.09	3.93	12.65
A20, B20	6.33	13.44	13.15	32.32	18.35	5.41
A12, B12	8.58	23.48	14.97	22.08	7.73	17.25
A3, B3	3.92	19.95	6.59	8.63	7.48	11.66
A5, B5	8.77	27.72	19.65	15.81	4.20	14.01
A17, B17	4.49	21.83	30.81	40.13	3.54	16.61
A10, B10	6.21	11.60	19.48	44.61	7.21	11.08
A11, B11	15.67	35.31	22.24	23.92	4.63	15.04
A4, B4	10.70	13.66	6.92	11.55	6.20	11.42
A16, B16	5.48	12.82	10.75	23.86	4.28	13.52
A19, B19	2.93	8.81	9.49	14.59	7.10	5.96
A7, B7	5.37	8.09	5.89	17.55	3.33	10.45
A9, B9	18.98	29.74	24.56	31.82	7.43	21.55
A8, B8	9.50	47.70	25.65	21.04	5.08	10.66
A1, B1	0.73	17.52	11.30	8.08	2.98	5.55
A13, B13	11.29	19.07	8.79	12.36	2.04	12.14
A18, B18	11.14	29.09	21.27	31.61	2.99	9.89
A14, B14	14.04	20.10	6.79	14.32	6.44	18.26

Major results (Case study)

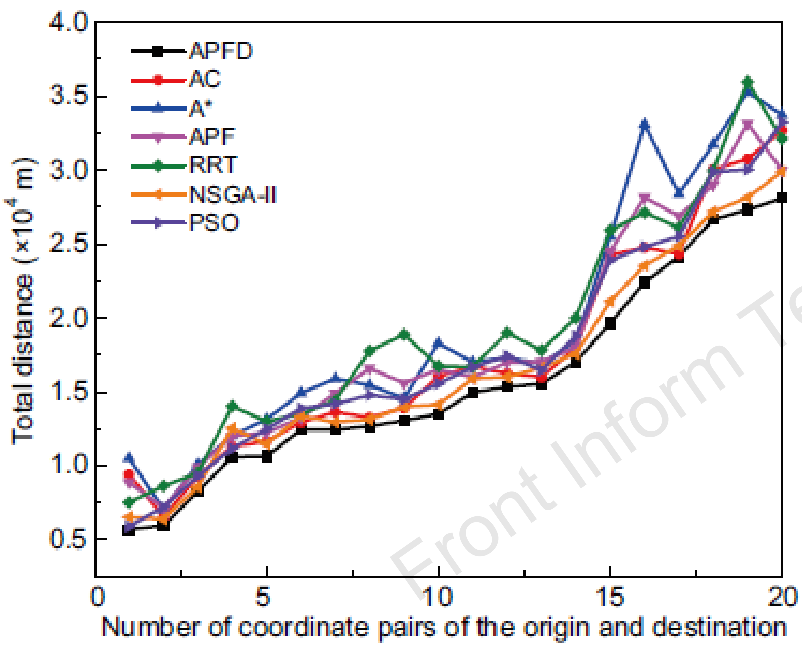

Fig. 15 The shortest route comparison among several methods under random coordinates in the real-word map

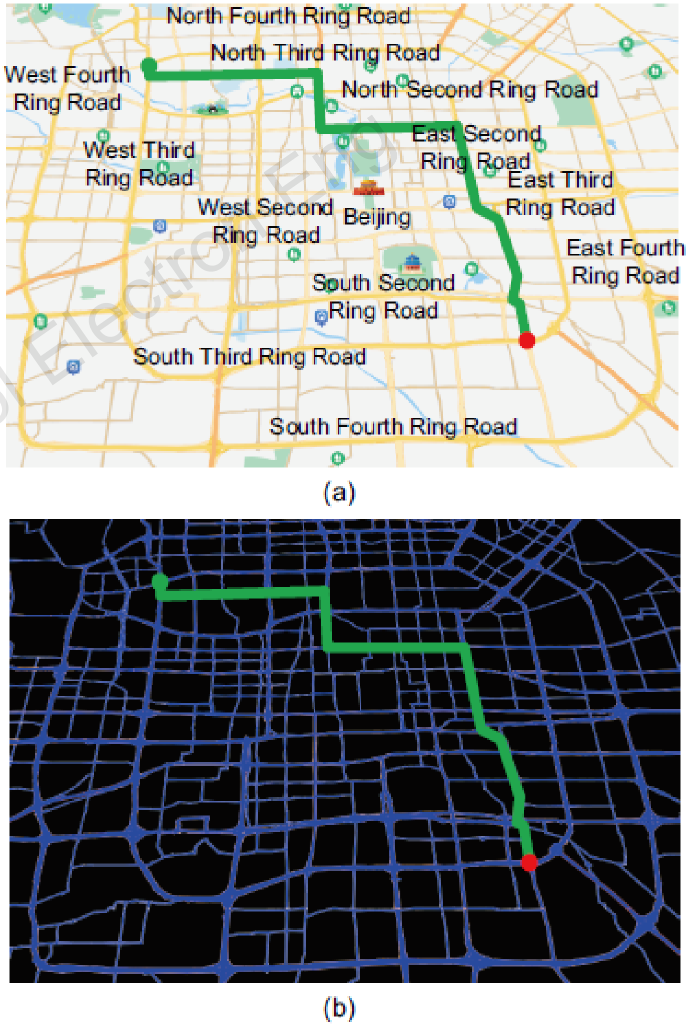

Fig. 16 The shortest route in the real-world road network: (a) road network on the Fourth Ring Road in Beijing; (b) extracted road network

Major results (Case study)

Table 7 Efficiency comparison between APFD and Dijkstra in the real-world map

Number	Coordinate pair	Execution time (ms)		η
		APFD	Dijkstra	
1	A15, B15	11	97	8.82
2	A6, B6	11	79	7.18
3	A2, B2	22	87	8.50
4	A20, B20	32	199	6.22
5	A12, B12	55	299	5.44
6	A3, B3	52	211	4.06
7	A5, B5	35	170	4.86
8	A17, B17	70	228	3.26
9	A10, B10	68	190	2.79
10	A11, B11	51	382	7.49
11	A4, B4	111	358	3.23
12	A16, B16	87	343	3.94
13	A19, B19	84	388	4.62
14	A7, B7	101	402	3.98
15	A9, B9	130	394	3.03
16	A8, B8	242	437	1.81
17	A1, B1	289	459	1.59
18	A13, B13	190	458	2.41
19	A18, B18	406	451	1.11
20	A14, B14	411	485	1.18

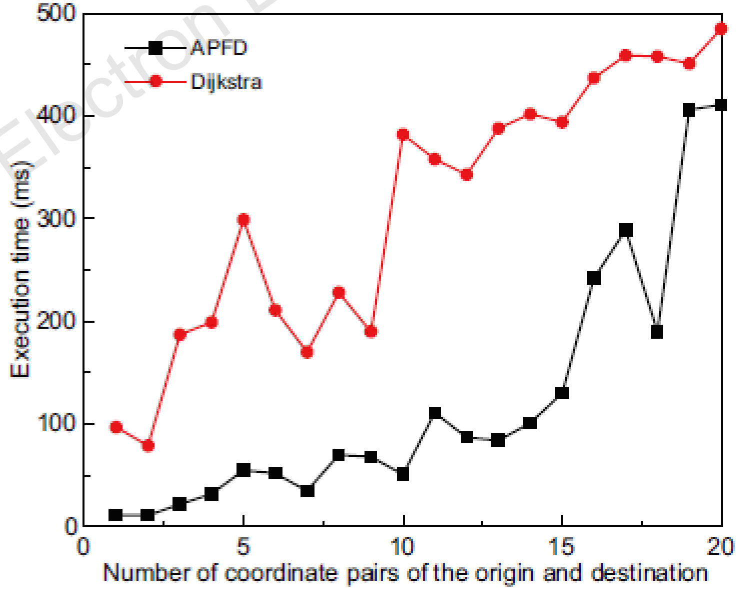

Fig. 17 Comparisons of the shortest route between APFD and Dijkstra in the real-world map

Conclusions

To quickly and efficiently recommend optimal route for taxis, we proposed an efficient route recommendation method (named APFD) based on the APF and Dijkstra methods. We first extracted the regions in which the optimal route via the origin and destination would likely be found, and then used the APF method to remove network nodes that might not form the optimal route, namely, redundant points. Finally, the Dijkstra method was used to complete the optimal route recommendation.

Wenyong ZHANG is currently an experimentalist at the College of Data Science and Information Engineering & Key Laboratory of Pattern Recognition and Intelligent Systems of Guizhou Province, Guizhou Minzu University, Guiyang, China. He received his MS degree at the College of Data Science and Information Engineering, Guizhou Minzu University, Guiyang, China, in 2020. His research interests include big data analytics, artificial intelligence, and machine learning.

Dawen XIA is currently a professor at the College of Data Science and Information Engineering & Key Laboratory of Pattern Recognition and Intelligent Systems of Guizhou Province, Guizhou Minzu University, Guiyang, China. He received his PhD degree from the College of Computer and Information Science & College of Software, Southwest University, Chongqing, China, in 2016. From November 2019 to November 2020, he was a visiting scholar supported by China Scholarship Council with the Rutgers, the State University of New Jersey, USA. His research interests include big data analytics, artificial intelligence, and data mining.

Huaqing LI is currently a professor at the College of Electronic and Information Engineering & Key Laboratory of Nonlinear Circuits and Intelligent Information Processing of Chongqing, Southwest University, Chongqing, China. He received his BS degree in information and computing science from Chongqing University of Posts and Telecommunications, Chongqing, China, in 2009, and his PhD degree in computer science and technology from Chongqing University, Chongqing, China, in 2013. He was a post doctorate with the School of Electrical and Information Engineering, University of Sydney, Sydney, NSW, Australia, from September 2014 to September 2015, and the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, from November 2015 to November 2016. His main research interests include artificial intelligence, nonlinear dynamics and control, multi-agent systems, and distributed optimization. He serves as a regional editor for *Neural Computing and Applications* and an editorial board member for *IEEE Access*.