

User Guide of T-Drive Data

Version 1.0

Updated on August 12, 2011

1 Data Description

This dataset contains the GPS trajectories of 10,357 taxis during the period of Feb. 2 to Feb. 8, 2008 within Beijing. The total number of points in this dataset is about 15 million and the total distance of the trajectories reaches to 9 million kilometers. Fig. 1 plots the distribution of time interval and distance interval between two consecutive points. The average sampling interval is about 177 seconds with a distance of about 623 meters. Each file of this dataset, which is named by the taxi ID, contains the trajectories of one taxi. Fig. 2 visualized the density distribution of the GPS points in this dataset.

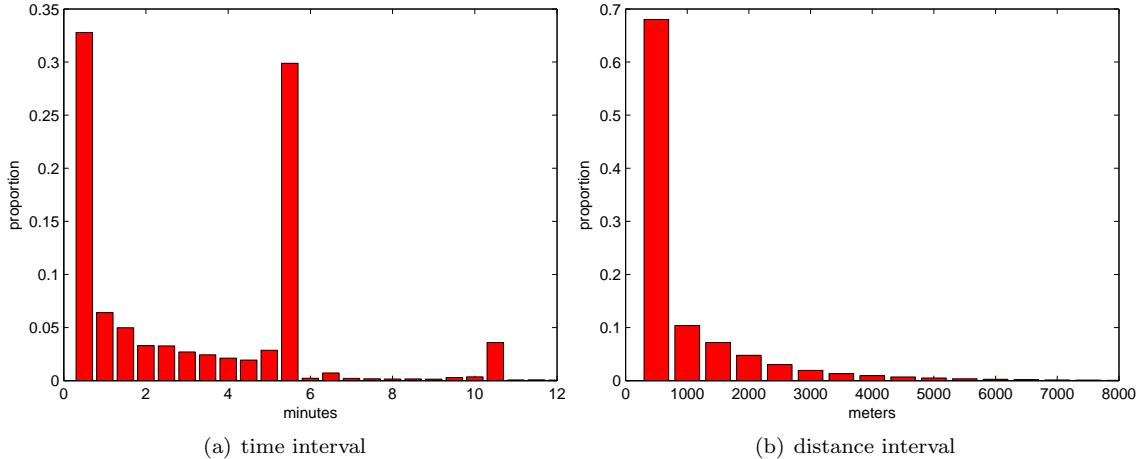


Figure 1: Histograms of time interval and distance between two consecutive points

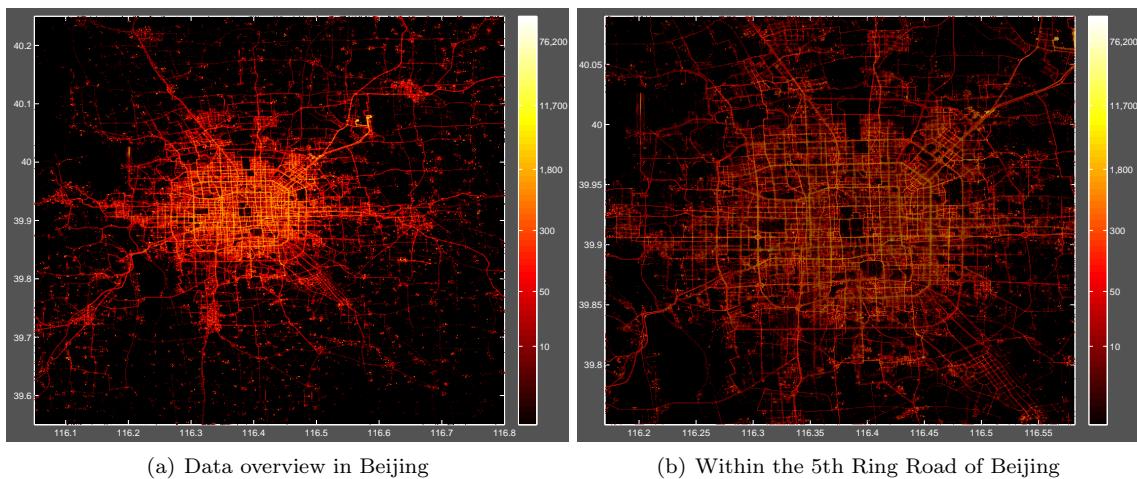


Figure 2: Distribution of GPS points, where the color indicates the density of the points

2 Data Format

Here is a piece of sample in a file:

1,2008-02-02 15:36:08,116.51172,39.92123
1,2008-02-02 15:46:08,116.51135,39.93883
1,2008-02-02 15:46:08,116.51135,39.93883
1,2008-02-02 15:56:08,116.51627,39.91034
1,2008-02-02 16:06:08,116.47186,39.91248
1,2008-02-02 16:16:08,116.47217,39.92498
1,2008-02-02 16:26:08,116.47179,39.90718
1,2008-02-02 16:36:08,116.45617,39.90531
1,2008-02-02 17:00:24,116.47191,39.90577
1,2008-02-02 17:10:24,116.50661,39.9145
1,2008-02-02 20:30:34,116.49625,39.9146

Each line of a file has the following fields, separated by comma:

`taxi id, date time, longitude, latitude`

3 Contact

Yu Zheng

Tel: 86-10-59173038 Email: yuzheng@microsoft.com

Homepage: <http://research.microsoft.com/en-us/people/yuzheng/default.aspx>

Address: Microsoft Research Asia, Tower 2, No. 5 Danling Street, Haidian District, Beijing, P.R. China 100080

4 Paper Citation

Please cite the following papers when using the dataset:

- [1] Jing Yuan, Yu Zheng, Xing Xie, and Guangzhong Sun. Driving with knowledge from the physical world. In *The 17th ACM SIGKDD international conference on Knowledge Discovery and Data mining*, KDD '11, New York, NY, USA, 2011. ACM.
- [2] Jing Yuan, Yu Zheng, Chengyang Zhang, Wenlei Xie, Xing Xie, Guangzhong Sun, and Yan Huang. T-drive: driving directions based on taxi trajectories. In *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, GIS '10, pages 99–108, New York, NY, USA, 2010. ACM.

5 Microsoft Research License Agreement

T-Drive Taxi Trajectories

This Microsoft Research License Agreement, including all exhibits ("MSR-LA") is a legal agreement between you and Microsoft Corporation (Microsoft or we) for the software or data identified above, which may include source code, and any associated materials, text or speech files, associated media and "online" or electronic documentation and any updates we provide in our discretion (together, the "Software").

By installing, copying, or otherwise using this Software, you agree to be bound by the terms of this MSR-LA. If you do not agree, do not install copy or use the Software. The Software is protected by copyright and other intellectual property laws and is licensed, not sold. SCOPE OF RIGHTS:

You may use this Software for any non-commercial purpose, subject to the restrictions in this MSR-LA. Some purposes which can be non-commercial are teaching, academic research, public demonstrations and personal experimentation. You may not distribute this Software or any derivative works in any form. In return, we simply require that you agree:

1. That you will not remove any copyright or other notices from the Software.
2. That if any of the Software is in binary format, you will not attempt to modify such portions of the Software, or to reverse engineer or decompile them, except and only to the extent authorized by applicable law.

3. That Microsoft is granted back, without any restrictions or limitations, a non-exclusive, perpetual, irrevocable, royalty-free, assignable and sub-licensable license, to reproduce, publicly perform or display, install, use, modify, post, distribute, make and have made, sell and transfer your modifications to and/or derivative works of the Software source code or data, for any purpose.

4. That any feedback about the Software provided by you to us is voluntarily given, and Microsoft shall be free to use the feedback as it sees fit without obligation or restriction of any kind, even if the feedback is designated by you as confidential.

5. THAT THE SOFTWARE COMES "AS IS", WITH NO WARRANTIES. THIS MEANS NO EXPRESS, IMPLIED OR STATUTORY WARRANTY, INCLUDING WITHOUT LIMITATION, WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE, ANY WARRANTY AGAINST INTERFERENCE WITH YOUR ENJOYMENT OF THE SOFTWARE OR ANY WARRANTY OF TITLE OR NON-INFRINGEMENT. THERE IS NO WARRANTY THAT THIS SOFTWARE WILL FULFILL ANY OF YOUR PARTICULAR PURPOSES OR NEEDS.

6. THAT NEITHER MICROSOFT NOR ANY CONTRIBUTOR TO THE SOFTWARE WILL BE LIABLE FOR ANY DAMAGES RELATED TO THE SOFTWARE OR THIS MSR-LA, INCLUDING DIRECT, INDIRECT, SPECIAL, CONSEQUENTIAL OR INCIDENTAL DAMAGES, TO THE MAXIMUM EXTENT THE LAW PERMITS, NO MATTER WHAT LEGAL THEORY IT IS BASED ON.

7. That we have no duty of reasonable care or lack of negligence, and we are not obligated to (and will not) provide technical support for the Software.

8. That if you breach this MSR-LA or if you sue anyone over patents that you think may apply to or read on the Software or anyone's use of the Software, this MSR-LA (and your license and rights obtained herein) terminate automatically. Upon any such termination, you shall destroy all of your copies of the Software immediately. Sections 3, 4, 5, 6, 7, 8, 11 and 12 of this MSR-LA shall survive any termination of this MSR-LA.

9. That the patent rights, if any, granted to you in this MSR-LA only apply to the Software, not to any derivative works you make.

10. That the Software may be subject to U.S. export jurisdiction at the time it is licensed to you, and it may be subject to additional export or import laws in other places. You agree to comply with all such laws and regulations that may apply to the Software after delivery of the software to you.

11. That all rights not expressly granted to you in this MSR-LA are reserved.

12. That this MSR-LA shall be construed and controlled by the laws of the State of Washington, USA, without regard to conflicts of law. If any provision of this MSR-LA shall be deemed unenforceable or contrary to law, the rest of this MSR-LA shall remain in full effect and interpreted in an enforceable manner that most nearly captures the intent of the original language.

Copyright (c) Microsoft Corporation. All rights reserved.