

大数据下的灾难行为分析和城市应急管理

宋 轩

日本东京大学

关键词：应急管理 灾难信息学 时空数据挖掘

近年来，在人类改造自然的同时，自然灾害（如地震、海啸、台风等）的破坏程度和发生频率也随之有了不同程度的增加。而且，这种趋势在未来数年恐怕还将持续下去。因此，如何开展快速有效的灾难救援、灾后重建和城市应急管理已成为世界各国政府及学术界日益关心的研究课题。

日本是自然灾害多发国，其国土面积仅占世界陆地面积的0.25%，但地震等自然灾害却占世界灾害总量的20%。根据日本内阁府统计，自2001年来，日本受灾总额占世界受灾总额的16%^[1]。特别是2011年3月11日，日本东部海域发生了里氏9.0级地震，自有人类文明记载以来，该地震是日本历史上最大的地震，也是人类近代史上五大地震之一^[2]。地震引发了巨大海啸，造成了重大人员伤亡和财产损失。地震和海啸还袭击了福岛核电站，引发了继切尔诺贝利核

泄漏以来，最大规模的核泄露事故^[3]。这场人类历史上最大的复合型灾难之一，至今仍对日本人民的生活有着极为深远的影响。本文将以东日本大地震和福岛核事故为背景，着重分析和介绍在大数据时代下，新一代灾难救援、灾后重建和城市应急管理核心技术及其应用。

大数据下的灾难应急管理

灾难应急管理往往以人为中心，在最大程度挽救生命的同时，如何快速有效地组织救援、安置疏散灾民、科学指导灾民避难、快速恢复城市交通和秩序成为灾后应急管理重中之重。因此，如何分析灾民的行为模式，如何对受灾城市可能发生的大规模灾民避难、迁移行为进行建模、预测和模拟，如何对灾民的撤离路线进行有效的推荐成为新一代灾后

应急管理技术的核心研究课题。

然而，由于缺乏有效的观测手段和支撑数据，以上研究课题的开展变得十分困难。传统的分析方法通常是：政府或民间对灾区进行实地考察，发放问卷调查，对灾民进行访谈；或在此基础上，使用空间遥感数据、户籍管理数据、灾害统计数据等对灾民的行为进行分析和建模^[4]。然而，这些方法往往成本巨大、时效性差、覆盖面小、数据量有限，而且获得的部分数据难以保证客观有效。

近年来，随着个人数字助理、智能手机、移动互联网、社交网络的发展与普及，各类基于手机定位的移动数据、GPS轨迹数据、地铁公交刷卡数据、社交网络数据等呈“爆炸式”增长，人类进入了大数据时代。在此背景下，这些海量的、基于位置的移动轨迹数据和社交网络数据在灾难发生的情况下，为人流移动的分析、建模、预测以及灾难应急管理提

供了新的方法^[5-8]。例如,日本东京大学的研究人员^[6]通过对推特数据进行分析和挖掘,对地震进行了实时预警。瑞典斯德哥尔摩大学的研究人员^[5]收集了190万海地手机用户的移动数据,通过这些数据来分析2010年1月12日海地大地震发生后,灾民的移动行为模式。此外,我们^[7,8]建立了一个160万人在日本一年中的GPS移动轨迹数据库,利用这个数据库对东日本大地震和福岛核事故发生后的灾民移动、避难行为进行了建模、预测和模拟。图1显示了在东日本大地震发生前后,日本首都圈的人流移动情况。从图中可以看到,大地震发生后,人流移动明显减缓,整个首都圈的公共交通完全陷入了瘫痪状态。

灾难行为模式分析和挖掘

根据自然灾害类型、规模的不同,灾难发生后,灾民的行为模式也不尽相同。研究表明^[8],大型灾难后的灾民避难行为通常是群体性的,并受到多方面因素影响,如新闻舆论、政府行政命令、救灾措施与援助、灾民的社会关系以及社交网络传播等等。

按照响应时间快慢,可分为即刻避难和延迟避难;按照避难周期长短,可以分为短期避难和中长期避难;按照目的地不同,可分为自行避难和安置避难等等。因此,在大数据时代下,分析和挖掘灾民的灾后避难行为模式,对于灾难救援和应急管理有着非常重要的意义。

显著地点挖掘 在海量时空轨迹数据中,挖掘不同类型的灾难行为模式,通常需要先对个体数据在灾难发生前长时间的移动行为模式进行分析,找出一些显著地点,如住所、工作单位、经常光顾的超市和商店等^[9,10]。通过对比灾难发生前后显著地点分布的变化,挖掘出个体中长期的避难行为和返回情况^[8]。图2显示了东日本大地震和福岛核事故发生后,福岛县、宫城县和岩手县灾民的中长期避难行为和一些主要受灾城市的人口变化情况。

基于位置的社交网络挖掘 研究发现^[5],在2010年海地大地震中,灾民的移动和避难行为模式通常和他在正常情况下的移动高度关联,个体灾民的避难模式常常取决于他在正常情况下的社交联系。因此,基于个体的长期移动轨迹数据,对其在灾难发

生前后的社交网络挖掘变得尤为重要。近年来,学者们^[11]通过建立移动个体间基于位置的图模型、基于用户的图模型以及基于用户和位置的图模型对个体间的社交网络进行建模。同时,基于位置的社交网络挖掘也成为数据挖掘和城市计算等研究领域的热门课题^[11]。

避难行为模式的因果推断 研究发现^[8],在灾难发生后,城市受灾程度、城市类型、交通状况、各种新闻媒体报道、社交网络数据传播、政府避难建议以及安置点处置等等各类复杂因素都会对灾民的避难行为产生影响。如图2所示,世界各地包括日本在内的各类媒体、社交网络等对福岛核事故进行广泛的报道讨论后,在福岛县及其临近的县市产生了大规模的人流移动和中长期避难行为。因此,分析各类复杂因素是如何影响灾民的避难行为的,并对其建模和进行因果推断就显得尤为重要。此类研究也将成为大数据下城市应急管理的重要研究方向。

群体避难迁移行为的建模、预测和模拟

研究表明,2010年的海地

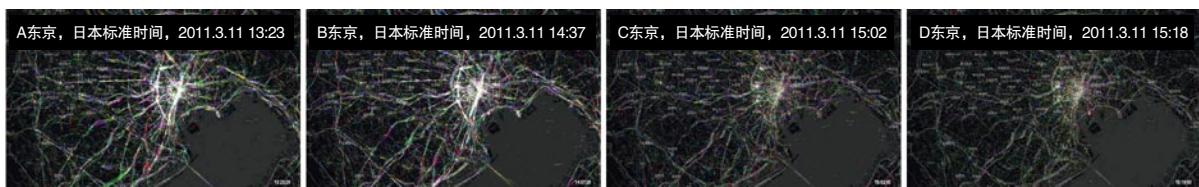


图1 东日本大地震发生时首都圈的人流移动

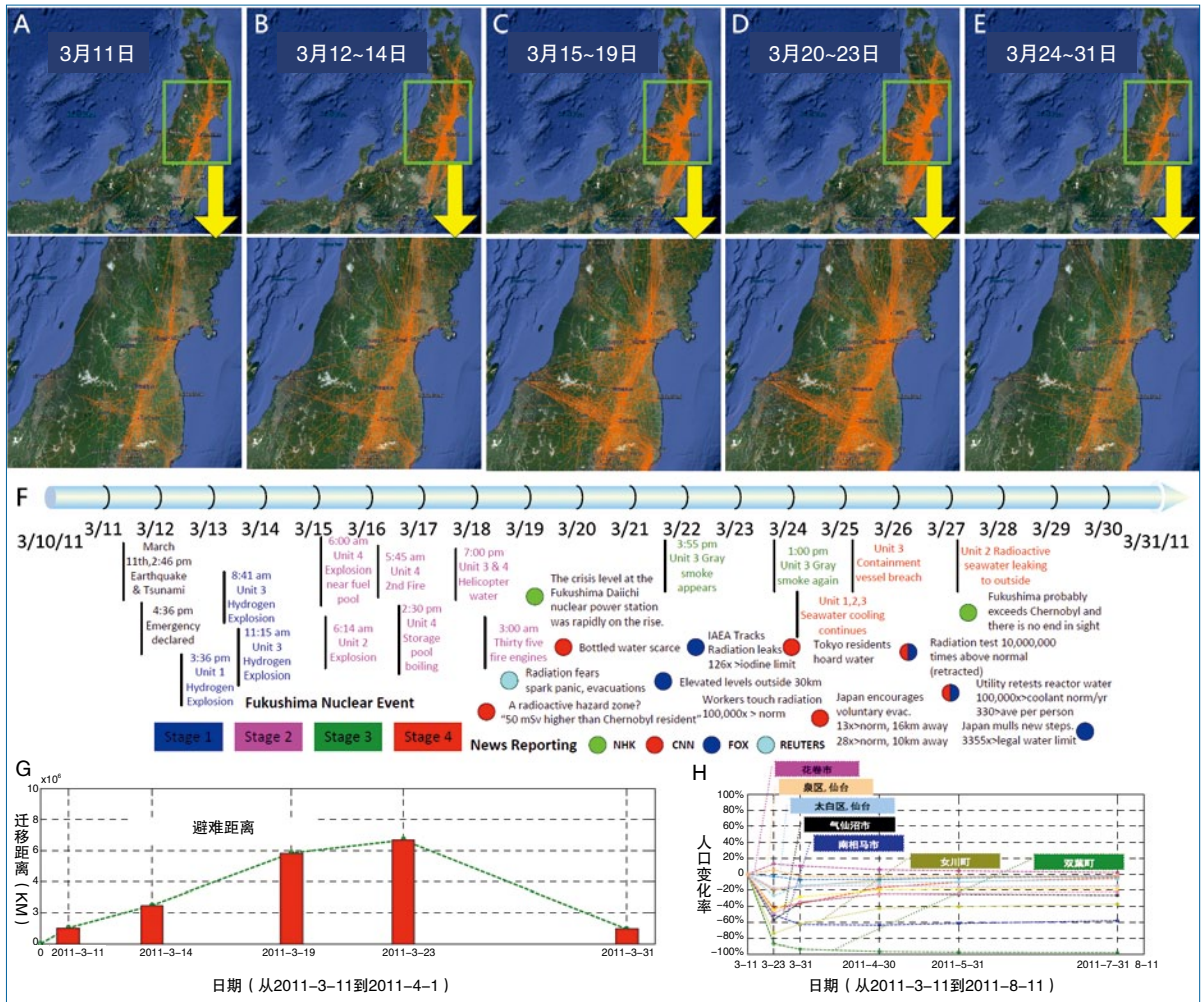


图2 东日本大地震和福岛核事故中挖掘出的中长期避难行为（图A~E显示了挖掘出的中长期避难行为（避难起点到终点的迁移连线）。图F显示了相关媒体报道及时间轴。图G显示了在不同时间点挖掘出的避难行程距离。图H显示了主要受灾城市不同时间点的人口变化率）

大地震、2011年的东日本大地震以及福岛核事故都导致了受灾城市大规模的群体性迁移避难行为^[5,7,8]。因此，群体迁移行为的建模、预测和模拟对未来可能的灾后人道救援、灾后安置以及城市应急管理十分重要。

群体避难迁移行为建模

对海量灾后移动轨迹数据进行建模，通常需要一个简明模型来描述这些群体的避难迁移行为。常

规想法是将这些轨迹数据映射到城市交通网络上，并对其进行参数学习。然而，由于大型灾难发生后，城市交通网络通常会发生瘫痪（如图1所示），因此，此类方法通常难以达到理想效果。另一个选择是使用这些避难迁移行为（海量移动轨迹数据）来建立群体避难图模型 (evacuation graph)^[8]，并用这些避难轨迹来学习图的节点和边的参数以及图的

结构连接^[8]。图3显示了学习得到的东日本大地震和福岛核事故中，福岛县、宫城县和岩手县灾民的避难图模型。在该图模型中，节点表示一些重点受灾区域，边描述受灾区域间灾民的移动信息以及区域间的交通、社会和经济联系。通过该模型，我们可以很容易地获得对各个受灾区域人流迁移的推断。图3中，暖色的节点代表人流的迁出，冷色的节点

代表人流的迁入。分析发现，受海啸袭击严重的城市居民通常会选择去与其临近的大城市避难。

群体避难迁移行为的预测和模拟 对特定群体避难迁移行为建模的实现也使得对给定城市未来可能发生的、灾难下的大规模人流移动进行预测和模拟成为可能。由于城市区域间的交通、社会和经济联系通常是相对稳定的，而这些联系和一些外部因素（如灾难类型、受灾程度、政府媒体避难建议等等）又决定了灾难下大规模人流移动情况。因此，可以使用贝叶斯推断 (Bayesian inference) 对特定城市内的大规模人流移动进行模拟和预测^[8]。训练完成的避难图模型可提供特定

区域或城市的大规模人流移动的先验概率分布，同时对观测到的短时间内（如灾难发生后的很短时间内）部分人流的移动轨迹进行计算，得到人流移动的后验概率，并对其进行最大后验概率估计，从而预测和模拟出可能的大规模人流移动^[8]。图4显示了假设当日本首都圈的城市交通再次陷入大规模瘫痪时（如未来发生大型灾难时），一些重点城市区域可能发生大规模人流移动的模拟和预测。

避难应急路线推荐

大型灾难发生后，需要尽快转移和安置灾民，恢复城市交通，

科学地指导灾民撤离避难。因此，如何推荐安全、快速、有效的避难路线是重要的研究课题。

近年来，研究人员通过对大量城市车辆的GPS轨迹进行分析和建模，设计出了基于车辆运行模式的路线推荐系统^[12,13]，并取得了很好的效果。然而，在大型灾难情况下，常规的城市交通系统往往会部分或全部瘫痪。此外，大型灾难情况下的路线推荐，除了要考虑行程长短和所需时间，往往还要考虑路线的安全性及可能发生的大规模交通拥堵。通常可以在避难图模型上，使用马尔科夫决策过程 (Markov decision process)^[14]对灾民的撤离行为进行分析建模，通过训练模型参数、

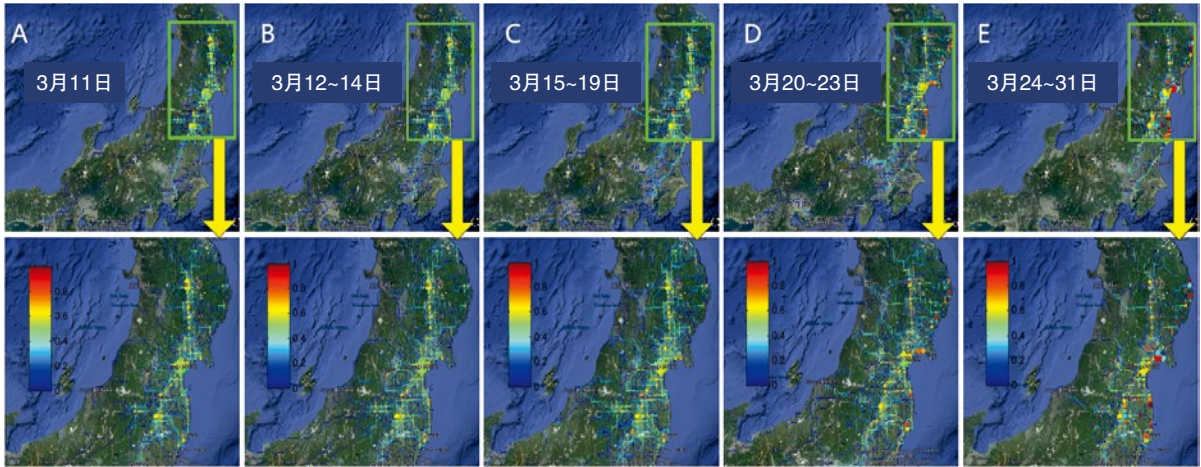


图3 东日本大地震和福岛核事故中福岛县、宫城县以及岩手县的避难图模型

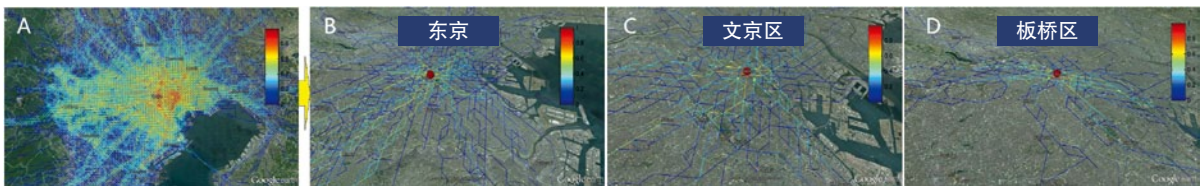


图4 大型灾难发生时日本首都圈人流移动的模拟和预测（其中图A为灾难情况下首都圈的人流移动图模型。B~D为对重点城市区域人流移动的模拟和预测。其中暖色轨迹代表高发生概率，冷色轨迹代表低发生概率）

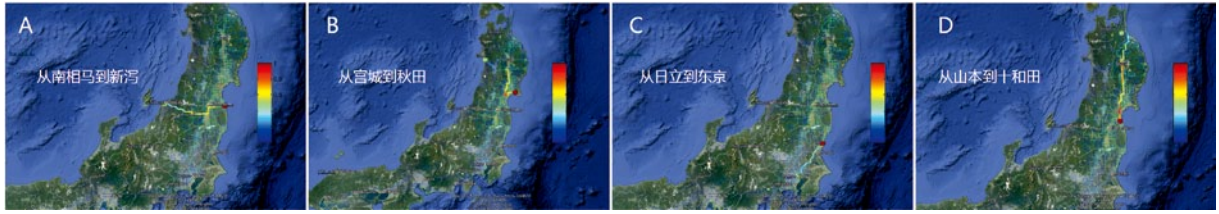


图5 大灾难发生时的撤离路线推荐 (其中深红色圆点表示起点, 淡绿色圆点表示终点。线路中的冷暖色表示路线的相对安全程度)

概率推理完成对最优撤离路线的推荐^[8,15]。图5显示的是通过分析东日本大地震和福岛核事故中灾民的撤离行为, 假设大型灾难再次发生时, 在给定起点和终点情况下一些最优路线的推荐。

大数据时代下的灾难救援和城市应急管理作为一个全新的研究领域, 还面临着诸多挑战。我们相信, 随着数据的积累和技术革新, 该领域会催生更多的研究课题和应用, 帮助人类进一步减少灾难损失, 挽救更多的生命。■



宋 轩

日本东京大学特任助理教授。主要研究方向为数据挖掘、机器感知、智能系统设计、应急管理和灾难信息学。
songxuan@csis.u-tokyo.ac.jp

参考文献

[1] Cabinet Office, government of Japan, <http://www.cao.go.jp/index.html>

[2] O. Norio, T. Ye, Y. Kajitani, P. Shi, and H. Tatano. The 2011 eastern Japan great earthquake disaster: overview and comments, *International Journal of Disaster Risk Science*, 2 (2011), 34~42

[3] M. Holt, R. J. Campbell, and M. B. Nikitin. Fukushima nuclear disaster, *Congressional Research Service Reports*, (2012)

[4] V. Hristidis, S. Chen, T. Li, S. Luis, and Yi Deng. Survey of data management and analysis in disaster situations, *The Journal of Systems and Software*, 83 (2010), 1701~1714

[5] X. Lu, L. Bengtssona, and P. Holme. Predictability of population displacement after the 2010 Haiti earthquake, *Proc. of the National Academy of Sciences of USA (PNAS)*, 109 (2012), 11576~11581

[6] T. Sakaki, M. Okazaki, and Yutaka Matsuo. Tweet analysis for real-time event detection and earthquake reporting system development. *IEEE Trans. Knowl. Data Eng.* 25(4) (2013), 919~931

[7] X. Song, Q. Zhang, Y. Sekimoto, T. Horanont, S. Ueyama, R. Shibasaki. An intelligent system for large-scale disaster behavior analysis and reasoning, *IEEE Intelligent Systems*, (2013)

[8] X. Song, Q. Zhang, Y. Sekimoto, T. Horanont, S. Ueyama, R. Shibasaki. Modeling and probabilistic reasoning of population evacuation during large-scale disaster, *Proc. of 19th SIGKDD conference on Knowledge Discovery and Data Mining (KDD)*, (2013)

[9] S. Isaacman, R. Becker, R. Caceres, S. Kobourov, M. Martonosi, J. Rowland and A. Varshavsky, identifying important places in peoples lives from cellular network data, *Pervasive Computing* (2011), 133~151

[10] J. Zheng, L. Zhang, X. Xie and W. Y. Ma, Mining interesting locations and travel sequences from GPS trajectories, in *Proc. 18th Int. Conf. on World Wide Web (ACM, 2009)*, 791~800

[11] Yu Zheng. Location-based social networks: users. in *computing with spatial trajectories*, Zheng, Y and Zhou, X, Eds. Springer, 2011

[12] J. Yuan, Y. Zheng, C. Zhang, W. Xie, X. Xie, G. Sun, and Y. Huang. T-Drive: driving directions based on taxi trajectories, *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM GIS 2010)*, 99~108

[13] B. D. Ziebart, A. Maas, J.A. Bagnell, and A. K. Dey. Navigate like a cabbie: probabilistic reasoning from observed context-aware behavior, *Proc. of Ubicomp.*, (2008), 322~331

[14] M. L. Puterman. *Markov decision processes: discrete stochastic dynamic programming*, Wiley-Inter-science, (1994)

[15] B. D. Ziebart, A. Maas, J.A. Bagnell, and A. K. Dey. Maximum entropy inverse reinforcement learning, *Proc. of AAAI Conference on Artificial Intelligence (AAAI)*, (2008), 1433~1438