

Towards Parallel Detection of Moving Flock Patterns in Large Spatio-temporal Datasets

Final report

Andres Calderon

December 7, 2016

1 Introduction

Spatio-temporal data is ubiquitous nowadays. Thanks to new technologies and location devices (such as Internet of Things, Remote Sensing, Smart phones, GPS, RFID, etc.), the collection of huge amount of spatio-temporal data is now possible. With the appearance of this datasets also appears the need of new techniques which allow the analysis and detection of useful patterns.

Applications for this kind of information are diverse and interesting, in particular if they come in the way of trajectory datasets. Case of studies range from transportation system management (Di Lorenzo et al. [2016]; Johansson and others [2015]) to Ecology (Johnston et al. [2015]; La Sorte et al. [2016]). For instance, Turdukulov et al. [2014] explore the finding of complex motion patterns to discover similarities between tropical cyclone paths. Similarly, Amor et al. [2016] use eye trajectories to understand which strategies people use during a visual search. Also, Holland et al. [1999] track the behavior of tiger sharks in the coasts of Hawaii in order to understand their migration patterns.

Recently, there has been an increasing interest in exploiting more complex movement patterns in spatio-temporal datasets. Traditional range and nearest neighbor queries do not capture the collective behavior of moving objects. Moving cluster (Kalnis et al. [2005]), convoys (Jeung et al. [2008]) and flock patterns (Benkert et al. [2008]; Gudmundsson and van Kreveld [2006]) are new movement patterns which unveil how entities move together during a minimum time interval.

In particular, a moving flock pattern show how objects move close enough during a given period of time. A better understanding on how entities move in space is of special interest in areas such as sports (Iwase and Saito [2002]), surveillance and security (Makris and Ellis [2002]; Piciarelli et al. [2005]), urban development (Huang et al. [2016]; Long and Thill [2015]) and socio-economic geography (Frank et al. [2000]).

Despite the fact that much more data become available, state-of-the-art techniques to mine complex movement patterns still depict low scalability and poor performance in big spatial data. The present work aims to find an initial solution to implement a parallel method to discover moving flock patterns in large spatio-temporal datasets. It is thought that new trends in distributed in-memory framework for spatial operations could help to speed up the detection of this kind of patterns.

The following section will state the related work in the area. Section 3 will explain the details of the implementation of the proposed solution while section 4 will present their experimental results. Finally, section 5 will discuss some conclusions and future work.

2 Related work

Recently increase use of location-aware devices (such as GPS, Smart-phones and RFID tags) has allowed the collection of a vast amount of data with a spatial component linked to them. Different studies have focused in analyzing and mining this kind of collections [Leung, 2010][Miller and Han, 2001]. In this area, trajectory datasets have emerged as an interesting field where diverse kind of patterns can be identified [Zheng and Zhou, 2011][Vieira and Tsotras, 2013]. For instance, authors have proposed techniques to discover motion spatial patterns such as moving clusters[Kalnis et al., 2005], convoys[Jeung et al., 2008] and flocks [Benkert et al., 2008][Gudmundsson and van Kreveld, 2006]. In particular, [Vieira et al., 2009] and [Turdukulov et al., 2014] propose two novel algorithms to find moving flock patterns in very large spatio-temporal datasets.

A flock pattern is defined as a group of entities which move together for a defined lapse of time [Benkert et al., 2008]. Applications to this kind of patterns are diverse and range from surveillance to integrated transport systems. For example, [Turdukulov et al., 2014] explore the finding of this class of patterns to discover similarities between tropical cyclone paths. Also, [Calderon Romero, 2011] finds moving flock patterns in iceberg trajectories to understand their movement behavior and how they related to changes in ocean's currents.

The algorithms proposed by [Vieira et al., 2009] and [Turdukulov et al., 2014] share the same initial strategy to detect flock patterns. In that, first they find clusters of points which could be close enough to initiate a flock for each time interval. This is a costly operation due to the large number of points and intervals to be analyzed. The technique uses a grid-based index and a stencil (see figure 1) to speed up the process but the complexity is still high.

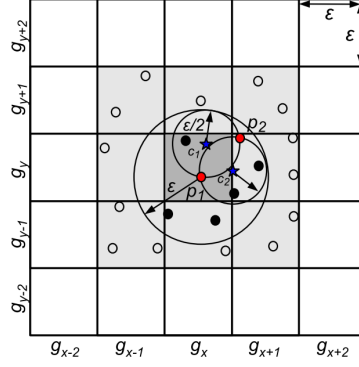


Figure 1: Grid-based index used in [Vieira et al., 2009].

3 Parallelizing flock detection

4 Experiments

5 Conclusions and future work

I will share the lessons learned during the project and some future ideas to continue with the research.

References

- Tatiana A. Amor, Saulo D. S. Reis, Daniel Campos, Hans J. Herrmann, and José S. Andrade. Persistence in eye movement during visual search. *Scientific Reports*, 6: 20815, February 2016. ISSN 2045-2322. doi: 10.1038/srep20815. URL <http://www.nature.com/articles/srep20815>. bibtex: amor_persistence_2016.
- Marc Benkert, Joachim Gudmundsson, Florian Hübner, and Thomas Wolle. Reporting flock patterns. *Computational Geometry*, 41(3):111–125, November 2008. ISSN 09257721. doi: 10.1016/j.comgeo.2007.10.003. URL <http://linkinghub.elsevier.com/retrieve/pii/S092577210700106X>.
- Andres Oswaldo Calderon Romero. *Mining moving flock patterns in large spatio-temporal datasets using a frequent pattern mining approach*. mastersthesis, University of Twente, 2011. URL https://www.itc.nl/library/papers_2011/msc/gem/calderon.pdf. bibtex: calderon_romero_mining_2011.
- G. Di Lorenzo, M. Sbodio, F. Calabrese, M. Berlingerio, F. Pinelli, and R. Nair. AllAboard: Visual Exploration of Cellphone Mobility Data to Optimise Public Transport. *IEEE Transactions on Visualization and Computer Graphics*, 22(2):1036–1050, February 2016. ISSN 1077-2626. doi: 10.1109/TVCG.2015.2440259. URL <http://ieeexplore.ieee.org/document/7117451/>. bibtex: di_lorenzo_allaboard_2016.

- Andrew Frank, Jonathan Raper, and J. P. Cheylan, editors. *Life and Motion of Socio-Economic Units*. CRC Press, London ; New York, 1 edition edition, December 2000. ISBN 978-0-7484-0845-0.
- Joachim Gudmundsson and Marc van Kreveld. Computing Longest Duration Flocks in Trajectory Data. In *Proceedings of the 14th Annual ACM International Symposium on Advances in Geographic Information Systems*, GIS '06, pages 35–42, New York, NY, USA, 2006. ACM. ISBN 1-59593-529-0. doi: 10.1145/1183471.1183479. URL <http://doi.acm.org/10.1145/1183471.1183479>. bibtex: gudmundsson_computing_2006.
- K. N. Holland, B. M. Wetherbee, C. G. Lowe, and C. G. Meyer. Movements of tiger sharks (*Galeocerdo cuvier*) in coastal Hawaiian waters. *Marine Biology*, 134(4):665–673, 1999. URL <http://link.springer.com/article/10.1007/s002270050582>. bibtex: holland_movements_1999.
- Xiaoke Huang, Ye Zhao, Chao Ma, Jing Yang, Xinyue Ye, and Chong Zhang. TrajGraph: A Graph-Based Visual Analytics Approach to Studying Urban Network Centralities Using Taxi Trajectory Data. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):160–169, January 2016. ISSN 1077-2626. doi: 10.1109/TVCG.2015.2467771. URL <http://ieeexplore.ieee.org/document/7192687/>.
- Sachiko Iwase and Hideo Saito. Tracking Soccer Player Using Multiple Views. In *MVA*, pages 102–105, 2002. URL http://hvrl.ics.keio.ac.jp/paper/pdf/international_Conference/2002/sachiko_mva2002.pdf.
- Hoyoung Jeung, Man Lung Yiu, Xiaofang Zhou, Christian S. Jensen, and Heng Tao Shen. Discovery of Convoys in Trajectory Databases. *Proc. VLDB Endow.*, 1(1):1068–1080, August 2008. ISSN 2150-8097. doi: 10.14778/1453856.1453971. URL <http://dx.doi.org/10.14778/1453856.1453971>. bibtex: jeung_discovery_2008.
- Karl Henrik Johansson and others. An efficiency measure for road transportation networks with application to two case studies. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 5149–5155. IEEE, 2015. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=7403025. bibtex: johansson_efficiency_2015.
- Alison Johnston, Daniel Fink, Mark D. Reynolds, Wesley M. Hochachka, Brian L. Sullivan, Nicholas E. Bruns, Eric Hallstein, Matt S. Merrifield, Sandi Matsumoto, and Steve Kelling. Abundance models improve spatial and temporal prioritization of conservation resources. *Ecological Applications*, 25(7):1749–1756, 2015. URL <http://onlinelibrary.wiley.com/doi/10.1890/14-1826.1/full>. bibtex: johnston_abundance_2015.
- Panos Kalnis, Nikos Mamoulis, and Spiridon Bakiras. On Discovering Moving Clusters in Spatio-temporal Data. In Claudia Bauzer Medeiros, Max J. Egenhofer, and Elisa Bertino, editors, *Advances in Spatial and Temporal Databases*, Lecture Notes in Computer Science, pages 364–381. Springer Berlin Heidelberg, August 2005. ISBN

- 978-3-540-28127-6 978-3-540-31904-7. URL http://link.springer.com/chapter/10.1007/11535331_21. DOI: 10.1007/11535331_21 bibtex: kalnis_discovering_2005.
- Frank A. La Sorte, Daniel Fink, Wesley M. Hochachka, and Steve Kelling. Convergence of broad-scale migration strategies in terrestrial birds. *Proceedings of the Royal Society B: Biological Sciences*, 283(1823):20152588, January 2016. ISSN 0962-8452, 1471-2954. doi: 10.1098/rspb.2015.2588. URL <http://rspb.royalsocietypublishing.org/lookup/doi/10.1098/rspb.2015.2588>. bibtex: la_sorte_convergence_2016.
- Yee Leung. *Knowledge Discovery in Spatial Data*. Springer Science & Business Media, March 2010. ISBN 978-3-642-02664-5. bibtex: leung_knowledge_2010.
- Ying Long and Jean-Claude Thill. Combining smart card data and household travel survey to analyze jobs–housing relationships in Beijing. *Computers, Environment and Urban Systems*, 53:19–35, September 2015. ISSN 0198-9715. doi: 10.1016/j.compenvurbsys.2015.02.005. URL <http://www.sciencedirect.com/science/article/pii/S0198971515000356>.
- Dimitrios Makris and Tim Ellis. Path detection in video surveillance. *Image and Vision Computing*, 20(12):895–903, October 2002. ISSN 0262-8856. doi: 10.1016/S0262-8856(02)00098-7. URL <http://www.sciencedirect.com/science/article/pii/S0262885602000987>.
- Harvey J. Miller and Jiawei Han. *Geographic Data Mining and Knowledge Discovery*. Taylor & Francis, Inc., Bristol, PA, USA, 2001. ISBN 0-415-23369-0. bibtex: miller_geographic_2001.
- C. Piciarelli, G. L. Foresti, and L. Snidaro. Trajectory clustering and its applications for video surveillance. In *IEEE Conference on Advanced Video and Signal Based Surveillance, 2005.*, pages 40–45, September 2005. doi: 10.1109/AVSS.2005.1577240.
- Ulanbek Turdukulov, Andres Oswaldo Calderon Romero, Otto Huisman, and Vasilios Retsios. Visual mining of moving flock patterns in large spatio-temporal data sets using a frequent pattern approach. *International Journal of Geographical Information Science*, 28(10):2013–2029, October 2014. ISSN 1365-8816. doi: 10.1080/13658816.2014.889834. URL <http://dx.doi.org/10.1080/13658816.2014.889834>. bibtex: turdukulov_visual_2014.
- Marcos R. Vieira and Vassilis Tsotras. *Spatio-Temporal Databases: Complex Motion Pattern Queries*. Springer Science & Business Media, October 2013. ISBN 978-3-319-02408-0. bibtex: vieira_spatio-temporal_2013.
- Marcos R. Vieira, Petko Bakalov, and Vassilis J. Tsotras. On-line Discovery of Flock Patterns in Spatio-temporal Data. In *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, GIS '09*, pages 286–295, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-649-6. doi:

10.1145/1653771.1653812. URL <http://doi.acm.org/10.1145/1653771.1653812>.
bibtex: vieira_line_2009.

Yu Zheng and Xiaofang Zhou. *Computing with Spatial Trajectories*. Springer Science & Business Media, October 2011. ISBN 978-1-4614-1629-6. bibtex: zheng_computing_2011.