

## Learning from our movements – Mobility data analytics

Instructor: prof. Yannis Theodoridis (Univ. Piraeus)

ACM Europe Summer School on Data Science, Athens, Jul. 17, 2018

### Supplementary bibliography

#### 1) Getting to know your data

##### a) Nature of mobility data; sources; applications

- Alvares LO, Bogorny V, Kuijpers B, de Macedo JAF, Moelans B, and Vaisman A (2007) A model for enriching trajectories with semantic geographical information. In Proceedings of GIS.
- Andersen O, Jensen CS, Torp K, Yang B (2013) EcoTour: reducing the environmental footprint of vehicles using eco-routes. In Proceedings of MDM.
- Andrienko G, Andrienko N, Wrobel S (2007) Visual analytics tools for analysis of movement data, ACM SIGKDD Explorations, 9(2), pp. 38-46.
- Andrienko N, Andrienko G, Pelekis N, Spaccapietra S (2008) Basic concepts of movement data. In Giannotti F, Pedreschi D (eds.) Mobility, Data Mining and Privacy - Geographic Knowledge Discovery, pp. 15-38.
- Andrienko G, Andrienko N, Fuchs G, Wood J (2017a) Revealing patterns and trends of mass mobility through spatial and temporal abstraction of origin-destination movement data. IEEE Trans. Vis. Comput. Graph. 23(9), pp. 2120-2136.
- Andrienko G, Andrienko N, Chen W, Maciejewski R, Zhao Y (2017b) Visual analytics of mobility and transportation: state of the art and further research directions. IEEE Trans. Intelligent Transportation Systems, 18(8), pp. 2232-2249.
- Bajaj R, Ranaweera SL, Agrawal DP (2002) GPS: Location-tracking technology. IEEE Computer, 35(4), pp. 92-94.
- Bogorny V, Avancini H, De Paula BL, Kuplish CR, Alvares LO (2011) Weka-STPM: a Software Architecture and Prototype for Semantic Trajectory Data Mining. Transactions in GIS, 15(2), pp. 227-248.
- Bogorny V, Kuijpers B, Alvares LO (2009) ST-DMQL: a semantic trajectory data mining query language. International Journal of Geographical Information Science, 23(10), pp. 1245-1276.
- Brakatsoulas S, Pfoser D, Salas R, Wenk C. (2005) On map-matching vehicle tracking data. In Proceedings of VLDB.
- Brandt T, Grawunder M (2018) GeoStreams: a survey. ACM Computing Surveys, 51(3), article no. 44.
- Brinkhoff T. (2002) A framework for generating network-based moving objects. Geoinformatica 9(1), pp. 153-180.
- Claramunt C, Ray C, Camossi E, Joussemme AL, Hadzagic M, Andrienko GL, Andrienko NV, Theodoridis Y, Vouros GA, Salmon L (2017) Maritime data integration and analysis: recent progress and research challenges. In Proceedings of EDBT.
- Frentzos E, Gratsias K, Theodoridis Y (2007) Towards the next generation of location-based services. In Proceedings of W2GIS.
- Frentzos E, Theodoridis Y, Papadopoulos AN (2009) Spatio-temporal trajectories. In: Liu L, Özsu TM (eds) Encyclopedia of Database Systems, pp. 2742-2746. Springer.
- Guo C, Ma Y, Yang B, Jensen CS, Kaul M (2012) EcoMark: evaluating models of vehicular environmental impact. In Proceedings of SIGSPATIAL/GIS.
- Hightower J, Borriello G (2001) Location systems for ubiquitous computing. IEEE Computer, 34(8), pp. 57-66.
- Parent C, Spaccapietra S, Renso C, Andrienko G, Andrienko N, Bogorny V, Damiani ML, Gkoulalas-Divanis A, Macedo JA, Pelekis N, Theodoridis Y, Yan Z (2013) Semantic trajectories modeling and analysis. ACM Computing Surveys, 45(4), Article no. 42.
- Pelekis N, Theodoridis Y (2014) Mobility data management and exploration. Springer.
- Ray C, Dreo R, Camossi E, Joussemme AL (2018) Heterogeneous integrated dataset for maritime intelligence, surveillance, and reconnaissance (Version 0.1) [Data set]. Zenodo. <http://doi.org/10.5281/zenodo.1167595>.
- Renso C, Baglioni M, de Macedo JAF, Trasarti R, Wachowicz M (2012) How you move reveals who you are: Understanding human behavior by analyzing trajectory data. Knowledge and Information Systems, 37(2), pp. 331-362.
- Renso C, Spaccapietra S, Zimányi E, eds. (2013) Mobility data: modeling, management, and understanding. Cambridge Univ. Press.
- Spaccapietra S Parent C (2011) Adding meaning to your steps. In Proceedings of ER.
- Spinsanti L, Celli F, Renso C (2010) Where you stop is who you are: Understanding peoples' activities. In Proceedings of BMI.
- Vouros GA, Vlachou A, Santipantakis GM, Doulkeridis C, Pelekis N, Georgiou HV, Theodoridis Y, Patroumpas K, Alevizos E, Artikis A, Claramunt C, Ray C, Scarlatti D, Fuchs G, Andrienko GL, Andrienko NV, Mock M, Camossi E, Joussemme AL, Cordero Garcia JM (2018) Big data analytics for time critical mobility forecasting: recent progress and research challenges. In Proceedings of EDBT.

- Yan Z, Chakraborty D, Parent C, Spaccapietra S, Aberer K (2011) SeMiTri: A Framework for Semantic Annotation of Heterogeneous Trajectories. In Proceedings of EDBT.
- Yan Z, Chakraborty D, Parent C, Spaccapietra S, Aberer K (2012) Semantic trajectories: Mobility data computation and annotation. *ACM Transactions on Intelligent Systems and Technology*, 9(4), Article no. 49.
- Zheng Y, Chen Y, Xie X, Ma WY (2010) Understanding transportation modes based on GPS data for Web applications. *ACM Transactions on the Web*, 4(1), Article no. 1.
- Zimmermann M, Kirste T, Spiliopoulou M, (2009) Finding stops in error-prone trajectories of moving objects with time-based clustering. In Proceedings of IMC.

### **b) Measuring data (dis-)similarity**

- Bollobas B, Das G, Gunopulos D, Mannila H (2001) Time-series similarity problems and well-separated geometric sets. *Nordic Journal of Computing*, 8(4), pp. 409-423.
- Buchin K, Buchin M, van Kreveld M, Luo J (2009) Finding long and similar parts of trajectories. In Proceedings of SIGSPATIAL-GIS.
- Chen L, Ng RT (2004) On the marriage of lp-norms and edit distance. In Proceedings of VLDB.
- Chen L, Ozsü MT, Oria V (2005) Robust and fast similarity search for moving object trajectories. In Proceedings of SIGMOD.
- Frentzos E, Gratsias K, Theodoridis Y (2007) Index-based most similar trajectory search. In Proceedings of ICDE.
- Lee JG, Han J, Whang KY (2007) Trajectory clustering: a partition-and-group framework. In Proceedings of SIGMOD.
- Nanni M, Pedreschi D (2006) Time-focused clustering of trajectories of moving objects. *Journal of Intelligent Information Systems*, 27(3), pp. 267–289.
- Pelekis N, Andrienko G, Andrienko N, Kopanakis I, Marketos G, Theodoridis Y (2012) Visually exploring movement data via similarity-based analysis. *Journal of Intelligent Information Systems*, 38(2), pp. 343-391.
- Tiakas E, Papadopoulos AN, Nanopoulos A, Manolopoulos Y, Stojanovic D, Djordjevic-Kajan S (2009) Searching for similar trajectories in spatial networks. *Journal of Systems and Software*, 82(5), pp. 772-788.
- van Kreveld M, Luo J (2009) The definition and computation of trajectory and subtrajectory similarity. In Proceedings of GIS.
- Vlachos M, Kollios G, Gunopulos D (2002) Discovering similar multidimensional trajectories. In Proceedings of ICDE.

## **2) Pre-processing your data**

### **a) Data curation (cleansing, transformation, simplification, enrichment, etc.)**

- de Boor C (1978) A practical guide to splines. Springer-Verlag.
- Buchin M, Driemel A, van Kreveld M, Sacristán V (2010) An algorithmic framework for segmenting trajectories based on spatio-temporal criteria. In Proceedings of GIS.
- Chen Y, Jiang K, Zheng Y, Li C, Yu N (2009) Trajectory simplification method for location-based social networking services. In Proceedings of LBSN.
- Douglas D, Peucker T (1973) Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *The Canadian Cartographer*, 10(2), pp. 112–122.
- Georgiou H (2017) 1-pass fixed-rate linear resampler in Matlab/Octave. <http://doi.org/10.5281/zenodo.232221>.
- Meratnia N, de By RA (2004) Spatiotemporal compression techniques for moving point objects. In Proceedings of EDBT.
- Panagiotakis C, Pelekis N, Kopanakis I, Ramasso E, Theodoridis Y (2012) Segmentation and sampling of moving object trajectories based on representativeness. *IEEE Transactions on Knowledge and Data Engineering*, 24(7), pp. 1328-1343.
- Patroumpas K, Artikis A, Katzouris N, Vodas M, Pelekis N, Theodoridis Y (2015): Event Recognition for Maritime Surveillance. In Proceedings of EDBT.
- Patroumpas K, Alevizos E, Artikis A, Vodas M, Pelekis N, Theodoridis Y (2017) Online event recognition from moving vessel trajectories. *GeoInformatica*, 21(2), pp. 389-427.
- Pelekis N, Kopanakis I, Panagiotakis C, Theodoridis Y (2010) Unsupervised trajectory sampling. In Proceedings of ECML-PKDD.
- Potamias M, Patroumpas K, Sellis TK (2006) Sampling trajectory streams with spatiotemporal criteria. In Proceedings of SSDBM.
- Quddus MA, Ochieng WY, Zhao L, Noland R (2003) A general map matching algorithm for transport telematics applications. *GPS Solutions*, 7(3), pp. 157-167.
- Quddus MA, Ochieng WY, Noland R (2007) Current map-matching algorithms for transport applications: state-of-the-art and future research directions. *Transportation Research Part C: Emerging Technologies*, 15(5), pp. 312-328.
- Santipatakis GM, Glenis A, Patroumpas K, Vlachou A, Doukeridis C, Vouros GA, Pelekis N, Theodoridis Y (2018) SPARTAN: Semantic integration of big spatio-temporal data from streaming and archival sources. *Future Generation Computer Systems*, to appear.

### **b) Data storage and searching**

- de Almeida VT, Güting RH, Behr T (2006) Querying moving objects in secondo. In Proceedings of MDM.
- Bakalov P, Hadjieleftheriou M, Keogh E, Tsotras V (2005) Efficient trajectory joins using symbolic representations. In Proceedings of MDM.
- Düntgen C, Behr T, Güting RH (2009) BerlinMOD: a benchmark for moving object databases. *The VLDB Journal*, 18(6), pp. 1335-1368.
- Frentzos E, Gratsias K, Theodoridis Y (2007) Index-based most similar trajectory search. In Proceedings of ICDE.
- Frentzos E, Gratsias K, Theodoridis Y (2009) On the effect of location uncertainty in spatial querying. *IEEE Transactions on Knowledge and Data Engineering*, 21(3), pp. 366-383.
- Gryllakis F, Pelekis N, Doulkeridis C, Sideridis S, Theodoridis Y (2017) Searching for spatio-temporal-keyword patterns in semantic trajectories. In Proceedings of IDA.
- Jiang J, Bao H, Chang EY, Li Y (2011) MOIST: a scalable and parallel moving object indexer with school tracking. *Proceedings of the VLDB Endowment*, 5(12), pp. 1838-1849.
- Lange R, Dürr F, Rothermel K (2011) Efficient real-time trajectory tracking. *The VLDB Journal*, 20(5), pp. 671-694.
- Mokbel MF, Aref WG (2005) PLACE: A scalable location-aware database server for spatio-temporal data streams. *IEEE Data Engineering Bulletin*, 28(3), pp. 3-10.
- Papadias D, Zhang J, Mamoulis N, Tao Y (2003) Query processing in spatial network databases. In Proceedings of VLDB.
- Pelekis N, Frentzos E, Giatrakos N, Theodoridis Y (2014) HERMES: A Trajectory DB Engine for mobility-centric applications. *International Journal of Knowledge-based Organizations*, 4(1).
- Pelekis N, Ntrigkogiannis C, Tampakis P, Sideridis S, Theodoridis Y (2013) Hermoupolis: a trajectory generator for simulating generalized mobility patterns. In Proceedings of ECML-PKDD.
- Pelekis N, Sideridis S, Tampakis P, Theodoridis Y (2016) Simulating our LifeSteps by example. *ACM Trans. Spatial Algorithms and Systems*, 2(3), pp. 11:1-11:39.
- Pfoser D, Jensen CS (1999) Capturing the uncertainty of moving-object representations. In Proceedings of SSD.
- Sakr MA, Güting RH (2011) Spatiotemporal pattern queries. *Geoinformatica*, 15(3), pp. 497-540.
- Sideridis S, Pelekis N, Theodoridis Y (2016) On querying and mining semantic-aware mobility timelines. *Journal of Data Science & Analytics*, 2(1), pp. 29-44.
- Tan H, Luo W, Ni LM (2012) CloST: a Hadoop-based storage system for big spatio-temporal data analytics. In Proceedings of CIKM.
- Theodoridis Y, Nascimento M (2000) Generating spatiotemporal datasets on the WWW. *ACM SIGMOD Record*, 29(3), pp. 39-43.

### 3) Analyzing your data

#### a) Cluster analysis and outlier detection; Group behavior

- Ankerst M, Breunig MM, Kriegel HP, Sander J (1999) OPTICS: Ordering points to identify the clustering structure. In Proceedings of SIGMOD.
- Cadez V, Gaffney S, Smyth P. (2000) A general probabilistic framework for clustering individuals and objects. In Proceedings of SIGKDD.
- Ester M, Kriegel HP, Sander J, Xu X (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In Proceedings of KDD.
- Gaffney S, Smyth P (1999) Trajectory clustering with mixtures of regression models. In Proceedings of KDD.
- Ge Y, Xiong H, Zhou ZH, Ozdemir H, Yu J, Lee KC (2010) TOP-EYE: Top-k evolving trajectory outlier detection. In Proceedings of CIKM.
- Gudmundsson J, van Kreveld MJ (2006) Computing longest duration flocks in trajectory data. In Proceedings of GIS.
- Gudmundsson J, van Kreveld MJ, Speckmann B (2007) Efficient detection of patterns in 2d trajectories of moving points. *Geoinformatica*, 11(2), pp. 195-215.
- Jensen CS, Lin D, Ooi BC (2007) Continuous clustering of moving objects. *IEEE Transactions on Knowledge and Data Engineering*, 19(9), pp. 1161-1174.
- Jeung H, Yiu ML, Zhou X, Jensen CS, Shen HT (2008) Discovery of convoys in trajectory databases. In Proceedings of VLDB.
- Kalnits P, Mamoulis N, Bakiras S (2005) On discovering moving clusters in spatio-temporal data. In Proceedings of SSTD.
- Lan R, Yu Y, Cao L, Song P, Wang Y (2017) Discovering evolving moving object groups from massive-scale trajectory streams. In Proceedings of MDM.
- Laube P, Imfeld S, Weibel R (2005) Discovering relative motion patterns in groups of moving point objects. *International Journal of Geographical Information Science*, 19(6), pp. 639-668.
- Lee JG, Han J, Whang KY (2007) Trajectory clustering: a partition-and-group framework. In Proceedings of SIGMOD.
- Lee JG, Han J, Li X (2008) Trajectory outlier detection: A partition-and-detect framework. In Proceedings of ICDE.
- Li Y, Han J, Yang J (2004) Clustering moving objects. In Proceedings of KDD.

- Li X, Han J, Kim S, Gonzalez H (2007) ROAM: Rule- and motif-based anomaly detection in massive moving object data sets. In Proceedings of SDM.
- Li Z, Ding B, Han J, Kays R (2010) Swarm: Mining relaxed temporal moving object clusters. Proceedings of VLDB, 3(1), pp. 723-734.
- Masciari E (2011) Trajectory outlier detection using an analytical approach. In Proceedings of ICTAI.
- Nanni M, Pedreschi D (2006) Time-focused clustering of trajectories of moving objects. Journal of Intelligent Information Systems, 27(3), pp. 267-289.
- Pelekis N, Kopanakis I, Kotsifakos E, Frenzos E, Theodoridis Y (2009) Clustering trajectories of moving objects in an uncertain world. In Proceedings of ICDM.
- Pelekis N, Kopanakis I, Kotsifakos E, Frenzos E, Theodoridis Y (2011) Clustering uncertain trajectories. Knowledge and Information Systems, 28(1), pp. 117-147.
- Pelekis N, Tampakis P, Vodas M, Panagiotakis C, Theodoridis Y (2017a) In-DBMS sampling-based sub-trajectory clustering. Proceedings of EDBT.
- Pelekis N, Tampakis P, Vodas M, Doulkeridis C, Theodoridis Y (2017b) On temporal-constrained sub-trajectory cluster analysis. Data Mining and Knowledge Discovery, 31(5), pp. 1294-1330.
- Rosén O, Medvedev A (2012) An on-line algorithm for anomaly detection in trajectory data. In Proceedings of ACC.
- Roy J (2008) Anomaly detection in the maritime domain. In Proceedings of SPIE - Optics and Photonics in Global Homeland Security IV.
- Spiliopoulou M, Ntoutsis I, Theodoridis Y, Schult R (2006) MONIC: Modeling and monitoring cluster transitions. In Proceedings of KDD.
- Wang Y, Lim E P, Hwang S Y (2006) Efficient mining of group patterns from user movement data. Data and Knowledge Engineering, 57(3), pp. 240-282.
- Yu Y, Cao L, Rundensteiner EA, Wang O (2014) Detecting moving object outliers in massive-scale trajectory streams. In Proceedings of KDD.

#### **b) Frequent route / location discovery**

- Cao H, Mamoulis N, Cheung DW (2005) Mining frequent spatio-temporal sequential patterns. In Proceedings of ICDM.
- Cao H, Mamoulis N, Cheung DW (2006) Discovery of collocation episodes in spatiotemporal data. In Proceedings of ICDM.
- Cao H, Mamoulis N, Cheung DW (2007) Discovery of periodic patterns in spatiotemporal sequences. IEEE Transactions on Knowledge and Data Engineering, 19(4), pp. 453-467.
- Chen Z, Shen HT, Zhou X (2011) Discovering popular routes from trajectories. In Proceedings of ICDE.
- Giannotti F, Nanni M, Pinelli F, Pedreschi D (2007) Trajectory pattern mining. In Proceedings of KDD.
- Giannotti F, Nanni M, Pedreschi D, Pinelli F, Renso C, Rinzivillo S, Trasarti R (2011) Unveiling the complexity of human mobility by querying and mining massive trajectory data. The VLDB Journal, 20(5), pp. 695-719.
- Li X, Han J, Lee JG, Gonzalez H (2007) Traffic density-based discovery of hot routes in road networks. In Proceedings of SSTD.
- Li Z, Ding B, Han J, Kays R, Nye P (2010). Mining periodic behaviors for moving objects. In Proceedings of KDD.
- Palma AT, Bogorny V, Kuijpers B, Alvares LO (2008) A clustering-based approach for discovering interesting places in trajectories. In Proceedings of ACM-SAC.
- Sacharidis D, Patroumpas K, Terrovitis M, Kantere V, Potamias M, Mouratidis K, Sellis T (2008) On-line discovery of hot motion paths. In Proceedings of EDBT.

#### **c) Classification and Prediction**

- Ashbrook D, Starner T (2003) Using GPS to learn significant locations and predict movement across multiple users, Personal Ubiquitous Computing, 7, pp. 275-286.
- Georgiou H, Karagiorgou S, Kontoulis Y, Pelekis P, Petrou P, Scarlatti D, Theodoridis Y (2018) Moving objects analytics: survey on future location & trajectory prediction methods. Technical Report. arXiv:1807.04639.
- Lee JG, Han J, Li X, Gonzalez H (2008b) TraClass: trajectory classification using hierarchical region-based and trajectory-based clustering. Proceedings of VLDB, 1(1), pp. 1081-1094.
- Li Y, Shahabi C (2018) A brief overview of machine learning methods for short-term traffic forecasting and future directions. SIGSPATIAL Special, 10(1), pp. 3-9.
- Monreale A, Pinelli F, Trasarti R, Giannotti F (2009) WhereNext: a location predictor on trajectory pattern mining. In Proceedings of KDD.
- Morzy M (2007) Mining frequent trajectories of moving objects for location prediction. In Proceedings of MLDM.
- Tao Y, Sun J, Papadias D (2003) Analysis of predictive spatio-temporal queries. ACM Transactions on Database Systems, 28(4), 295-336.
- Tao Y, Faloutsos C, Papadias D, Liu B (2004) Prediction and indexing of moving objects with unknown motion patterns. In Proceedings of SIGMOD.
- Trasarti R, Guidotti R, Monreale A, Giannotti F (2017) MyWay: location prediction via mobility profiling. Inf. Syst. 64, pp. 350-367.

Yavas G, Katsaros D, Ulusoy O, Manolopoulos Y (2005) A data mining approach for location prediction in mobile environments. *Data and Knowledge Engineering*, 54(2), pp. 121–146.