Analyzing Semantic Indoor Trajectories for understanding Museum visitors' movement (A-SITM)

Supervisors

- Claudia MARINICA (co-supervisor, 30%) Claudia.Marinica@univ-nantes.fr
- Fabien PICAROUGNE (co-supervisor, 30%) Fabien.Picarougne@univ-nantes.fr
- Fabrice GUILLET (thesis director, 40%) Fabrice.Guillet@univ-nantes.fr

Partners

This PhD subject is built on an international collaboration between DUKe research team of LS2N lab, Nantes, France, and the Museology Research Laboratory of the Department of Archives, Library Science and Museology at the Ionian University in Corfu, Greece, together with museums in Corfu, Greece, including the Corfu Museum of Asian Art, the Mon Repos Museum in Corfu, and the Corfu art Gallery. Visitors' data already collected and to be collected through LBS or systems from the aforementioned institutions comprise datasets to be analyzed.

Participants from LS2N-DUKE: *Claudia Marinica, Fabien Picarougne, Fabrice Guillet.*

Participant from Museology Research Laboratory, Department of Archives, Library Science and Museology, Ionian University in Corfu, Greece: *Georgios Papaioannou, Mat Damigos, Eleutherios Kalogeros.*

Participants from Museums in Corfu: *staff to be appointed once the project starts.*

Abstract. Numerous museums use location-based services (LBS) to track visitors' movement (e.g. the Louvre, British Museum, etc.) to better understand visitors. In addition to the localization information, LBS can provide contextual or semantic information about the visitors. Moreover, museums can provide supplementary data regarding the topological description of their indoor space. Thus, the main objective of this PhD is to develop new techniques for visitors' movement analysis through trajectory analysis approaches, where positioning information in the trajectories is enriched with topological and semantic information. To this end, this PhD works towards three challenges: (1) providing a standard for enriched trajectories based on Semantic Indoor Trajectory Model (SITM) that we proposed previously and IndoorGML standard, (2) developing descriptive analysis to extract patterns from SITM-based trajectory data, and (3) developing trajectory time series classification algorithms to provide movement predictions.

Description of the subject

1. Context and global objectives

Museums have been studying their visitors for decades to understand why visitors go to museums, what they do there, how they learn, and what their engagement and satisfaction may be. Indeed, in addition to their role of cultural mediation, museums desire to enhance visitors' experiences [1]. More precisely, a better understanding of their visitors might allow museums, first, to personalize their experiences, promote accessibility, and/or propose dynamic museum tours (as presented in our previous work in [2]). Second, studying visitors might encourage museums to take managerial decisions such as improved evacuation routes [2].

Nowadays, location-based services (LBS) allow to quite easily capture information on people or object movement. Numerous museums use them to track visitors' movement and visiting habits: e.g. the Metropolitan Museum of Arts - New York, the Smithsonian Institution - Washington D.C., the British Museum - London, the Louvre, and the Rijksmuseum - Amsterdam. In the Louvre, for example, special audio guides (e.g. the Nintendo 3DS for the Louvre Museum) and smartphone applications (e.g. the Louvre My Visit) are used. As this technology becomes more sophisticated, we can expect to see even more museums using this technology to improve the visitor experience. Moreover, these devices can provide, with the localization data, some contextual or semantic information about the visitor (visitor' profile, visitor' activity in the

application during the visit, etc.).

The main objective of this PhD subject is to develop new techniques for visitors' movement analysis, by building for each visitor its *trajectory* inside the museum. The originality of this work comes from *enhancing visitors' trajectories* with (1) *indoor space constraints* restraining the visitor's movement (e.g. position of doors, corridors, etc.), and (2) *contextual and/or semantic information* related to the museum or the visitor.

2. Scientific challenges

The PhD subject addresses three scientific challenges that will be explained hereafter.

Challenge 1. Towards a Standard for Semantic Indoor Trajectories based on SITM.

Data collection is intended to be done during the PhD which allows us to modify the collection protocol as needed to collect qualitative data. Three types of information should be collected for a museum: topological information on the museum's building (corridor, door, elevator, etc.), visitors' localization information, and contextual/semantic information related to visitors' movement (e.g. activity on the application, intention for the visit, etc.) and points of interest (e.g. nature and semantics associated with the museum works).

This challenge will benefit from the work we already completed in [3] where we proposed a new model for spatiotemporal indoor trajectories enriched with semantic annotations, called *Semantic Indoor Trajectory Model (SITM)*. SITM trajectory model is composed of:

- An *indoor space model*, based on the Multi-Layered Space Model of IndoorGML[1], under the form of a tree where on the vertical we have a *contains* relation and on the horizontal we have an *accessible* relation.
- A *formalism for a semantic trajectory* composed of traces:
  $T_{IDmo,\ tstart,\ tend} = (trace_{IDmo},\ t_{start},\ t_{end},\ A_{traj})$, where $trace_{IDmo,\ tstart,\ tend} = (e_k,\ v_k,\ t_{startk},\ t_{endk},\ A_k)$ $k \in [1,n]$
- A set of *semantic information* enriching the trajectories and the traces, represented as sets *A*.

Here the challenge is to express the collected data under the proposed formalism by providing, if needed, a series of adaptations. For example, in our previous work we limited the relations between two nodes to *contains* and *accessible*. The question here is if it is suited and possible to use other relations between nodes already integrated in IndoorGML. Moreover, we can see that every use case needs to adapt the existing SITM model. The presented PhD subject aims at proposing the SITM model as a *standard for Semantic Indoor Trajectories representation*, based on IndoorGML.

Challenge 2. Descriptive analysis of Semantic Indoor Trajectories (SITM)

For decades, museums have been interested in *understanding* their visitors. Thus, in this second challenge, we aim to develop a set of *descriptive analysis* over trajectory data represented by means of the semantic indoor trajectory model presented in [3]. Thus, we focus here on trajectory data mining techniques that met an important interest in the last decade [5], and the goal is to *extract trajectory patterns* describing the visitors' movement. The patterns would allow us to understand among which category/categories the visitors of the museum take place according to the form of their visit: ant, butterfly, grasshopper or fish as described in [4], or discover new categories.

The previous studies were interested in mining temporally annotated sequential patterns [6], in pattern extraction from semantic-enriched human movement [7] or in integrating indoor topology in the pattern mining process [8]. To our knowledge, none of the existing pattern mining approaches allow to mine patterns from trajectory data enhanced with specificities of the indoor space and with semantic information concerning the visitor and the visit.

Challenge 3. Predictive approaches for Semantic Indoor Trajectories (SITM)

While the previous challenge aims to help museums to enhance visitors' experience, this third

[1] http://www.indoorgml.net/

challenge aims to encourage the museums to take *managerial decisions* (such as deriving improved evacuation routes) by providing *movement predictions*.

To this end, we propose to formalize SITM trajectories as *trajectory time series* and to work towards developing *trajectory time series classification algorithms*. Our choice is motivated by the existence of multi-dimensional time series allowing us to represent the semantics and the topology information within SITM. In recent years, Deep Learning methods joined Time Series Classification (TSC) process [9] and InceptionTime algorithm [10] outperforms non-deep models for TSC.

For SITM trajectory time series classification, the difficulty is to find how to integrate the topology information and the semantic information in the deep learning algorithms. One solution that we consider is the use of Convolutional neural networks (CNNs) that are used for spatial and/or temporal data processing. A second solution consists in the attention mechanism (transformers) that is used for sequential data processing and enables the model to dynamically adjust the importance of certain elements from the input sequence to the output.

Your profile. We expect to welcome a candidate fulfilling the following requirements:

- Completion of an excellent master or diploma in Computer Science
- Strong programming skills and experience
- Background knowledge in the following areas are highly appreciated: data mining, deep learning, time series classification.
- Ability to develop methods and concepts.
- Willingness to contribute in interdisciplinary projects.
- Organizational and analytical skills.
- Ability to work in a team, problem-solving skills, and creative thinking.
- Excellent spoken and written communication skills in English.

We offer. The PhD will take place in the DUKe research team of LS2N lab, Nantes, France. The most of the DUKe team members work in the Polytech Nantes (Engineering School of the Nantes University), thus the PhD Student will have an office there. We are a dynamic research team composed of 20 permanent researchers and around 10 PhD students and postdocs.
Given the context of the international collaboration of the PhD subject, the PhD student will maybe have to make some research stays in the Museology Research Laboratory at the Ionian University in Corfu, Greece and also visit some of the Museums providing the data.

How to apply. Interested candidates can include in their applications the following documents:
- Curriculum vitae
- Letter of Motivation specific to the given position
- Abstract of master thesis
- Degree certificates for the last years
- At least two recommendation letters
- List of publications (if any)

Interested candidates should send their application to:
Assistant Professor Claudia Marinica (Claudia.Marinica@univ-nantes.fr)

Milestones

T+0 - T+2: Subject understanding and state of the art on the application domain: analysis of museum's visitors movement

T+2 - T+8: Understanding SITM model and state of the art on enriched trajectories modeling (based on existing state of the art)

State of the art on frequent pattern extraction from trajectories

State of the art on multivariate time series classification with deep learning methods, and trajectories classification with time series techniques

T+8 - T+12: Collecting and cleaning the visitor movement data from partner museums. Modeling by means of SITM formalism.
Proposal of a standard for indoor enriched trajectories
T+12 - T+18: Proposal of new approaches for descriptive analysis of SITM trajectories based on trajectory mining techniques
T+18 - T+20: Evaluation of the proposed approaches
T+20 - T+21: Publication of the results
T+21 - T+27: Proposal of new algorithms for the classification of the trajectories modeled as time series. The new algorithms will be based on deep learning techniques.
T+27 - T+29: Evaluation of the proposed algorithms.
T+29 - T+30: Publication of the results.
T+30 - T+36: Thesis writing.

References

[1] John H. Falk and Lynn D. Dierking. The Museum Experience Revisited. *Routledge*, 1st edition, 2013.
[2] Alexandros Kontarinis, **Claudia Marinica**, Dan Vodislav, Karine Zeitouni, Anne Krebs. Towards a Better Understanding of Museum Visitors' Behavior through Indoor Trajectory Analysis. *Seventh International Conference on Digital Presentation and Preservation of Cultural and Scientific Heritage (DiPP2017)*, Sep 2017, Burgas, Bulgaria. pp.19-30
[3] Kontarinis, A., Zeitouni, K., **Marinica, C**. et al. Towards a semantic indoor trajectory model: application to museum visits. *Geoinformatica* 25, 311–352 (2021).
[4] Lanir, J., Kuflik, T., Sheidin, J. et al. Visualizing museum visitors' behavior: Where do they go and what do they do there?. *Pers Ubiquit Comput* 21, 313–326 (2017).
[5] Mazimpaka, J.D.; Timpf, S. Trajectory data mining: A review of methods and applications. *J. Spat. Inf. Sci.* 2016, 2016, 61–99
[6] Fosca Giannotti, Mirco Nanni, Fabio Pinelli, and Dino Pedreschi. 2007. Trajectory pattern mining. *In Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '07)*. Association for Computing Machinery, New York, NY, USA, 330–339.
[7] Renso, C., Baglioni, M., de Macedo, J.A.F. et al. How you move reveals who you are: understanding human behavior by analyzing trajectory data. *Knowl Inf Syst* 37, 331–362 (2013).
[8] L. Radaelli, D. Sabonis, H. Lu and C. S. Jensen, Identifying Typical Movements among Indoor Objects - Concepts and Empirical Study, 2013 *IEEE 14th International Conference on Mobile Data Management*, Milan, Italy, 2013, pp. 197-206.
[9] Ismail Fawaz, H., Forestier, G., Weber, J. et al. Deep learning for time series classification: a review. *Data Min Knowl Disc* 33, 917–963 (2019).
[10] Ismail Fawaz, Hassan, et al. "InceptionTime: Finding alexnet for time series classification." *Data Mining and Knowledge Discovery* 34.6 (2020): 1936-1962.