Towards Engineering Drones' Semantic Trajectories as Knowledge Graphs

Efthymia Moraitou[a], Sotiris Angelis[a], Konstantinos Kotis[a], George Caridakis[a], Ermioni - Eirini Papadopoulou[b], Nikolaos Soulakellis[b]

a. *Intelligent Systems Lab, Dept. of Cultural Technology and Communication, University of the Aegean, Mytilene, 81100 Greece*
b. *Cartography and Geoinformation Lab, Dept. of Geography, University of the Aegean, Mytilene, 81100 Greece*

Abstract

The information related to the movement of vehicles can be enriched with data beyond latitude, longitude, and timestamp, enhanced with complementary segmentations, constituting what is called a semantic trajectory. Semantic Web (SW) technologies have already been used for the modeling and enrichment of semantic trajectories. Our work-in-progress focuses on the engineering of semantic trajectories of drones as knowledge graphs (KG). Particularly, the work is motivated by a use case that focuses on UAV (Unmanned Aerial Vehicles) drones with a mission to document specific regions/points of interest (a petrified forest in a GeoPark). This research work aims to develop a) a methodology for the engineering of semantic trajectories as KGs (STaKG), b) a toolset for the management of KG-based semantic trajectories and c) a repository of semantically annotated GIS recording missions and the corresponding produced documentation records. In this paper, we present work-in-progress related to a) the STaKG engineering methodology, b) the STaKG management toolset for supporting the methodology, and c) the semantic model for representing knowledge related to drones' semantic trajectories and the related documentation recordings.

Keywords[1]

Geoinformatics, drone, knowledge graph, semantics, trajectory

1. Introduction

Today, Geospatial Linked Data (GLD) is vital for emerging research and development areas such as autonomous/unmanned aerial vehicles and (UAV) related services, e.g., delivery, surveillance, and documentation. The next generation of spatial knowledge graphs (KGs) will integrate numerous spatial and general datasets, such as weather data, points and regions of interest (POI/ROI). GLD and KGs principles and tools could contribute to the building of next-generation spatial data applications, facilitating the processing and management of data related to moving objects' trajectories.

The segments of an object's movement track, which have been defined based on the interest that they present for some application (e.g., a drone's movement in an area for a given recording mission), are called trajectories of the moving object [1]. A trajectory can be enriched with additional data (beyond latitude, longitude, and timestamp information), and/or enhanced with several complementary segmentations, constituting a semantic trajectory (ST) [2]. In terms of deployment, a ST may be useful for applications that require the interpretation of the trajectory of a moving object (e.g., points or regions of interest that a drone has documented).

GeoLD2022 - ESWC2022, May 29 - June 2, 2022, Hersonissos, Greece
EMAIL: e.moraitou@aegean.gr (A.1); sotiris@aegean.gr (A.2); kotis@aegean.gr (A.3); gcari@aegean.gr (A.4); epapa@geo.aegean.gr (A.5), nsoul@aegean.gr (A.6)
ORCID: 0000-0001-9384-1105 (A.1); 000-0002-4588-0713 (A.2); 0000-0001-7838-9691 (A.3); 0000-0001-9884-935X (A.4); 0000-0003-4122-8450 (A.5); 0000-0002-5328-2342 (A.6)

CEUR Workshop Proceedings (CEUR-WS.org) Proceedings

Semantic Web (SW) technologies have been used for the modeling and enrichment of STs, since they facilitate the modeling and interlinking of data that could enhance a trajectory of raw movement data, as well as the segmentation of the trajectories themselves, based on semantic data, in a standardized and meaningful way [2, 3]. In this direction, Santipatakis et al. [2] proposed the datAcron ontology for representing STs at varying levels of spatiotemporal analysis. Mobility analysis tasks are based on a wealth of disparate and heterogeneous sources of information that need to be integrated. Additionally, Gao et al. [3] proposed a representation of STs that considers domain knowledge, in addition to spatiotemporal data, to achieve improved retrieval of STs. The proposed approach measures similarities among vectors and emphasizes on the context of trajectories to extract semantic relations among target objects.

KGs incorporate semantic models (in many cases they can be considered as populated ontologies in the form of directed graphs) utilized for the structured and formal representation of heterogeneous data, as well as for reasoning with multiple integrated views of it [4, 5]. Therefore, they could be exploited for the representation and processing of STs.

Our research is motivated by use cases related to drones' mission of documenting (with photos recording events) specific POIs and ROIs, such as the GeoPark of petrified forest in Lesvos Island. Particularly, we aim to develop a KG-based approach for transforming trajectories of drones (usually operating in a swarm) - and particularly, UAV drones - into STs that can be effectively managed, visualized, and analyzed. The main objectives of the approach are to facilitate a) the modeling of STs of drones and swarm of drones, their flights and recordings per mission (e.g., volume and frequency of recording episodes), b) the visualization and analysis of STs, c) the retrieval of semantic information of flights/missions (e.g., drone position, recording position, episodes' date/time, weather data), and, d) the retrieval of records (e.g., photos) which have been produced during different recording events of trajectories related to a flight/mission, based on parameters such as the type or location of recording events (e.g., nearby recording positions, photograph recording, the object of interest that has been recorded, etc.).

Based on the aforementioned objectives, this research aims to contribute a) a methodology for the engineering of drones' STs as KGs (STaKG), b) an integrated toolset for the management of KG-based drones' STs, and c) a repository of semantically annotated GIS recording missions and the corresponding produced documentation records. Specifically, in this paper, we present work-in-progress related to a) the proposed STaKG methodology, b) the under-development STaKG management toolset for supporting the methodology, and c) the developed semantic model for representing knowledge related to drones' STs and the related documentation recordings.

The remainder of the paper is structured as follows: Section 2 presents the proposed methodology for engineering drones' STs as KGs. In Section 3 the architecture and technological choices for the toolset that supports the proposed methodology are presented. In Section 4, the developed semantic model and the expected results of the work-in-progress implementations are discussed. The paper concludes with a discussion summarizing the research work conducted so far.

2. Engineering drones' semantic trajectories as knowledge graphs

Ontologies constitute the backbone of KGs since they provide the formal and explicit semantics that KGs need for the effective modeling of Linked Data (LD) on the Web of Data. Ontology engineering methodologies (OEMs) define specific methodological phases, processes, and tasks for the engineering of ontologies, including feasibility analysis, identification of goals, requirements specification, implementation, evaluation, and maintenance. Those steps present - to some extent - an analogy to KG building steps. As suggested in related work [6], the ontology and the KG that is built on top of it, can both be developed following the general principles and similar/analogous tasks and steps of an OEM (e.g., DILIGENT [7], HCOME [8]). In this direction, the proposed hybrid (human-center/top-down and data-driven/bottom-up) methodology, namely Semantic Trajectories as Knowledge Graphs (STaKG) methodology, borrows and adapts principles and tasks of the collaborative and iterative OEM HCOME [8], and merges with principles from KG engineering approaches [9, 10].

Figure 1. Phases and processes of STaKG methodology.

The three phases of the STaKG methodology are briefly described below (and depicted in Figure 1):

- The first phase, namely Specification, includes the specification of the involved stakeholders of the engineering team (who is doing what), as well as the aim, scope, and requirements in terms of data, semantic annotations, segmentations of the trajectories, and the model that will capture the required knowledge.
- The second phase, namely Development, includes the creation of the explicit knowledge related to the STaKGs, i.e., the extension and specialization of reused ST models (e.g., an extension of existing ST ontology) based on the requirements of the first phase. It also includes the creation of instance data, i.e., spatiotemporal, and contextual data about the recorded trajectories. In the same phase, storage, publishing, retrieval, and visualization of the STaKGs are included.
- The third phase, namely Evaluation and exploitation, includes a) the evaluation of the quality of the modeled STaKGs, in terms of correctness, completeness, and bias, b) the cleaning and enrichment of the STaKGs. Enrichment refers to the discovery and linking to additional/external knowledge sources (e.g., from the Web). In the same phase, deployment and maintenance procedures are included.

3. The STaKG management toolset

To support STaKG methodology with an engineering environment, we have designed a management toolset based on state-of-the-art technology for LD and KGs. Its interconnected components exchange data through a pipeline process (see Figure 2) that involves a) preprocessing of position/movement data (data cleaning, data compression), b) the enrichment of raw trajectories for the engineering of STs (semantic segmentation, semantic annotation, utilization of application domain and geographical data, linking to external data sources), c) conversion of ST to KGs (ST management, retrieval), d) analysis of STaKG (classification, clustering, aggregation, comparison of STaKGs). The analysis may result in the discovery of previously unknown behaviors of moving entities where there is no a priori knowledge for them, or behavior detection/reasoning, which refers to the recognition of an already known moving behavior of a moving entity.

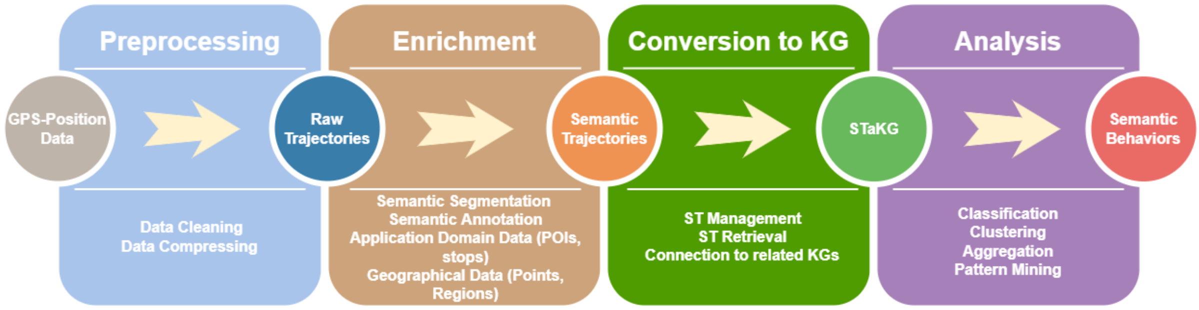

Figure 2. STaKG pipeline processing

The high-level architecture of the designed toolset includes: a) a tool for raw trajectory data cleaning and RDFization based on automated/semi-automated mapping to related semantic models, utilizing specialized tools (OpenRefine [11], Karma [12]), b) a tool for trajectory data summarization, trajectories' enrichment with Linked Open Data by performing linking tasks to external data sources, recording metadata, weather data, and structured data of POI/ROI shapefiles, c) a tool for ST management (split, merge, combine, analyze), and d) a web-based tool for ST browsing and visualization. The tools described in (b), (c), and (d) will be developed using the GRAND [13] technology stack which includes GraphQL, React.js, Apollo, Node.js, and Neo4j. Furthermore, a graph database, namely Neo4j [14], supports the web-based tool, and stores the managed data (STs, GIS recording missions, and produced records). Especially for RDF store technology, although noteworthy alternatives exist, specialized in spatiotemporal RDF data storing, such as Strabon RDF store [15], Neo4j was selected due to its integration in the GRAND technological stack, and due to the integrated graph analytics solutions, that it provides. Although Neo4j is not a native spatial database, it includes data types for geospatial and temporal values and provides spatial and temporal functions as well as spatial indexing and complex polygon representations.

A high-level architectural design of the interconnected tools of the STaKG toolset, and the related exchanged data, is depicted in Figure 3.

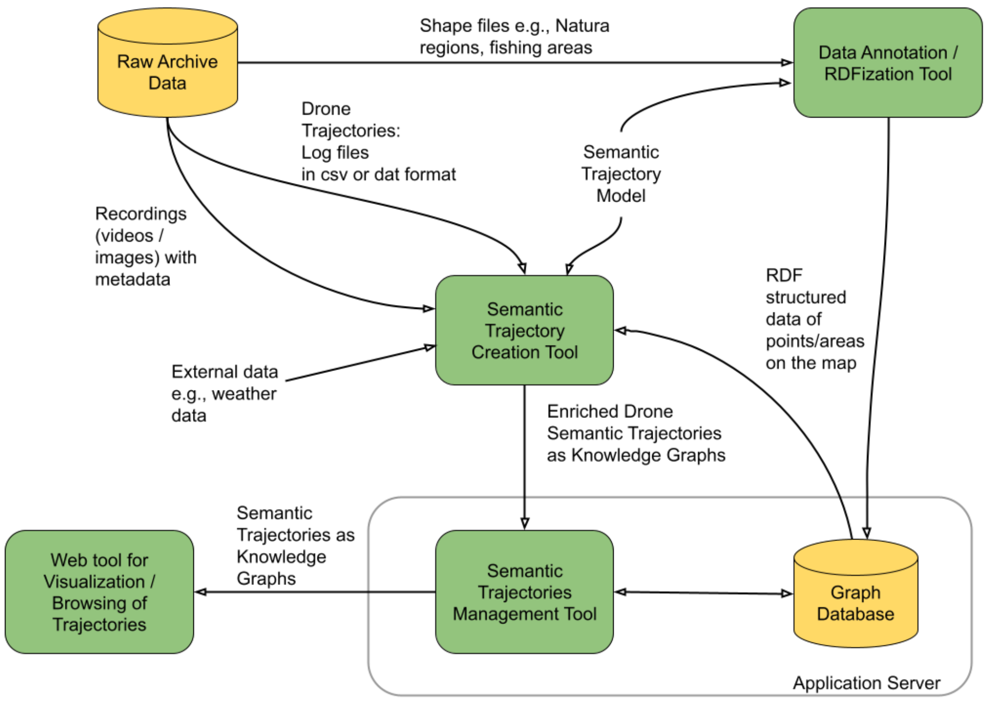

Figure 3. STaKG toolset high-level architecture and data exchange.

4. Implementation and expected results

The data and knowledge that the STaKG model aims to integrate, considering the motivation use-cases, includes a) flight data, derived from flight log files which are the records of a flight, automatically generated by a drone (usually in CSV format), b) equipment data, reported by the flight operator, describing the characteristics of a drone (e.g., model, serial number), c) recording/mission data, reported by the flight operator in the context of the mission planning procedure (e.g., the purpose of the mission), d) records data (aerial and terrestrial), provided either by exif (Exchangeable Image File Format) files of the records (photos, videos, lidar data) or directly from the acquired or processed records, e) geographic names and elements, about the POI/ROIs (e.g., location of the recorded object), f) weather data, data (e.g., temperature) recorded by weather monitoring devices or services.

For the development of the semantic model that is required to represent knowledge related to StaKGs, existing related ontologies have been studied. The datAcron ontology has been selected for reusing the main conceptualization of a ST in the aviation domain. At this stage, a first version of the semantic model (Onto4drone [16]) has been developed (version 1.0.0) and it is available in OWL. It is directly based on the datAcron ontology, and indirectly on the DUL [17], SKOS [18], SOSA/SSN [19], SF [20, 21], GML [22], and GeoSPARQL [23] ontologies. The model was developed following the HCOME collaborative engineering methodology, supported by Protégé 5.5 (for personal space engineering), and WebProtégé (for shared space engineering) tools respectively. In addition, Google Docs and Meet have been used for further collaborative engineering tasks. The ontology has been populated by individuals in order to be evaluated at this initial stage, while SPARQL queries have been executed for evaluation purposes [16] (aligned to the competency questions specified in the Specification phase of the STaKG methodology).

5. Conclusion and future work

This paper presents a KG-based approach for transforming trajectories of drones into STs managed by an integrated toolset. Particularly, it presents STaKG engineering methodology, a semantic model for representing STaKGs, and the architecture and implementation choices of a management toolset (its implementation is a work-in-progress) based on state-of-the-art technology for LD and KGs. The engineered STaKGs using the proposed methodology, model, and toolset, are expected to constitute a Geospatial LD knowledge-base available for a) utilization by drone-related applications, and b) the deployment of related services which will facilitate the work of experts/stakeholders in the Geoinformatics domain. First and foremost, the STaKG knowledge-base will be exploited for advanced map-based visualization of trajectories, flights, missions, recording events, timelines, and records, e.g., a geographic map that visualizes different flights and individual photographic records of a drone in the form of a ST, along with the related data-recording episodes recorded during a specific mission. Additionally, the STaKG knowledge-base will be exploited for management and analytics tasks. Such tasks include the merging of two or more STaKGs that are related to the same recording mission, the splitting, and the refinement of a STaKG to specific episodes e.g., splitting the recording episodes of the moving trajectory of a drone to sub-episodes of camera-shooting position set at up-shooting-departure for the next shooting position.

Acknowledgments

This research was funded by the Research e-Infrastructure [e-Aegean R&D Network], Code Number MIS 5046494, which is implemented within the framework of the “Regional Excellence” Action of the Operational Program “Competitiveness, Entrepreneurship and Innovation”. The action was co-funded by the European Regional Development Fund (ERDF) and the Greek State [Partnership and Cooperation Agreement 2014–2020].

References

[1] C. Parent, S. Spaccapietra, C. Renso, G. Andrienko, N. Andrienko, V. Bogorny, M. Damiani, A. Gkoulalas-Divanis, J. Macedo, N. Pelekis, Y. Theodoridis, and Z. Yan, "Semantic trajectories modeling and analysis," *ACM Computing Surveys (CSUR)*, vol. 45, no. 4, Aug. 2013, doi: 10.1145/2501654.2501656.

[2] G. M. Santipantakis, G. A. Vouros, A. Glenis, C. Doulkeridis, and A. Vlachou, "The datAcron Ontology for Semantic Trajectories," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10577 LNCS, pp. 26–30, May 2017, doi: 10.1007/978-3-319-70407-4_6.

[3] Q. Gao, F. Zhou, K. Zhang, F. Zhang, and G. Trajcevski, "Adversarial Human Trajectory Learning for Trip Recommendation," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–13, 2021, doi: 10.1109/TNNLS.2021.3058102.

[4] A. Hogan, E. Blomqvist, M. Cochez, C. D'Amato, G. de Melo, C. Gutierrez, S. Kirrane, J. Gayo, R. Navigli, S. Neumaier, A. Cyrille N. Ngomo, A. Polleres, S. Rashid, A. Rula, L. Schmelzeisen, J. Sequeda, S. Staab, and A. Zimmermann, "Knowledge Graphs," *ACM Computing Surveys (CSUR)*, vol. 54, no. 4, Jul. 2021, doi: 10.1145/3447772.

[5] L. Ehrlinger and W. Wöß, "Towards a definition of knowledge graphs," in *CEUR Workshop Proceedings*, 2016, vol. 1695.

[6] V. A. Carriero, A. Gangemi, M. L. Mancinelli, A. G. Nuzzolese, V. Presutti, and C. Veninata, "Pattern-based design applied to cultural heritage knowledge graphs," *Semantic Web*, vol. 12, no. 2, 2021, doi: 10.3233/SW-200422.

[7] H. S. Pinto, S. Staab, and C. Tempich, "DILIGENT: Towards a fine-grained methodology for DIstributed, Loosely-controlled and evolvInG Engineering of oNTologies," in *Frontiers in Artificial Intelligence and Applications*, 2004, vol. 110.

[8] K. Kotis and G. A. Vouros, "Human-centered ontology engineering: The HCOME methodology," *Knowledge and Information Systems*, vol. 10, no. 1, 2006, doi: 10.1007/s10115-005-0227-4.

[9] D. Fensel, Fensel, Dieter, U. Şimşek, K. Angele, E. Huaman, E. Kärle, O. Panasiuk, I. Toma, J. Umbrich, and A. Wahler, "Introduction: What Is a Knowledge Graph?," *Knowledge Graphs*, pp. 1–10, 2020, doi: 10.1007/978-3-030-37439-6_1.

[10] "How-to: Building Knowledge Graphs in 10 Steps | Ontotext Fundamentals." https://www.ontotext.com/knowledgehub/fundamentals/how-to-building-knowledge-graphs-in-10-steps/.

[11] "OpenRefine." https://openrefine.org/.

[12] "Karma: A Data Integration Tool." https://usc-isi-i2.github.io/karma/.

[13] "Build Fullstack GraphQL Applications With Ease | GRANDstack." https://grandstack.io/.

[14] "Graph Data Platform | Graph Database Management System | Neo4j." https://neo4j.com/.

[15] K. Kyzirakos, M. Karpathiotakis, and M. Koubarakis, "Strabon: A Semantic Geospatial DBMS", [Online]. Available: http://www.strabon.di.uoa.gr

[16] "GitHub - KotisK/onto4drone: An ontology for representing knowledge related to drones and their semantic trajectories." https://github.com/KotisK/onto4drone.

[17] "DOLCE+DnS Ultralite ontology." http://www.ontologydesignpatterns.org/ont/dul/DUL.owl.

[18] "SKOS Simple Knowledge Organization System Reference." [Online]. Available: https://www.w3.org/TR/skos-reference/

[19] "Semantic Sensor Network Ontology." https://www.w3.org/ns/ssn/

[20] "Simple features geometries." http://schemas.opengis.net/sf/1.0/simple_features_geometries.rdf.

[21] "Simple Feature Access - Part 1: Common Architecture | OGC." https://www.ogc.org/standards/sfa.

[22] "GML geometries." http://schemas.opengis.net/gml/3.2.1/gml_32_geometries.rdf.

[23] "OGC GeoSPARQL 1.0." http://schemas.opengis.net/geosparql/1.0/geosparql_vocab_all.rdf.