An ontology-based approach for handling explicit and implicit knowledge over trajectories

Rouaa Wannous[1], Cécile Vincent[2], Jamal Malki[1], and Alain Bouju[1]

[1] Laboratoire L3i, Pôle Sciences & Technologie
Avenue Michel Crépeau, 17042 La Rochelle, France
[2] Centre d'Etudes Biologiques de Chizé
CNRS UMR 7372, 79360 Villiers-en-Bois, France

Abstract. The current information systems manage several, different and huge databases. The data can be temporal, spatial and other application domains with specific knowledge. For these reasons, new approaches must be designed to fully exploit data expressiveness and heterogeneity taking into account application's needs. As part of ontology-based information system design, this paper proposes an ontology modeling approach for trajectories of moving objects. Consider domain, temporal and spatial knowledge gives a complexity to our system. We propose optimizations to annotate data with these knowledge.

Keywords: Trajectory data model, Spatial data model, Time data model, Ontology inference, Domain Spatial Temporal knowledge, Ontology rules.

1 Introduction

At present, most information systems are facing the problem of analysis, interpretation and querying huge masses of data. In these systems, knowledge plays a central role in such processes. This knowledge can be expressed in form of explicit and implicit semantics. An important part of the explicit semantics comes from the application domain. Also, we need to consider intrinsic data properties. Otherwise, there are others explicit semantics resulting from the nature of the data. Furthermore, users needs or requirements also form another part of explicit semantics that must be taken into account. These different forms of knowledge are involved at different levels of data modeling. In general, any approach considering this issue, should provide answers to three issues: data models; explicit semantics models; connecting explicit semantics and data models.

This article presents an ontological based approach to solve this problem by respecting the previous three requirements. The central idea of our approach is data and semantics models transformation into ontologies, considering two steps:
- declarative mapping: concepts and relationships of the resulting ontologies are mapped to a higher and generic predefined ontologies;
- imperative mapping: user defined rules on generic predefined ontologies are defined to decide semantics of new created data.

In this case, input models can be formalized and transformed independently from the generic ontologies which can be extended to accomplish the declarative mapping. The imperative mapping leads to implicit semantics and then to new knowledge.

Application domain considered in this work is trajectories. So, we present a trajectory modeling framework to enrich a high-level data layer. To meet challenges imposed by a semantic trajectory notion, we present a modeling approach based on a generic trajectory ontology centered on a triple (object, trajectory, activity). Considering this workspace, we show how we can use our modeling approach in a particular domain application: marine mammal trajectories. This paper provides solutions to several scientific problems to be considered in a trajectory's modeling approach based on ontologies taking into account spatial, temporal and domain knowledge.

2 Related work

Data management approaches including modeling, indexing, inferencing and querying large data have been actively investigated during the last decade [3,11]. Most of these techniques are only interested in representing and querying moving object trajectories [15,14,4]. The problem of using implicit or explicit knowledge associated with these data, or the domain context is not considered. In [12], authors propose a conceptual view on trajectories. In this approach, a trajectory is a set of stops, moves. Each part contains a set of semantic data or knowledge. Based on this conceptual model, several studies have been proposed [3,15]. In [3], authors proposed a trajectory data preprocessing method to integrate trajectories with spatial knowledge. Their application concerned daily trips of employees from home to work and back. However, this work is limited to the formal definition of semantic trajectories with spatial and time knowledge without any implementation and evaluation. In [15], authors proposed a trajectory computing platform which exploits a spatio-semantic trajectory model. One of the layers of this platform is a data preprocessing layer which cleanses the raw GPS feed, in terms of preliminary tasks such as outliers removal and regression-based smoothing. Based on a space-time ontology and events approach, in [5] authors proposed a generic meta-model for trajectories to allow independent applications. They processed trajectories data benefit from a high level of interoperability, information sharing. Their approach is inspired by ontologies, however the proposed resulting system is a pure database approach. This work elaborated a meta-model to represent moving objects using an ontology mapping for locations. In extracting information from the instantiated model, this work seems to rely on a pure SQL-based approach not on semantic queries.

3 Modeling approach

3.1 Design and methodology

Our work is based on moving objects trajectories. This requires a trajectory data model and a moving object model. Moreover, to enrich data with knowledge, a semantic model should be taken into consideration. Therefore, we need a generic model to consider the trajectory, moving object and semantic models simultaneously. This is represented by a semantic trajectory model shown in Figure 1. This model can consume captured data of trajectories and other external data as shown in Figure 1 link (1). These data are related to an application domain. This requires an application domain trajectory model

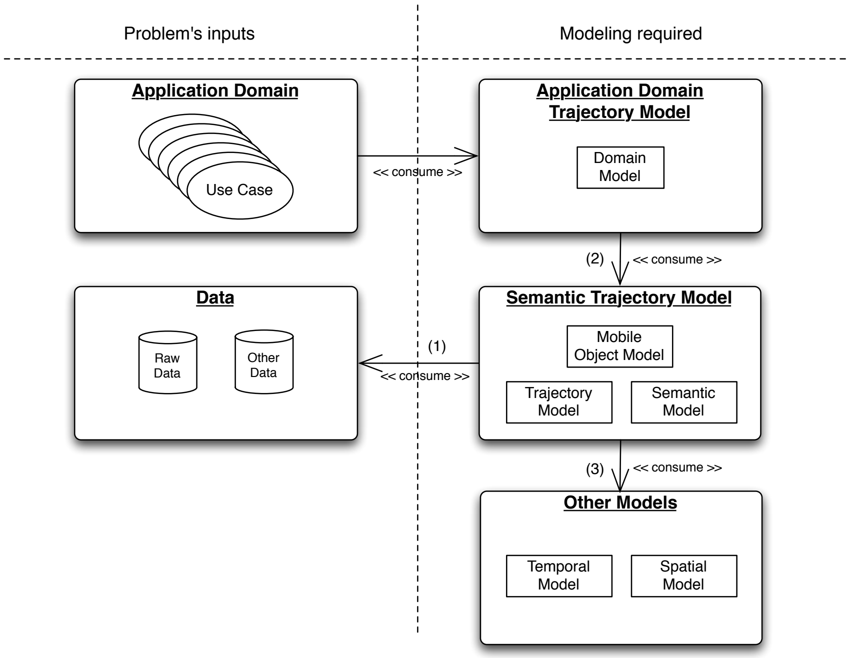

Fig. 1. Problem and its modeling required

which consists of domain model, as shown in Figure 1 link (2). The latter will support semantics related to users' needs. In the domain model, we also find the necessary semantics related to the real moving object, its trajectories, its activities and others. This semantics is often designed by a domain expert. In general, considering various facets of data, the semantic trajectory model must be extended by other models: application domain, temporal and spatial models. Then, the main issue is to build and design the semantic trajectory model with its required components.

The semantic trajectory modeling approach is tightly related to the problem of a semantic gap between this model and raw data. Link (1) in Figure 1 presents this gap. Moreover, our approach involves multiple models and then must establish semantic mappings among them, to ensure interoperability. In Figure 1, links (2) and (3) match the domain, temporal and spatial models with the semantic trajectory model. This matching extends the capabilities of our approach. For more efficient semantic capabilities, we want to annotate the data with domain, temporal and spatial knowledge. These knowledge are defined by experts representing users' needs. Annotating data with these knowledge could be done automatically or manually. We cannot use a manual annotation over huge data. Therefore, we choose an automatic annotation which can be accomplished by an ontology inference mechanism. This inference mechanism derives new semantics from existing information using additional knowledge. Later in this paper, we will present this inference mechanism as sets of rules.

3.2 Semantic trajectory ontology

In [13], we proposed a methodology for modeling trajectory data. This methodology focused on several real cases. For each case, we define a context, data capture, an analysis process of these data, and a domain model. From these models, we define a trajectory pattern also called generic trajectory model, Figure. 2. A trajectory is a set of sequences of spatio-temporal path covered by a moving object and has an activity.

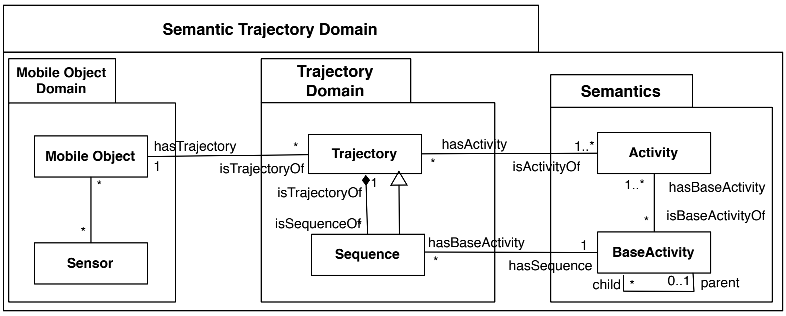

Fig. 2. Semantic trajectory modeling approach

The declarative part of our semantic trajectory ontology is presented in Figure. 3, named `owlSemanticTrajectory`. This ontology contains three models: mobile object, trajectory and semantic. By definition, a trajectory is a set of spatio-temporal concepts. Spatial and temporal models can be reused to enrich description of the concepts in the trajectory ontology to represent their spatial and temporal localization.

3.3 Reusing time ontology

The requirements of an ontology of time highlight the temporal concepts: `instant` and `interval`. The identification of temporal relationships leads to consider Allen temporal algebra [2]. In our approach, we consider an ontology of temporal concepts named `OWL-Time` [3] [8] developed by the World Wide Web Consortium (W3C).

`OWL-Time` ontology has a precise specification of temporal concepts and relationships as defined in the theory of Allen, formalized in OWL. An extract of the declarative part of this ontology is shown in Figure. 4 and described in detail in [9].

3.4 Reusing spatial ontology

The requirements of spatial ontology highlight spatial concepts, such as `point`, `line` and `polygon` concepts and others. The identification of spatial relationships leads to consider spatial relationships such as: `Equals, Within, Touches, Disjoint, Intersects, Crosses, Contains` and `Overlaps`.

[3] `http://www.w3.org/2006/time`

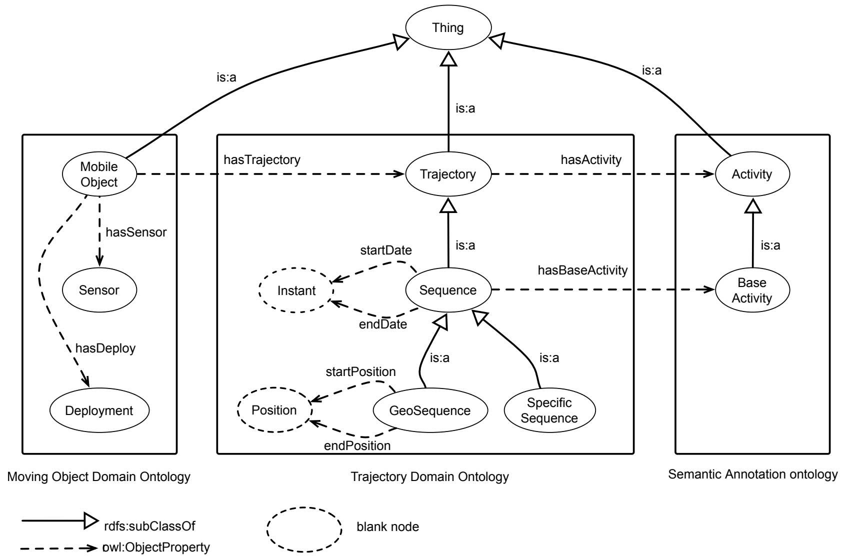

Fig. 3. A view of the semantic trajectory ontology `owlSemanticTrajectory`

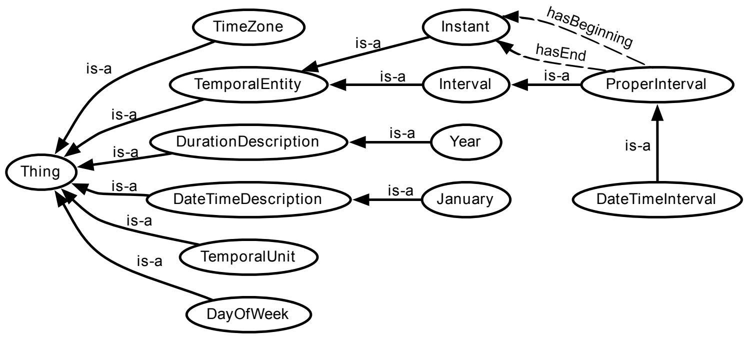

Fig. 4. A view of time ontology `OWL-Time`

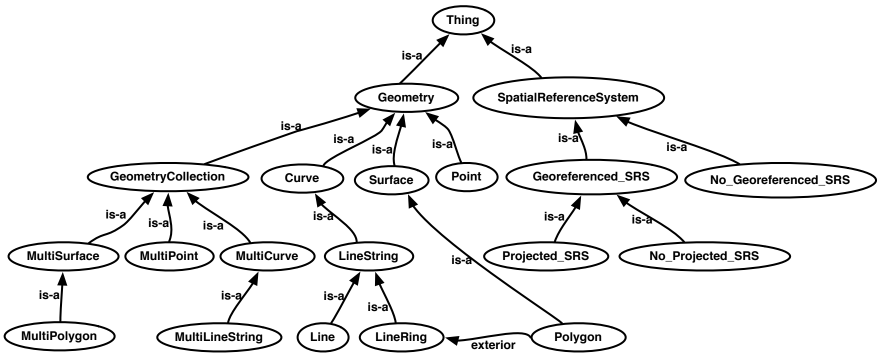

Fig. 5. A view of spatial ontology `owlOGCSpatial`

The standard `OGC OpenGIS` [10] presents spatial objects and functions over these objects. This standard contains a precise definition of spatial classes and reference systems. We transform this model into a formal ontology called `owlOGCSpatial` ontology. Figure 5 presents an extract of this ontology.

4 Trajectory ontology inference

Inference is the ability to make logical deductions based on ontologies, and optionally individuals. It derives new knowledge based on rules. A rule's definition, Figure. 6, has an antecedent, filters and a consequent. If knowledge are represented using RDF triples, then the antecedent is a set of triples, filters apply restrictions, and finally consequent is a new derived triple.

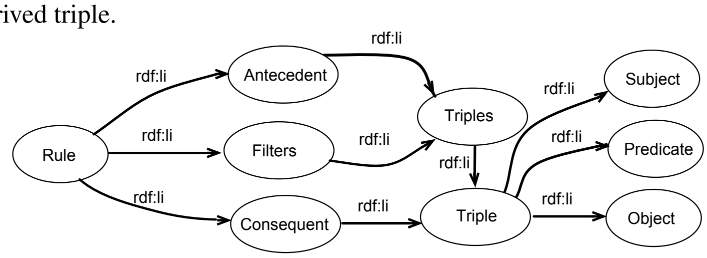

Fig. 6. Rule's definition

Inference using standard rules . Our trajectory ontology `owlSemanticTrajectory` is based on RDF, RDFS and OWL constructs. Inference mechanism associates with each construct a rule. The results sets are called standard rules. An example of standard rules is `OWLPrime` in Oracle RDF triple store [1].

Inference using temporal rules . Our trajectory ontology `owlSemanticTrajectory` uses temporal relationships as defined by Allen's algebra [2]. Each relationship is defined as a rule, `intervalAfter, intervalBefore, intervalDuring`, etc.

Inference using spatial rules . Our ontology `owlSemanticTrajectory` uses spatial relationships as defined by the Dimensionally Extended Nine-Intersection Model (DE-9IM) [6,7]. Each relation is defined as a rule, `Contains, Overlaps`, etc.

5 Trajectory ontology inference using domain rules

Our application domain is seals' trajectories, where a seal is considered as a mobile object. The captured data comes from the LIENSs laboratory[4] in collaboration with SMRU[5]. We consider three main states of a seal : `Dive, Haulout` and `Cruise`. Every state is related to a seal's activity, like `Resting, Traveling` and `Foraging`.

[4] Lab. CNRS/University of La Rochelle - `http://lienss.univ-larochelle.fr`
[5] SMRU: Sea Mammal Research Unit - `http://www.smru.st-and.ac.uk`

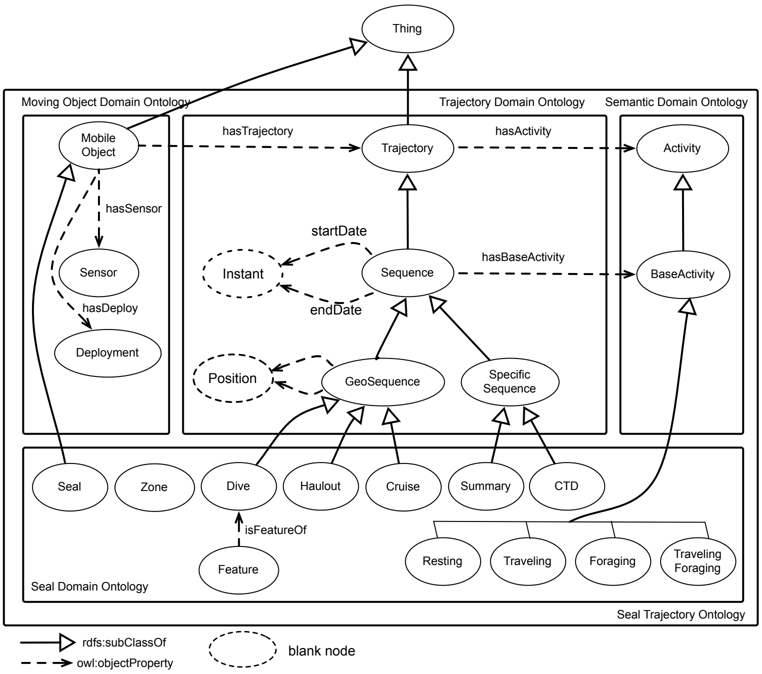

Fig. 7. Overview of the seal trajectory ontology with their activities

The captured data can also contain some meta-data `CTD` (Conductivity-Temperature-Depth) about the marine environment such as water conductivity, temperature and pressure. Starting from our trajectory ontology `owlSemanticTrajectory`, we define the seal trajectory ontology, named `owlSealTrajectory`, Figure 7. Formally, each activity is declared in the ontology and associated to a domain rule.

6 Implementation

Our implementation framework uses Oracle RDF triple store [1]. Based on a graph data model, RDF triples are persisted, indexed and queried, like other object-relational data. In this framework, we create the following models and rulebases (a set of rules):

- `owlTrajectory`, `owlTime`, `owlOGCSpatial`, `owlSealTrajectory`: declarative part of the trajectory, time, spatial and seal ontologies;
- `OWLPrime`: rulebase of the standard rules;
- `Time_Rules`, `Spatial_Rules`, `Seal_Rules`: rulebase of the temporal, spatial and seal rules, respectively.

In our framework, inference mechanism creates a rule index, Figure 8. A rule index (entailment) is an object containing pre-computed triples from applying a specified set of rulebases to a specified set of models. If a graph query refers to any rulebases, a rule index must exist for each rulebase-model combination in the query. The `USER_RULES=T` option is required while applying user-defined rules. The default number of rounds that the inference engine should run is `SEM_APIS.REACH_CLOSURE`.

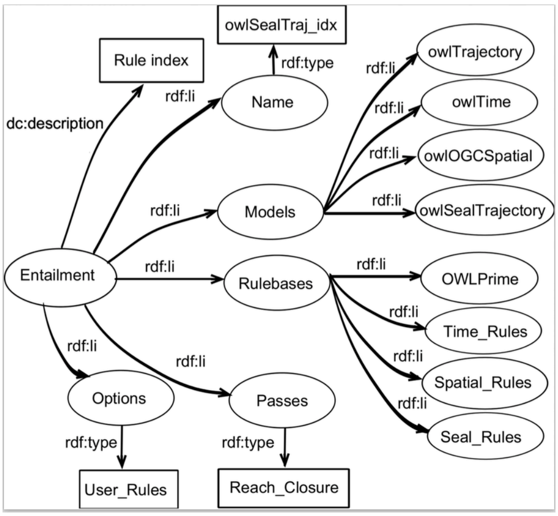

Fig. 8. Seal ontology inference: entailment

7 Trajectory ontology inference refinement

The inference mechanism computes data relationships and annotates each one with activities of the moving object. This mechanism is needed for queries on the spatio-temporal trajectory ontologies. Our objective is to enhance the inference mechanism as much as possible, particularly, when using user-defined rules.

The inference engine has many computation cycles. The inference time increases seemingly out of proportion when using user-defined rules. To control this number of cycles, we define an inference refinement, named **passes refinement**. In this refinement, algorithm 1 limits these cycles into one pass. However, we keep the assurance of the results quality. The idea is to persist the inference results into a database from the first pass and use these results for the other passes. The algorithm 1 considers the list of objects (L_SP) and list of relationships (L_R). For each two objects, we check the existence of a relationship between them in the database, otherwise we compute it and persist it in the database.

8 Evaluation

Experiments and evaluations are performed over real seal trajectories, around 410 690 raw data as seal's dives and 1 255 raw data as seal's haulout. So, in our system we have at least 5 749 660 triples in the triple store. In this work, we will focus on spatial rules. We evaluate the proposed passes refinement. The relationships, considered in algorithm 1, are the spatial rules. The experimental results are shown in Figure 9. The impact results are shown by the following experiments:

1. Spatial rules: presents the number of executions of the spatial rules;
2. Spatial rules - passes refinement: displays the number of executions of the spatial rules when applying the passes refinement.

input : List of the objects: *L_SP*
input : List of Relationships: *L_R*
initialization;
for $S_r, S_a \in L\_SP$ **do**
  if *L_R between* $(S_r, S_a) \notin database$ **then**
    Res := calculate *L_R* between S_r and S_a;
    Save *Res* in the *database*;
  end
end

Algorithm 1: Passes refinement algorithm

In this experiment, the number of spatial rules executions decreases while applying the passes refinement. For example, considering 300 dives, the spatial rules are executed 2 000 000 times. However, in the passes refinement case, the spatial rules are executed 500 000 times. The passes refinement enhances 4 times in terms of executions and time computations.

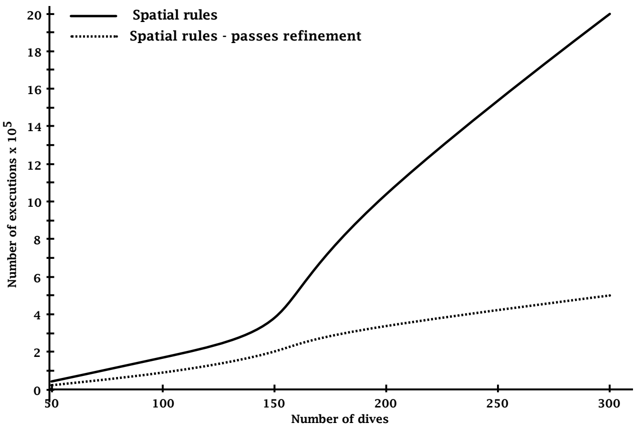

Fig. 9. Evaluation of the spatial ontology inference over the passes refinement

9 Conclusion

Trajectories are usually available as raw data. The data lack semantics which is fundamental for their efficient use. Our work is based on an ontology modeling approach for semantic trajectories. However, the domain part of the trajectory ontology focuses on mobile object's characteristics and its trajectory's activities. Our model considers a trajectory as a spatio-temporal concept, then we map it to temporal and spatial models. We reuse the W3C OWL-Time ontology and the spatial ontology based on the OpenGIS Simple Features Interface Standard (SFS). We apply our modeling approach to an application domain: marine mammal tracking, in particularly seal trajectories. In this implementation, we consider RDF triple store. Technically, we use Oracle Semantic Data Technologies. The objective is to annotate data with the considered knowledge: the domain, temporal and spatial knowledge. Over huge data, an ontology inference mechanism can do this automatic annotation. The inference mechanism therefore becomes an expensive mechanism in terms of time computations. To reduce the inference complexity, we propose a passes refinement. We evaluate it on real-world trajectory data. The experimental results highlight the positive impact of the passes refinement. We approximately reduce half of the time computation of the inference mechanism.

References

1. Oracle Database Semantic Technologies Developers guide 11G Release 2. Technical report, 2012.
2. James F. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, pages 832–843, 1983.
3. L O Alvares, V Bogorny, B Kuijpers, A F Macedo, B Moelans, and A Vaisman. A model for enriching trajectories with semantic geographical information. In *Proceedings of the 15th annual ACM international symposium on Advances in geographic information systems*, pages 22:1–22:8. ACM, 2007.
4. M Baglioni, J Macedo, C Renso, and M Wachowicz. An ontology-based approach for the semantic modelling and reasoning on trajectories. In *Advances in Conceptual Modeling - Challenges and Opportunities*, pages 344–353. Springer Berlin/Heidelberg, 2008.
5. Azedine Boulmakoul, Lamia Karim, and Ahmed Lbath. Moving object trajectories metamodel and spatio-temporal queries. In *International Journal of Database Management Systems (IJDMS)*, pages 35–54, 2012.
6. Eliseo Clementini, Paolino Di Felice, and Peter van Oosterom. A small set of formal topological relationships suitable for end-user interaction. In *Proceedings of the Third International Symposium on Advances in Spatial Databases*, pages 277–295. Springer-Verlag, 1993.
7. Eliseo Clementini, Jayant Sharma, and Max J. Egenhofer. Modelling topological spatial relations: Strategies for query processing. *Computers and Graphics*, pages 815–822, 1994.
8. J. R. Hobbs and P. Fang. Time ontology in OWL. W3C recommendation, 2006.
9. R. H Jerry and P Feng. An ontology of time for the semantic Web. In *ACM Transactions on Asian Language Information Processing*, pages 66–85, 2004.
10. Herring John R. OpenGIS implementation standard for Geographic information - simple feature access - part 2: SQL option. *Open Geospatial Consortium Inc.*, 2011.
11. J Malki, A Bouju, and W Mefteh. An ontological approach modeling and reasoning on trajectories. taking into account thematic, temporal and spatial rules. In *TSI. Technique et Science Informatiques*, volume 31/1-2012, pages 71–96, 2012.
12. S Spaccapietra, C Parent, M Damiani, J Demacedo, F Porto, and C Vangenot. A conceptual view on trajectories. *Data and Knowledge Engineering*, pages 126–146, 2008.
13. Rouaa WAnnous. *Trajectory ontology inference considering domain, temporal and spatial dimensions. Application to marine mammals*. PhD thesis, La Rochelle university, 2014.
14. Rouaa Wannous, Jamal Malki, Alain Bouju, and Ccile Vincent. Time integration in semantic trajectories using an ontological modelling approach. In *New Trends in Databases and Information Systems*, pages 187–198. Springer Berlin Heidelberg, 2013.
15. Z Yan, C Parent, S Spaccapietra, and D Chakraborty. A hybrid model and computing platform for spatio-semantic trajectories. In *The Semantic Web: Research and Applications*, pages 60–75. Springer Berlin/Heidelberg, 2010.