Time Integration in Semantic Trajectories Using an Ontological Modelling Approach

Rouaa Wannous, Jamal Malki, Alain Bouju, Cécile Vincent

Time integration in semantic trajectories using an ontological modelling approach

A case study with experiments, optimization and evaluation of an integration approach

Rouaa WANNOUS, Jamal MALKI, Alain BOUJU and Cécile VINCENT

Abstract Nowadays, with a growing use of location-aware, wirelessly connected, mobile devices, we can easily capture trajectories of mobile objects. To exploit these raw trajectories, we need to enhance them with semantic information. Several research fields are currently focusing on semantic trajectories to support queries and inferences to help users for validating and discovering more knowledge about mobile objects. The inference mechanism is needed for queries on semantic trajectories connected to other sources of information. Time and space knowledge are fundamental sources of information used by the inference operation on semantic trajectories. This article presents a case study of inference mechanism on semantic trajectories. We propose a solution based on an ontological approach for modelling semantic trajectories integrating time information and rules. We give experiments and evaluations of the proposed approach on generated and real data.

1 Introduction

Over the last few years, there has been a huge collection of real-time data of mobile objects. These data are obtained by satellite based systems like GNSS[1], GPS[2] or ARGOS, phone location or radio-frequency identification. This opens new perspectives for several applications like road traffic supervision and animals tracking. Therefore, it becomes necessary to provide mechanisms for storage, modelling, ef-

WANNOUS, MALKI and BOUJU
University of La Rochelle, L3i laboratory, EA 2118, F-17000 La Rochelle, France

VINCENT
University of La Rochelle, LIENSs laboratory, UMR 7266, F-17000 La Rochelle, France
e-mail: rwanno01, jmalki, abouju, cvincent@univ-lr.fr

[1] GNSS : Global Navigation Satellite System

[2] GPS: Global Positioning System

ficient analysis and knowledge extraction from these data. The raw data captured, commonly called trajectories, traces moving objects from a departure point to a destination point as sequences of pairs (sample points captured, time of the capture). In [12], authors give a general definition of a trajectory: "*A trajectory is the user defined record of the evolution of the position (perceived as a point) of an object that is moving in space during a given time interval in order to achieve a given goal*". Trajectories can be constrained to existing networks [10], or be unconstrained like in our study. Raw trajectories don't contain contextual information about moving objects like goals of travelling nor activities accomplished [3]. To consider these semantic information, semantic trajectories are defined as a result of the annotation process of raw data with semantic annotations [12]. This annotation process can be done automatically or manually. Semantic trajectories can be seen as a high-level information layer on raw trajectories [14]. In [8], to model semantic trajectories, a domain ontology is constructed to represent domain concepts and rules. In the continuation of this paper [8], we discuss strategies for time integration with evaluation on synthetic and real data. We study seal trajectories and focus on semantic annotations for there activities such as foraging, travelling and resting. The inference mechanism on semantic trajectories which is connected to time knowledge has time and space storage complexity problem. This work addresses these two problems and gives some ideas for improving the complexity of the proposed approach.

This paper is organized as follows: section 2 presents the state of the art on semantic trajectories and some recent related work. Section 3 details our domain application and queries we aim to answer. Section 4 gives the two ontologies needed, seal trajectory and time ontologies. Section 5 presents our domain ontology rules and the temporal ontology rules. Section 6 defines the connection between trajectory and time ontologies. Section 8 discusses the evaluation of the proposed approach. Finally, section 9 concludes this paper and presents ideas for the future work.

2 Related work

Data management techniques including modelling, indexing and querying large spatio-temporal data are actively investigated during the last decade [13]. Most of these techniques are only interested in raw trajectories. Projects like GeoPKDD [4] and MODAP[3] emphasized the need to address and to use semantic information about moving objects for efficient trajectories analyses. Recently, new projects are born like MOVE[4] who aims to improve methods for knowledge extraction from massive amounts of moving objects data. As an example, in birds migration project [12], trajectories are analysed for better understanding birds behaviours. Scientists try to answer queries such as: where, why and how long birds stop on their travels, which activities they do during their stops, and which weather conditions the

[3] MODAP: Mobility, Data Mining and Privacy

[4] Move: European Cooperation in Science and Technology - http://move-cost.info/

birds face during their flight. Considering these new requirements, new researches have emerged offering data models that can easily be expanded to take into account semantic data. In [12], the trajectory is seen as a user defined time-space function from a temporal interval to a space interval. To consider semantics of trajectories, a conceptual view is defined by three main concepts: stops, moves, and begin-end of a trajectory. Each part contains a set of semantics data. This model is implemented and evaluated on a relational database. Most domain and temporal operations are SQL based and use elementary data comparators. Based on this conceptual model of trajectories, several works have been proposed such as [3].

Using ontologies for semantic spatio-temporal data modelling is a new research field. In [9], authors work on a military application domain with complex queries that require sophisticated inferences methods. For this application, they present an upper-level ontology defining a general hierarchy of thematic and spatial entity classes and associating relationships to connect these entity classes. They intend for application-specific domain ontologies in the thematic dimension to be integrated into the upper-level ontology through subclassing of appropriate classes and relationships. Temporal information is integrated into the ontology by labelling relationship instances with their valid times. In this work, the temporal and spatial dimensions are included in the global ontology. Moreover, the ontology is formalised by the RDFS vocabulary and implemented on a relational database. Consequently, the inference mechanism is based on several domain specific table functions. The inference mechanism defined uses only the RDFS rules indexes. In [14], authors design a conceptual model of trajectories based on the approach introduced in [12]. This model represents trajectories from low-level real-life GNSS data to different semantically abstracted levels. Their application concerns daily trips of employees from home to work office and coming back. So, they start from basic abstractions (e.g. stops, moves) to enriched higher-level abstractions (e.g. office, shop). In [6], Malki et. al define an ontological approach modelling and reasoning on trajectories. This approach takes into account thematic, temporal and spatial rules. The ontologies constructed are formalised using both RDFS and OWL vocabulary. The inference mechanism is based on rules defined as entailments.

Finally, in [12], domain or time inference mechanism are not proposed neither spatio-temporal data mining are not investigated. However, in this paper, we focus on time knowledge integration and use inference mechanisms on semantic trajectories based on the approach introduced in [6]. Nevertheless, authors did not address the evaluation of the proposed approach. For all of that, this article gives experiments and evaluates the performance problems of time integration on generated and real data.

3 Domain application

In this section, we present the seal trajectory modelling approach. We introduce the seal trajectory data model and its semantic associated layer.

3.1 Seal trajectory data model

As in [6], this paper considers trajectories of seals. The data comes from the LIENSS[5] (CNRS/University of La Rochelle) in collaboration with SMRU [11]. These laboratories work on marine mammals ecology. Trajectories of seals between their haulout sites along the coasts of the English Channel or in the Celtic and Irish seas are captured using GNSS systems provided by SMRU Instrumentation. We use trajectories data coming from GPS/GSM tags. The captured spatio-temporal data of seals trajectories can be classified into three main states: haulout, cruise and dive. The Fig. 1 shows the three states, the transitions and their guard conditions [8].

3.2 Semantic seal trajectory

We focus on studying seals' activities in order to identify their foraging areas. The main activities of seal, like foraging, resting and travelling, occur in the dive parts of the trajectory. So we aim at answering queries, such as:

1. foraging activities;
2. foraging activities during a given time interval;
3. foraging activities performed after travelling during a given time interval.

For all queries, we have to define a domain rule called "foraging" on the seal trajectory model. However, for the last two queries, time rules must be defined between trajectory's parts. For example, the query 3 needs the two domain rules "foraging, travelling" and the time rule "during" as illustrated by the Table 1.

4 Modelling approach

The need of a time model with temporal relationships appears in Table 1. As in [8], we consider separated and independent data models using an ontological approach.

4.1 Seal trajectory ontology

The seal trajectory ontology, `owlSealTrajectory`, is a result of a transformation model of the semantic seal trajectory. An extract of this ontology is in Fig. 2. This ontology defines the main following concepts:

- `Seal`: represents the animal equipped with a tag;

[5] http://lienss.univ-larochelle.fr

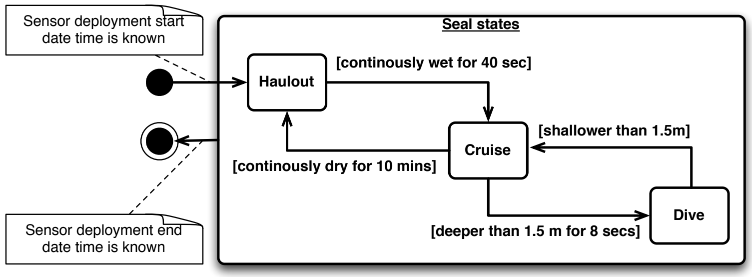

Fig. 1 The three states of seal trajectory

Table 1 Domain and time concepts and rules needed by the query 3

Concepts and rules			**Description**
Concepts	Domain	Dive	specific part of the seal trajectory
	Time	Temporal-interval	the given temporal interval
Rules	Domain	Travelling	seal activity
		Foraging	seal activity
	Time	After	temporal relationship between the two activities
		During	temporal relationship between activity and time interval

- `Sequence`: captures in the form of temporal intervals with a spatial part called `GeoSequence` and can be `Haulout`, `Cruise` or `Dive`. The other parts are metadata called `Summary` and `CTD` (Conductivity-Temperature-Depth);
- `Trajectory`: is the logical form to represent a set of sequences;
- `Activity`: is the seal activity for a sequence or for a trajectory.

Besides these concepts, `owlSealTrajectory` defines these relationships:

- `seqHasActivity`: is the object property between an activity and a sequence;
- `TAD` (Time Allocation at Depth): is the data property calculated to define the shape of a seal's dive, as mentioned in [7].

4.2 Time ontology

Table 1 clearly highlights the need for temporal concepts as well as temporal relationships between these concepts. In our approach, we chose `owlTime` ontology [5] developed by the World Wide Web Consortium (W3C) thanks to the definition of the temporal concepts and relationships as defined by Allen algebra [1]. An extract of the declarative part of this ontology [6] is given in Fig. 3.

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

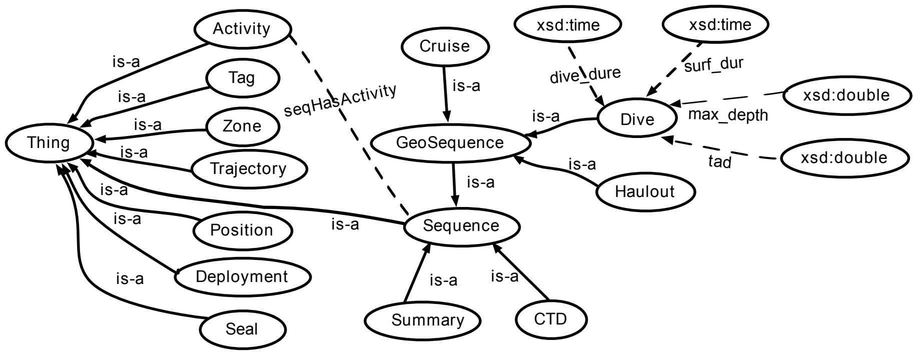

Fig. 2 An extract of the ontology `owlSealTrajectory`

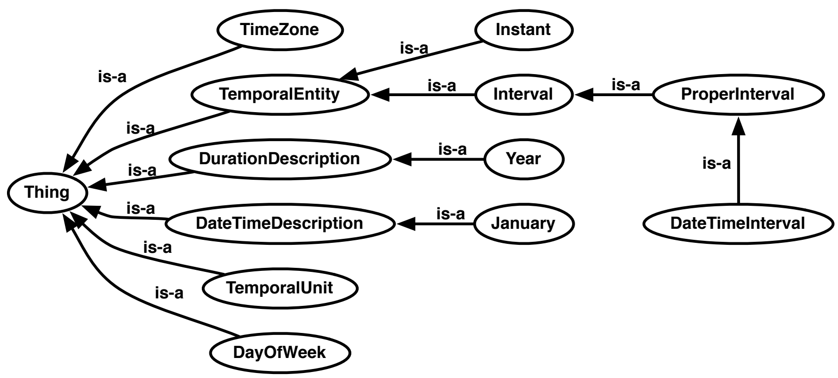

Fig. 3 A view of the `owlTime` ontology

5 Ontological rules

5.1 Seal trajectory ontology rules

We define four seals' activities during dives: resting; travelling; foraging; travelling-foraging. In our approach, each seal activity is defined in the ontology and has both a declarative and an imperative corresponding parts. The Fig. 4 shows the declarative parts. The imperative parts are based on the decision Table 2 which determines the maximum dive depth, the dive's shape (TAD) and the surface ratio dividing the surface duration by the dive duration. We implement the imperative parts using the Oracle database supporting semantic technologies. We create the rule base `sealActivities_rb` to hold the activities' implementation as domain rules. The Code 1 gives the implementation of `foraging_rule` (line 3) in the rule base `sealActivities_rb`. In this code, the line 6 checks the maximum dive depth `d` to be more than 3 meters, the TAD `t` to be 0.9 and the surface duration `s` divided by the dive duration `v`, to be smaller than 0.5.

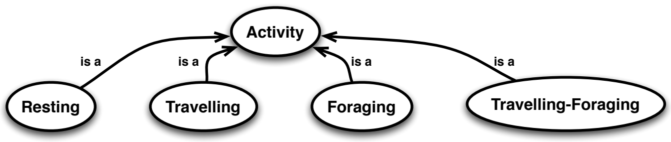

Fig. 4 Declarative part of seal activities

Table 2 Decision table associated to seal activities

Rules	**Maximum dive depth**	**Dive shape** or **TAD**	**Surface ratio = surface duration/dive duration**
Resting	< 10	all	> 0.5
Travelling	> 3	> 0 & < 0.7	< 0.5
Foraging	> 3	> 0.9 & < 1	< 0.5
Travelling_Foraging	> 3	> 0.7 & < 0.9	< 0.5

```
EXECUTE SEM_APIS.CREATE_RULEBASE('sealActivities_rb');
INSERT INTO mdsys.semr_sealActivities_rb
VALUES( 'foraging_rule',
'(?x rdf:type ost:Dive)
(?x ost:tad ?t)(?x ost:max_depth ?d)(?x ost:surf_dur ?s)(?x ost:dive_dur ?v)
    ',
'(d > 3) and (t > 0.9) and (s/v < 0.5)',
'(?x ost:seqHasActivity ?activiteJ) (?activiteJ rdf:type ost:Foraging)',
SEM_ALIASES(SEM_ALIAS('ost', 'http://l3i.univ-larochelle.fr/Sido/
    owlSealTrajectory#')));
```

Code 1 The imperative part of the seal activity `foraging`

5.2 *Time ontology rules*

The `owlTime` ontology declares 13 relationships based on Allen algebra [1]. These are: intervalEquals, intervalBefore, intervalMeets, intervalOverlaps, intervalStarts, intervalDuring, intervalFinishes, intervalAfter, intervalMetBy, intervalOverlappedBy, intervalStartedBy, intervalContains, intervalFinishedBy. Allen temporal relationships are implemented inside the rule base `owlTime_rb`. For example, the Code 2 presents the implementation of the `intervalAfter_rule`.

```
EXECUTE SEM_APIS.CREATE_RULEBASE('owlTime_rb')
INSERT INTO mdsys.semr_owltime_rb
VALUES('intervalAfter_rule',
'(?x rdf:type owltime:ProperInterval)(?y rdf:type owltime:ProperInterval)
(?x owltime:hasEnd ?xEnd)(?xEnd :inXSDDateTime ?xEndDateTime) (?y owltime:
    hasBeginning ?yBegin)(?yBegin :inXSDDateTime ?yBeginDateTime)',
 '(yBeginDateTime > xEndDateTime)',
'(?y owltime:intervalAfter ?x)',
SEM_ALIASES(SEM_ALIAS('owltime','http://www.w3.org/2006/time#')));
```

Code 2 The imperative part of Allen temporal relationship `intervalAfter`

6 Semantic integration by ontological mapping

The need of a semantic integration clearly appears while considering separated and independent sources of information, like seal trajectory and time ontologies. This ontological mapping may lead to discover more semantic trajectory patterns. The property rdfs:subClassOf is not appropriate in separated ontologies. Even more the property owl:sameAs means that the two connected classes have the same intention meaning, however it does not go further for their properties. Consequently the property owl:equivalentClass is the most appropriate connection in our case. The mapping process is shown in Fig. 5 following these steps:

1. `owlSealTrajectory:Sequence` is the mapping concept by OWL construct `owl:equivalentClass` to `time:ProperInterval`;
2. `owlSealTrajectory:s_date` is the mapping object property by OWL construct `owl:equivalentProperty` to `owlTime:hasBeginning`;
3. `owlSealTrajectory:e_date` is the mapping object property by the OWL construct `owl:equivalentProperty` to `owlTime:hasEnd`.

In particular, the reasoner considers the owl property "owl:equivalentClass" which allows the inference of a "Sequence" instance as a "ProperInterval" instance. Therefore, the interval temporal rules are also valid for sequences of trajectories, which means valid for dives also.

7 Temporal rule extension

The inference mechanism is needed for queries on the semantic trajectory `owlSealTrajectory` mapped to the time ontology `owlTime`. Calculating the inference between all sequences of trajectories considering all time rules takes a huge amount of time and space storage capacity. To enhance the inference mechanism, we define a refinement called **temporal neighbour inference**: "*A temporal neighbour is when a sequence happened within a conceptual distance to another*". The goal of this refinement, algorithm 1, is to consider the distance between two sequences in order to calculate the corresponding temporal relationships. The temporal rules must comply with this refinement. It is still difficult to determine the best candidate for the temporal neighbour distance, and even then, there is uncertainty on its usefulness.

8 Evaluation and Analysis

The experiments aim at checking the usefulness of the temporal neighbour refinement. We create a synthetic temporal interval data and examine the validity

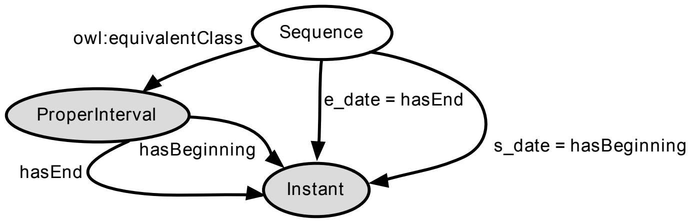

Fig. 5 Integration of `owlSealTrajectory` with `owlTime` ontology

input : Two sequences: a referent S_r and an argument S_a
input : A neighbour of S_r
output: Temporal rule between S_r and S_a
if $S_a \in$ *to the neighbour of* S_r **then**
  calculate the temporal rule between S_r and S_a;
  go the next sequence S_a
else
  go the next sequence S_a
end

Algorithm 1: Temporal neighbour inference algorithm

of our proposal. Then we evaluate it's efficiency on real GPS-GSM seal trajectory data. We use Oracle 11g Release 2. For the experiments, we consider the query 3 (§3.2). The Code 3 gives the SQL code of this query.

```
SELECT D1, D2
FROM TABLE ( SEM_MATCH(
'(?D1 rdf:type ost:Dive) (?D1 ost:sequenceHasActivity ?activiteD1) (?
     activiteD1 rdf:type ost:Travelling)
(?D2 rdf:type ost:Dive) (?D2 ost:sequenceHasActivity ?activiteD2) (?
     activiteD2 rdf:type ost:Foraging)
(?D1 ot:intervalBefore ?D2)
(?I rdf:type ot:ProperInterval) (?I ot:hasBeginning ?beginI) (?beginI ot:
     inXSDDateTime "2003-08-02T00:00:00"^^xsd:datetime) (?I ot:hasEnd ?endI
     )(?endI ot:inXSDDateTime"2003-08-09T23:59:00"^^xsd:datetime)
(?D1 ot:intervalDuring ?I) (?D2 ot:intervalDuring ?I)',
SEM_Models('owlSealTrajectory','owlTime'),
SEM_Rulebases('OWLPRIME','sealActivities_rb','owlTime_rb'),
SEM_ALIASES(SEM_ALIAS('ost', 'http://l3i.univ-larochelle.fr/Sido/
     owlSealTrajectory#'),
SEM_ALIAS('ot','http://www.w3.org/2006/time#')),
null));
```

Code 3 The SQL code of the query 3

In Fig. 6, the inference mechanism is done on semantic seal trajectory before and after the mapping. The experiment is done for different numbers of dives shown in the horizontal axis multiple by 100. In Fig. 6(a), the vectorial axis shows the time multiple by 10 000 needed for the inference mechanism. In Fig. 6(b), the vectorial axis shows the number of triples multiple by 100 000 related to the space storage. From analysing Fig. 6, the problem is obvious comparing the time and the space storage needed from the inference mechanism. In other words, after using temporal rules, calculating the inference becomes very expensive in terms of time and the space storage. In our point of view, this problem is related to the temporal rules integration without any constraints. With biological feedback, we define the temporal neighbour distance for seal trajectories to five minutes (300 seconds). Then we modify the implementation of temporal rules considering the temporal neighbour refinement. For instance, the modification of the temporal `intervalAfter` rule with the temporal neighbour refinement, called `intervalAfterRefined` rule, is given by the Code 4.

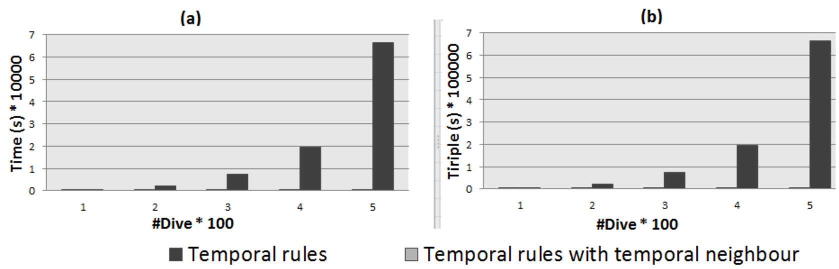

Fig. 6 Compare the time and space storage taken from the inference mechanism on the semantic seal trajectory integrated with/without the temporal rules

```
EXECUTE SEM_APIS.CREATE_RULEBASE('owlTime_rb')
INSERT INTO mdsys.semr_owltime_rb VALUES(
'intervalAfterRefined_rule',
'(?x rdf:type owltime:ProperInterval)(?x owltime:hasEnd ?xEnd)(?xEnd :
     inXSDDateTime ?xEndDateTime)(?y rdf:type owltime:ProperInterval)(?y
     owltime:hasBeginning ?yBegin)(?yBegin :inXSDDateTime ?yBeginDateTime)'
     ,
'(yBeginDateTime > xEndDateTime) and ((timeIntervalLengthInSeconds(
     dateTime2TimeStamp(xEndDateTime),dateTime2TimeStamp(yBeginDateTime))
     <300)',
'(?y owltime:intervalAfterTime ?x)',
SEM_ALIASES(SEM_ALIAS('owltime','http://www.w3.org/2006/time#')));
```

Code 4 Create the temporal `intervalAfterRefined_rule`

Then we validate the usefulness of the temporal neighbour on the synthetic data integrated first with the temporal rules and later with the extended temporal rules, as shown in Fig. 7. The vectorial axis, Fig. 7(a) shows the time and Fig. 7(b) shows the number of triples, where both time and space needed for the inference mechanism. Figure 7 shows the useful impact of applying the refined temporal rules in reducing the time and the space storage.Finally we apply the experiment on a real GPS/GSM data integrated first with the temporal rules and then with the refined temporal rules, as shown in Fig. 8. In Fig. 8(a), the vectorial axis shows the time (multiple by 10 000) needed for the inference mechanism for each number of dive. In Fig. 8(b), the vectorial axis shows the number of triples (multiple by 100 000) related to the space storage taken from the inference mechanism for each number of dive. Figure 8 shows the improvement made on calculating the inference after applying the temporal neighbour concept. In

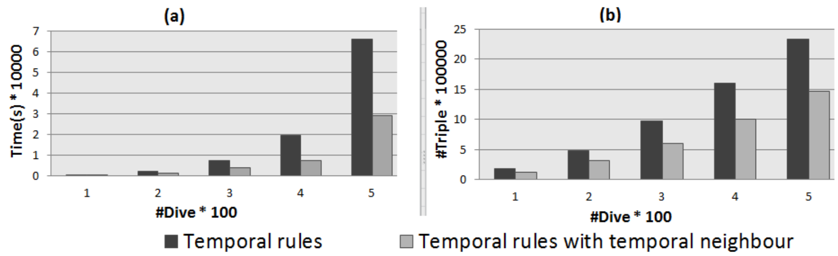

Fig. 7 The contribution of using temporal neighbour reduced the time and space storage taken form the inference mechanism on the synthetic data

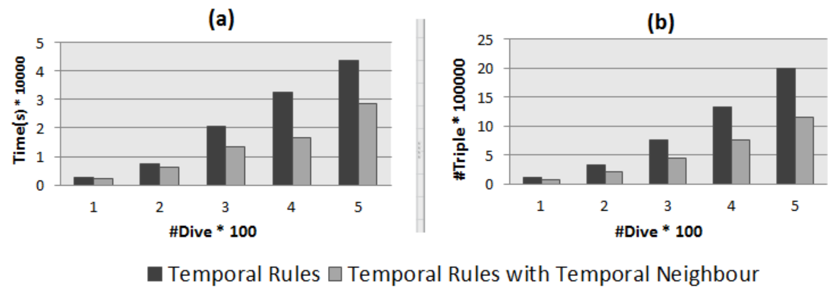

Fig. 8 The impact of using temporal neighbour refinement in the time and space storage taken by the inference mechanism on the GPS/GSM data

other words, the temporal neighbour is efficient in reducing the complexity of the inference mechanism on the GPS/GSM data.

9 Conclusion and future work

Trajectory data are usually available as sample points, and lake of semantic information, which is of fundamental importance for the efficient use of these data [2]. In this paper, we present a case study on the use of an ontological based approach for modeling semantic trajectories integrated with time rules. The main goal is to apply the inference mechanisms on semantic trajectories and to present a solution to reduce the complexity reasoning and querying on these semantic trajectories. We define a new condition on temporal rules which is *temporal neighbour refinement*. Then, we evaluate our approach on synthetic data as well as on real GPS/GSM data. The experiment's results verify the positive impact of the temporal neighbour refinement on reducing the complexity of the inference mechanism. The influence of one condition positively appears nevertheless the reasoning complexity still exists. So applying more domain conditions on rules is therefore very important for reducing time and space storage inference complexity. As future work, we aim at applying more conditions on temporal and spatial rules. We also intend to assess the impact of using incremental inference to improve the inference mechanism complexity and the contribution of temporal and spatial relationships composition.

References

1. J. F. Allen. Maintaining Knowledge about Temporal Intervals. *Communications of the ACM*, 26(11):832–843, 1983.
2. L. O. Alvares, J. A. Fernandes, D. Macedo, V. Bogorny, B. Moelans, B. Kuijpers, and A. Vaisman. A Model for Enriching Trajectories with Semantic Geographical Information. In *in ACM-GIS, ACM*. Press, 2007.
3. 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*, volume 5232 of *Lecture Notes in Computer Science*, pages 344–353. Springer Berlin / Heidelberg, 2008. 10.1007/978-3-540-87991-6.
4. GeoPKDD. Geographic Privacy-aware Knowledge Discovery and Delivery. Coordinator: KDDLAB, Knowledge Discovery nad Delivery Laboratory, ISTI-CNR and University of Pisa. http://www.geopkdd.eu.
5. J. R. Hobbs and F. Pan. An Ontology of Time for the Semantic Web. In *ACM Transactions on Asian Language Information Processing*, volume 3, pages 66–85, 2004.
6. J.Malki, A.Bouju, and W.Mefteh. An Ontological Approach Modeling and Reasoning on Trajectories. taking into account Thematic, Temporal and Spatial Rules. *Revue des sciences et technologies de l'information*, 31/1 - 2012(doi:10.3166/tso.31.71-96):71–96, 2012.
7. I. D. Jonsen, R. A. Myers, and M. C. James. Identifying Leatherback Turtle Foraging Behaviour from Satellite Telemetry using a switching State-space Model. *Marine Ecology Progress Series*, 337:255–264, May 2007.
8. J. Malki, W. Mefteh, and A. Bouju. Une Approche Ontologique pour la Modélisation et le Raisonnement sur les Trajectoires. Prise en compte des règles métiers, spatiales et temporelles. In *JFO 2009 3ème édition des journées Francofones sur les Ontologies*, pages 157–168, France, Dec. 2009.
9. M. Perry. *A Framework to Support Spatial, Temporal and Thematic Analytics over Semantic Web Data*. PhD thesis, Wright State University, June 10, 2008.
10. I. Sandu Popa, K. Zeitouni, V. Oria, D. Barth, and S. Vial. Indexing in-Network Trajectory Flows. *The VLDB Journal*, 20(5):643–669, Oct. 2011.
11. SMRU. Sea Mammal Research Unit. Collaborative ventures between marine biologists and systems engineers, University of St Andrews, UK. http://www.smru.st-and.ac.uk/.
12. S. Spaccapietra, C. Parent, M. Damiani, J. Demacedo, F. Porto, and C. Vangenot. A Conceptual View on Trajectories. *Data and Knowledge Engineering*, 65(1):126–146, Apr. 2008.
13. Z. Yan, D. Chakraborty, C. Parent, S. Spaccapietra, and K. Aberer. SeMiTri: A Framework for Semantic Annotation of Heterogeneous Trajectories. In *Proceedings of the 14th International Conference on Extending Database Technology*, EDBT/ICDT '11, pages 259–270, New York, NY, USA, 2011. ACM.
14. Z. Yan, C. Parent, S. Spaccapietra, and D. Chakraborty. A Hybrid Model and Computing Platform for Spatio-semantic Trajectories. In L. Aroyo, G. Antoniou, E. Hyvnen, A. ten Teije, H. Stuckenschmidt, L. Cabral, and T. Tudorache, editors, *The Semantic Web: Research and Applications*, volume 6088 of *Lecture Notes in Computer Science*, pages 60–75. Springer Berlin / Heidelberg, 2010. 10.1007/978-3-642-13486-9.