Trajectory data warehouse modeling based on a Trajectory UML profile: Medical example

Wided Oueslati[1] and Jalel Akaichi[2]

[1]Computer scienceDepartment, Higher Institute of Management of Tunis

[2]Computer scienceDepartment, Higher Institute of Management of Tunis

widedoueslati@atclive.fr; jalel.akaichi@isg.rnu.tn

Abstract. Trajectory data is a new concept appeared with the continuous mobility imposed by social and professional daily life activities and the advance of mobile devices used to accomplish missions while moving. Trajectories data can be used either for transactional and analysis purposes in various domains (heath care, commerce, environment, etc.). For this reason, modeling trajectory data at the conceptual level is an important rung leading to flourishing implementations. However, there is not till now a recognized model answering all the requirements for modeling trajectory data, and the trajectory data conceptual modeling remains an open line of research. Hence, we propose in this paper a rich conceptual schema modeling for trajectory data and trajectory data warehouse which is the basis of future efficient storage system and consequently meaningful interpretations and efficient analysis in ubiquitous applications.

Keywords: Trajectory data, Conceptual modeling, trajectory data warehouse, Moving objects, UML extension.

1 Introduction

New requirements of our social and professional life impose continuous mobility of people (suppliers, customers, commercial agents, maintenance agents, doctors, teachers…). This omnipresent phenomenon demands the use of mobile devices (mobile phone, PDA…) allowing routing information in communication systems and accomplishing missions. These mobile devices are equipped with sensors tracking the movements of their mobile users, which leaves digital traces in information systems of organizations providing services through wireless networks. These traces describe the evolution of the position of the moving object in the geographical space, during a certain interval of time recently called trajectory. Many applications belonging to social, economic, ecological or biological fields are becoming based on the analysis of movement data belonging to objects still in motion such as people, vehicles, animals, birds, viruses, planets… and profiting from to address many ubiquitous issues such as new scientific discoveries or explanation of complex and unusual behaviors. For example in biology, tracking and studying the movement of virus or cellules in the human or animal corpus could explain causes of some diseases, and lead to science and medical discoveries.

In our case study the moving object is the mobile hospital. This latter was emerged in order to decrease the geographical disparities related to health care services. In fact, in many countries, rural populations undergo various inadequacies related to health care necessities. This is due to health care problems, since they have to cross long distances to reach medical institutions such as hospitals. This, obviously, may generate loss of lives and/or medical complications. Mobile hospitals, which may contribute to decrease such misery, can move to populations, following declared emergencies and/or planned activities.

Modeling the trajectory and the trajectory data warehouse of the mobile hospital is a real challenge. Indeed, witnessing the trajectory of a moving hospital implies the recording of different information leading to get a finite set of observations destined to current transactional and future analysis activities. In fact, efficient analysis and pertinent results are based on correct and complete view on trajectory that is means on a good trajectory data conceptual model. However, currents modeling tools fail to fulfill specific moving objects, such as mobile hospital, activities requirements. In this paper, we propose a new profile based on UML in order to enhance the conceptual modeling of trajectory data and trajectory data warehouse related to mobile objects by adding new stereotypes, icons and constraints.

This paper is organized as follows: In section 2, we will present different research works related to moving objects, trajectory data conceptual model, trajectory data warehouse and positioning technologies. In section 3, we will propose a trajectory data conceptual modeling for mobile hospitals. In section 4, we will propose a mobile hospitals trajectory data warehouse conceptual modeling. In section 5, we will present the development of the trajectory data profile based on UML. In section 6, we will go over the main proposed points and give some perspectives for future works.

2 State of the art

Some works [1, 2, and 3] define a trajectory as a set of stops and moves. Each element of a set has a “begin” and an “end”. The “stop” is considered as an important component of the trajectory which is characterized by a non-empty time interval. This means that the moving object can perform many actions but it has to be immovable in a given “stop”. A move is delimited by two stops between which a moving object is in movement. Authors in [4] distinguish many types of trajectories according to relationships between them or with their environments. For the first type of relationship, we can cite those relations: Intersect (two trajectories cross themselves), Equal (a trajectory is near another trajectory), Near (a trajectory is near another trajectory) and Far (a trajectory is far from another trajectory).

For the second type of relationship, trajectories can be in relation with other spatial objects such as infrastructure elements (roads, buildings, etc.) or virtual entities such as border of a city. Among those relationships we can cite: Stay within (the trajectory is always in a sector of interest), Bypass (the trajectory pass by the sector of interest), Leave (the trajectory leaves the sector of interest), and Enter (the trajectory enters the sector of interest) and Cross (the trajectory crosses the sector of interest).

To collect data resulting from moving objects activities, positioning and sensors technologies, and services associated to them and technologies with sensors seem to be a necessity. In fact, they permit to gather, to process and to broadcast localization information and data generated in it. Among those technologies we can mention GPS (Global Positioning System), RFID (Radio Frequency IDentification tags), Bluetooth, WiFi, ZigBee, etc., which permit the localization and the follow up of persons, goods and animals on move.

Positioning technologies can be classified into two categories according to their mobility degree [5]: fixed positioning technologies and mobile ones.

Positioning technologies [5, 6, 7] can have many functions such as authentication, data transfer via smart cards for example, object localization, etc. Services associated to them are responsible of the acquisition of data for a given system which is able to collect data coming from his environment.

Many research works were interested in the conceptual modeling of data and data warehouses [8, 9, and 10], spatial and temporal data and data warehouses [11, 12, and 13] but few of works were interested in the conceptual modeling of trajectory data and trajectory data warehouses [1, 14]. In [1], the requirements for the trajectory modeling take into account the characterization of trajectories and their components and the different types of constraints (semantic, topologic, etc.) in order to fix a conceptual view of the concept of trajectory. The conceptual model is seen as a direct support and an explicit representation of trajectories' components (stops, moves, begins, ends). The two solutions proposed in [1] are driven by different modeling goals. The first solution suggests a trajectory design pattern, and the second one recommends some trajectory data types. The two approaches can be combined to offer richer modeling tools when needed. The design pattern is a predefined generic schema that can be connected to any other database schema by the designer and can be adjustable. In fact, the designer can modify the design pattern by adding new elements or semantic attributes, and/or deleting other ones to adapt it to the requirements of the new applications. The second approach is based on trajectory data types. It holds the idea that semantic information is specific to an application and cannot be encapsulated into a data type, but has to be defined by the database designer. Moreover, authors in [1] define a generic data types to hold the trajectories' components such as the "begin", the "end", the "stops" and the "moves", etc., and functions of interpolation. We have to mention that authors in [1], extended those data models for trajectories from MADS model [15], since this latter supports spatial and temporal objects and relationships. Indeed, they announce icons defined in it, in their trajectory data relational conceptual model to make it more understandable by users.

3 Trajectory data conceptual modeling

The mobility of a hospital, supported by adequate terrestrial transportation mean, is constrained by the road network. Mobile hospitals manager defines for each mobile hospital its trajectory which is composed by a set of stops and moves. For each trajectory, the mobile hospital is equipped with medical hardware, physicians, nurses, drivers, and a mobile hospital manager. This latter is equipped with a PDA used to manage trajectory data. He has to indicate the beginning and the end of the trajectory, each stop and move. He also sends messages to the responsible of the mission or receives messages from him. Physicians are equipped with PDA in order to send data related to patients and examinations to a trajectory data warehouse. Such data are important since they allow detecting prospective epidemic peaks.

3.1 Mobile hospitals' trajectory components constraints modeling

In this section, we propose the pre-conceptual modeling of a trajectory and its components that are trajectory-sections, stops and moves in order to provide assistance to designers in defining the conceptual model of a moving object's trajectory. To do this, some relationships' constraints between a given trajectory and their components have to be taken into account in the conceptual modeling of moving object's trajectory data step. For this reason we propose the following table containing the terminology that we will use to present the mathematical description of the moving object's trajectory pre-conceptual constraints.

Table 1: Our terminology for trajectory components

$\mathbf{T_k}$	Trajectory number k
$\mathbf{Tr\text{-}s_{ik}}$	Trajectory section number i of the Trajectory number k
$\mathbf{S_j}$	Stop number j
$\mathbf{M_i}$	Move number i
$\mathbf{Tr\text{-}s_i}$	Trajectory section number i

- $T_k = \sum_{i=1}^{n} Tr\text{-}sik$ **;** This equation means that a given trajectory k is composed of the set of its trajectory sections.
- $Tr\text{-}s_i = \sum_{j=i}^{i+1} Sj + Mi$ **;** This equation means that a given trajectory section i is composed of two successive stops and only one move.
- $Tr\text{-}s_i \cap Tr\text{-}s_{i+1} = S_{j+1}$ with $j=i$**;** This means that the intersection between two successive trajectory sections ***Tr-si*** and $Tr\text{-}s_{i+1}$ is the stop S_{i+1}.

Let $M_i = 1$ if the move ***i*** is affected to a trajectory section, ***0*** else.

Let $Tr\text{-}s_{ij} = 1$ if the move ***i*** is affected to the trajectory section ***j, 0*** else.

- $\sum_{i=1}^{n} Tr - Sij = 1$ **;** This equation means that a trajectory section j has only one move.
- $\sum_{j=1}^{n} Tr - Sij = 1$ **;** This equation means that the move ***i*** is affected to only one trajectory section ***j*** of the given trajectory then $\boldsymbol{Tr\text{-}s_{ij} = Mi}$ must be verified.

Proof: if the move i is not yet affected to any trajectory section, it will have the value 0, then the Tr-sij will have the value 0 too; the condition Tr-sij = Mi is verified. If the move i is affected to another trajectory section k, it will have the value 1 but the trajectory section Tr-sij will have the value 0; the condition Tr-sij = Mi is verified. If the move i is affected to the trajectory section j, it will have the value 1, then the Tr-sij will have the value 1 too; the condition Tr-sij = Mi is verified.

3.2 Mobile hospital trajectory data conceptual modeling

The trajectory concept can be defined at three levels. Those latter are the raw level, the structured level and the semantic level. At the raw level, the trajectory is defined as a set of points. Each point is the triple (X, Y, T) captured by a positioning device, where (X,Y) are spatial coordinates and T represents the time instant. At the structured level, the raw data are transformed into significant units that are trajectory-section (Tr-S), stops (S) and moves (M). A stop is the spatio-temporal object having a time interval [t-begin-stop,t-end-stop], duration and location where the stop is happened. A move is the spatio-temporal path of the moving object detected by capturing position technology. This path has a non empty time interval delimited by a beginning instant and an ending one and characterized with a duration derived from the time interval. At the semantic level, the trajectory is rich with semantic information obtained from the application domain such as the geographical label of places where stops happened, the activity done at each stop…Such additional information allow more meaningful analysis based on concrete queries. To manipulate a trajectory as a rich semantic object, a conceptual model is needed.

The conceptual modeling of trajectory data resulting from mobile hospitals is represented by a class diagram of UML. This latter describes the classes that make up a system and the static relationships between them. Classes are defined in terms of their names, attributes (or data) and behaviors. The class diagram shows the system's internal structure; it offers a static view because it ignores the temporal fact in its behaviors. It shows the different classes, their characteristics and relationships between them. The following class diagram is related to trajectory data resulting from mobile hospital. We created a new UML profile called Trajectory-UML, in which we added some constraints, pictograms and some stereotypes to identify each class (entity). We used pictograms of MADS project [15] only for some entities which are representing trajectory and trajectory components such as "move", "stop", "tr-section", and "location" . The same idea was done in [1] but in a relational model.

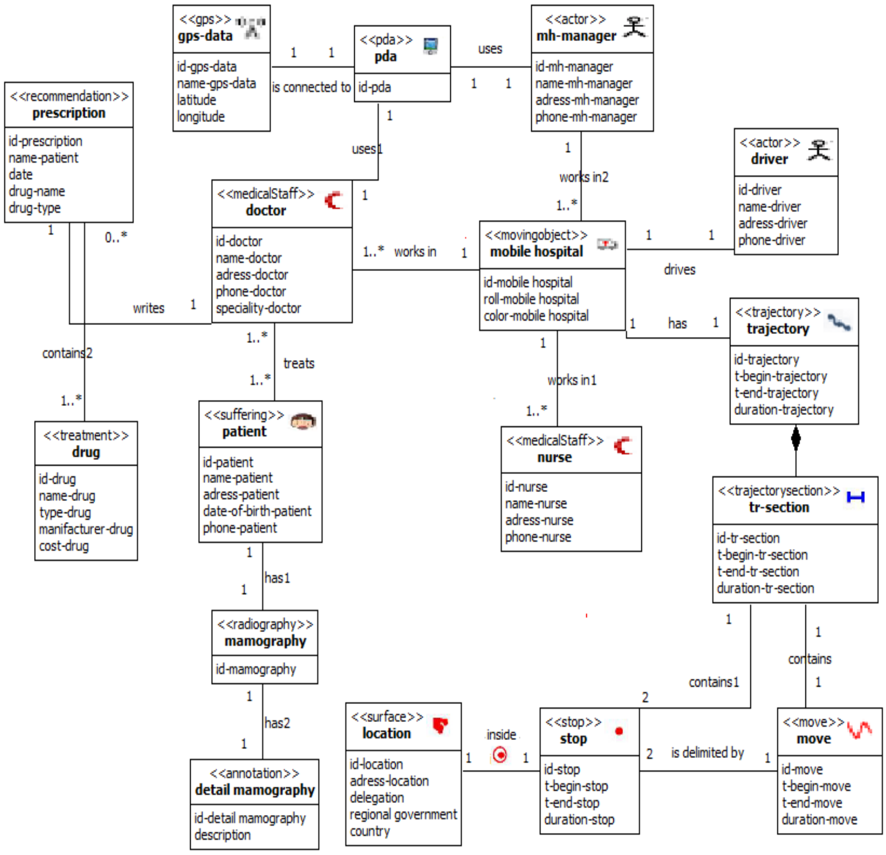


Figure 1. Mobile hospital's trajectory conceptual modeling

We propose the following constraints to be respected when modeling the trajectory data conceptual modeling:

- Constraint 1: the class <<trajectory-section>> can be connected only to the class <<trajectory>> or <<stop>> or <<move>>. self.allOppositeAssociationEnds->forAll(participant.oclIsTypeOf(Trajectory)oroclIsTypeOf(Stop)or oclIsTypeOf(Move)).
- Constraint 2: the class << move >> can be connected only to the class << trajectory-section >> or <<stop>>. self.allOppositeAssociationEnds>forAll(participant.oclIsTypeOf(Trajectory-section) or oclIsTypeOf(Stop)).
- Constraint 3: the class << stop >> can be connected only to the class << trajectory-section >> or <<move>> or <<surface>>. self.allOppositeAssociationEnds->forAll(participant.oclIsTypeOf(Trajectory-section)or oclIsTypeOf(Surface)or oclIsTypeOf(Move)).

- Constraint 4: the relationship << inside >> can bconnected only to the class << stop >> to the class <<surface>>. Self.associationEnd.participant->forAll(oclIsTypeOf(Stop)orclIsTypeOf(Surface)).

4 Trajectory data warehouse conceptual modeling

The trajectory data warehouse conceptual modeling of our running example is represented as a trajectory fact class in a shared aggregation relationship with n dimension classes. Those latter are composed of trajectory components classes and other domain application classes. The following table represents the main elements of a mobile hospital trajectory data warehouse conceptual model.

Table 2: Trajectory data warehouse elements' description

Element	Description
Trajectory fact	The trajectory is the subject of analysis in our application, thus the fact table is the trajectory. This latter has an identifier composed of primary keys of all dimensions attached to it and some measures that are mainly: duration-trajectory, t-begin-trajectory, t-end-trajectory.
Mobile hospital dimension	The mobile hospital represents the moving object that is an axe of trajectory analysis.
Stop dimension	A stop is a basic axe of analysis. In fact, the moving object that is the mobile hospital accomplishes its mission in a given stop.
Move dimension	A move is a part of the whole trajectory and is proper to the mobility phenomena. An important characteristic of the moving object is the continuity of movement. Let is mention that it is impossible to store an infinite set of positions of a continued movement but from a minimal set of representative points of continued movement we can deduce all other points in a rough way.
Doctor dimension	The doctor plays an important role in such application. In fact, he is the actor who determines if a woman in a given trajectory has the breast cancer or not. Such result is very important in the analysis step.
Patient dimension	The patient is a main actor in our application. In fact, according to their mammography results, a set of decisions can be taken by decision makers.

Time dimension	Time refers to the date where the trajectory and its components took place. This dimension is of big importance in the analysis of a mobile hospital.
location dimension	Location refers to the place where the stops and moves happened. Typical hierarchies are delegation, regional government and country.

The following schema represents the mobile hospital trajectory data warehouse conceptual modeling.

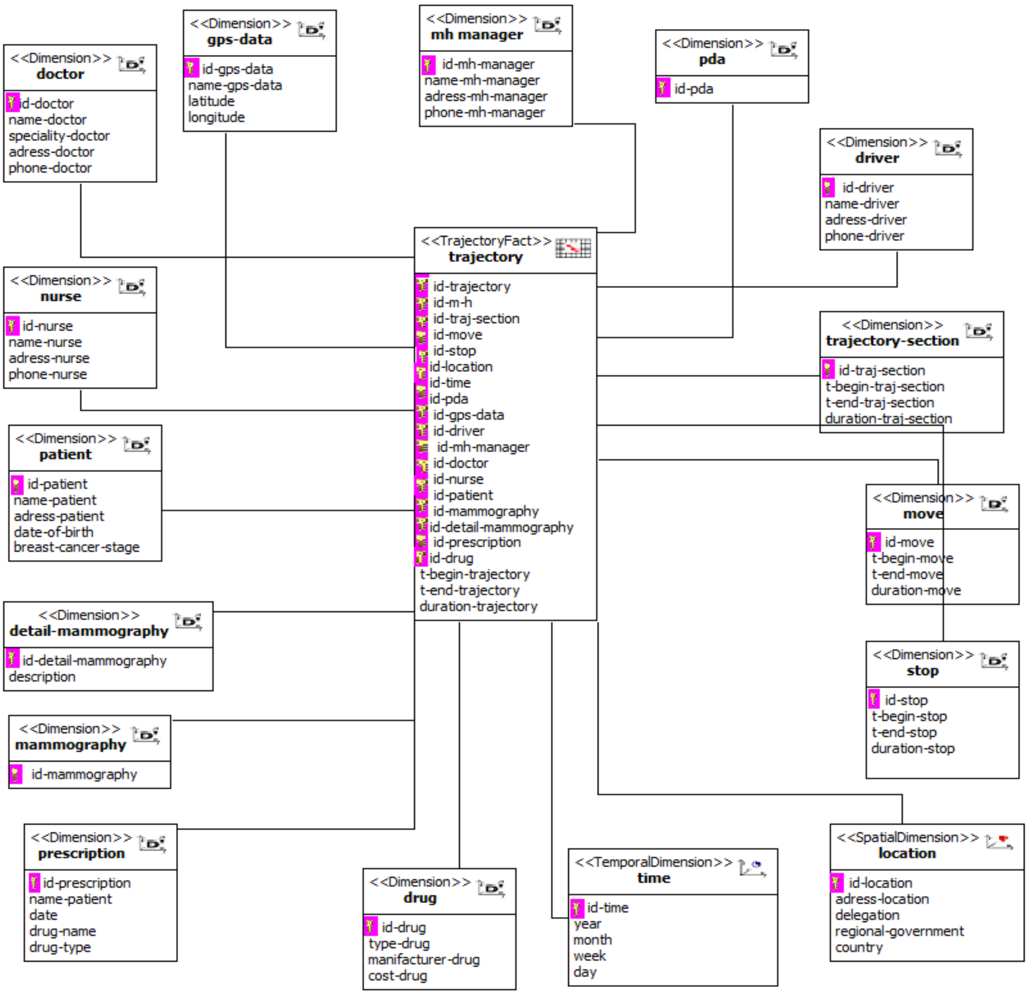


Figure 2. Mobile hospital's trajectory data warehouse conceptual modeling

We defined some constraints for the stereotypes of the trajectory data warehouse diagram in order to use it correctly.

- All attributes of a <<TrajectoryFact>> class must be of type Object IDentifier {OID} or Fact Attribue {FA};

self.feature->select (oclIsKindOf(Attribute))->forAll(oclIsTypeOf(OID) or oclIsTypeOf(FA)). Consequently a fact attribute belongs only to the trajectory fact dimension; self.owner.oclIsTypeOf(Fact).

- A <<TrajectoryFact>> class connects only dimension or temporal dimension or spatial dimension class self.allOppositeAssociationEnds->forAll(participant.oclIsTypeOf(Dimension)oroclIsTypeOf(TemporalDimension)or oclIsTypeOf(SpatialDimension)).
- A <<Dimension>> stereotype indicates that the dimension class is related to the application domain. We defined a constraint for the stereotype <<Dimension>> in order to use it rightly. In fact, all dimension class must be attached only to trajectory fact class;

self.allOppositeAssociationEnds->forAll(participant.oclIsTypeOf(TrajectoryFact).

- All attributes of a <<Dimension>> class must be of type Object IDentifier {OID} or Dimension Attribue {DA};

self.feature->select (oclIsKindOf(Attribute))->forAll(oclIsTypeOf(OID) or oclIsTypeOf(DA)). Consequently a dimension attribute belongs only to the dimension; self.owner.oclIsTypeOf(Dimension).

- A <<TemporalDimension>> stereotype indicates that the dimension represents a temporal component such as date, day, month, year and time. We defined a constraint for the stereotype <<TemporalDimension>> in order to use it right. In fact, all temporal dimension class must be attached only to trajectory fact class;

self.allOppositeAssociationEnds->forAll(participant.oclIsTypeOf(TrajectoryFact).

- A <<SpatialDimension>> stereotype indicates that the dimension represents a spatial component such as country, delegation, regional government and location. We defined a constraint for the stereotype <<SpatialDimension>> in order to use it correctly. In fact, all spatial dimension class must be attached only to trajectory fact class;

self.allOppositeAssociationEnds->forAll(participant.oclIsTypeOf(TrajectoryFact).

- A <<DimensionLevel>> stereotype indicates that the level represents a dimension hierarchical level or a granularity level. For example for the dimension country we have delegation, district and regional government as dimension levels. We defined a constraint for the stereotype <<DimensionLevel>> in order to use it correctly. In fact, all dimension level class must be attached only to dimension or temporal dimension or spatial dimension class;

self.allOppositeAssociationEnds->forAll(participant.oclIsTypeOf(Dimension)or oclIsTypeOf(TemporalDimension)or oclIsTypeOf(SpatialDimension)).

5 Trajectory UML profile realization

We propose in this section to add some extensions (stereotypes, icons) to UML diagrams (class diagram, sequence diagram). This leads to our trajectory profile which is based on UML. We were inspired by the work of [1, 15] to extend the UML profile in order to bring to light the trajectory and its environment.

An UML profile allows specializing UML in a precise domain. Our trajectory UML profile is composed of two diagrams which are the trajectory data sequence diagram and the trajectory data class diagram. The first diagram has the aim to show trajectory data classes and interactions between them. In the following table, we describe our new stereotypes and icons for each class which is in interaction with the trajectory class and the moving object class.

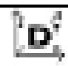
Table 1. Stereotypes and icons of the trajectory data class diagram

Elements	***Stereotypes***	***Icons***
Trajectory	«trajectory»	
Trajectory-section	«trajectory-section»	
Stop	«stop»	
Move	«move»	
Pda	«pda»	
Gps	«gps data»	
Location	«surface»	
Mobile hospital	«moving object»	
Doctor/nurse	« Medical staff »	
Driver/manager	«actor»	
Patient	«suffering»	

The second diagram has the aim to show trajectory data warehouse classes. In the following table, we describe our new stereotypes and icons for each class.

Table 2. Stereotypes and icons of the trajectory data warehouse class diagram

Stereotypes	***Class type***	***Icons***
«trajectoryFact»	UML class	

<<Dimension>>	UML class	
<<TemporalDimension>>	UML class	
<<SpatialDimension>>	UML class	
<<DimensionLevel>>	UML class	

To implement the trajectory data diagram and the trajectory data warehouse class diagram, we used the open source platform called StarUml. This latter is extensible since it uses the XML. In fact, StarUML allows adding new functions which are adaptable to users' needs. We extended a new approach of UML called tdw (trajectory data warehouse) approach and a new profile. The "tdw approach" defines new types of diagrams (trajectory data sequence and class diagrams) and their order of appearance. The trajectory UML profile is used to widen the capabilities of UML to express specific elements in a certain domain.

6 Conclusion

In this work, we presented a pre-conceptual modeling of a trajectory and its components that are trajectory-sections, stops and moves in order to provide assistance to designers in defining the conceptual model of a moving object's trajectory. We proposed also a new profile called trajectory profile which is based on UML. We defined in this latter a trajectory data and a trajectory data warehouse class diagrams with new stereotypes and icons to enhance the conceptual trajectory data level. At this level we can say that the success of trajectory data warehousing process rests on a good conceptual modeling of schema. In fact, the trajectory data warehouse schema will determine the analysis possibilities. However the schema of the trajectory data warehouse can become undefined due to the fact that the trajectory data warehouse integrates heterogeneous information sources which can often change their content and their structure. Resolving such problem is a must and will be treated in future works.

References

1. S .Spaccapietra., C. Parent, M. L. Damiani, J. A. de Macedo, F. Porto, C. Vangenot.: A Conceptual View on Trajectories. Research Report. Ecole Polytechnique Fédérale, Database Laboratory, Lausane, Switzerland. May, 29th, (2007)
2. J .Fernando. Braz. :Trajectory data warehouse proposal of design and application to exploit data. IX Brazilian Symposium on GeoInformatics, Campos do Jordão, Brazil, November 25-28, 2007, INPE, p 61-72 (2007)
3. J .Trujillo., M. Palomar et J. Gomez, I.Y. Song. : Designing data warehouses with OO conceptual models. IEEE Computer 34 (12) 66–75 (Special issue on Data Warehouses)

(2001)
4. S .Brakatsoulas., D. Pfoser, and N. Tryfona.:Modeling, storing and mining moving objects databases. In International Database Engineering and Applications Symposium (IDEAS), pp. 68-77 (2004)
5. I.Sandu Popa,Ahmed Kharrat Karine Zeitouni. : Un système de gestion de données de capteurs à Localisation Mobile ». Conférence SAGEO.(2008)
6. A .Bouju., F. Bertrand, V. Mallé-Noyon, S. Servigne, T. Devogele, C. Ray, H Martin et J. Gensel. : Gestion de données spatio-temporelles au sein de bases de données capteurs ». Conférence sur les Technologies de l'Information, de la Communication et de la Géolocalisation dans les Systèmes de Transports Bretagne (2009)
7. E. Gaudreau, Bruno Agard, Martin Trepanier et Pierre Baptiste. : Pilotage réactif des systèmes de production à l'aide de capteurs intelligents ». 6e Congrès international de génie industriel- Besançon (France). 7-10 juin (2005)
8. S. Orlando, Renzo Orsini, Alessandra Raffaetà, Alessandro Roncato and Claudio Silvestri. "Trajectory Data Warehouses: Design and Implementation Issues". Journal of computing science and engineering, vol.1, no.2, pages 211-232. (2007)
9. X .Meng. Z. Ding. "DSTTMOD: A Discrete Spatio-TemporalTrajectory Based Moving Object Databases System". LNCS 2736, (Springer verlag). (2003)
10. S. Lujàn-Mora, Juan Trujillo et Il-Yeol Song. "A UML profile for multidimensional modelling in data warehouses".: Data & Knowledge Engineering (DKE), 59(3), 725-769. (2006)
11. G .Pestana. M. Mira da Silva.: Multidimensional Modeling Based on Spatial, Temporal and Spatio-Temporal Stereotypes". ESRI International User Conference San Diego Convention Center, California. (2005)
12. Exemple d'application en foresterie. : Ingénierie des systèmes d'information, pages 89-111. (2002)
13. S. Nebojsa, Han Jiawei et Kosperski Krzysztof.: Object-Based Selective Materialization for Efficient Implementation of Spatial Data Cubes. In: IEEE Transactions on Knowledge and Data Engineering, Vol. 12, n° 6, p 938-958. (2000)
14. C .Parent., S. Spaccapietra and E. Zimanyi.: Conceptual Modeling for Traditional and Spatio-Temporal Applications - The MADS Approach. In Springer Verlag, (2006)
15. W .Oueslati., J. Akaichi. : Mobile information collectors' trajectory data warehouse design. International journal of managing information technologies. P 1-20.(2010)