International Journal of Innovation and Scientific Research
ISSN 2351-8014 Vol. 10 No. 2 Oct. 2014, pp. 536-551

http://www.ijisr.issr-journals.org/

Mobile Data Warehousing: A Survey

Ferdaous Jenhani and Jalel Akaichi

Bestmod Laboratory,
Departement of Computer Science,
University of Tunis,
Higher Institute of management of Tunis,
Bouchoucha, Tunis, Tunisia

ABSTRACT: Capturing mobile data is becoming feasible thanks to the frequent and widespread use of mobile devices embedded with positioning technologies. This phenomenon raises the need to handle efficiently, for decision making purposes, data becoming more and more mobile.
The data warehouse technology constitutes the core of modern decision making systems, however, it must be adapted to handle mobile data specificities which bear the mobile data warehousing subfield.
The goal of this paper is to explore, through a survey, mobile data warehousing domain in order to provide to research community an overview about main issues handled in this new area by concentrating on conceptual modeling one.

KEYWORDS: data warehousing, mobile data, data warehouse, mobile data warehousing, conceptual modeling.

1 INTRODUCTION

Nowadays, information is available with huge quantities in many containers such as transactional and operational data bases, web, organization's information systems…these containers are rich in term of data but poor in term of information. In fact, information sources are numerous, heterogeneous and dispersed containing a lot of detailed data difficult to read, to manage and to analyze with human's abilities and traditional analytical tools which results on the loss of the informational data needed for decision making purposes. The solution to this omnipresent phenomenon is to perform a unique repository containing only consistent data extracted from these diverse sources.

The data warehouse is the adequate solution, it is a new concept introduced by Bill Inmon and defined as 'a subject oriented, integrated, time variant, non volatile collection of data in support of management's decision making process' [22]. It is the core of decision making system and used to support difficult applications such as planning and forecasting since it provides the useful information always ready, easily and directly accessible for querying as well as analysis. It stores summarized data, and provides the meaningful information, still missing in operational data bases containing millions records of detailed and individual transactions. It is based on a multidimensional structure that supports OLAP and data mining and other exploitation techniques dedicated to transform summarized data into useful knowledge for decision making, planning tasks, prediction... Therefore, since its appearance and development, it becomes the solution in many fields aiming to ameliorate their activities based on the best decision at the right time taken by the right person.

However, today, we are face to a new challenge which is how to store and analyze dynamic data. In fact, the continuous mobility of professionals make the use of mobile technologies as a necessity in order to accomplish their missions and facilitate business. Also, scientists and ecologists are interested always to the comprehension of natural phenomena closely related to the movement of animals, stars, and even planets, mountains, oceans… tracked by satellites. Vehicles are also equipped with GPS receivers capturing their positions while moving. In summary, sensors embedding used technologies make production of huge amounts of spatio-temporal data. These later, commonly called trajectory data are cumulating every moment with huge quantities in information systems of ubiquitous applications which raise the need to efficiently store and analyze them in order to extract the useful knowledge. This later is still missing but always required for decision making purposes in order to understand complex phenomena such as seasoned migrations, traffic jams, climate changes, people movements, customer behaviors, causes of some unusual habits, causes of diseases…. The data warehouse is the solution since it is widely recognized as the engine of decision support systems. However, it is not conceived to support the complex continuous spatio-temporal nature of mobile data which imposes its adaptation to the mobile context point of view conceptual and logical modeling, indexing, aggregation functions, ETL process, querying and analysis techniques and tools, implementation utilities....

In this paper we aim to investigate main works handling mobile data warehousing issues, and we focus our survey on conceptual modeling proposals following the subsequent organization. Section 2 provides an overview on data warehousing technology and most famous conceptual models proposed in the literature. In section 3, we explore spatial and spatio-temporal data warehousing and its main conception requirements. In section 4, we handle trajectory data warehousing field, which is a particular kind of spatio-temporal data warehouse demanding specific interest. Finally, in section 5, we conclude our work.

2 DATA WAREHOUSING

To move toward an integrated environment and abandon the operational and non organized one, a data warehousing process should be conducted. The data warehousing is defined as the process which encompasses architectures, algorithms and tools for bringing together selected data from multiple data bases or other information sources into a single repository called a data warehouse, suitable for direct querying or analysis [66]. Authors in [11] affirmed that the 'Data warehousing is a collection of decision support technologies, aimed at enabling the knowledge worker (executive, manager, analyst) to make better and faster decision'. They described a data warehousing architecture with the process of Extract, Transform and Load data (ETL) from operational data bases and external sources into a unique integrated repository called the data warehouse. Since its appearance, the data warehouse becomes the major interest of the database industry from design to implementation.

2.1 THE FOUNDATION OF THE DATA WAREHOUSE TECHNOLOGY

The multidimensional structure is the modeling foundation of the data warehouse technology. It consists on data storage in a cube permitting visualization and analysis of data from different perspectives. The glossary of the multidimensional paradigm is composed of fact, dimensions, measures and hierarchies. In the following a concise definition of each term of the multidimensional glossary:

- **Fact**: The fact table represents the fact of interest or the subject of analysis including numerical attributes called measures and other descriptive ones generally the keys of related dimensions.
- **Dimensions**: The dimension tables represent attributes closely related to the focus of analysis (fact) and different axes of measures analysis (Product, store and time dimensions are crucial elements in sales-DW where the fact table represent sales themselves). A dimension is organized in members; each member is a specific attribute or a group of attributes which values are defined in a specific domain. For example the time dimension could be organized on day, week, month and year members. Members are organized in hierarchies allowing the exploitation of measures under different granularities by applying roll-up and drill-down OLAP operations on them. Example of hierarchies: City<State<Country for space and Day<Week<Month<Quarter<Year for time. Hierarchies over dimensions are useful in constraining the values of members. A dimension member could include also descriptive attributes such as year-type in year member. The Measures; Called also fact attributes having numeric values such as number of clients, number of transactions, quantity sold… which are analyzed along the different referenced dimensions and their levels by applying aggregate functions on their values kind of Sum, Average, Min, Max,… in the case of no measure affected to the fact table, this later records a default measure which is the number of occurrence of the object fact type satisfying a given condition (count).
- **Hierarchies**. The organization of dimension' members in hierarchies is necessary for the analysis of data at different granularities which enable the precise detection of the source of a given problem, for example the province in which the quantity of sales is the lower. This hierarchical organization is necessary for applying OLAP roll-up and drill-down operations exploring different details of data.

A multidimensional model is implemented in the well known logical designs namely the star, snowflake or constellation scheme defined in the following:

- **The Star Schema:** A star schema is characterized with a unique central fact table and surrounding single dimension tables. Each dimension is rooted in a unique entry (key), and each reference in the fact table points to the root of the corresponding dimension. In this schema, the hierarchy along dimension tables is not explicitly represented.
- **The Snowflake Schema**: This schema is the refinement of the star one in which dimensions are normalized and members are explicitly modeled. Indeed, each dimension from the star schema is broken into set of members organized in an application-defined hierarchy. Beside the increasing number of tables, this structure provides facilities of maintenance.
- **The Fact-Constellation Schema:** In this structure, we have more than one fact table sharing some dimensions. This structure is adopted for huge data warehouses.

2.2 DATA WAREHOUSE CONCEPTUAL MODELS

To respond to the DW multidimensional modeling requirements, designers use often the abovementioned relational models. Concerning the conceptual modeling, there is not till now an agreement about a conceptual model offering both graphical representation and formal definition [31]. In fact, conceptual modeling phase in the world of data warehousing is of crucial importance especially in reducing costs of maintenance of the system in the case of any technological changes. However, the well known conceptual models, namely the Entity-Relationship (ER) and the Object Oriented (OO) model, don't represent the semantic of multidimensional paradigm. Nevertheless, many attempts were proposed to adopt them, for instance, the ME/R-Multidimensional Entity Relationship Model introduced in [58]. It is a specialization of the ER model to represent multidimensional properties. The authors propose a minimal set of extensions which are powerful enough to represent the semantic of multidimensional paradigm (facts, measures, dimensions). The approach consists on two extensions; a special entity set and two relationship sets. The entity set defines the dimension level grouping qualifying data. The two relationship sets are, first the fact relationship set defining the n-ary relationship between dimensions and models at the same time the quantifying data. The second is a special binary relationship set connecting dimension levels between each others called also the rolls-up to relationship set and define a DAG-directed acyclic graph between them. Seeking for a clear and distinguished graphical representation to these new concepts, authors propose special graphical components.

The above contribution favorites the explicit representation of dimension levels by introducing the notion of dimension level set defining hierarchical classification thanks to rollup relationships between levels. Moreover, thanks to the fact relationship set, it is possible and easy to model complex structure called 'multi-cube model' by including additional facts.

Tryfona and colleagues introduced the starER conceptual model to meet main data warehouses conceptual modeling requirements revealed by authors in the following points [62]:

- Represent facts and their characteristics;
- Connect the temporal dimension to facts which are always closely related to the aspect time;
- Represent application objects, their properties and relationships among them (specialization/generalization, aggregation, membership);
- Establish semantic links between facts and application objects modeled;
- Detect dimensions and their possible and indispensable hierarchies;

Based on the following requirements and with respect to the semantic of the ER model, authors proposed some concepts to adapt it to design efficiently the data warehouse. StarER is the result of combining the ER model, to which most designers and users are familiar, and the efficient structure of the star schema dominant in the world of data warehouses. Main constructs of StarER are the fact set, the entity set, relationship set and attributes. The major weakness of this approach is that each component of the model is represented separately with its own graphical shape, and the variety of relationships between entity sets which complicate the readability and understandability of the final conceptual or semantic design especially in the case of large data warehouse.

These weaknesses are coped in an important contribution in the literature of DW conceptual modeling which is of Malinowski [34] called MultiDimER. In this later, a minimum set of relationships is used which is of the ER formalism (many-to-many, many-to-one…). It is a conceptual model based on the ER graphical notation and defined by authors as a finite set of dimensions and fact relationships. It consists on a multidimensional adaptation of ER model to suit the DW philosophy. Indeed, they integrate specific symbols to the ER notation (entity types, attributes, relationship types) to represent the fact relationship, the dimensions and their possible levels as well as the relationships among them, the cardinalities and the analysis criteria. The strength of this contribution is that the MultiDimER is based on a simple graphical notation and clear formal definition to which most approaches lack.

However, this is not the case of CGMD-Conceptual Generalized MultiDimensional model proposed in [23] which is based on a strong formal definition. In fact, the author have extended the ER model (EER) with simple and multidimensional aggregated entities, and combined it with the GMD-Generalized MultiDimensional model which is a logic-based multidimensional model to obtain a new DW design called CGMD. It is based on the ER semantic since it includes the notion of entities, relationships, attributes, associations and cardinalities. According to the GMD abstracts, each entity is a fact, each relationship is a dimension, hierarchical levels relationships are defined by is-a link and measures are attributes directly associated to the fact entity. Nevertheless, the major limitation of this model resides in its complex graphical representation.

In [14] and [15] writers build a conceptual model by using the ER schemas, implemented in existing information systems of the enterprises, to derive a DFM-Dimensional Fact Model. The proposed model is a tree-structured fact schemes designed as a directed acyclic and weakly connected graph. Authors defined a methodology to be followed in a data warehouse development which consists on an iterative process to transform the E/R model into a fact scheme. The process begins with identifying facts then constructing the main components of their fact tree namely attribute tree, dimensions, fact attributes and hierarchies. The DFM is designed with simple graphical notations and based on clear formal definition.

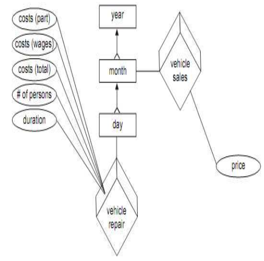
In [20] authors developed the same idea by defining a complete process for data warehouse design starting from user requirements analysis passing with the conceptual and logical design until reaching the implementation phase, which resembles the traditional data base modeling process. Concerning the phase of conceptual design, main issues to be handled are the definition of measures, the design of dimensional hierarchies (path aggregation) and the determination of summarizability constraints, with always determining functional dependencies from dimension levels to measures, and between dimension levels each others. They distinguish also between simple and multiple hierarchy, optional and alternative groups of aggregation paths. Their procedure shows how to derive a graphical multidimensional diagram in a multidimensional normal form (MNF) from an operational data base designed in E/R schema.

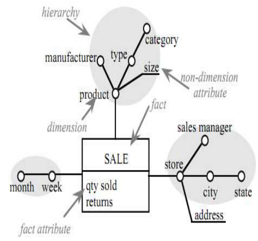
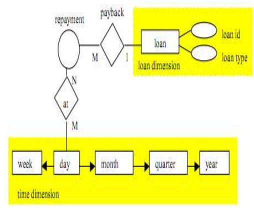
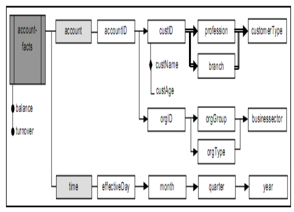
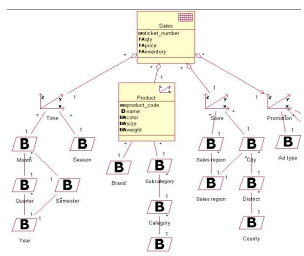
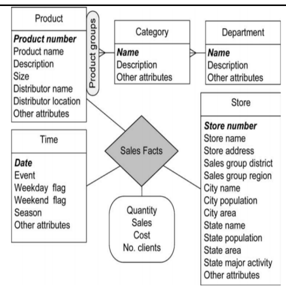
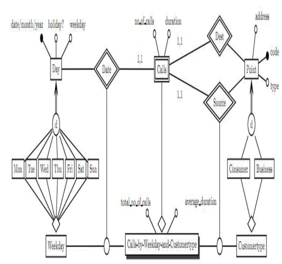
Within UML-based conceptual models, the most famous approaches are of Trujillo and colleagues. In [29] authors proposed UML extensions for object-oriented multidimensional modeling. This extension is performed thanks to stereotype mechanism, tagged values and constraints expressed in OCL-Object Constraint Language. In addition to a set of Well-Formedness rules managing new elements added and determining the semantic of the model. Stereotypes and icons allow an expressive representation of different constituent elements of a multidimensional model namely fact classes, dimension classes, hierarchy levels and attributes. Dimension level classes (stereotyped Base classes) should define a directed acyclic graph rooted in dimension class. Concerning relationships, the aggregation relates facts to dimensions, and association/generalization relates dimension levels to each other.

In another work of the same team [30], an UML package is proposed to facilitate the modeling of large data warehouse systems. In fact, they suggest a set of UML diagrams (package) extended with the aforementioned stereotypes, icons and constraints (OCL) to cope with multidimensional modeling and consequently designers will not be limited only to the class diagram.

In the following table we compare between the abovementioned conceptual approaches by considering the major criteria diversely handled in all of them mainly: the graphical representation, the multidimensional structure foundation of a data warehouse (facts, dimensions, measures, and hierarchies), fact/dimension relationships, and dimension levels relationships. The study is accompanied with an example to visualize the conceptual design since the conceptual modeling is an art rather than a science:

Table1. A comparative study between DW conceptual models

DW Conceptual Model	Graphical representation	Multidimensional Structure Description	Fact/ Dimension relationship	Dimension Levels relationship	Example of the Conceptual Model
Sapia et al., 1998 (ME/R)	Based on ER model	<u>*Facts:*</u> *an n-ary relationship set (d*iamonds) <u>*Dimensions:*</u> a dimension is defined with a set of dimension levels (rectangles) <u>*Hierarchies:*</u> explicitly represented with 'roll-up to' relationship set <u>*Measures*</u> : related to the fact relationship (ovals)	n-ary relationship	-to-one	

Golfarelli et al., 1998 (DFM)	Tree-structured fact schemes	*Facts*: important event in the enterprise world (box) *Dimensions*: conceptualization of discrete attributes (vertex attached directly to fact) *Hierarchies*: explicitly represented with sub-trees rooted in dimensions	Many- to-one	-to-one	
Tryfona et al., 1999 (StarER)	Based on ER model	*Fact set*: real world data generated over time as facts thus connected always to time (circles) *Entity set*: real world objects (rectangles) *Hierarchies*: explicitly represented with membership relationships *Attributes*: static properties of fact set, entity set and relationship set.	many-to-many many-to-one	strict, non-strict, complete or non-complete memberships	
Husemann et al., 2000 (MNF diagram)	Multidimensional Normal Form (MNF)	*Fact schemata:* atomic information elements *Dimension levels*: derived by determining all functional dependencies *Hierarchies*: explicit aggregation paths on dimension levels	Edges	Simple arrows	
Lujan-Mora and Trujillo, 2001	Based on UML notation	*Fact classes* : contain identifier (OID) and Fact attributes representing measures *Dimension classes*: contain identifier, Dimension attributes and descriptors *Hierarchies* : Base classes forming a DAG	Aggregation: Many-to-many Many-to-one	Association/ Generalization	
Malinowski et al., 2006 (MultiDim ER)	Based on ER model	*Fact relationships:* the focus of analysis (diamond) *Dimensions*: abstraction for grouping data having same semantic characteristics *Hierarchies*: explicitly represented with a parent/child relationship	Zero-to-many generally omitted from the schema	-to-many	
S.Kamble, 2008 (CGMD)	Based on ER model	*Facts* : aggregated weak entities (double lined rectangle) *Dimensions*: weak relationships (double lined diamond) *Hierarchies*: Aggregated entities links between levels explicitly represented	-to-one -to-many	Aggregation	

3 Spatio-Temporal Data Warehousing

Spatio-temporal data are captured by positioning technologies (GPS) used by an object in motion (people…) or connected to an entity in continuous movement (stars, animals, vehicles…). Moving objects are defined through the literature as follows:

- **Definition 1**: 'A mobile object can be defined as an object that changes its geographical position as time passes. Such object integrates spatial and temporal characteristics' [53].
- **Definition 2**: 'when we try an integration of space and time, we are dealing with geometries changing over time' [12].
- **Definition 3**: 'Objects that change position or extent continuously termed moving objects' [18].

Objects moving continuously over time are of particular interest and hard handling compared with discretely moving objects. In fact, discrete changes are widely treated and easily modeled in classical data bases, however, continuous changes (continuous updates of each change on moving object state in the data base) are difficult to support.

In fact, nowadays tracking the movement of a spatial object is becoming easy, possible and cheap thanks to the popularity of location-aware devices used by ordinary people to communicate between each others, or by scientists to collect data necessary in some experiences, or in other applications belonging to different fields. The frequent use of such devices allows capturing huge amounts of movement data, stored in the information systems of organizations providing services through wireless networks. These data are precious resource that must be rightly exploited in order to extract interesting patterns bringing the knowledge helping managers to make decision in many mobility based applications such as transportation planning, traffic management, location based services, mobile marketing… as well as to understand various social and natural phenomena for instance unknown causes of certain diseases or natural disasters, animal and birds movements behavior, customer comportment…

The data warehouse provides the adequate support for such analysis, however, it is not conceived and developed to store and manage mobile data regarding its dynamic nature. Therefore, traditional data warehouse technology must be adapted to mobility context, point of view modeling and design, ETL process, aggregation of spatial and temporal measures, analysis and querying tools and techniques, implementation utilities… which are opened issues in the mobility data warehousing field. The first and basic question is how to incorporate space and time dimensions in the data warehouse model with always keeping its coherence, semantic, and dimensional structure? Most of applications solve the problem by adopting directly the well-known relational models in the world of data warehouse namely the star, snowflake and fact constellation schemes providing a concrete vision of the target application. However, conceptual modeling of such data is very important and not straightforward regarding their particularity.

3.1 Spatial Data Warehouse Conceptual Models

Let's start from the beginning of the recent history, with the development of GIS-Geographic Information Systems and the advance of geospatial data, a real need to analyze spatial data and transform it into knowledge is born, phenomena closely related to the importance of spatial data in the decision making process. However, GIS allow to store, analyze and visualize geospatial data as points, lines and polygons defined with x and y numeric coordinates, and cannot support a decision process based on geo-localization of some problems. Therefore, works on integrating spatial dimension in the data warehouse was largely developed. In fact, spatial data are geometric components defined with X and Y coordinates and other semantic characteristics that should be modeled (names of places). This spatial data is often stored in data bases in a string format like 'street address' having generally implicit hierarchy. In the new situation, spatial objects' hierarchy should be explicitly represented. The spatial data types have facilitated the task of modeling spatial data warehouses by extending existing models by the stated above new data types as well as their properties. The SDW are the result of coupling the spatial technology (GIS) with the non spatial-technology namely the data warehouse. In [4] fundamentals of spatial data warehousing are defined. Indeed, a SDW is based on spatial dimensions and spatial measures. Authors consider three types of spatial dimensions:

- Non geometric spatial dimension: dimensions contain only nominal data to locate a spatial phenomenon such as name of municipalities whose generalization are also nominal such as district and province. Such dimensions could be implemented in non spatial technologies where the cartographic representation is not of interest.
- Geometric-to-non geometric spatial dimension: this kind of dimension is geometric at the finest level but its generalization is non geometric. Thus, the finest level is represented with its geometric shape e.g. polygon. For example, a province represented with a polygon in a map.

- Fully geometric spatial dimension: A dimension whose primitive as well as each of the generalizations levels are geometric.

Concerning measures, a spatial data cube contain numerical and spatial measures:

- Numerical measures: represents numerical data for example the quantity sold, the revenue…
- Spatial measures: a spatial measure is a pointer to spatial objects having similar characteristics.

Han is the pioneer of the SDW concept [19] by extending the traditional data warehouse and OLAP operators with spatial character. For the SDW design, he used the star schema and tried to extend it with spatial dimensions and measures. Spatial dimensions are modeled in two cases in a spatial data cube; spatial-to-non spatial dimension or spatial-to-spatial dimension. Spatial measures are measures containing one or more pointers to spatial objects. He aimed along this work to prove the weakness of existing systems such as GIS to support spatio-temporal decision process, and to find the suitable design and implementation of a spatial data warehouse as well as efficient analysis of spatial data stored in a multidimensional structure (Spatial OLAP). In fact, regarding the importance of spatial data in decision support systems and the disability of GIS to support such process, new dedicated and adopted tools must be developed. Face to this need, OLAP systems was integrated with GIS to bear SOLAP systems. While OLAP technology is defined in [10] as “a category of software technology that enables analysts, managers and executives to gain insight into data through fast, consistent, interactive access to a wide variety of possible views of information that has been transformed from raw data to reflect the real dimensionality of the enterprise as understood by the user”, SOLAP tools are identified in [56] as ‘a visual platform built especially to support rapid and easy spatio-temporal analysis and exploration of data following a multidimensional approach comprised of aggregation levels available in cartographic displays as well as in tabular and diagram displays’.

Motivated with the contribution of Han, subsequent works in the SDW subfield appeared, for instance [40]. In the later paper authors proposed two alternative solutions to model heterogeneous data in multidimensional structures where the goal is to support a geographic knowledge discovery process in the field of forestry. Before the DW design process, they defined the hierarchy of spatial dimension which depends on the forest management rules, and of descriptive dimensions (age, species, disease, density, height). In forestry, the facts are measured area thus the measure is volume. Concerning multidimensional modeling of forestry spatio-temporal data, two solutions are proposed. The first is a temporally integrated multi-epoch cube involving only three descriptive dimensions (age, height and species) and the spatial one. The temporal variation is integrated within these later. Second solution consists on three spatially-specific mono-epoch cubes, one for each inventory based on forest stands, and one spatially-unified multi-epoch cube based on regular cells. The design is represented with a fact constellation schema composed of four fact tables sharing some dimensions. Indeed, it optimizes the data storage besides its complex structure composed of four multidimensional cubes. A prototype of the SOLAP application is developed and a performance comparative study between the two structures is performed. In [33] authors extended the MultiDimER model with MADS graphical notations to represent spatial dimensions, spatial hierarchies and spatial measures. Spatial dimensions are defined as entities having at least one spatial hierarchy. Spatial hierarchy is hierarchy having at least one spatial level, and a spatial level is the one to which the spatial characteristics should be kept. Spatiality is represented using geometry attribute represented with pictograms, having as values spatial data types namely point, area, line… and topological relations linking spatial level such as contains, equals, intersect, adopted from MADS. Non spatial dimensions are called thematic. Spatial fact relationships are fact relationships with the symbol of topological relation (pictogram) to join one or more spatial dimensions. Spatial measures are associated to fact relationships independently if whether it is spatial or not. They are either represented with geometry and define a spatial function used for further aggregation among hierarchies, or numerical values result of spatial or topological computations. An amelioration of this work is in [35] in which authors introduced several kinds of spatial hierarchies. They distinguish simple hierarchies; symmetric, asymmetric and generalized, non-strict hierarchies, multiple alternative hierarchies and parallel spatial hierarchies.

Still in the context of adaptation of multidimensional concepts to spatial data, Spaccapietra and Damiani in [60] propose, in addition to spatial dimensions and facts, the exploitation of granularity on spatial measures also, not only on spatial dimensions. They introduced the multi-level spatial measure concept in a new model called MuSD (Multi-granular Spatial Data warehouse). A multi-level spatial measure is a measure that is represented by multiple geometries at different levels of details. They include also in their proposed model a set of spatial OLAP operators in order to analyze the different levels of granularity of spatial data over dimensions as well as measures.

Authors in [57] proposed a formal description of a spatial multidimensional model. The model describes the main features of a spatial application in a multidimensional space represented with a base cube. Facts and dimensions are defined as entity schemas related between each others with n to n relationships. Writers introduced also the concept of entity instances representing tuples of values associated to attributes of the entity schemas. Entity schema and instances are organized in hierarchies (hierarchy schemas and hierarchy instances). Measures are supported by the model as complex entities and spatial measures as geographical objects thanks to aggregate ad-hoc user defined functions between two entities; one representing the detailed measure and the second representing the measure after the aggregation process. This model is just a formal description of a spatial multidimensional model and don't adopt any specific graphical representation (design). Also, the spatial data are not fully exploited in this formal definition such as topological relationships and type of spatial objects treated. In [34], Malinowski addressed another problem in the field of spatial data warehousing. In fact, to avoid future problems in SDW related to its design and feeding with spatial data, a methodology should be established and followed starting from an efficient user requirements specification. They proposed three approaches for requirements specification; one based on user' requirements, other based on data available in source systems and the third is based on both. In the present work, Malinowski refers to the methodology followed when developing ordinary data bases (requirements specification, conceptual modeling, logical modeling, and implementation). Recent work on this field is [13] in which authors consider that spatiality is a personalization feature since each decision maker has its spatial information need. Their vision is that a unique static spatial multidimensional schema may not cover some needs or may be very large since it supports all needs of decision makers which complicate the schema and harden the task of accessing the spatial information needed. To better satisfy each specific need of decision makers, a personalization process is defined to give a personalized version of spatial data warehouse for each need. In [55] authors present an index structure for spatial data to facilitate the work of data warehouse designer. They extended the traditional set-grouping hierarchy structure to provide grouping schema of spatial data in multidimensional space. In this context, a spatial index tree is introduced organizing spatial data in hierarchies. Such structure permits an automatic generation of spatial hierarchy on spatial dimensions necessary for further aggregations. In this contribution, Rao and his team introduced a structure facilitating the logical modeling step and schema generation of the data warehouse (star schema), and don't provide any graphical notation or formal definition of a multidimensional conceptual model.

3.2 Spatio-temporal Data Warehouse Conceptual Models

In most applications acting in the mobile context and manipulating spatio-temporal data, the space and time are closely related characteristics, so they have to be considered together in the multidimensional modeling task. In other words, dealing with spatio-temporal data means dealing with spatial objects evolving over time, and find a concise formal definition and graphical representation for them. MADS-Modeling of Application Data with Spatio-temporal features [52], [50], [51], [60] is a famous conceptual model supporting spatio-temporal features in an ER-like model. It maintains the semantic and graphical notations of the ER model (Object type is an entity, relationship type is an association, descriptive and key attributes, aggregation links, constrained relationships, cardinalities). It supports spatial and temporal features representation for objects, attributes and relationships. In fact, authors represent spatial feature by the predefined attribute Geometry whose value is a spatial abstract data type (SADT) namely line, point, area, point set, line set, complex area, and represented graphically with icons to enable visual notation. The temporal feature of objects, attributes or relationships is represented with Lifecycle attribute having as values a temporal abstract data types (TADT) such as instants, time intervals…and represented with icons also. MADS supports also constraining relationships which are mainly; topological relationship (disjoint, adjacent, intersect, cross, inside, equal…) between at least two objects which can be either spatial or non-spatial object types. Synchronization relationship (before, during, overlap, meets, starts, finishes…) constraining the binary link between temporal or non-temporal object types. Concerning the granularity of information, MADS define the unit (hour, day, minute…) for time and the resolution for space. In this model mobility of objects is represented with varying data types (space varying attributes such as weather, time varying attribute such as population) represented with functions from time to space or the opposite and represented graphically with special notations. Hence, MADS supports the conceptual representation of spatial and temporal features of objects, attributes and relationships based on visual notation, and consider the movement as a time varying geometry. Besides MADS design is not dedicated to the data warehousing since it don't support constructors of a multidimensional paradigm (dimensions, fact, hierarchy, measures), but, its notations are used in many multidimensional models thanks to their power in expressing space and time features and time-varying attributes and objects. For instance, [63] in which authors extend the spatial MultiDimER model by adopting the MADS abstract data types. Moving objects are considered as moving types defined to describe the evolution over time of spatial objects and attributes by applying the constructor moving () on them. Authors aimed along their work to find a concise multidimensional conceptual model for spatio-temporal data warehouse supporting spatio-temporal data types and moving objects representation. To enable temporal querying, authors performed the extension of their model with temporal features represented with LS and VT pictograms representing temporal levels and temporal attributes, also LS pictogram is used to represent synchronization (temporal relationship) between parent-child levels. Those pictograms (LS and VT) are introduced in [34] in the context of developing Temporal DW. Still with MultiDimER model, in [42] an attempt to adopt the spatial MultiDimER to spatio-temporal context by supporting partial containment and tracking temporal relationships between dimensions. They introduced the concept of Total/Partial containment relationship of a child spatial level in the parent one represented with specific symbols. They proposed also 'µ' symbol allowing an expressive representation of time in their model; main features are time-varying geometries and temporal relationships between spatial levels expressing mobility of objects over time. From Entity-Relationship based spatio-temporal models to UML ones. In [59] authors developed a multidimensional conceptual model to represent people trajectories using UML notation. They performed an approximation to represent trajectory defined by the pair (id zone, id time-interval) based on discrete space and time. The levels of detail are performed by zone subdivision for space and time intervals for time. For the design issue of their spatio-temporal data warehouse, writers used the snowflake schema for modeling population displacement in space and time. The fact is ST-Journey class and two dimensions; space and time and other thematic dimensions. Their solution was implemented to support the HEARTS project aiming at studying human motion by exploring his daily trips and journeys in order to produce statistics on flexible groups by performing querying on the resulting Spatio-Temporal DW. The UML model used at the conceptual level is not expressive enough to represent the multidimensional paradigm of the DW (fact, dimensions, and hierarchies), space-time features and relationships between spatial and temporal entities. Unlike [8], who extended UML diagrams using PVL graphical notations defined by Perceptory visual tool (conceptual modeling tool). PVL is a visual language depicting geometric and temporal properties of objects and attributes. It uses five basic constructs represented by pictograms (point, line, area, punctual time, durable time) introduced to conceptual model schemas using UML stereotype mechanism. This work aimed to find a geospatial repository reflecting as much as possible spatiality and temporality of data, and don't interested on finding a dimensional model suitable for the DW design. In [27], authors proposed a general object-oriented framework for spatio-temporal data modeling in which each object is characterized with when, where and what attributes enriching the model with semantic interpretation of real world dynamic. The approach is based on a class of spatio-temporal objects defined with an identifier, space/time/theme attributes and behaviors. Concerning theme, space and time attributes, they are represented with three tables related with one-to-one, one-to-many and many-to-many relationships between them in order to model the continuous changes of geometry of the theme over time. However, behaviors are divided into spatial changes (stay-in, transform-between, splitting, moving..., spatial relationships (topological: include, intersect…and metric: distance, and temporal relationships (topological: overlap, during…). The proposed model supports also relationships between objects namely association, generalization, specialization and aggregation. This contribution proposes a solution for continuous movement changes modeling but not in the multidimensional environment. Indexing issue was also the interest in the data warehousing field, for instance Papadias and colleagues in [49]. In the mentioned paper, authors proposed an index structure aiming to replace the data cube. They indexed static spatial dimensions, considered as a static set of road segments, using an aggregated RB-Tree. Each region, considered as the finest granularity of spatial hierarchy, is stored only once and indexed in the R-tree, and each entry has a pointer to a B-tree containing its historical aggregated data. They dealt also with dynamic spatial dimensions in which the region can change its extent over time or new regions can appear or disappear (volatile regions). The dynamism of regions alters the spatial tree structure, to cope with this problem, Papadias and his group proposed alternative multi-tree indexes namely the Historical RB-Tree and the 3DRB-Tree. The goal of their work is to save memory space occupied by proposing alternative structures smaller than the data cube.

We propose the following table summarizing and comparing some of the most famous contributions in the spatio-temporal DW conceptual modeling according to these criteria: graphical representation, space feature modeling, time feature modeling, spatial relationship modeling, temporal relationship modeling, and object movement modeling:

Table 2. A comparative study between spatio-temporal DW conceptual models

STDW conceptual Model	Graphical Representation	Space Feature Modeling	Time Feature Modeling	Spatial Relationship Modeling	Temporal Relationship Modeling	Object' Movement Modeling
Brodeur et al., 2000	Based on UML notation	Geometry (point, line, area)	Temporality (punctual time, durable time)	Not treated	Not treated	Not treated
Savary et al., 2005	Based on UML notation	Spatial zones	Time intervals (Units)	Not treated	Not treated	*Journey*, defined with (id-zone, id-TimeInterval)
Pestana and Mira da Silva, 2005	Based on UML notation	Geometry (point, line, area)	Temporality (punctual time, durable time)	Topological relationship	Not treated	Not treated
Vaisman and Zimanyi, 2008 (MultiDimER)	Based on ER model	Spatial Abstract Data Type (point, points, line, region)	Time type (instant, periods)	MADS topological relationship	MADS Synchronization relationship	*Moving types*: *Moving (spatial data type)*
Moreno et al., 2010 (STMultiDimER)	Based on ER model	Geometry	The unit (time instant, ex: *day*)	MADS topological relationship	Track' keeping of assignments between spatial levels	Time-varying geometries

4 TRAJECTORY DATA WAREHOUSING

Spatio-temporal models support a powerful representation of space and time features thanks to abstract data types, nevertheless, the movement of spatial objects defined with data continuously changing is still not supported, they can handle only discrete space and time. In fact, data continuously changing are defined recently with the concept of trajectory data. Through the literature it was diversely defined, some of them consider it a pure spatio-temporal concept and others link it to a certain semantic allowing its high-level interpretation:

- **Definition 1**: "A trajectory is the record of the movement of some object i.e. the record of the positions of the object at specific moments in time." [17].
- **Definition 2**: "A trajectory is a continuous variation in space and time." [65]
- **Definition 3**: "movement can be described as continuous change of position in the geographical space and through comparison between two different temporal instants."[36]
- **Definition 4**: 'A trajectory is the path made by the moving entity through the space where it moves. The path is never made instantly but requires a certain amount of time.' [17]

The above definitions describe the movement of an object (perceived as a point) as a spatio-temporal concept. Moving object's trajectory data is captured by a positioning technology as a sequence of raw spatio-temporal positions, each position is the triple <x,y,t>, Trajectory = {<x1,y1,t1>, < x2,y2,t2>, …, <xn,yn,tn>}.

- **Definition 5**: "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" [60].

Spaccapietra defined the trajectory of a traveling object in a limited and finite time interval, going from a starting instant to a finishing one. Thus, a trajectory can be assimilated to a function from a user-specified time interval to a geometric space: Trajectory: [tbegin, tend] →space. Through his definition, authors give also a supplementary semantic to the trajectory concept by linking it to the goal for which the moving object travels.

4.1 TRAJECTORY DATA MODELS

First attempts on handling trajectory data modeling, storage and management focus on MOD-moving object data bases technology, since traditional data base technology cannot represent adequately data continuously changing. MOD becomes an important line of research based on moving abstract data types mainly moving points and moving regions facilitating the task of modeling moving objects data.

In this context we find in [5] several issues related to trajectory data are handled involving data preprocessing, conceptual modeling, data storage thus logical modeling and data mining. At the conceptual level, MOD is built based on modeling spatial data (roads, buildings…), non-spatial data (thematic information), and trajectories of moving objects. They explore the semantic involved in trajectory data and model them in an object-oriented way. The semantic is captured in term of properties of a unique trajectory (speed, heading, covered area, traveled distance, traveled time), relationships with its spatial environment (stay within, by pass, leave, enter, cross) and relationships among trajectories (intersect, meet, equal, near, far). Then, this semantic is represented in an UML class diagram as functions or operations (example: GetHeading(), Cross()…). This later representation is proved by the authors of the paper as the right representation of continuous changes of moving object data (properties and relations related to trajectories). Besides efforts on modeling moving objects, the semantic related to trajectory data is not fully supported. Spaccapietra and colleagues in [60] addressed this problem for a moving point and considered a trajectory as a first-class object. They performed a semantic segmentation of the spatio-temporal path into trajectories and a trajectory into stops, moves. They expressed the semantic facet of a trajectory with BES concept, move phases delimited with two stops, the network constraints, and the links to the rest of application objects. Concerning conceptual modeling, authors used MADS notations in an ER like model, thus in a transactional environment. They proposed two alternative modeling approaches for trajectory data; one based on a design pattern and the other is based on dedicated data types. The Trajectory Design Pattern offers predefined sub-schema providing basic data structures for trajectory data modeling. It can be adjusted toward the application requirements and connected to the rest of its data base schema. Trajectory Data Type is the alternative solution consisting in encapsulating common components of trajectories a TrajectoryType data type and define methods enabling access to them. Both approaches may be combined to offer an efficient modeling system according to application needs. The purpose of this work is to realize a generic conceptual model permitting an expressive representation of trajectory characteristics, components and constraints succeeding the expression of semantic facet of trajectories. Authors in [44] used the UML formalism to model trajectories as classes defined with displacements over time, so displacements and time also are classes. The trajectory class contains attributes speed, direction, duration…, the displacement class contains starting time, ending time, first position, last position… and the time class contains the attribute temporal hierarchy, expressing the finest granularity of the time. The model is designed in order to support the comprehension of human mobility. In [47] a new UML profile is defined to cope with modeling the semantic of trajectory data. They extended UML class diagram with stereotypes and icons to allow their model certain flexibility needed to succeed the representation of semantic aspects of trajectories and its related concepts inspired from the work of Spaccapietra mainly stops, moves, begin, and end. Their final objective is to model the movement of a mobile hospital vehicle while accomplishing its mission.

Concerning indexing moving objects' trajectories issue, we have [54] in which authors define a trajectory as a set of line segments defined in 3D space (X, Y, T). They proposed two access methods to trajectory data; STR-Tree which is an extension of the R-Tree coping with organizing line segments according to spatial properties (proximity between line segments) and belonging to the same trajectory. Thus, they introduce the problem of trajectory preservation. They proposed an insertion and split algorithm since those of the R-tree are not appropriate to trajectory data. The second method is TB-tree which is a method that strictly preserve trajectories so called trajectory bundle; thus this structure don't consider spatial properties but only line segments belonging to the same trajectory. These structures are proposed to address successfully pure spatio-temporal queries.

In summary, the abovementioned works aim to find an approximation abstraction for trajectories and a conceptual representation in non-dimensional models. However, trajectory data should efficiently exploited and analyzed in order to extract the hidden information useful in making the competitive decision in applications becoming invaded with trajectory data. Since the data warehouse technology is the foundation of decision support systems, new line of research handling trajectory data warehousing is handled. In the subsequent section, we investigate main proposals on trajectory data warehouse conceptual modeling.

4.2 TRAJECTORY DATA WAREHOUSE CONCEPTUAL MODELING

TrDW is a very recent concept introduced by Braz and colleagues [7]. It is a challenging new area facing three main issues; one related to conceptual and logical modeling, the second related to the loading phase (ETL process) and the last to the computation of measures (aggregations of trajectory-oriented measures) for future OLAP purposes. Concerning the conceptual modeling, there is not until now considerable interest beside the importance of such phase, and designers are satisfied by only representing the TrDW by adopting the well-known relational schemes at the logical level namely star, snowflake and fact constellations and extending them with space and time dimensions. In this [6][7][45]authors handle the ETL process. In fact, they proposed the sampling for trajectory reconstruction; they dealt with stream of samplings assuming that a trajectory is a finite set of observations. They adopted the classic star schema for TrDW design defined with one fact table containing the main trajectory-oriented measures and three dimension tables (X and Y for space and T for time fed with intervals of sample points). Their major focus is in measure aggregation problem namely the distinct count solved by proposing two alternative aggregate non-holistic functions; Presence Distributive and Presence Algebraic consisting on combining a bounded amount of measures stored in the base cells of the data cube. Writers handled also the loading phase consisting on feeding the data warehouse with sub-aggregate measures by proposing loading based on updating the data warehouse on the basis of single trajectories.

Similar work is of Marketos and his group, in [36][37][38][39] who used the classic star schema multidimensional model to design the TrDW defined with one central fact table and three dimensions (Object-Dim for the user profile, Space-Dim for the space and Time-Dim for the time), in which the authors integrate some semantic information and not only the geometric data captured by a positioning technology in contrary to the work of Braz and colleagues . The process of TrDW development begins with reconstruction trajectories consisting on filtering the raw time-stamped locations that arrive in bulk sets to decide if the data is to be appended to an existing trajectory or not, then the transformed trajectories are stored in MOD.

Concerning the loading phase, Marketos proposed two alternative approaches; the cell-oriented and the trajectory-oriented ETL processes. Authors coped also with the problem of holistic aggregate function by adopting the solution of Braz.

These two later works focused on storing trajectory data in a data warehouse to enable multidimensional analysis (OLAP operations) on them. In fact, they concentrated on trajectory reconstruction, ETL and aggregation issues in a trajectory data warehouse, and along their work they considered trajectory data as set of sample points thus, a pure spatio-temporal concept. However, there are other very recent works handling the conceptual modeling problem and taking into account the complex structure of trajectory data and its semantic character, for instance, [2][3]. In this later paper authors are interested to modeling herd movement data in a trajectory data warehouse. They adopt the idea of stop and move proper to Spaccapietra to model the movement of a herd of animals perceived as a point in order to simplify the modeling task. The same idea employed to enhance the commerce investment activity in [47] by modeling trajectory data generated by a mobile information collector. Indeed, mobile professionals, communicating via mobile devices and moving in a road network thanks to means of transport furnished with sensors, are in charge of collecting huge amount of mobile from planned and not planned observations. These data are modeled in a trajectory data warehouse, implemented and analyzed in order to help investors deciding about the best investment to be made. In fact, the trajectory data warehouse was proven as a solution in many fields, therefore, it invaded all domains. In [43] the trajectory data warehouse is conceived to enhance the supply chain. The proposed model also employs the famous and successful approach of Spaccapietra to succeed the modeling of continuously changing data thanks to move and stop concepts. The medical filed is very important in which the data warehouse could prove its results, and is an indispensable repository for historical data about patient' states. It is necessary to enhance treatments and facilitate the detection of causes of disease as well as other scientific discoveries. In this context Akaichi et al. in [1] proposed a conceptual model for storing streams of facial nerve trajectory data in the purpose of studying the evolution of the states of patients affected by facial paralysis. This work is enhanced with a visualization algorithm of the facial nerve; FAN. Regarding the efficiency of data warehouse technology, and the tendency of moving from operational and transactional environments to multidimensional one, in [48] a system is proposed to assist users in designing a data warehouse called DWADS-Data Warehouse Assistant Design System. The system is based on the clover model and works on two steps. The first consists on choosing the schema model (star, snowflake, or constellation). The second finalize the task. An evolved step in the world of trajectory data warehouse is tackled in [46] consists on updating the trajectory data warehouse according to structure changes in information sources. The proposed solution is a schema versioning approach keeping the track of TrDW evolution. To succeed their approach, authors proposed a set of algorithms handling dimension updates (add dimension, delete dimension, add attribute to a given dimension, delete an attribute from a given dimension) by always respecting integrity constraints. Also, updating dimension levels (add dimension level, delete dimension level) by always considering integrity constraints. Concerning fact updates, measures could change, and therefore algorithms to add or delete a measure are proposed.

In [25], authors handled the logistic management field in particular truck delivery of goods as a very interesting branch in the economy domain. They proposed a solution for modeling and analyzing trajectory data generated by a traveling truck accomplishing its dedicated mission. Their final goal is to improve logistic management field. They consider a trajectory, at the conceptual level, as manageable and a decomposable object into meaningful events (each event is defined by begin, end and a semantic information) based on the idea of moves and stops introduced by Spaccapietra. Indeed, writers suppose that the trajectory' semantic depends on the application. They considered a spatio-temporal path of a moving object a sequence of trajectories, and this later is a sequence of triples (latitude, longitude, and timestamp) generated by a positioning technology. Concerning conceptual modeling, authors proposed a trajectory design pattern representing trajectory data using UML graphical notation and stereotype mechanism to distinguish facts from dimensions and from thematic information. They treat trajectories and events as first-class objects allowing an explicit modeling of trajectory' geometry

(TrajectoryType fact table) and its semantic facet (EventType fact table) represented with its components (moves, stops). For implementation purpose, the model is implemented in an object-relational Data Base Management System DBMS. Campora et al. in [9] adopt the same previous approach but using an ER formalism for conceptual design. In fact, they considered a trajectory as a first class semantic object having an identity, sub-components (stops, moves, and episodes) and other related thematic information. They used the Episode concept offering a high level and a more general unified representation of a trajectory element (move, stop). They developed a semantic trajectory model based on the MultiDimER notation to design their TrDW. Dimension tables are organized around a central fact table Episode. Spatial and temporal objects (dimensions), level hierarchies, measures and relationship types are expressively represented thanks to the visual icons used. In this work a complete framework, called St-Toolkit, was developed to design, store and query trajectory data by supporting OLAP, SOLAP and STOLAP tools. In [41] authors employed the MultiDimER model to conceive a public transportation application consisting on taxi travel. They defined a trajectory as a set of observations; each observation is the triples (location, time, semantic). Authors proposed two alternative modeling approaches. The first one is based on composed multi-valued time-stamped measures in which a trajectory is considered as a measure composed of observations (stops and moves) which are themselves composed multi-valued measures, in the context of application adopted in this paper they are fill-up, pick-up and move. However, it presents some drawbacks related to aggregation functions, the relationship between time dimension and observations' timestamps, and time consistency checks. Therefore, authors proposed a second approach which is based on composition of facts. It consists on considering each observation a separate fact relationship linked to the central fact relationship, representing the trajectory, with zero-to-many relationship. Thus, the trajectory becomes a derived measure from the isolate fact relationships. Those later are linked directly to the time dimension to cope with the above stated drawback and so assuring time hierarchy navigation. This later approach is the amelioration of the previous one, but it presents also some difficulties; first one is the need for an operation linking fact relationships, example, combining a taxi-journey with its observations. Second, the resulting number of fact-relationships is becoming large which could complicate the querying task.

To summarize all the described works concerning trajectory data warehouse conceptual modeling, we elaborate the following table representing most famous conceptual representation of trajectory data in the literature. As evaluating characteristics, we consider in our comparative study: the graphical representation on which the model is based, trajectory's geometric facet modeling, trajectory's semantic facet modeling, links constraining the trajectory's relationships with the rest of application objects. Finally, we depict if the basis of multidimensional structure are considered:

Table 3. A comparative study between trajectory data warehouse conceptual models

TrDW Conceptual Model	Graphical Representation	Trajectory' Geometric Modeling	Trajectory' Semantic Modeling	Links to Application Objects	Multidimensional Paradigm
C.Leal et al., 2011	Based on UML notation	Trajectory = (*longitude, latitude, instant*)	Stops and moves each one is defined with *begin*, *end* and *semantic information:* (*e = b, d, a*),	No constraints	*Facts*: *TrajectoryType* and *EventType* classes *Dimensions*: Time, Region, MovingObject, TrajectoryCategory and EventCategory *Hierarchies* not explicitly represented
Campora et al., 2011	Based on ER model (MultiDimER notation)	Trajectory = segment of the spatio-temporal path covered by a moving object	moves and stops represented with a unified concept called *Episode*	Topological and synchronization relationships	*Fact*: *Episode* fact relationship *Dimensions*: time, Space, Events and Trajectory Group tables *Hierarchies*: explicitly represented and annotated with MADS icons
Moreno and Arango, 2011	Based on ER model (MultiDimER notation)	trajectory = sequence of observations ($o_i = (l_i, t_i, s_i)$)	Stops and moves modeled as composed multi-valued or derived measures	No constraints	*Fact*: Journey fact relationship *Dimensions*: Space and Time dimensions *Hierarchies*: explicitly represented

5 CONCLUSION

The data warehouse technology is the revolution in the world of data base industry, recognized as the best repository providing the ready, interesting and directly accessible data for support of knowledge extraction and consequently decision-making process. Today, with the increasing volume of mobile data having a complex space-time character, many attempts to adopt data warehouse technology to store and manage them was appeared through the literature.

In this paper, we investigated main issues in the field of mobile data warehousing aiming to extend traditional data warehouse technology in order to handle mobility data requirements. We focused our survey on conceptual modeling issue regarding its importance in the process of building information systems. Some of these models are based on UML notation and others on ER formalism. Hence, we divided our investigation into two parts, first one focuses on spatial and spatio-temporal data warehouse conceptual models considering trajectory or movement of mobile objects as time-varying spatial attribute. Whereas, the second part is dedicated to trajectory data warehousing TrDW subfield representing the trajectory as a first-class object having a complex structure and semantic facet that cannot be supported by the spatio-temporal models.

ACKNOWLEDGMENT

I think my professor Jalel AKAICHI and every one, just every one.

REFERENCES

[1] J. Akaichi, H. Bouali, Z. Dhouioui (2010), “Facial Nerve Stream Trajectory Data Modeling and Visualization”, *Information Reuse and Integration, IEEE International conference*, LAS VEGAS, 2010.
[2] N. Arfaoui and J. Akaichi, “A Data Warehouse Assistant Design System Based On Clover Model”, *International Journal of Database Management Systems*, Vol. 2 No. 2, 2010.
[3] N. Arfaoui and J. Akaichi, “Modeling Herd Trajectory Data Warehouse”, *International Journal of Engineering Trends and Technology*- May to June Issue, 2011.
[4] Y. Bédard, T. Merrett, J. Han, “Fundamentals of Spatial Data Warehousing for Geographic Knowledge Discovery”, *Geographic, Data Mining and Knowledge Discovery*, Taylor and Francis, Vol. Research Monographs in GIS No. Chap. 3, pp. 53-73, 2001.
[5] S. Brakatsoulas, D. Pfoser, N. Tryfona, “Modeling, Storing and Mining Moving object Databases”, *International Database Engineering and Applications Symposium* (IDEAS), IEEE proceedings, pp. 68- 77, 2004.
[6] F. Braz, “Trajectory Data Warehouses: Proposal of Design and Application to Exploit Data”, 9th GeoInfo conference, Campos do Jordão, Brazil, pp. 61-72, 2007.
[7] F. Braz, S. Orlando, R. Orsini, A. Raffaeta, A. Roncato, C. Silvestri, “Approximate Aggregations in Trajectory Data Warehouses”, Proceedings of ICDE Workshop STDM, pp. 536–545, 2007.
[8] J. Brodeur, Y. Bédard, M. Proulx, “Modeling Geospatial Application Databases using UML-Based Repositories”, International Standards in Geomatics’, 2000.
[9] S. Campora, J. De Macedo, L. Spinsanti,”St-Toolkit: A Framework for Trajectory Data Warehousing”, AGILE 2011, 18-22 April 2011.
[10] E. Caron, and H. Daniels, “Diagnosis in the OLAP Context”, *Erasmus University Rotterdam, ERIM Institute of Advanced Management Studies,* July 2004.
[11] S. Chaudhuri and U. Dayal, “An Overview of Data Warehousing and OLAP Technology”, *SIGMOD Record* 26(1), pp. 65-74, 1997.
[12] M. Erwig, R.H. Guting, M. Schneider, M. Vazirgiannis. “Spatio-temporal Data Types: An Approach to Modeling and Querying Moving Objects in Database”, *GeoInformatica*, pp 269-296, 1999.
[13] O. Glorio, J. Mazon, I. Garrigos, J. Trujillo, “A Personalization Process for Spatial Data Warehouse Development”, *Decision Support Systems* 52, pp. 884–898, 2012.
[14] M. Golfarelli, D. Maio, S. Rizzi, ‘Conceptual design of data warehouses from E/R schemes’, *Proceedings of the Hawaii International Conference on System Sciences*, Hawaii, pp. 334–343, 1998.
[15] M. Golfarelli, D. Maio, S. Rizzi, “The Dimensional Fact Model: A Conceptual Model For Data Warehouses”, *International Journal of Cooperative Information Systems*, 1998.
[16] R.H. Guting, “An Introduction to Spatial Database Systems”, *VLDB Journal*, Vol. 3 No. 4, 1994.
[17] F. Gianotti and D. Pedreschi, Mobility, data mining and privacy, Ed. Springer Verlag, Berlin, Heidelberg, 2008.
[18] R.H. Guting, M.H. Bohlen, M. Erwig, C.S. Jensen, N.A. Lorentzos, M. Schneider, M. Vazirgiannis, “A Foundation for Representing and Querying Moving Objects” *ACM Transactions on Database Systems*, Vol. 25 No. 1, pp. 1–42, March 2000.
[19] J. Han, N. Stefanovic, K. Koperski, “Selective Materialization: An Efficient Method for Spatial Data Cube Construction”, *School of Computing Science, Simon Fraser University*, Burnaby, Canada, 1998.
[20] B. Huseumann, J. Lechtenborger, G. Vossen, “Conceptual Data Warehouse Design”, DMDW’2000: *Proceedings of the International Workshop on Design and Management of Data Warehouses*, Stockholm, Sweden, June 5-6, 2000.
[21] W.H. Inmon, “Building the Data Warehouse: Getting Started”, *John Wiley*, 2000.
[22] W.H. Inmon, “What is a Data Warehouse?”, *John Wiley*, 2000.

[23] A.S. Kamble, “A Conceptual Model for Multidimensional Data”, *5th Asia-Pacific Conference on Conceptual Modeling (APCCM 2008)*, Wollongong, NSW, Australia, 2008.

[24] K. Koperski, J. Adhikary, J. Han, “Spatial Data Mining: Progress and Challenges Survey Paper”, *School of Computing Science*, Simon Fraser University, Burnaby, Canada, 1996.

[25] B.C. Leal, J.A.F. de Macedo, V.C. Times, M.A. Casanova, V.M.P. Vidal, M.T.M. de Carvalho, “From Conceptual Modeling to Logical Representation of Trajectories in DBMS-OR and DW Systems”, *Journal of Information and Data Management,* Vol. 2 No. 3, pp. 463–478, 2011.

[26] L. Leonardi, G. Marketos, E. Frentzos, N. Giatrakos, S. Orlando, N. Pelekis, A. Raffaet`a, A. Roncato, C. Silvestri, Y. Theodoridis, “T-Warehouse: Visual OLAP Analysis on Trajectory Data”, 2009.

[27] B. Li and G. Cai, “A General Object-Oriented Spatial Temporal Data Model”. Ottawa, 2002.

[28] M. Luisa Damiani, S. Spaccapietra, “Spatial Data Warehouse Modeling Processing and Managing Complex Data for Decision Support”, 2006.

[29] S. Lujan-Mora and J. Trujillo, “Extending UML for Multidimensional Modeling”, 2001.

[30] S. Lujan-Mora, J. Trujillo, Y. Song, “Multidimensional Modeling with UML Package Diagrams”, 2002.

[31] E. Malinowski and E. Zimanyi, “A Conceptual Solution for Representing Time in Data Warehouse Dimensions”. *Third Asia-Pacific Conference on Conceptual Modelling (APCCM),* Hobart, Australia, 2006.

[32] E. Malinowski and E. Zimanyi, “Hierarchies in a Multidimensional Model: From Conceptual Modeling to Logical Representation”. *Data and Knowledge Engineering*, 348–377, 2006.

[33] E. Malinowski and E. Zimanyi, “Representing Spatiality in a Conceptual Multidimensional Model”. *Proceedings of the 12th Annual ACM International Workshop on Geographic Information Systems,* Washington, DC, USA, pp.12–22, 2004.

[34] E. Malinowski and E. Zimanyi, “Requirements Specification and Conceptual Modeling for Spatial Data warehouses”, *Private university of Bruxel*, 2006.

[35] E. Malinowski and E. Zimanyi, “Spatial Hierarchies and Topological Relationships in the Spatial MultiDimER Model”, *Private university of Bruxel*, 2005.

[36] G. Marketos, “Data Warehousing and Mining Techniques for Moving Objects Data Bases”, PhD Thesis, Department of Informatics, University of Piraeus, December, 2009.

[37] G. Marketos, “Mobility Data Warehousing and Mining”. *VLDB ’09 PhD Workshop*, August 24, Lyon, France, 2009.

[38] G. Marketos, E. Frentzos, N. Giatrakos, I. Ntoutsi, N. Pelekis, A. Raffaeta, Y. Theodoridis, “A Framework for Trajectory Data Warehousing”, *the Proceedings of ACM SIGMOD Workshop on Data Engineering for Wireless and Mobile Access (MobiDE'08),* 2008.

[39] G. Marketos, E. Frentzos, I. Ntoutsi, N. Pelekis, A. Raffaeta, Y. Theodoridis, “Building Real-World Trajectory Data Warehouses”. *MobiDE’08,* Vancouver, British Columbia, Canada, June 13,2008.

[40] M. Miquel, Y. Bédard, A. Brisebois, J. Pouliot, P. Marchand, J. Brodeur, « Modeling Multidimensional Spatio-temporal Data Warehouses in a Context of Evolving Specifications”. Ottawa, 2002.

[41] F. Moreno and F. Arango, “A Conceptual Trajectory Multidimensional Model: An Application to Public Transportation”, *Dyna,* Year 78, No.166, pp. 142-149, Medellin, April, 2011.

[42] F. Moreno, J.A.E. Arias, B.M. Losada, “A Conceptual Spatio-temporal Multidimensional Model”. *Universal Engineering Review of Medellin*, Vol. 9, No. 17, pp. 175-184, ISSN 1692-3324, July-December, Medellin, Colombia, 2010.

[43] Z. Nakhla and J. Akaichi, “Intensification of the Supply Chain by the Storing of Trajectories Data”, *International Journal of Managing Value and Supply Chains*, Vol. 1, No. 1, September, 2010.

[44] A. Olteanu, T. Couronné, J. Fen-Chong, Z. Smoreda, «Modélisation des Trajectoires Spatio-temporelles Issues des Traces Numériques de Téléphones Mobiles », SAGEO, Paris, July 7, 2011.

[45] S. Orlando, R. Orsini, A. Raffaeta, A. Roncato, « Trajectory Data Warehouse s: Implementation Issues”. Journal of Computing Science and Engineering, Vol. 1, No. 2, pp 211–232, December, 2007.

[46] W. Oueslati and J. Akaichi, “A Multiversion Trajectory Data Warehouse to Handle Structure Changes”, International Journal of Database Theory and Application, Vol. 4, No. 2, June, 2011.

[47] W. Oueslati and J. Akaichi, “A Trajectory UML Profile For Modeling Trajectory Data: A Mobile Hospital Use Case”, 2010.

[48] W. Oueslati and J. Akaichi, “Mobile Information Collectors Trajectory Data Warehouse Design”, *International Journal of Managing Information Technology (IJMIT)* Vol.2, No.3, August, 2010.

[49] D. Papadias, Y. Tao, P. Kalnis, J. Zhang, «Indexing Spatio-temporal Data Warehouses”, 2001.

[50] C. Parent, S. Spaccapietra, E. Zimanyi, « Spatio-temporal Conceptual Models: Data Structure+ Space + Time”. *Proceedings of the 7th ACM Symposium on Advances in Geographic Information Systems, ACM GIS'99*, pp. 26-33, 1999.

[51] C. Parent, S. Spaccapietra, E. Zimanyi, “The Mur Mur Project: Modeling and Querying Multi-Representation Spatio-temporal Databases”, In Information Systems, vol. 31, pp 733-769, 2006.

[52] C. Parent, S. Spaccapietra, E. Zimanyi, P. Donini, C. Plazanet, C. Vangenot, « Modeling Spatial Data in the MADS Conceptual Model”. *Proceeding of the International Symposium on Spatial Data Handling*, SDH 98, Vancouver, Canada, July 11-15, 1998.

[53] J.H. Pena, M.Y. Santos, “Representing, Storing and Mining Moving Objects Data”, *Proceeding of the Congress on Engineering*. Vol, 3, WCE July 6-8, 2011.

[54] D. Pfoser, C.S. Jensen, Y. Theodoridis, “Novel Approaches to the Indexing of Moving Object Trajectories”. *Proceedings of the 26th VLDB Conference*, Cairo, Egypt, pp. 395-406, September 10-14, 2000.

[55] F. Rao, L. Zhang, X.L. Yu, Y. Li, Y. Chen, “Spatial Hierarchy and OLAP-Favored Search in Spatial Data Warehouse”. DOLAP’03, New Orleans, Louisiana, USA, November 7, 2003.

[56] S. Rivest, Y. Bédard, M.J. Proulx, M. Nadeau, “SOLAP: A New Type of User Interface to Support Spatio-Temporal Multidimensional Data Exploration and Analysis”. Workshop ISPRS, Quebec, Canada, October 2-3, 2003.

[57] S. Bimonte, A. Tchounikine, M. Miquel, “Towards a Spatial Multidimensional Model”, *DOLAP’05*, Bremen, Germany, November 4–5, 2005.

[58] C. Sapia, M. Blaschka, G. Hofling, B. Dinter, “Extending the E/R Model for the Multidimensional Paradigm”. *Proceeding 1st International Workshop on Data Warehousing and Data Mining (DWDM 98), vol. 1552, Lecture Notes in Computer Science, Springer- Verlag*, New York, pp. 105-116, 1998.

[59] L. Savary, T. Wan, K. Zeitouni, “Spatio-temporal Data Warehouse Design for Human Activity Pattern Analysi”. PRISM Laboratory, Versailles University, 2003.

[60] S. Spaccapietra, C. Parent, M.L. Damiani, J. Antonio de Macedo, F. Porto, C. Vangenot, “A conceptual view on trajectory”, *Data and Knowledge Engineering* 65(1) pp. 126-146, 2007.

[61] S. Spaccapietra, C. Parent, E. Zimanyi, “Spatio-temporal and Multi-Representation Modeling : A Contribution to Active Conceptual Modeling”, 2007.

[62] N. Tryfona, F. Busborg, J. Christiansen, “StarER: A Conceptual Model for Data Warehouse Design”, *Proc. of the ACM 2nd Intl. Workshop on Data warehousing and OLAP (DOLAP'99*), Kansas City, Missouri, USA, 1999.

[63] A. Vaisman, E. Zimanyi, “What is Spatio-Temporal data Warehousing?”, 2008.

[64] H.J. Watson, “Recent Developments in Data Warehousing”, Communications of the Association for Information Systems, Vol. 8, 2011.

[65] T. Wan and K. Zeitouni, “Modélisation d’objets mobiles dans un entrepôt de données’, Laboratoire PRISM, Université de Versaille, 2005.

[66] J. Widom, “Research Problems in Data Warehousing”, *the Proceeding of 14th International Conference on Information and Knowledge Management*, November, 1995.