# Enriching Dimension Hierarchies with Topological Relations to Improve the Development of Spatial Data Warehouse

Sana Ezzedine

LTSIRS Laboratory, National School of Engineers,
University of Tunis El Manar, Tunisia

sana.ezzedine@gmail.com

Sami Yassine Turki and Sami Faiz

LTSIRS Laboratory, National School of Engineers,
University of Tunis El Manar, Tunisia
yassine.turki@isteub.rnu.tn and
sami.faiz@insat.rnu.tn

*Abstract*—**The design of a spatial data warehouse depends on operational data, spatial and non spatial requirements in order to support the decision-making process required by final users. It is crucial to consider decision maker requirements in the conceptual level of the construction of a spatial data warehouse. Furthermore, updating a spatial data warehouse and especially the addition of a new user's requirement after the construction of a spatial data warehouse is a great need for users. In this paper, to overcome this problem, dimension hierarchies will be specified in the Spatial Data Warehouse using topological relationships among spatial objects. Dimension hierarchies added show spatial requirements which are necessary to improve decision-making process. Decision makers thus will be able to achieve their information needs for analysis. Finally, we show the benefits of our approach by providing a case study, which defines an enriched conceptual model of a spatial data warehouse.**

***Keywords-spatial data warehouse; updating; spatial requirements; decision-making.***

## I. INTRODUCTION

If we consider the definition proposed by Inmon [1], a Spatial Data Warehouse (SDW) is "a subject oriented, integrated, non-volatile, and time variant collection of spatial data in support of management's decision".

The design of a SDW is based on a Multidimensional Model, which contains facts and dimensions. Facts contain the business metrics (i.e., measures) and dimensions describe facts and context to analyze these facts using dimension attributes organized in hierarchies.

Several approaches are proposed to model the design of a SDW. They did not define formal and standard transformations between the design and the implementation of SDWs in a specific platform. Moreover, they did not suggest an automatic transformation from the conceptual model to the possible logical representation. In addition, they did not consider needs related to the spatial Decision Maker (DM)'s requirements. To overcome these problems, we defined an approach [2] based on the Model Driven Architecture (MDA) models and the Unified Modeling Language (UML). This approach considers both spatial and non spatial requirements, described by a Geographic Computation Independent Model (Geo CIM), the first MDA Model. The Geo CIM is integrated by means of transformation rules into a Geographic Platform Independent Model (Geo PIM), the second MDA model, which defines the conceptual Multidimensional model of a SDW.

Within this approach, once user requirements are correctly captured, we obtain automatically the corresponding Multidimensional and conceptual model of a SDW. Nevertheless, in this approach, we find that the required multidimensional model does not take account of the updating requirements of a DM. Therefore, the final SDW will not completely satisfy final user requirements.

Our aim is to improve the quality of dimension hierarchies by means of adding new hierarchy aggregation levels, which allow SDW DMs to achieve their analysis information needs [2]. Dimension hierarchies enable also to the adding of new requirements to better support the decision-making process.

In this paper, we present an approach that treats updating in terms of adding new spatial requirements. We propose to enrich dimension hierarchies by adding new levels of aggregation in order to obtain the required hierarchies.

To accomplish this, we propose the use of semantic relations among spatial concepts provided by topological relationships [3]. The initial hypothesis is that both SDWs and Topological relationships present hierarchical structures: dimension hierarchies in SDWs show the relationships between value domains from different dimension attribute (set by levels of aggregation) [4], while topological relations present hierarchical semantic relations between spatial concepts, such as adjacency or inclusion or intersection, etc. [3]. Therefore, our approach is based on using these topological relations to add new levels to dimension hierarchies in order to obtain the required hierarchies. Figure. 1 summarizes this scenario.

The remainder of this paper is structured as follows. Section 2 presents an overview of works about the development of SDWs and the addition of dimension Hierarchies in the conceptual level. Section 3 defines our approach for enriching dimension hierarchies using topological relationships. In Section 4, a case study is presented. Finally, we point out our conclusions and sketch some future work in Section 5.

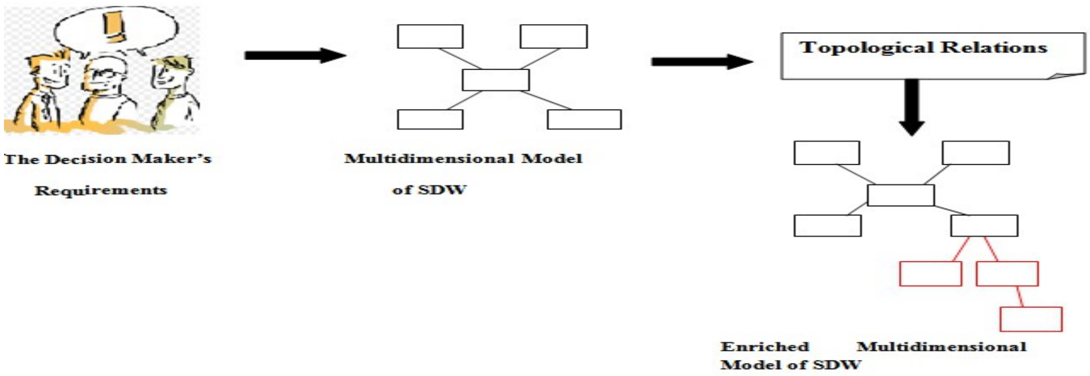



Figure1: Using Topological Relations to enrich the Multidimensional Model of SDW

## II. RELATED WORK

It is widely accepted that the development of SDWs must be based on a conceptual model. Therefore, in this section, we focus on briefly describing the most relevant approaches for the conceptual modeling of SDWs and, more generally, the addition of dimension Hierarchies in conceptual modeling.

### A. *Conceptual Modeling of SDWs*

Various approaches for the conceptual design of SDW systems have been proposed in the last few years. In this section, we present a brief discussion about some of the most well-known approaches.

The first attempt integrated spatial information and ensured correct aggregation over spatial [4][5]. Other works defined a multidimensional analysis tool that modeled spatial Data in a SDW [6][7][8][9]. Alternatively, authors defined a query language [10][11] that allowed the use of multidimensional and spatial and topological operators such as GeoMDQL [12]. All these approaches did not present an unequivocal and automatic transformation to every possible logical representation from a conceptual model. In addition, they did not consider needs related to the DM's requirements.

More recently, some approaches have tried to overcome these limitations, especially the problem of the automatic transformations. These approaches used standards framework as MDA. MDA provides a set of guidelines to structure specifications expressed as models. An alignment of multidimensional spatial model with MDA is proposed in [13]. The same approach is extended [14] to include spatial data in the SDW design level. It allows DM to define his geographical queries independently of the logical presentation. [15] proposed to consider the DM's aims and defined [16] some spatial elements describing the top DM's goals. In the same context, a Case tool based on Unified Modeling Language (UML) standard is used by [17] to model both spatial and non spatial data in the SDW design. [18] focused on the use of transformations based on MDA to automatically generate the data and the analysis models.

Every of the above-described approaches presented conceptual models lacked the integration of all DMs' spatial needs in the SDW design. To overcome this problem, [2] proposed an approach, which aims to integrate DM's requirements in the SDW Design. It presented an approach that automatically generates the design of SDW from a requirement's model.

This approach does not consider updating requirements in terms of adding new contexts to the requirements model after the development of SDW. This must be taken into account in stages of the development process, i.e., the conceptual modeling of the SDWs.

### B. *Adding Dimension Hierarchies in SDWs*

Mazón and Trujillo [19] suggested enriching dimension hierarchies in terms of structure and data. They considered dimension hierarchy as semantic relationships between values and they proposed to exploit the hypernymy/hyponymy relationships ("is-a-kind-of") and Meronymy/Holonymy ("is-a-part-of") WordNet. In this approach, levels of granularity are created at the end of hierarchy.

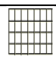
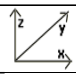
Favre et al. [20] proposed to enhance the dimension hierarchies by exploiting the knowledge of users. This knowledge is represented by a meta-aggregation rule and different rules. A meta-aggregation rule represents the structure of the link aggregation between two levels of granularity. And rules "if-then" represent the link at the instances. The levels created can be inserted into a hierarchy or created at the end thereof.

## III. USING TOPOLOGICAL RELATIONSHIPS TO ENRICH DIMENSION HIERARCHIES

Dimension hierarchies in SDWs show the relationships between domains of values from different dimension attributes (set in levels of aggregation). Topological relationships also present hierarchical relationships between spatial concepts, such as adjacency and connectivity. Thereby, we use topological relationships to automatically complete dimension hierarchies in a conceptual model of a SDW.

In this paper, we use classes' stereotypes defined in [2]. These classes are based on Unified Modeling Language (UML) as shown in Table1.

TABLE 1. STEREOTYPES USED TO DEVELOP THE CONCEPTUAL MODEL OF A SDW

| Stereotype | Description | Presentation |
|---|---|---|
| Fact Class | Facts contain business measures |  |
| Dimension Class | Dimensions describe Facts |   |
| Base Class | Base represent Dimension Hierarchy with their attributes |   |

In this paper, we define another stereotype based also on UML named Spatial Hierarchy, as shown in Table 2.

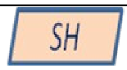
Spatial Hierarchy is added in the conceptual model of a SDW when the DM needs to take account of a new context of spatial requirements in the developing of the SDW.

Our proposal consists of identifying topological relationships between existing dimensions and bases in the conceptual model of SDW and the new added spatial requirements given by the user.

With each identified topological relationship, we create a Spatial Hierarchy, which is named with the same name of the identified topological relationship and has as attributes the characteristics of the added requirement.

Following, we explain the main steps of our approach (an overview is shown in Figure. 2).

TABLE 2. SPATIAL HIERARCHY STEREOTYPE USED TO DEVELOP THE CONCEPTUAL MODEL OF A SDW

| Stereotype | Description | Presentation |
|---|---|---|
| Spatial Hierarchy | Spatial Hierarchy present spatial Dimension hierarchy with their attributes. |   |

Prerequisite 1. A dimension attribute is chosen from the initial conceptual model of the SDW. The spatial hierarchy will be enriched starting from this attribute.

Prerequisite 2. A new spatial requirement has been added by the DM, which is in relation with one of existing dimensions in the initial conceptual model.

Step 1. Extract different instances from the dimension attribute chosen from the initial conceptual model.

Step 2. Identify topological relationships between spatial objects recently required with spatial objects existing in the dimension attribute chosen.

Step 3. If there are relationships between the required spatial objects and the existed ones, a spatial hierarchy for every relationship is created having the same name as the topological relationship.

Step 3'. If there are no relationships between the required spatial objects and the existed ones, a new record is inserted in the selected dimension without creating a new hierarchy.

In Figure. 2, every step of our approach is illustrated. From a dimension or a dimension hierarchy in a Multidimensional model, which not accomplishes all user requirements, a dimension attribute is chosen. Then the topological relationships are identified between instances of the dimension attribute chosen and the new requirement in order to create a new level of the spatial dimension hierarchy. If there are no relationships between added requirement and existed dimensions, a new record is inserted. Iterations are repeated until all required spatial objects are classified.

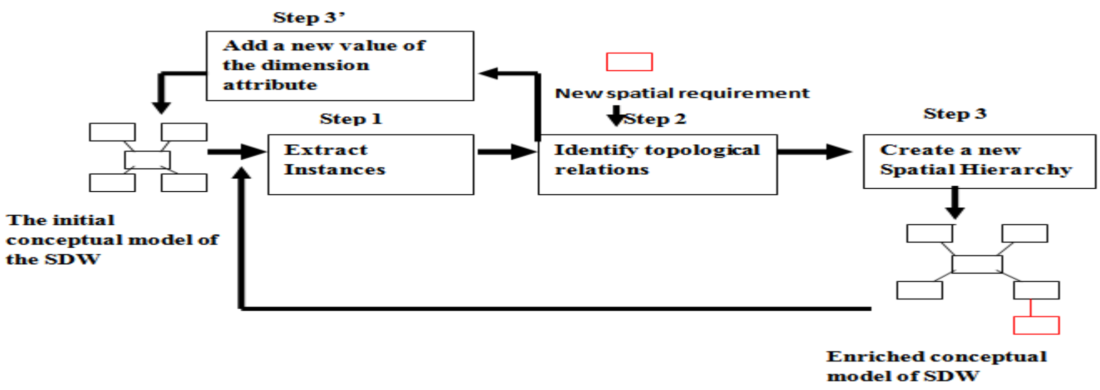



Figure 2. Overview of our approach

## IV. CASE STUDY

In this section, we show the benefits of our approach by providing a case study, in which Spatial hierarchies are enriched in the conceptual Model of a SDW. Our case study consists of defining a conceptual model of a sales manager who wants to analyze sales operations in stores situated 2 km around the airport. Then, he needs to extend the analysis region by adding others streets.

The initial conceptual model before adding spatial requirements and applying our approach is presented in Figure. 3. In this case, the added requirements are the extended streets.

As is described previously, we should choose firstly a dimension attribute. In this case, we choose the dimension spatial cover and the dimension attribute spatial objects. Then we identify relations between spatial objects and streets added. We found that some streets have an intersection relationship with existed streets; others have an inclusion relationship and others streets having no relationships with existed spatial objects.

The conceptual model is extended according to the different stages of the approach as shown in Figure. 4.

The initial conceptual model contains facts and dimensions presenting the spatial and non spatial requirements. The facts presented in Figure. 3 are: Sale and Spatial Cover. The dimensions are: Product, Time, Operational, DMcharacteristic, Presentation and Semantic.

Figure.3 presents the spatial and non-spatial data that are necessary for the decision maker, the sales manager, to make the right decision. We use classes of the Unified Modeling Language UML to model requirements. Each class is a one of the requirements expressed by the sales manager.

After the addition of new spatial requirements, we identify two types of relations, as shown in Figure.4, inclusions' relations and intersections' relations between spatial objects and added spatial objects.

Consequently, two spatial hierarchies are added according to the two relationships identified.

## V. CONCLUSION AND FUTURE WORK

Spatial Dimension hierarchies are important to support the decision making process, since they allow the analysis of data at different levels of detail (i.e., levels of aggregation). Then, obtaining the required spatial hierarchies captured from decision maker is crucial for specifying a successful SDW.

In this paper, we propose the application of topological relationships to obtain the required hierarchies. The advantage of our proposal is clear: the enrichment of the conceptual model of the SDW by adding new aggregation levels in order to satisfy the required DM requirements.

These required hierarchies allow SDW users to satisfy their information analysis needs, since they better support the decision-making process.

Our proposal can be generalized to generate a star scheme or a snowflake scheme of a conceptual model of a SDW appropriate for a group of DMs.

 ISBN: 978-1-61208-334-6

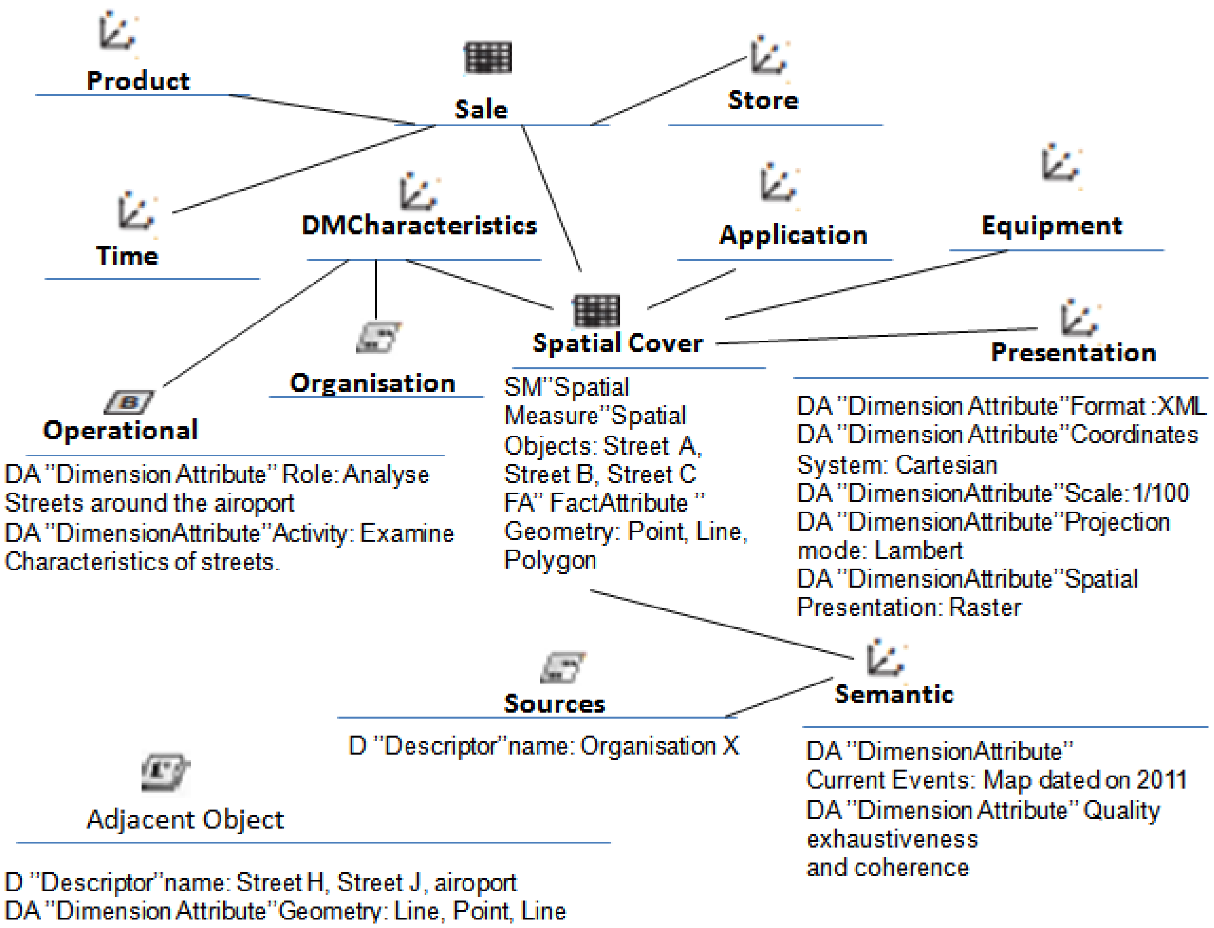



Figure 3. The conceptual model of the SDW used by the sales manager [2]

## REFERENCES


[1] W. Inmon, "Building the Data warehouse". In John Wiley & Sons, Chichester , 1996.

[2] S. Ezzedine, S. Y. Turki, and S. Faiz "An approach based on Model Driven Engineering for data integration in a spatial data warehouse', in CODIT'13, 2013.

[3] J. Egenhofer and R. D.Franzosa "Point-Set Topological Spatial Relations", in International Journal for Geographical Information Systems, vol. 5, pp. 161-174, 1991.

[4] S. Bimonte, A. Tchounikine, and M. Miquel, "Geocube, a multidimensional model and navigation operators handling complex measures: Application in spatial olap', Advances in Information Systems (ADVIS)", pp. 100–109, Berlin / Heidelberg, Germany, 2006.

[5] S. Bimonte, A. Tchounikine, and M. Bertolotto "Integration of geographic information into multidimensional models", in ICCSA, pp. 316–329, 2008.

[6] A. Escribano, L. Gomez., B. Kuijpers, and A. Vaisman "Piet: a gis-olap implementation", in DOLAP '07: Proceedings of the ACM tenth international workshop on Data warehousing and OLAP, pp 73–80, New York, NY, USA, 2007.

[7] L. Gomez, S. Haesevoets, B. Kuijpers, and A. Vaisman "Spatial aggregation: Data model and implementation", in CoRR, abs/0707.4304, 2007.

[8] E. Malinowski and E. Zimanyi "Representing spatiality in a conceptual multidimensional model", in GIS '04: Proceedings of the 12th annual ACM international workshop on Geographic information systems, pp. 12–22, New York, NY, USA, 2004.

[9] E. Malinowski and E. Zimanyi "Implementing spatial datawarehouse hierarchies in object-relational dbmss", in ICEIS, pp. 186–191, 2007.

[10] J. Da Silva, V. C. Times, A. C. Salgado, C. Souza, R. do Nascimento Fidalgo, and A.G. de Oliveira "Querying geographical data warehouses with geomdql", in SBBD, pp. 223–237, 2007.

[11] N. Stefanovic, J. Han, and K. Koperski "Object-based selective materialization for efficient implementation of spatial data cubes", in IEEE Transactions on Knowledge and Data Engineering, vol. 12. N. 6, pp. 938–958, 2000.

[12] S. Rivest, Y. Bedard, and P. Marchand "Toward better support for spatial decision making: Defining the

characteristics of spatial on-line analytical processing", in Geomatica, vol. 55, N. 4, pp.539–555, 2001.

[13] O. Glorio and J. Trujillo, "An MDA Approach for the Development of Spatial Data Warehouses", In DaWaK '08, 2008.

[14] O. Glorio and J. Trujillo, "Designing Dhata Warehouses for Geographic OLAP querying by using MDA", in ICCSA '09, 2009.

[15] J. Mazon and J. Trujillo, "A Hybrid Model Driven Development Framework for the Multidimensional Modeling of Data Warehouses", in SIGMOD, 2007.

[16] O. Glorio, J. Mazón, I. Garrigós, and J. Trujillo, "Using Web-based Personalization on Spatial Data Warehouses", in EDBT '10, 2010.

[17] R. Fidalgo and A. Cuzzocrea, "An Enhanced Spatial Data Warehouse Metamodel", in CAiSE Forum, Vol. 855 of CEUR Workshop Proceedings, pp.32-39, 2012.

[18] A. Cuzzocrea, J. Mazon, J. Trujillo, and J. Zubcoff "Model-driven data mining engineering: from solution-driven implementations to 'composable' conceptual data mining models", in Int. J. of Data Mining, Modelling and Management, 2011 Vol.3, No.3, pp.217 – 251, 2011.

[19] J.N. Mazón, and J. Trujillo, "Enriching Data Warehouse Dimension Hierarchies by Using Semantic Relations", in XXIIIrd British National Conference on Databases (BNCOD 2006), Vol. 4042, pp. 278–281. Springer, 2006.

[20] C. Favre, F. Bentayeb, and O. Boussaïd, "Dimension Hierarchies Updates in Data Ware-houses : a User-driven Approach", in IXth International Conference on Enterprise Informa-tion Systems (ICEIS 07), Funchal, Madeira, Portugal, 2007.

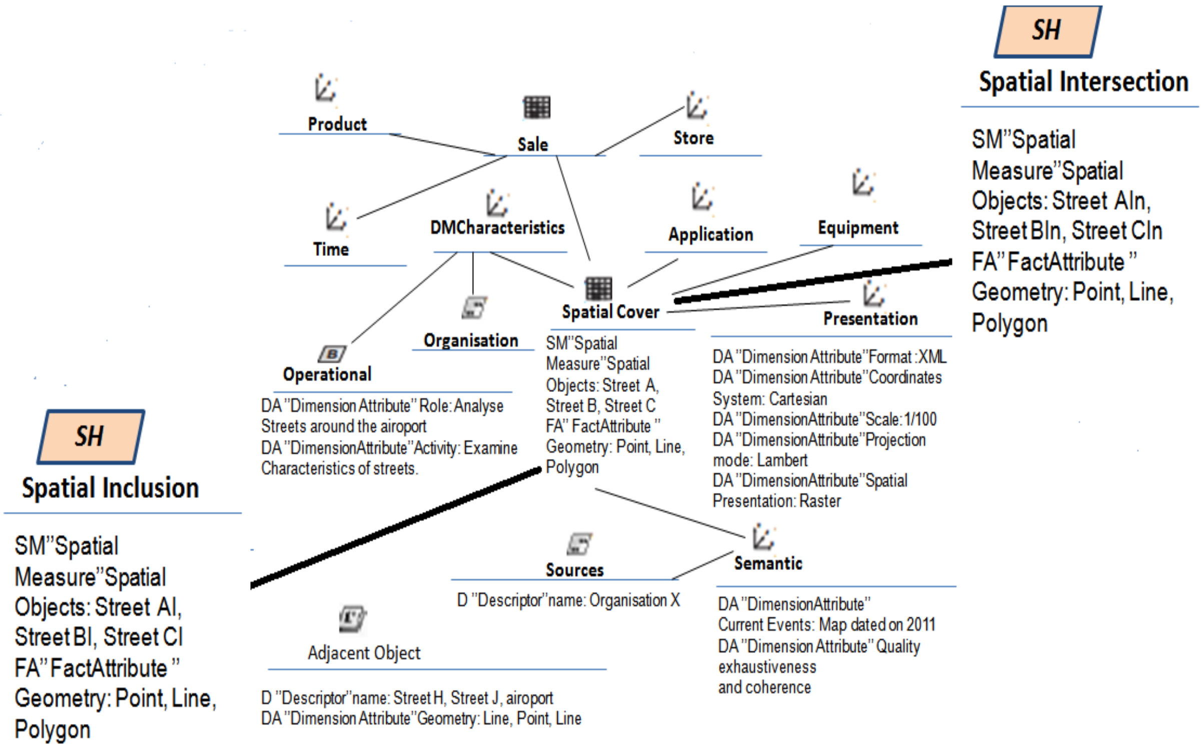



Figure 4. The extended conceptual model of the SDW after adding spatial requirements