



Ingeniare. Revista Chilena de Ingeniería

ISSN: 0718-3291

facing@uta.cl

Universidad de Tarapacá
Chile

Carrera Sepúlveda, Sabina; Varas Contreras, Marcela; Urrutia Sepúlveda, Angélica
TRANSFORMACIÓN DE ESQUEMAS MULTIDIMENSIONALES DIFUSOS DESDE EL NIVEL
CONCEPTUAL AL NIVEL LÓGICO

Ingeniare. Revista Chilena de Ingeniería, vol. 18, núm. 2, agosto, 2010, pp. 165-175
Universidad de Tarapacá
Arica, Chile

Disponible en: <http://www.redalyc.org/articulo.oa?id=77216407003>

- ▶ Cómo citar el artículo
- ▶ Número completo
- ▶ Más información del artículo
- ▶ Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

TRANSFORMACIÓN DE ESQUEMAS MULTIDIMENSIONALES DIFUSOS DESDE EL NIVEL CONCEPTUAL AL NIVEL LÓGICO

FUZZY MULTIDIMENSIONAL SCHEMA TRANSFORMATION FROM CONCEPTUAL TO LOGICAL LEVEL

Sabina Carrera Sepúlveda¹ Marcela Varas Contreras¹ Angélica Urrutia Sepúlveda²

Recibido 9 de abril de 2009, aceptado 17 de agosto de 2010

Received: April 9, 2009 Accepted: August 17, 2010

RESUMEN

En el ámbito de las bases de datos, diversos métodos han sido propuestos para la generación del diseño lógico de un almacén de datos o data warehouse (DW), por otra parte, existen estudios que permiten el almacenamiento y operación de datos imprecisos o difusos en las bases de datos relacionales. Este artículo presenta una propuesta metodológica con un enfoque arquitectura dirigida por modelos (MDA), la cual permite generar el diseño lógico de un data warehouse difuso (DWD). La técnica utilizada consistió en identificar los elementos básicos del modelado multidimensional (MD) que se extienden al manejo de atributos difusos en las medidas a nivel conceptual, a través de la aplicación de una secuencia de transformaciones con el propósito de obtener el diseño lógico multidimensional difuso. Uno de los elementos claves para esta transformación es la extensión del metamodelo common warehouse metamodel (CWM) OLAP con estereotipos difusos, considerando que actualmente no existe un método formal que nos permita realizar este tipo de transformaciones.

Palabras clave: Almacenes de datos, almacenes de datos difusos, transformación de modelos, arquitectura dirigida por modelos.

ABSTRACT

In the databases field, several methods have been proposed for the generation of the logical design of a datawarehouse, on the other hand there are also studies that allow the storage and operation of fuzzy data in relational databases. This paper presents a methodological proposal with an MDA approach, which allows to generate the logical design of a fuzzy DW. The technique was used to identify the basic elements of multidimensional modelling and extends them to manage fuzzy attributes in the measures at a conceptual level, through the application of a sequence of transformations to avoid the generation of a logical fuzzy multidimensional design. A key element for this process is the extension of the CWM OLAP metamodel with fuzzy stereotypes. Currently there is no formal method that allows to realize this kind of transformations.

Keywords: Data warehouses, fuzzy data warehouses, model transformation, model driven architecture.

INTRODUCCIÓN

Los almacenes de datos o data warehouse (DW) recogen información proveniente de diferentes bases de datos

operacionales de la empresa, la filtran y procesan para su almacenamiento, proporcionando una plataforma de datos consolidados e históricos para el posterior análisis del negocio orientados a la toma de decisiones.

¹ Departamento de Ingeniería Informática y Ciencias de la Computación. Universidad de Concepción. Edmundo Larenas 215. Concepción, Chile. E-mail: mvaras@udec.cl

² Departamento de Computación e Informática. Universidad Católica del Maule. Avenida San Miguel 3605. Talca, Chile. E-mail: aurrutia@spock.ucm.cl

Las bases de datos difusas permiten el almacenamiento de datos que contienen incertidumbre o imprecisión en su información tratados con la teoría de conjuntos difusos [23, 24].

En la literatura se pueden encontrar diferentes propuestas metodológicas para el modelamiento multidimensional de un DW; ninguna de ellas ha sido aceptada completamente. Las propuestas pueden clasificarse como las basadas en los requerimientos [5, 22, 16, 26, 15], las basadas en las bases fuentes [10, 4, 21, 13, 11] y en el enfoque compuesto [12, 9, 3, 2, 27].

En [7] se presenta una propuesta para el modelamiento multidimensional de un DW que permita el almacenamiento de datos que contengan incertidumbre o imprecisión en su información, sin explicitar la metodología para llegar a estos esquemas.

Por otra parte, existe una arquitectura multidimensional difusa propuesta en [1], la cual no considera mecanismos formales de transformación de modelos.

En [14] se propone un perfil UML para modelado multidimensional en almacenes de datos, el cual consiste en una extensión al perfil del lenguaje unificado de modelado (UML). Este perfil es definido por un conjunto de estereotipos, restricciones y valores etiquetados para representar las principales propiedades del MD a nivel conceptual. Utiliza el enfoque MDA para así poder separar el modelado conceptual de la plataforma. Además desarrollaron un algoritmo que a partir de los modelos MD utilizando el perfil UML genera la correspondiente aplicación en diferentes plataformas a través de una transformación vertical. En esta propuesta se crearon transformaciones formales para el modelado multidimensional, pero los autores no incorporaron datos con incertidumbre o imprecisión en el modelo multidimensional.

MDA es una arquitectura dirigida por modelos [20], propuesta por el OMG (object management group) [18]. Esta arquitectura considera tres niveles de abstracción de modelos: modelos independientes de computación (CIM), modelos independientes de plataforma (PIM) y modelos dependientes de plataforma (PSM).

La metodología de diseño conceptual de almacenes de datos propuesta en [27] permite especificar transformaciones formales y con una notación estándar entre metamodelos, basándose en el enfoque MDA; esta propuesta no consideró el diseño lógico multidimensional.

Lo anteriormente expuesto nos motiva a elaborar una propuesta metodológica con un enfoque MDA que permita

transformar un modelo conceptual multidimensional difuso en un modelo lógico multidimensional difuso. Se ha definido que el modelo lógico multidimensional difuso será implementado en una plataforma ROLAP, debido a la madurez de la tecnología, la disponibilidad de una arquitectura multidimensional difusa relacional [1] y del metamodelo CWM [17], lo que hace factible la extensión de este último y la generación de nuestra propuesta.

El orden en el cual se aplican las transformaciones lo determina la metodología propuesta. La definición de las reglas entre los metamodelos se expresa mediante lenguaje natural y posteriormente se formaliza mediante query view transformations (QVT), el cual es el lenguaje estándar de transformaciones propuesto por la OMG [17].

En la presente investigación se crearon transformaciones para los elementos básicos del modelo multidimensional y para los atributos difusos tipo 1, tipo 2 o tipo 3 [23], aplicados a las medidas de los cubos.

El artículo presenta un apartado que resume una serie de conceptos básicos como la arquitectura dirigida por modelos, el metamodelo CWM OLAP y el metamodelo CWM relacional. En el apartado transformación de modelos se presentan los tipos de transformaciones entre metamodelos, el metamodelo de entrada y el metamodelo de salida a las transformaciones. En el apartado transformaciones para generar el diseño lógico de un DWD se presenta la definición y formalización de las transformaciones necesarias para generar el diseño lógico. En el apartado siguiente se presenta la metodología para la aplicación de las transformaciones y en el apartado seis se presenta un caso de estudio. Finalmente, en el último apartado las conclusiones y trabajo futuro.

CONCEPTOS BÁSICOS

Arquitectura dirigida por modelos (MDA)

MDA [20] es una propuesta de la OMG para el desarrollo de software. La idea principal de MDA es especificar sistemas con modelos independientes de la plataforma, separando la especificación de la implementación en una plataforma específica. El modelo se convierte en el elemento más valioso, puesto que a partir de él y mediante una serie de transformaciones se puede obtener el código de la aplicación.

MDA de acuerdo al nivel de abstracción clasifica los modelos en:

CIM (computation independent model): modelo que tiene relación con el negocio, donde se representa una vista del sistema.

PIM (platform independent model): Modelo de un subsistema que no contiene información específica de la plataforma o tecnología específica que se utiliza, posee un alto nivel de abstracción, por lo cual no es necesario repetir el proceso de modelado cada vez que se introduzca una nueva tecnología al sistema y representa la estructura, funcionalidad y restricciones del sistema independientemente de la plataforma tecnológica sobre la que se va a implementar.

PSM (platform specific model): Modelo de un subsistema que incluye información sobre la tecnología específica que se utiliza; este modelo completa al PIM especificando cómo toma forma el sistema al ser implementado en una plataforma determinada. Los conceptos abstractos del PIM se detallan y describen de forma acorde a los recursos de la plataforma elegida y es la representación de un sistema, con detalles específicos de la plataforma en la que será implementado.

Common warehouse metamodel (CWM)

CWM [17] es un metamodelo de bases de datos propuesto por la OMG. Estandariza las bases para el modelado de los datos comunes a una organización, de forma que se habilite el intercambio de metadatos independientemente de su forma de almacenamiento físico.

Metamodelo CWM OLAP

El metamodelo CWM OLAP [17] es un estándar para representar una base de datos multidimensional. En este metamodelo un esquema (Schema) contiene dimensiones (Dimensions) y cubos (Cubes). Una dimensión se compone de niveles (Levels) y jerarquías (Hierarchies). Un cubo es una colección de medidas (Measures). La clase CubeDimensionAssociation relaciona un cubo con una dimensión. La clase HierarchyLevelAssociation relaciona un nivel con una jerarquía.

El metamodelo CWM OLAP se define por un conjunto de clases derivadas del metamodelo de UML

Metamodelo CWM Relacional

El metamodelo CWM Relacional [17] es un estándar para representar la estructura de una base de datos relacional, permitiendo especificar tablas, columnas, claves primarias, claves ajenas.

El metamodelo CWM Relacional se define por un conjunto de clases derivadas del metamodelo de UML

Atributos Difusos con Conjuntos Difusos

Los atributos difusos pueden clasificarse en tres tipos de acuerdo el tipo de referencial. El referencial es el

dominio subyacente o eje X donde se definen los conjuntos difusos. En todos ellos se incluyen los valores Unknown, Undefined y Null [8, 25]:

- **Tipo 1:** Son atributos precisos, sin imprecisión, crisp. Sin embargo, se permite definir etiquetas lingüísticas en su dominio y podremos usarlas en consultas difusas. Se almacenan igual que un atributo normal, pero puede ser transformado o manipulado usando condiciones difusas. Este tipo es útil para extender bases de datos tradicionales para permitir consultas difusas en sus dominios clásicos.
- **Tipo 2:** Son atributos “imprecisos sobre un referencial ordenado”. Admiten valores crisp y difusos, en forma de distribuciones de posibilidad o conjuntos difusos, sobre un dominio subyacente ordenado. Es una extensión del Tipo 1 que sí permite el almacenamiento de información imprecisa.
- **Tipo 3:** Son atributos difusos con datos “discretos sobre dominio subyacente no ordenado con analogía”. Aquí se pueden definir etiquetas (Rubio, Moreno...) que son escalares con una relación de similitud o proximidad definida sobre esas etiquetas, para indicar en qué medida se parece cada par de etiquetas. También se admiten distribuciones de posibilidad (o conjuntos difusos) sobre este dominio.

TRANSFORMACIÓN DE MODELOS

La transformación de modelos es el proceso de convertir un modelo en otro.

Las transformaciones entre los modelos definidos pueden ser: a) de CIM a CIM, b), de CIM a PIM, c) de PIM a PIM, d) de PIM a PSM, h), de PSM a PSM e i) de PSM a PIM.

Para la presente investigación se requiere transformar un esquema conceptual multidimensional difuso en un esquema lógico multidimensional difuso, por lo que el tipo de transformación utilizada será transformación de PIM a PSM, siendo utilizada cuando un PIM es refinado para ser llevado a una plataforma específica PSM.

Metamodelo de entrada a las transformaciones PIM

Para nuestro caso particular, la capa PIM representa las principales propiedades de los modelos multidimensionales difusos, sin tener en cuenta los detalles específicos de la tecnología de base de datos a utilizar. esto provee la independencia necesaria del modelo multidimensional difuso con respecto a la plataforma que lo implementará.

Data warehouse difuso (DWD). Lo definimos para efectos de esta investigación como: “Un data warehouse que permite almacenar y operar atributos difusos tipo 1, tipo 2 y tipo 3 en las medidas de los Cubos”.

Un DWD aporta información cuantitativa y cualitativa a un DW, lo cual permite analizar fenómenos y sistemas complejos, que las personas somos incapaces de describir con precisión.

Actualmente no existe un metamodelo que nos permita representar un DWD. La especificación del metamodelo CWM OLAP sólo permite representar las principales propiedades de los modelos multidimensionales, no permite representar atributos difusos definidos en esta propuesta. Esta carencia motiva a extender el metamodelo CWM OLAP incorporando formalmente en el metamodelo de CWM OLAP nuevas capacidades de modelado difuso, permitiéndole almacenar y operar atributos difusos tipo 1, tipo 2 y tipo 3 [23], en las medidas de los cubos, de acuerdo a la arquitectura para estructuras multidimensionales difusas propuesta en [1].

Para la presente investigación, PIM se utiliza en la extensión del metamodelo CWM OLAP, denominada metamodelo CWM OLAP difuso.

Descripción de la Extensión Propuesta

La extensión del metamodelo CWM OLAP define un conjunto de estereotipos, valores etiquetados y restricciones que permiten agregar atributos difusos tipo 1, tipo 2 y tipo 3 [23] en las medidas de los cubos. La extensión se definió de acuerdo a los mecanismos de extensión estándar provistos por UML: Estereotipos, Valores Etiquetados y Restricciones.

Definición de Nuevos Estereotipos, valores etiquetados y restricciones

Los estereotipos definidos utilizan la Arquitectura para Estructuras Multidimensionales Difusas propuesta en [1]. Estos estereotipos extienden a los estereotipos del metamodelo CWM OLAP. Todos los estereotipos definidos se desprenden de la clase Measure del metamodelo CWM OLAP (ver Figura 1).

Los estereotipos definidos son:

Estereotipo «MeasureFuzzyT1»: Los atributos de este estereotipo representan si una medida es un Atributo Difuso Tipo 1. Se extiende a partir de la Clase Base Measure y el icono que lo representa es T1.

Estereotipo «MeasureFuzzyT2»: Los atributos de este estereotipo representan si una medida es un Atributo Difuso Tipo 2. Se extiende a partir de la Clase Base Measure y el icono que lo representa es T2.

Estereotipo «MeasureFuzzyT3»: Los atributos de este estereotipo representan si una medida es un Atributo Difuso Tipo 3. Se extiende a partir de la Clase Base Measure y el icono que lo representa es T3.

Estereotipo «MetaAtribsTablas»: Los atributos de este estereotipo representan la metadata para representar si un atributo difuso es del Tipo 1, Tipo 2 o Tipo 3. Se extiende a partir de la Clase Base MeasureFuzzyT1 o MeasureFuzzyT2 o MeasureFuzzyT3 y el icono que lo representa es MAT.

Estereotipo «DboFuzzyT2»: Los atributos de este estereotipo representan la metadata para representar un atributo difuso Tipo2. Se extiende a partir de la Clase Base MeasureFuzzyT2 y el icono que lo representa es Dt2.

Estereotipo «DboFuzzyT3»: Los atributos de este estereotipo representan la metadata para representar un atributo difuso Tipo3. Se extiende a partir de la Clase Base MeasureFuzzyT3 y el icono que lo representa es Dt3.

Estereotipo «TablaFuzzy»: Los atributos de este estereotipo representan la tabla que posee el atributo difuso Tipo 2 o Tipo 3, para posteriormente agregar un campo con el grado o valor de pertenencia que los atributos en un conjuntos difuso. Se extiende a partir de la Clase Base MeasureFuzzyT2 o MeasureFuzzyT3 y el icono que lo representa es TF.

Se puede apreciar en la Figura 1 un extracto del metamodelo CWM OLAP difuso propuesto; los estereotipos definidos se extienden a partir de la Clase Base Measure, la cual se encuentra en color plomo claro y los estereotipos definidos se encuentran en color plomo oscuro.

Se puede apreciar en la Figura 2 un extracto del metamodelo CWM OLAP difuso propuesto, las clases propuestas son MeasureFuzzyT1, MeasureFuzzyT2, MeasureFuzzyT3, MetaAtribsTablas, DboFuzzyT2, DboFuzzyT3 y TablaFuzzy.

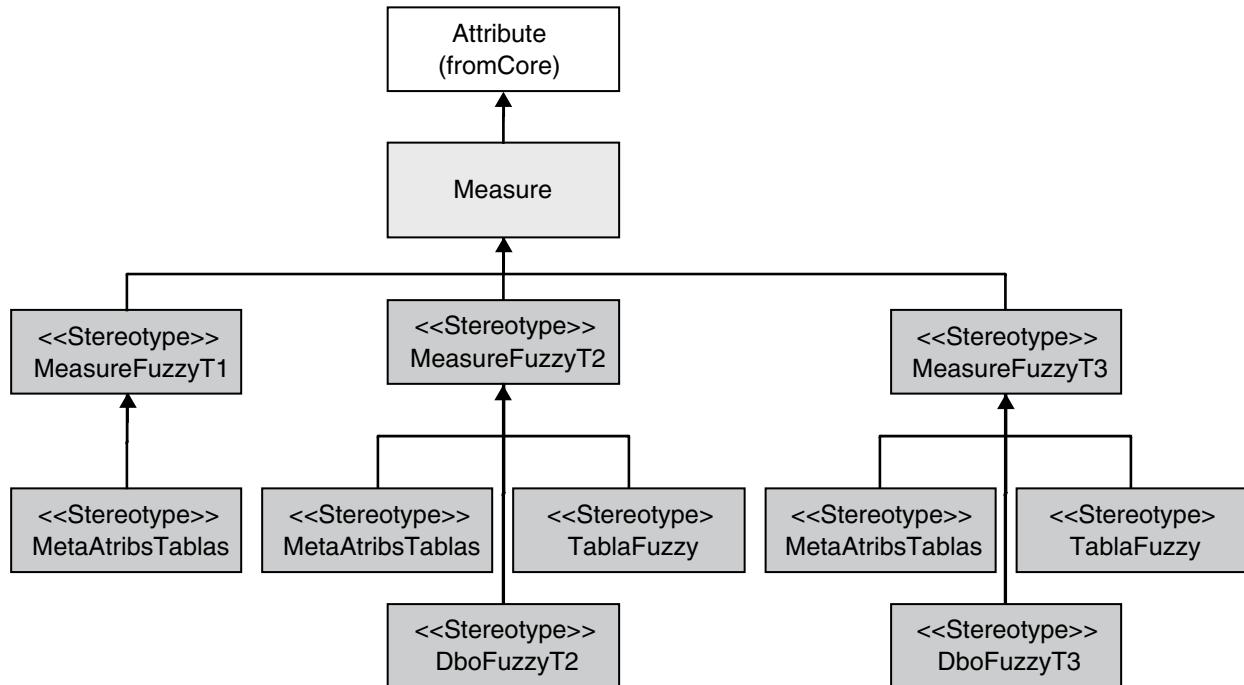


Figura 1. Extracto Extensión Metamodelo CWM OLAP con Estereotipos Difusos.

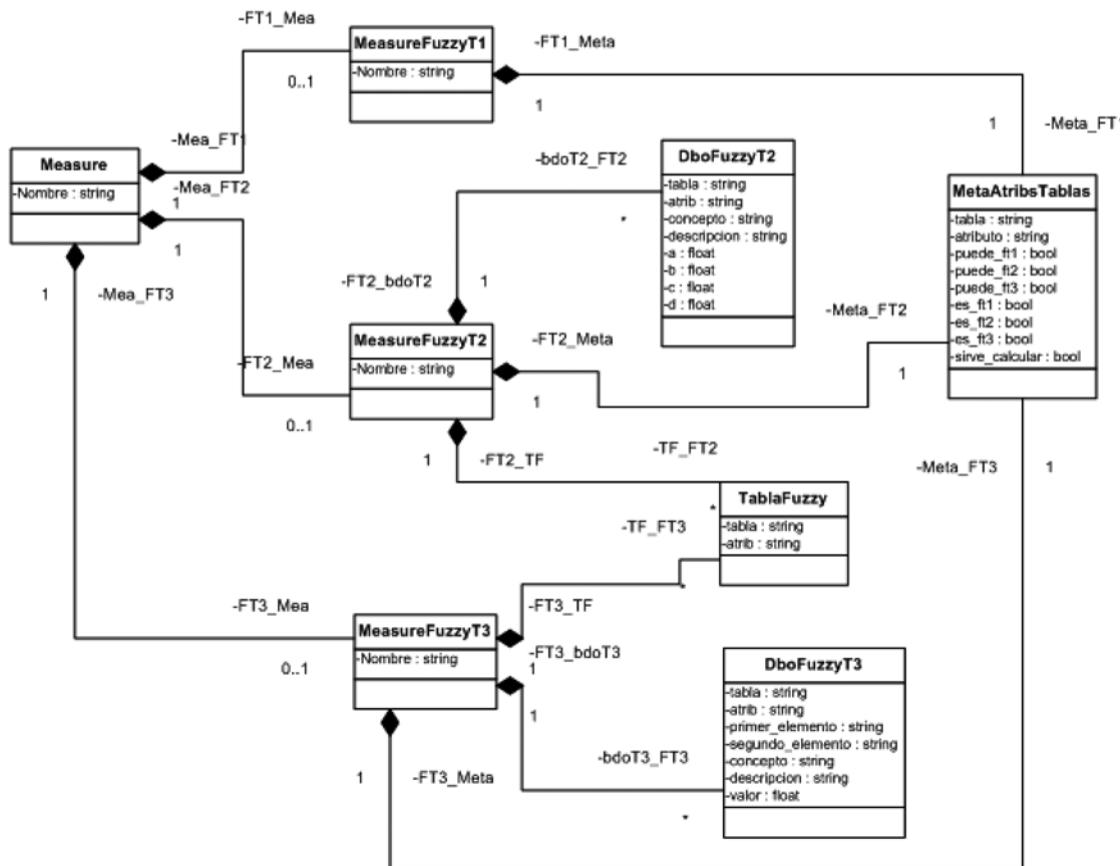


Figura 2. Extracto Metamodelo CWM OLAP Difuso.

Metamodelo de salida a las transformaciones PSM
Como PSM para la presente investigación se utiliza el metamodelo CWM Relacional, debido a que el propósito es transformar un esquema multidimensional conceptual difuso a un esquema lógico multidimensional difuso, a implementarse bajo tecnología ROLAP.

TRANSFORMACIONES PARA GENERAR EL DISEÑO LÓGICO DE UN DWD

La transformación es la generación de un modelo fuente en un modelo destino, de acuerdo a transformación claramente definida.

Una transformación está conformada por un conjunto de reglas que describen cómo el modelo en el lenguaje fuente puede ser convertido en un modelo en el lenguaje destino.

Una regla de transformación es una descripción de uno o más constructores en el lenguaje fuente que pueden ser transformados en uno o más constructores en el lenguaje destino.

A continuación se muestra una breve descripción de cada regla de transformación definida.

- (*RMain*) Principal: Todo esquema del metamodelo CWM OLAP difuso se transforma en esquema en el metamodelo CWM relacional.
- (*RCubTab*) CubeToTable: Todo cubo del metamodelo CWM OLAP difuso se transforma en una tabla en el metamodelo CWM relacional.
- (*RDimTab*) DimensionToTable: Toda dimensión del metamodelo CWM OLAP difuso se transforma en una tabla en el metamodelo CWM relacional.
- (*ADimCol*) AttributeDimensionToColumn: Todo atributo de una dimensión del metamodelo CWM OLAP difuso se transforma en una columna en el metamodelo CWM relacional.
- (*RLevTab*) LevelToTable: Todo nivel del metamodelo CWM OLAP difuso se transforma en una tabla en el metamodelo CWM relacional.
- (*ALevCol*) AttributeLevelToColumn: Todo atributo de un nivel del metamodelo CWM OLAP difuso se transforma en una columna en el metamodelo CWM relacional.
- (*RHlaFk*) HlaToForeignKey: Toda clase de Asociación de Nivel/Jerarquía del metamodelo CWM OLAP difuso se transforma en una clave foránea en el metamodelo CWM relacional.
- (*RDhFk*) DimensionHierarchyToForeignKey: Toda clase de asociación de dimensión/jerarquía del metamodelo CWM OLAP difuso se transforma en una clave foránea en el metamodelo CWM relacional.
- (*RCdaFk*) CdaToForeignKey: Toda clase de asociación del metamodelo CWM OLAP difuso se transforma en una clave foránea en el metamodelo CWM relacional.

- (*RMeaCol*) MeasureToColumn: Toda Medida del metamodelo CWM OLAP difuso se transforma en un atributo en el metamodelo CWM relacional.
- (*MatTab*) MetaAttrbsTablasToTable: Todo primer atributo difuso en el metamodelo CWM OLAP difuso, se transforma en la tabla meta_attrbs_tablas en el metamodelo CWM relacional.
- (*Dt2Tab*) DboFuzzyT2ToTable: Todo primer atributo difuso tipo 2 en el metamodelo CWM OLAP difuso, se transforma en la tabla dbo_fuzzyT2 en el metamodelo CWM relacional.
- (*Dt3Tab*) DboFuzzyT3ToTable: Todo primer atributo difuso tipo 3 en el metamodelo CWM OLAP difuso, se transforma en la tabla dbo_fuzzyT3 en el metamodelo CWM relacional.
- (*TfTab*) TablaFuzzyToTable: Toda TableFuzzy en el metamodelo CWM OLAP difuso, se transforma en una tabla más un campo grado en el metamodelo CWM relacional.

Formalización transformaciones

La OMG propuso un lenguaje estándar para la especificación de transformaciones llamado query view transformations (QVT [17]). QVT es un lenguaje de naturaleza híbrida (declarativa e imperativa). Su parte declarativa está dividida en dos capas, un lenguaje llamado *Relations*, el cual puede ser expresado mediante una notación gráfica y textual y el lenguaje llamado *Core*, el cual es utilizado internamente por las herramientas de transformaciones; el lenguaje *Core* posee una notación imperativa llamada operational mappings.

Un metamodelado es un mecanismo que permite construir formalmente lenguajes de modelado, como, por ejemplo, UML. El metamodelo de un lenguaje es una descripción de todos los conceptos que pueden usarse en el mismo, por ejemplo, los conceptos de package, clase, atributo y operaciones en UML.

QVT permite especificar transformaciones entre modelos y metamodelos, incluye una sintaxis de diagramas y una sintaxis textual para la definición de las transformaciones.

Para nuestra investigación las transformaciones se realizan a nivel de metamodelos, siendo el metamodelo CWM OLAP difuso el PIM fuente y el metamodelo CWM relacional el PSM destino, permitiendo instanciar un modelo multidimensional difuso a nivel lógico. El objetivo principal es mostrar la formalización en notación textual y gráfica de una de las transformaciones realizadas en esta investigación.

A continuación se muestra a modo de ejemplo la formalización de la transformación CubeToTable.

Formalización regla CubeToTable

El objetivo de esta transformación es convertir un cubo (c) del modelo CWM OLAP difuso en una tabla (t) del modelo CWM relacional (ver Figura 3). En la representación gráfica de la regla se muestra que una tabla (t) se crea por medio de un dominio obligado.

Esta regla de transformación se compone de la regla SchemaToSchema(s_{olap},s_{rel}) en la cláusula When (Precondición), y de las reglas CdaToForeignKey (cda,fk) y MeasureToColumn(m,co) en la cláusula Where (Poscondición); estas reglas se encargan de trasformar clases de asociación (cda) y medidas (m) en claves foráneas y columnas, respectivamente.

Notación textual regla CubeToTable

maping Molap::Cube CubeToTable() :Mrel::Table

```
{
  object
  {
    nombre:='Cub_' + self.nombre;
    Fk_Rel:= self.Cda_Cub ->
      collect(cdalcda.CdaToForeignKey())->
      asOrderSet();
    Col_Rel := self.Mea_Cub ->
      collect(mlm.MeasureToColumn()) ->
      asOrderSet();
  }
}
```

Notación Gráfica Regla CubeToTable

Cube To Table

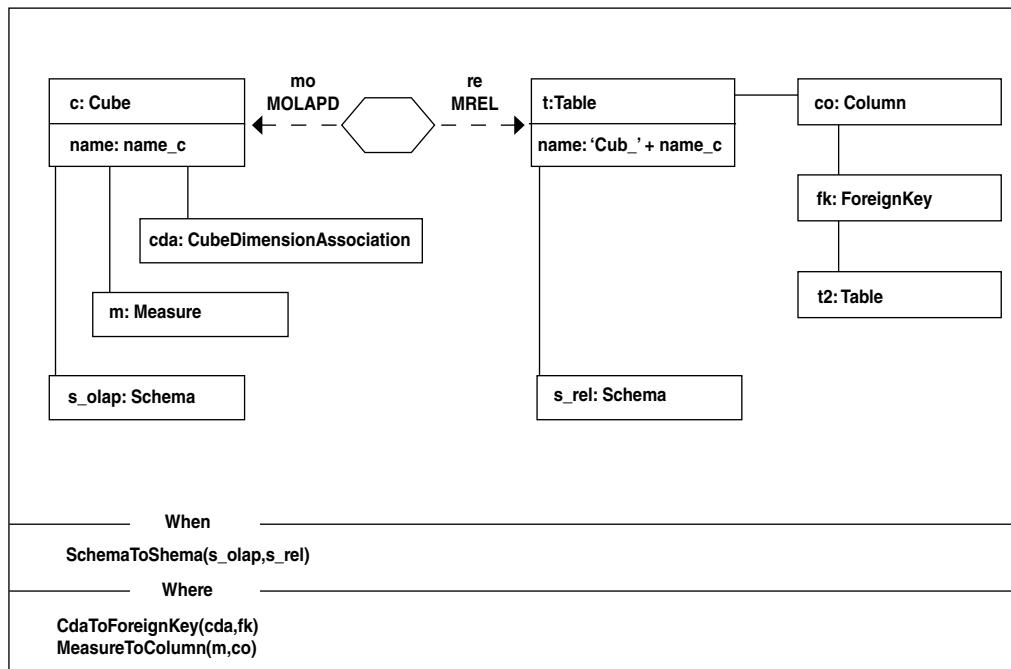


Figura 3. Regla CubeToTable.

METODOLOGÍA

La metodología para la generación del modelo lógico multidimensional difuso se dividió en tres pasos para una fácil aplicación y comprensión:

Paso 1. Generar el esquema conceptual multidimensional difuso mediante una instancia del metamodelo CWM OLAP Difuso.

Paso 2. En el esquema conceptual multidimensional difuso generado identificar los elementos básicos del modelado multidimensional difuso. Los elementos a identificar son los cubos, las medidas, las dimensiones, los niveles, las clases de asociación jerarquía,

las clases de asociación jerarquía nivel, las clases de asociación cubo dimensión y los atributos difusos tipo 1, tipo 2 y tipo 3.

Paso 3. Aplicar las reglas de transformación RMain, RCubTab, RDimTab, ADimCol, RLevTab, ALLevCol, RHlaFk, RDhFk, RCdaFk, RMeaCol, MatTab, Dt2Tab, Dt3Tab, TfTab. Para una correcta aplicación de las reglas se generó un diagrama con la secuencia que se debe seguir (ver Figura 4).

Cuando se apliquen las reglas de transformación se debe considerar el orden definido en la aplicación de las reglas.

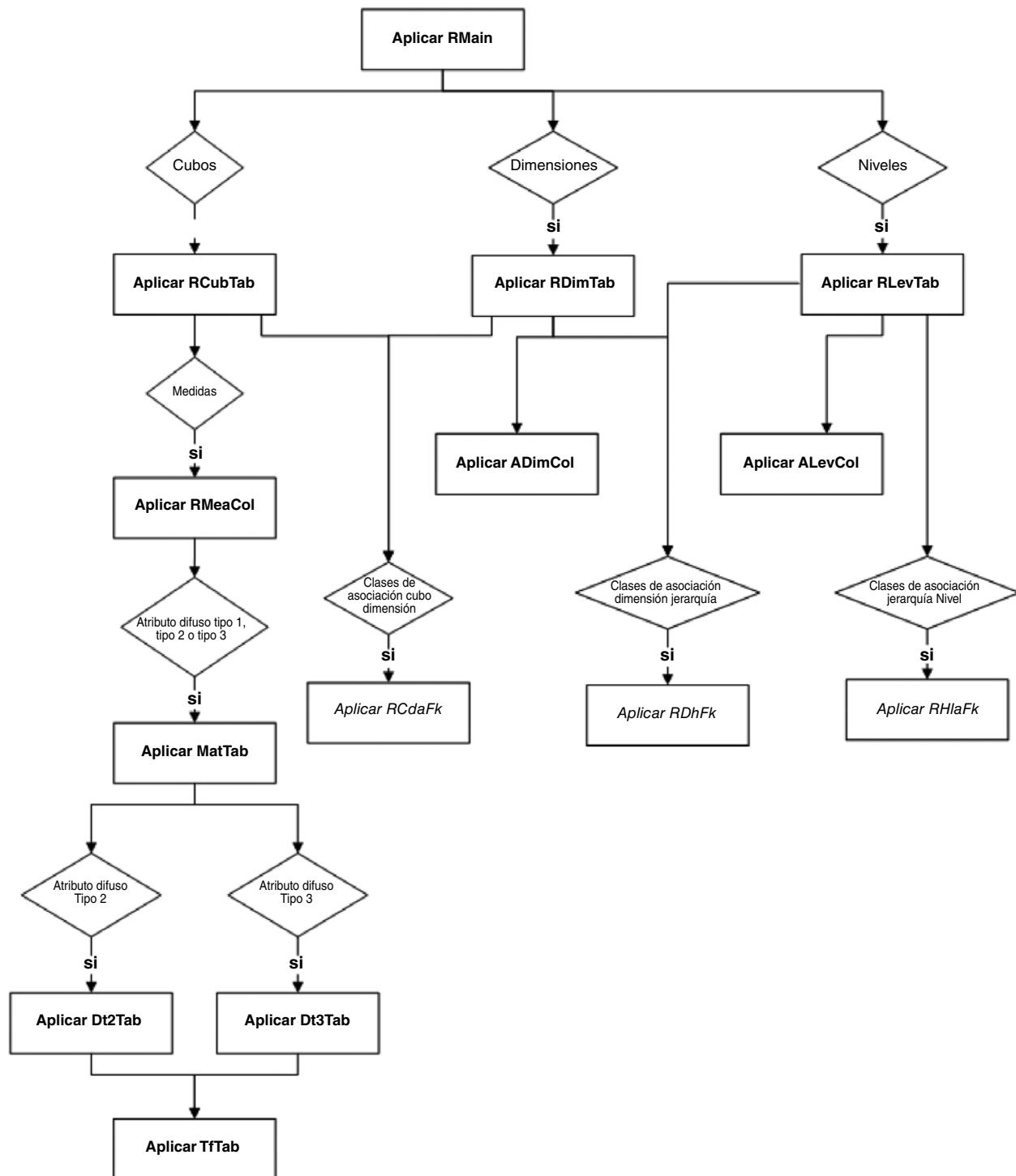


Figura 4. Diagrama con secuencia a seguir en la aplicación de las transformaciones.

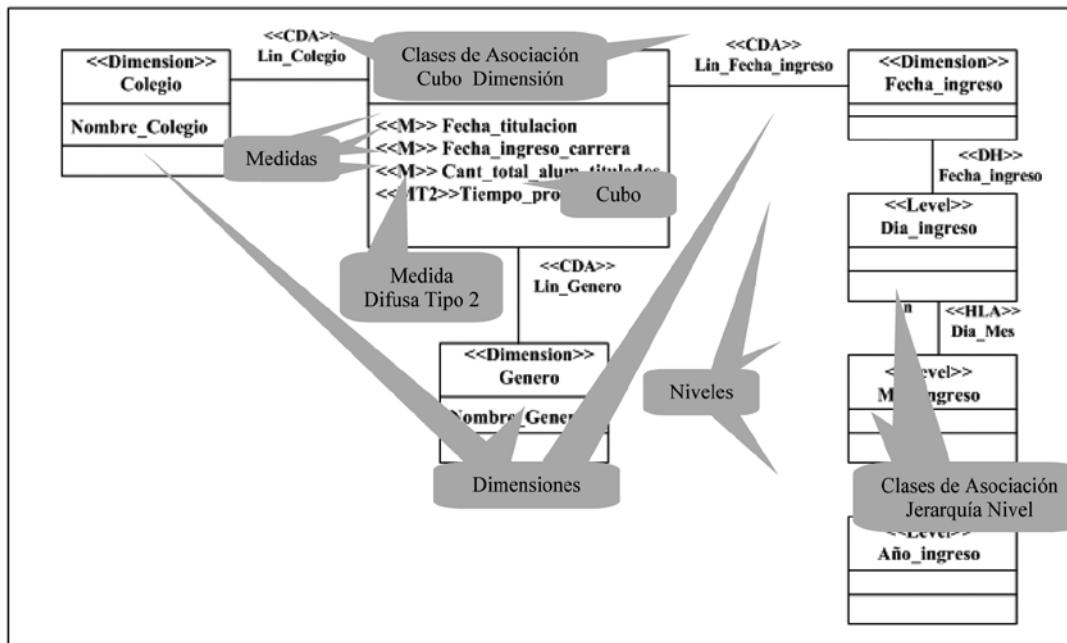


Figura 5. Esquema conceptual Tiempo Promedio de Titulación.

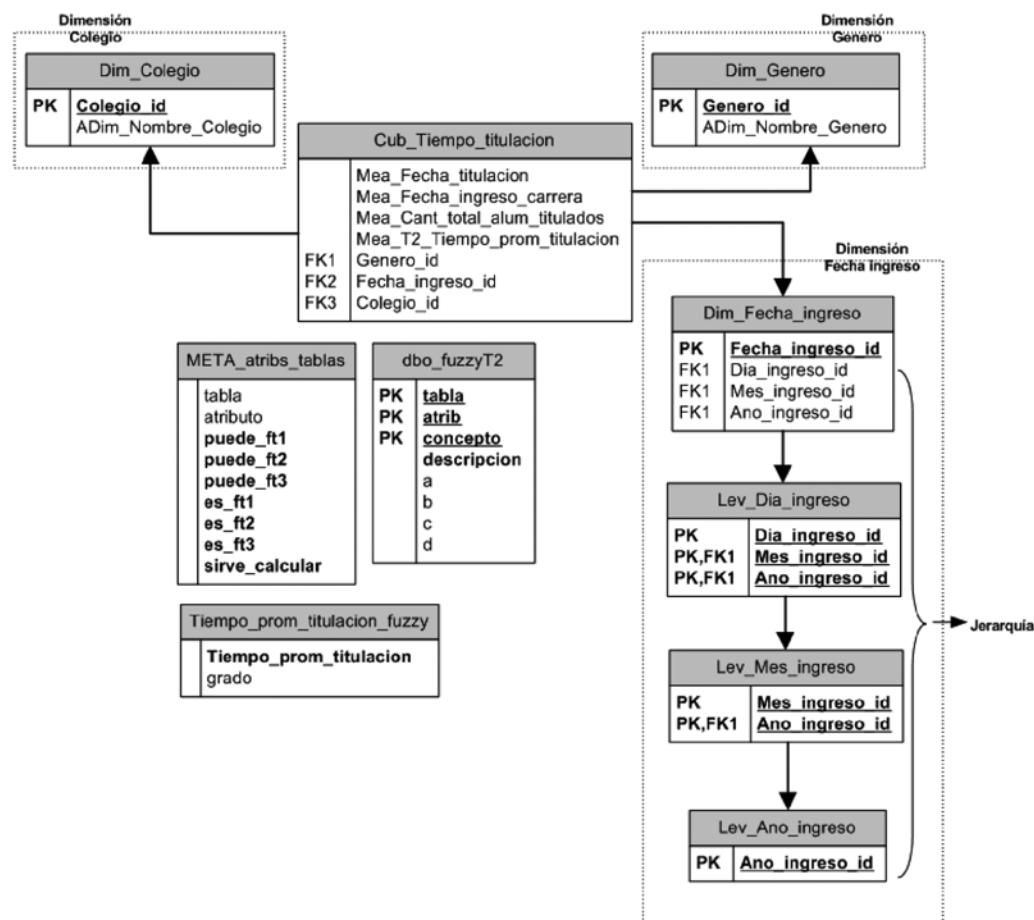


Figura 6. Esquema Lógico Tiempo Promedio de Titulación.

CASO DE ESTUDIO

En el presente caso de estudio se aplica la metodología propuesta, para lo cual se presenta un cubo para el indicador tiempo promedio de titulación [6].

Indicador Tiempo Promedio de Titulación

Se define el indicador correspondiente al Tiempo Promedio de Titulación, de los alumnos de la carrera Ingeniería Civil Informática de la Universidad de Concepción. El propósito de la implementación de este indicador es proporcionar información del tiempo en años que demoran los alumnos en terminar la carrera, haciendo un análisis por género y por colegio de origen, para así poder analizar si estas variables influyen en los tiempos de titulación de los alumnos; este indicador se crea con la finalidad de acelerar las consultas y permitir el análisis de los datos respecto a cada dimensión diseñada. Este caso de estudio será aplicado a los Atributos Difusos Tipo 2.

$$\text{Tiempo Promedio} = \frac{\sum (\text{Fecha Titulación} - \text{Fecha Ingreso})}{\text{Cantidad total alumnos titulados}}$$

El esquema conceptual del indicador se representa mediante una instancia del metamodelo CWM OLAP Difuso (esto corresponde al Paso 1 de la metodología propuesta). Ver representación gráfica en la Figura 5.

Según el paso 2 se identifican los Cubos, Medidas, Dimensiones, Atributos de dimensiones, Niveles, Atributos de niveles, Clases de asociación y los Atributos difusos Tipo 1, Tipo 2 y Tipo 3 de la instancia del metamodelo CWM OLAP Difuso (ver Figura 6).

Una vez identificados, se deben aplicar las reglas, generando así el Modelo Lógico Multidimensional Difuso (paso 3).

Al aplicar las reglas de transformación RMain, RCubTab, RDimTab, ADimCol, RLevTab, RHlaFk, RDhFk, RCdaFk, RMeaCol, MatTab, Dt2Tab, TfTab, se obtiene el esquema lógico, el cual se encuentra representado mediante una instancia del metamodelo CWM Relacional (ver Figura 6).

De esa manera se obtuvo el esquema lógico multidimensional difuso requerido. Sobre él se generaron los reportes asociados al indicador especificado.

CONCLUSIONES Y TRABAJO FUTURO

Las contribuciones más importantes de esta investigación son en el área del Diseño Multidimensional Difuso, más específicamente en el área de Transformación de Modelos Multidimensionales Difusos, las cuales son:

- Una extensión del metamodelo CWM OLAP, la cual permite representar datos difusos en las medidas de un Cubo, pudiendo así generar un Esquema Conceptual Lógico Multidimensional Difuso. En esta extensión se definieron estereotipos para los atributos difusos Tipo 1, Tipo 2 y Tipo 3 que permiten flexibilizar las medidas de un cubo y posibilitan su uso en otros metamodelos.
- La definición y formalización de un conjunto de transformaciones entre el metamodelo OLAP Difuso y el metamodelo CWM Relacional. Permitiendo modelar un Esquema Lógico Multidimensional Difuso a partir de un Esquema Conceptual Multidimensional Difuso.

Respecto a la aplicación de las reglas de transformación:

- Es importante dejar claro que si el diseño del esquema conceptual no se encuentra correcto, la aplicación de las transformaciones no va a solucionar estos problemas.
- Si el diseñador sigue el orden propuesto en la metodología, no debieran existir problemas con las transformaciones.

Como trabajo futuro se considera la implementación de datos difusos en los demás elementos del DW, como por ejemplo Dimensiones, Jerarquías. Incluyendo la formalización de reglas para estas nuevas implementaciones. También se está evaluando la implementación de reglas de transformación para las típicas funciones de agregación propias de los DW y la implementación automática de las transformaciones.

AGRADECIMIENTOS

Este trabajo fue parcialmente financiado por la Dirección de Investigación de la Universidad de Concepción, Chile, Proyecto 206.093.011-1 “Diseño Conceptual de Almacenes de Datos con Manejo de Incertidumbre Usando Lógica Difusa”.

REFERENCIAS

- [1] J. Acchiardo. “Implementando el soporte de estructuras multidimensionales difusas en sistemas OLAP”. Memoria Ingeniería Civil Informática. Dirigida por: Marcela Varas. Universidad de Concepción. Marzo 2008.
- [2] C. Ballard, D. Herreman, D. Schau, R. Bell, E. Kim and A. Valncic. “Data Modeling Techniques for Data Warehousing”. SG24-2238-00. IBM Red Book. ISBN number 0738402451. 1998.

- [3] A. Bonifati, F. Cattaneo, S. Ceri, A. Fuggetta and S. Paraboschi. "Designing Data Marts for Data Warehouses". *ACM Transactions on Software Engineering and Methodology*. Vol. 10, Issue 4, pp. 452-483. October, 2001.
- [4] L. Cabibbo and R. Torlone. "A Logical Approach to Multidimensional Databases". *Lecture Notes in Computer Science*. Vol. 1377. 1998.
- [5] F. Carpani. "CMDM: A conceptual multidimensional model for Data Warehouse". Master Thesis. Advisor: Raúl Ruggia. Pedeciba, Universidad de la República, Uruguay. URL: <http://www.lsi.upc.edu/~aabello/thesis.html>. 2000.
- [6] S. Carrera. "Transformación de Esquemas Multidimensionales Difusos del Nivel Conceptual al Nivel Lógico". Tesis para optar al grado de Magíster en Ciencias de la Computación. Universidad de Concepción. Chile. 2009.
- [7] M. Delgado, C. Molina, D. Sánchez, A. Vila and L. Rodríguez-Ariza. "A Fuzzy Multidimensional Model for Supporting Imprecision in OLAP". *Proceedings of IEEE International Conference on Fuzzy Systemns*. ISBN 0-7803-8353-2. 2004.
- [8] J. Galindo, A. Urrutia and M. Piattini. "Fuzzy Databases: Modeling, Design and Implementation". Idea Group Publishing Hershey. October, 2005.
- [9] P. Giorgini, S. Rizzi and M. Garzetti. "Goal-Oriented Requirements Analysis for Data Warehouse Design". In *Proceedings of the ACM 8th International Workshop on Data Warehousing and OLAP (DOLAP'05)*, pages 47-56. Bremen, Germany. November 4-5, 2005.
- [10] M. Golfarelli, D. Maio and S. Rizzi. "Conceptual Design of Data Warehouses from E/R Schemes". *Proceedings of the Thirty-First Hawaii International Conference on System Sciences*. 1998.
- [11] M. Gyssens and L. Lakshmanan. "A Foundation for Multi-Dimensional Databases". *23rd International Conference on Very Large Data Bases*, pp. 106-115. Atenas. Grecia. 1997.
- [12] B. Hüsemann, J. Lechtenbörger and G. Vossen. "Conceptual Data Warehouse Design". *DMDW'00*. Sweden. 2000.
- [13] Ch. Li and X.S. Wang. "A Data Model for Supporting On-Line Analytical Processing". *Conference on Information and Knowledge Management*. Maryland, USA. 1996.
- [14] S. Luján-Mora, J. Trujillo and I. Song. "A UML profile for multidimensional modeling in data warehouses", *Data and Knowledge Engineering*. Vol. 59, Issue 3. December, 2006.
- [15] J.N. Mazón, J. Trujillo, M. Serrano and M. Piattini. "Designing Data Warehouses: From Business Requirement Analysis to Multidimensional Modeling". In *Proceedings of the 1st Int. Workshop on Requirements Engineering for Business Need and IT Alignment*. Paris, France. September, 2005.
- [16] F. McGuff. "Data Modelling for Data Warehouses". URL: <http://members.aol.com/fmcguff/dwmodel>. 1998.
- [17] OMG. "Common Warehouse Metamodel CWM Specification". Versión 1.1. URL: <http://www.omg.org/cgi-bin/doc?formal/03-03-02>. March, 2003.
- [18] OMG. "Object Management Group". URL: <http://www.omg.org/>.
- [19] OMG. "MOF QVT Final Adopted Specification". URL: <http://www.omg.org/cgi-bin/doc?ptc/2005-11-01>. November, 2005.
- [20] OMG. "MDA Guide". Versión 1.0.1. URL: http://ormsc.omg.org/mda_guide_working_page.htm. Juny, 2003.
- [21] C. Phipps and K. Davis. "Automating Data Warehouse Conceptual Schema Design and Evaluation". In *Proceedings of the 4th International Workshop Design and Management of Data Warehouses (DMDW'02)*, pp. 23-32. Toronto. May, 2002.
- [22] C. Sapia, M. Blaschka, G. Höfling and B. Dinter. "Extending the E/R Model for the Multidimensional Paradigm". *DWDM'98*. Singapur, pp. 105-116. 1998.
- [23] A. Urrutia. "Definición de un Modelo Conceptual Para Bases de Datos Difusas". Tesis para optar al Grado de Doctor en Ingeniería Informática. URL: <http://www.ganimedes.ucm.cl/aurrutia/>. 2003.
- [24] A. Urrutia, M. Varas y J. Galindo. "Diseño de una base de datos difusa modelada con UML". *6^º Workshop Iberoamericano de Ingeniería de Requisitos y Ambientes Software. IDEAS*, pp. 145-155. Asunción, Paraguay. Mayo 2003.
- [25] A. Urrutia, L. Tineo and C. González. "FSQL and SQLf: Towards a Standard in Fuzzy Databases". In *Handbook of Research on Fuzzy Information Processing in Databases*. Vol. I, pp. 270-298. Information Science Reference. URL: <http://www.info-sci-ref.com>. 2008.
- [26] P. Westerman. "Data Warehousing: using the Wal-Mart Model". (The Morgan Kaufmann Series in Data Management Systems). Morgan Kaufmann, 1st edition. ISBN: 155860684X. 2000.
- [27] L. Zepeda. "Metodología para el Diseño Conceptual de Almacenes de Datos". Tesis Doctoral. Dirigida: Matilde Celma. Universidad Politécnica de Valencia. España. URL: <http://www.dsic.upv.es/docs/bib-dig/tesis/etd-04032008-120043/tesisprimeraversionLeopoldo.pdf>. 2008.