Integration of Business Intelligence Based on Three-Level Ontology Services

Longbing Cao[1], Chao Luo[2], Dan Luo[3], Chengqi Zhang[4]
[1,4]*Faculty of Information Technology, University of Technology Sydney, Australia*
[2,3]*Department of Electronics and Information, Liaoning Technical University, China*
{[1]lbcao, [4]chengqi}@it.uts.edu.au, {[2]chao.luo, [3]dan.luo}@mail.ia.ac.cn

Abstract

Usually, integration of business intelligence (BI) from realistic telecom enterprise is by packing data warehouse (DW), OLAP, data mining and reporting from different vendors together. As a result, BI system users are transferred to a reporting system with reports, data models, dimensions and measures predefined by system designers. As a result of survey, 85% of DW projects failed to meet their intended objectives. In this paper, we investigate how to integrate BI packages into an adaptive and flexible knowledge portal by constructing an internal link and communication channel from top-level business concepts to underlying enterprise information systems (EIS). An approach of three-level ontology services is developed, which implements unified naming, directory and transport of ontology services, and ontology mapping and query parsing among conceptual view, analytical view and physical view from user interfaces through DW to EIS. Experiments on top of real telecom EIS shows that our solution for integrating BI presents much stronger power to support operational decision making more user-friendly and adaptively compared with those simply combining BI products presently available together.

1. Introduction

The objective of building a Business Intelligence (BI) system [1] is to equip its users with ability of making decisions in realistic business production and operations scientifically and systematically. The ordinary approach from BI system integrators for building a BI system is to pack all BI packages commercially available together in order to provide analysis and reporting services.

However, almost all solutions usually available are prone to provide subjects, data models, analytical dimensions and measures, and reports for users predefined in design time. However, business requirements of operational analysis and internal structures of underlying EIS are often in a dynamically realistic environment. As a result, an existing BI system cannot adapt to changing or new requirements emergent in the problem domain daily. 85% of DW projects failed to meet their intended objectives, and 40% didn't even get off the ground [2].

In this paper, we report some of our explorations in integration of BI from business to DW and to EIS [3]. First we survey the complexities of integrating BI in telecom industry. We then introduce a three-level ontology space, which includes business ontology for user profile, DW ontology for the DW, and EIS ontology for underlying business/operation support systems. This external channel and its system architecture are proposed for providing business persons with a business-oriented rather than technology-centered analysis and reporting portal, which integrates BI from user profile, DW and EIS, and managing all business-oriented analysis and reporting. Ontology query and algorithm are also presented, which query and transform from key words or ontology concepts in source ontology domain to ontology concepts in target domain. This may be done among ontology domain of user profile, ontology domain of DW, and ontology domain of EIS from user interfaces through DW to EIS.

This work is part of our activities of building telecom BI system as a knowledge portal by integration of DW, OLAP, DM, and reporting systems commercially available [3]. Our experiments in the real world of telecom industry have shown it can support transparent integration of the above mentioned modules and provide user-friendly, flexible and adaptive capabilities for decision making scientifically and systematically.

2. Complexities in integrating BI

As a matter of fact, the process of integrating and mining business intelligence in real telecom industry is quite complicated. Complexities of integrating telecom BI may take forms of openness, heterogeneity, distribution, evolution, and emergence [4], which may come from multiple aspects.

First, the ubiquitous complexities are co-existing in the business environment and underlying operational systems: (i) so many distributed business and operation support systems providing services as Billing, Switches, Accounting, Customers Services, Operations and Maintenances, which usually are developed by different system integrators, respectively, (ii) many other related information systems, like Customer Relationship Management (CRM), Enterprise Resource Planning (ERP), Office Automation (OA), Management Information System (MIS) and external systems, (iii) varied of hardware and software platforms and infrastructures provided by multiple vendors, (iv) different specifications used for system analysis and design in different systems of BSS/OSS, (v) diversity of data structures hidden in the above systems, (vi) diversiform business models and taxonomies, operational workflows and information flows existing in both different systems and operators, (vii) hundred millions of customers with localized and personalized service requests, and (viii) daily changing and emergent behaviors both in systems and operations.

Second, a predefined BI system cannot adapt to changing or new requirements in run time both of data and analyses daily emergent in the real world. It is sometime helpful for users to utilize resources from EIS (usually an Operation Data Store rather than on EIS directly) when new changes haven't been added to data models by integrators.

Third, incapability of the present BI system by packing all related commercial BI products together also results from internal differences and incompatibility with specifications, interpretations of some similar terms and relations, interoperability in system architecture and design patterns, supports for information integration, metadata management, methodology and concrete design methods among different productivity systems, BI products vendors and BI system designers. All these make it very hard to construct a concentrated (or distributed as required) system (cluster) dealing with data warehousing, OLAP and data mining analyses smoothly and transparently. So, a BI system simply packing all related components cannot make its objective come true in a dynamic and live world.

Fourth, for users of BI system, what they want to interact usually is just a workplace on the network (namely a knowledge portal). It is expected that it can not only seamlessly adapt to their existing business definitions, processes and rules, but also dynamically bring them flexibility and intelligence for making decisions on finding interestingness and scientific evidences hidden in huge operational stream. So, it is a significant requirement for BI system architects to set up a bridge which helps their customers to cross the great gulf between their daily top-level concept-oriented business process and rules, and the multiform physical views inside BI and EIS.

Therefore, in order to set up a practical and productive BI system, from the above discussion, lexical and semantic [5] transformation and integration of heterogeneous information hidden in telecom business and operation systems has been a first important problem. As a matter of fact, it gets involved in the whole process of building BI system from data preparation to intelligence discovery. Another key issue is how to support dynamic and adaptive integration both of DW and EIS as required.

As a possibly fundamental solution, an embedded link and communication channel beyond the traditional linkage among reporting, DW and EIS should be set up from business portal to low level EIS for supporting run-time capabilities, helping users to adapt to live problem domain and requirements in a business-oriented rather than technology-centered way. This means business persons can interact with the system only in their favorite high level terms, it is not necessary for them to deal with symbols and jargons in the DW and EIS. As a one-stop interface to underlying colorful world, we should present BI users a unified knowledge portal, which supports online transparent integration of information, integrating and mining of business intelligence, and decision making from the view of business rather than technology.

3. Ontology namespace

In order that terms and identifiers can be used and understood without disagreement, we need a precise indication of what specific vocabularies are being used, and a precise declaration of how to organize elements and properties in a systematic way. This is what ontology namespace does. The basic strategy in this work about ontology namespace is to refers to specifications of XML namespace [6], RDF [7], and

OWL namespace [8]. This means that: (i) we use XML to describe/declare ontology and services, (ii) follow the similar rules, constraints, defaulting, overriding, scoping, etc. for XML namespace elements and attributes, and (iii) utilize rules and functions from OWL and RDF. The difference here is that it is not necessary for us to be w3.org references. These provide a means to unambiguously interpret identifiers and make the rest of the ontology presentation much more readable.

On the other hand, for the complexity and heterogeneity of naming in the underlying individual business and operation support system (BSS/OSS), and high level requirements from business and analysis, it is hard for us to enclose all terms hidden in levels from user profiles, through DW to the bottom BSS/OSS into one ontology namespace. In telecom industries in China, situations even get more complicated for so huge amount of customers, six operators providing overlapped services by using different technologies and specifications/metrics in the BSS/OSS and business analyses.

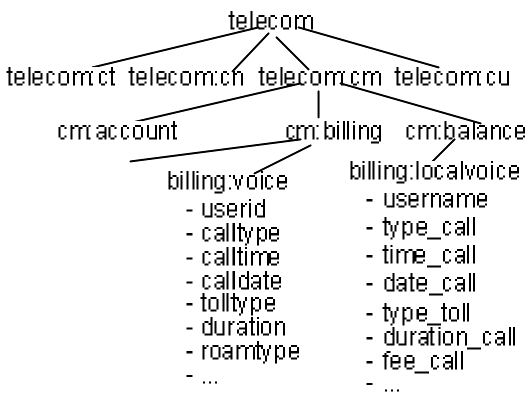

Figure 1. Partial business ontology domain of telecom in China (where ct refers to operator of China Telecom, cn directs to China Netcom, cm refers to China Mobile, and cu points to services provided in China Unicom)

Our strategy here is to build up a unifying business ontology namespace (here business ontology namespace means that it is abstracted from business analysts' view rather than from technicians designing EIS, so it mainly deals with interaction between users and portal, users and DW) for user profiles for all operators in telecom industry. Furthermore, the labels can be multiple, one in English for DW-related terms, another in Chinese for user interaction. For technical ontology namespaces (we mean that this ontology namespace is based on naming and reality in the underlying EIS, which is mainly abstracted from the BSS/OSS designed by some specific vendors in their favorable terms and naming specifications), different namespaces are designed for different EIS such as Billing system, Account system, Balance system used in different operators who rely on particular business models and telecommunications services (see Figure 1). However, for the technical ontology namespace there is meta-namespace hidden inside which can be implemented and instantiated specifically for each individual operator.

Another step after naming of ontology is to describe ontology concepts properly. This work also concerns semantic relationships of ontology concepts. The basic idea is to build up set of specifications for stating and managing ontology concepts and relations between each other. We have discussed about these issues briefly in another paper [9].

4. Structure of ontology integration

In the process of investigating how to integrate user profiles, DW, OLAP, DM and underlying heterogeneous and distributed business/operation information systems from different vendors, we proposed the following four-tier structure for integration of business intelligence as shown in Figure 2. The objective is to make a smooth mapping from top-level user defined key words/phrases to metadata items in DW or physical attributes/entities dispersed in operational tables of BSS/OSS.

There are four-level views coexisted in this ontology integration system from top down: a top-level Conceptual View supporting user portal interaction in user profiles, an Analytical View for DW-based data modeling and analysis, a low-level Physical View for EIS enclosing multiple business/operation support systems like billing system, accounting system, and others. The fourth view is a Ontology Mapping and Query Parsing mediator level, which provides supports for ontology mapping and query parsing. In the framework, relevant functionality components are also shown for individual tier. These components handle user interaction, data model, metadata management, mediation/transport/directory/naming of ontologies and services for each level, query transformation and parsing, and integration of data sources.

The key linkage between all four tiers is ontology concepts and services. Figure 3 further shows the structure of integration of ontologies located in different levels. The existing strategy commonly used in current BI systems, is to predefine metadata items in building data models on top of EIS, and link four-type analytical reports, DW and EIS by metadata and predefined mapping relationships between metadata items and entities, as shown by the line with double arrows. In order to support flexible and adaptive online customization, transformation, management and integration of user profiles, DW and EIS, we also set up ontology namespaces for user profiles, DW, and every sub-system of EIS, respectively. For instance, there is a global business ontology domain specified for user profiles according to business model and concept model, while for DW there is a DW ontology domain managing all concepts in the DW. But for Billing and other systems in BSS/OSS, for the complexity of heterogeneity and difficulty with building up a uniform ontology domain for all EIS, we alternatively set up specific ontology domains for Billing, Accounting and other systems, respectively.

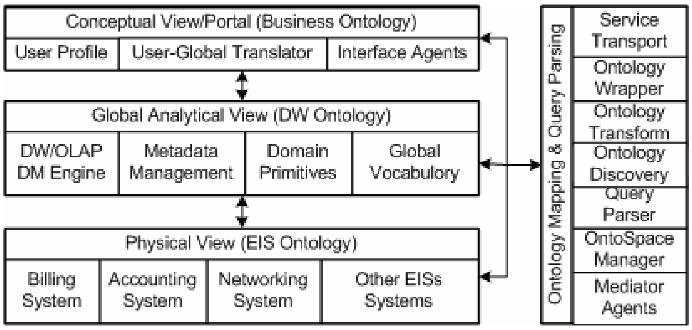

Figure 2. Structure for integrating business intelligence through ontology services

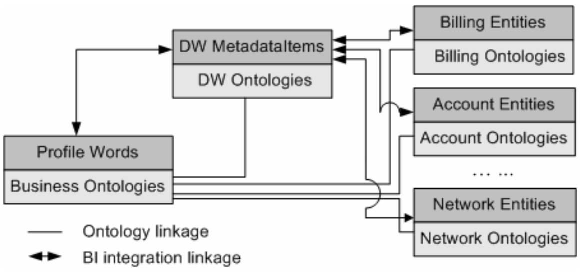

Figure 3. Three-level ontology space

5. Three-Level ontology space

Our major objective of heterogeneous information translation and integration includes: (i) providing transparent and seamless integration of the underlying heterogeneous resources among telecom operational systems, (ii) supporting smooth transformation from business concepts in user interface to low-level entities in specific resource systems, and (iii) furnishing online interactive techniques for transparent interoperability and smooth translation among levels.

5.1 Conceptual view and business ontology

The objective of Conceptual View is to present users with domain specific concepts, objects, business rules, and user interfaces in a conceptual profile user-friendly. The ontologies capture general knowledge about concepts, terminology and relationships from viewpoint of business in the world.

The output of this view is a conceptual ontology base, which includes a Concept Category Directory (CCD). The CCD, which is a hierarchical concept tree implementing telecom business namespace, lists and defines all terms and relationships abstracted in daily business, and generates a list of candidate concepts and expressions based on the business process and activities happened in the user views. Here, a concept rather than an attribute or entity is used to describe the world. For instance, a term of *Conditions…* rather than *Where…* is used in generating a query.

DEFINITION 1. Concept Category Directory Entry: A CCD entry consists of a unique Leading Item (LI, an identifier), and optionally multiple Substitute Items (SI, recommended candidate concepts) as follows:

{{{*Leading Item, MO*}:{*LI_Value*}},{{*Substitute Items, OM*}:{*SI_Value1, SI_Value2,…* }}},

for instance:

{{{*LI_Service_Provider*}:{*Service_Provider_Label*}},{{*SI_Service_Provider*}:{*Service_Provider_Name,Service_Provider_Nickname,Service_Provider_Description* }}}.

5.2 Analytical view and DW ontology

The Global Analytical View is a logical aggregated representation of underlying logical elements and relationships locating in DWs, OLAP server and DM engines. So, in terms of domain specific primitives in telecom information systems, this global side ontology wraps technical and business metadata items. These metadata items are defined in set of elements (attributes, dimensions and measures) in data model, source data, ETL, and also actions and rules of interaction between DW and data sources.

DEFINITION 2. Analytical Ontology Directory Entry: It consists of some domain-specific metadata items, which focuses on business and technical metadata required in the problem domain, rather than on business rules and concepts. The following elements are enclosed in it: globally unique identifier *(gui)*, recommended global name *(rgn)*, candidate substitute names *(csn)*, parent object *(po*, top-level coupled concept name*)*, child objects *(co*, low-level EIS instances*)*, analytical locator *(al*, where to find this entry from the bottom EIS resources, including related connection string, schema, metadata of resources, and so forth*)*, close associators *(ca*, including actions and relationships with other neighboring entries*)*. Furthermore, the cardinality property of an entry is shown in the following:

{{gui, MO}, {rgn, MO}, {csn, OM}, {po, MO}, {co, MM}, {al, MO}, {ca, OM}}

All item atom entries are stored into knowledge base and registered into ontology name database. The usage of KVT, KPP, the Pair-Element encoding system, and the introduction of elements *parent* and *child*, *locator* and *associators*, can help avoid conflicts of data type, scaling, generalization, naming and location.

5.3 Physical view and EIS ontology

The Low-level Physical View is a representation of physical entities and relationships related to transactions among underlying information systems. The most common form of Physical View is as tables and attributes located in EIS. EIS enclose multiple enterprise information resources in which store huge amount of operational data and information. On this level, telecom operational systems like BOSS, MIS, ERP, OA are all resource providers of the DW system.

In terms of technical implementation, the multiplicity of this level also brings us a colorful world of physical instances/attributes/relations, and so forth. For instance, the counterpoints of *Customer Name* on user conceptual view, may take names as *Customer_Name*, *Customer_Label*, *User_Name*, *User_Label* etc. in physical systems. These names may be distributed into the operational systems like Billing, Accounting, Switch, and Operation and Maintenance systems, etc.

6. Ontology services-based query algorithm

With key words and ontology concepts in user profile, DW and EIS, a key issue is how to do mapping, transformation, discovery and query of them between ontology domains [5, 9]. Particular algorithms and system components are required for dealing with these requirements. In this paper, we only discuss ontology query algorithm for limited space.

In order to handle query of ontology concepts, we propose an algorithm called OntologyQuery. The idea in OntologyQuery is to handle ontology query by a process combining ontology transformation and match as required, and a method combining automatic search and manual search if needed.

The details of OntologyQuery are in Figure 4. The first step (L1) is to check whether a user types in terms in native language (it exists and is not in English) rather than in English as the default, the ontology services S1 is quoted to output the key words according to transformation rules. If a user inputs key words or ontology concepts directly (L7), and wants to query the target ontology concepts from domain of user profile ontology (L8), then ontology services S2 is started up and looks for target concepts from user profile domain. If the user wants to query target concepts from DW ontology domain (L15), then service S3 is executed, and output found concepts. In case it is required to look for ontology concepts from EIS domains, then S3 will be run and export the target terms.

With regard to algorithms of the ontology services S1: UserProfileTransformer, S2: UserProfileMatcher, and S3: AutomaticOntologyMatcher, we will not discuss here in details for space limitation.

The advantages of ontology query and transformation using ontology services include (i) supporting user personification, (ii) supporting dynamic changes in both data model and EIS, which actually take place very commonly in real analytical requirements and operational environment.

7. Unified knowledge portal

To BI system users, what they care and want to interact is a one-stop entry. From this entry, they can capture information flow in the whole enterprise in multiple granularities as they like, undertake any online analytical work of the four-level analyses, obtain hidden information by DM technology, and finally are able to make reasonable and full-of-proof decisions to strengthen the efficiency and effectivity of daily analyses and decision making. This is the place what we called Unified Knowledge Portal (UKP) for enterprise decision making. Through the UKP, business intelligence emerges from the underlying EIS, DW and OLAP, and DM systems.

Algorithm OntologyQuery(input(s), output(s))

Input(s): $\{c_1, c_2, \ldots, c_m, p\}$

- $c_1, c_2, \ldots, c_m$: User-defined key words/phrases OR ontology concept(s) in user profile
- p: user profile property

Globals:

- langLabel: what language does a user use to input the key words or concepts, ‘EN’ for English, ‘CH’ for Chinese
- ontoSpace: which ontology namespace to search for the targets, ‘UP’ for ontology of user profile, ‘DW’ for data warehouse ontology space, ‘EIS’ for ontology spaces in BSS/OSS
- simValue: similarity value for the matching from one to another, which is in [0,1]
- outputMatchingRule: the current output resulting matching rules
- outputOntologyConcept: the current ouput ontology concept

Output(s): $\{o_1, o_2, \ldots, o_n\}$

- $o_1, o_2, \ldots, o_n$: recursively find and output resulting matching ontology concept(s) in target ontology space on Input with user profile property
- recursively store new matching rule if available into knowledge base

Method:

```
if langLabel <> en then //User input key terms in native language
    for i = 0; i < m; i++ //For each input item
        S1: UserProfileTransformer(c_i); //Query key words/phrases of user profile
        outputOntologyConcept(i); //Output key words/phrases
        outputMatchingRule(i); //Output query matching rule
    end for
else //User input key words/phrases or ontology concepts directly
    if ontoSpace == UP then //The target ontology space is in user profile
        for i = 0; i < m; i++ //For each input item
            S2: UserProfileMatcher //Search ontology concepts in ontology space of user profile
            outputOntologyConcept(i); //Output ontology concept
            outputMatchingRule(i); //Output query matching rule
        end for
        outputOntologyConcept(); //Output final ontology concepts
    else if ontoSpace == DW then //The target ontology space is in DW
        for i = 0; i < m; i++ //For each input item
            S3: AutomaticOntologyMatcher //Search ontology in DW ontology space
            outputOntologyConcept(i); //Output ontology concept
            outputMatchingRule(i); //Output query matching rule
        end for
        outputOntologyConcept(); //Output final ontology concepts
    else if ontoSpace == EIS then //The target ontology space is in EIS
        //start: Target ontology space of which BSS/OSS system(s) in EIS
            for i = 0; i < m; i++ //For each input item
                S3: AutomaticOntologyMatcher //Search ontology in BSS/OSS ontology space
                outputOntologyConcept(i); //Output ontology concept
                outputMatchingRule(i); //Output query matching rule
            end for
        //end: Target ontology space of which BSS/OSS system(s) in EIS
        outputOntologyConcept(); //Output final ontology concepts
    end if
```

Figure 4. Algorithm of ontology query transformation in the three-level ontology space

Technically, a UKP is an Internet-oriented platform for enterprise decision making. It also acts as single sign-on which seamlessly integrating all or most relevant verification, services, and applications in an enterprise. However, how to make the portal transparent and powerful to seamlessly embed all necessary components involved, and intelligent and active to flexibly emerge knowledge and decision-making evidences online, is still a big challenge to most of the commercial BI products available at the stage.

According to what we have discussed in the above sections, we have constructed a prototype of BI system called IOAS: Intelligent Operational Analysis System (see [3, 9] for details).

In the IOAS, IBM DB2 Universal database, DB2 DW, Oracle OLAP server, Cognos reporting, and Intelligent Miner are used. Under the IOAS, Informix, Oracle, Sybase, SQL Server are used in telecom EIS for storage of respective operational transactions. This platform is organized according to subjects and specials in the DW and marts, and additional functional modules for system management and decision supports. These subjects and specials and their related dimensions, measures, attributes and members are abstracted and matched from the above several heterogeneous distributed database servers in terms of ideas of three-level ontology mapping and query parsing among Conceptual View, Global View and Physical View.

Unified certification and single sign-on are lodged in the UKP; measures, dimensions, attributes and members are in a flexibly and efficiently integral presentation on the one-stop site in business concepts well known to users rather than in views of low-level source data. To users, complexities of multiple heterogeneous EIS resources, ETL process, ODS, DW, reporting presentation, and mapping from business concepts to underlying physical entities are shielded and hidden under the one-stop interface. They can easily launch analysis and observations without worries of underlying symbolization, authorization and information management from DW to bottom EIS.

8. Conclusions and future work

Implementation of a BI system has been attracting more and more interests from telecom industry. In this paper, we have proposed a new approach for integrating BI. This is a three-level ontology services space, which presents a business-oriented rather than technology-centered channel for synthesizing user profile, reporting, DW, OLAP and DM engines, and/or synthesizing telecom business operational systems as required, into a unified knowledge portal. The three Strategic ontology domains: a Business Ontology domain, a DW Ontology Domain, and possibly multiple EIS Ontology Domains, are presented to arrange the BI integration. Naming, system architecture synthesizing the above three strategic ontology domains, and ontology services-based query transformation between ontology concepts and its algorithm are introduced.

This work and some others are our explorations in building a practical and productive BI system, which is business-oriented rather than technology-centered, in China telecom industries. It has shown that it is more user-friendly, flexible and adaptive for telecom customers to online integrate and analyze business intelligence based on data in DW and/or from huge amount of business transactions, than simple packing BI components together.

This work is interesting both for research and industry. A systematic investigation is worth for the future in ontology representation, match, discovery, integration and application with existing systems of DW, OLAP, DM, and reporting in the real world.

References

[1] Mark Whitehorn, et al, *Business intelligence: the IBM solution*, New Springer, London, 1999.

[2] Usama M. Fayyad. Tutorial report. Summer school of DM. Monash Uni, Australia, July 2003.

[3] Longbing Cao, Dan Luo, Chao Luo, Chengqi Zhang. Systematic engineering in designing architecture of telecommunications business intelligence system. *Design and application of hybrid intelligent system*, A. Abraham et al (Eds.) IOS Press, 2003.

[4] Longbing Cao. *Studies on some problems in multi-agents-based open giant intelligent systems*. PhD thesis, Chinese Academy of Sciences, 2002.

[5] John Davies, Dieter Fensel, Frank van Harmelen. *Towards the Semantic Web: Ontology-Driven Knowledge Management*. John Wiley & Sons, 2003.

[6] Namespaces in XML 1.1, www.w3.org/TR/2004/REC-xml-names11-20040204/.

[7] http://www.w3.org/RDF/

[8] http://www.w3.org/TR/2004/REC-owl-test-20040210/

[9] Longbing Cao, Chao Luo, Dan Luo, Li Liu. Ontology services-based information integration in mining telecom business intelligence. *Proceeding of the 8th Pacific Rim Int. Conf. on Artificial Intelligence*, Auckland, New Zealand, 9-13 August, 2004.