Integrating Telecom CDR and Customer Data from Different Operational databases and Data warehouses into a Central Data Warehouse for Business Analysis

Sharmistha Mandal
A.K. Choudhury School of IT
University of Calcutta
Kolkata, India

Giridhar Maji
Department of Electrical Engineering
Asansol Govt. Polytechnic
Asansol, India

Abstract— **Mining operational telecom data has given many interesting insights about customer behavior, mobility, social interaction etc and has great business value. But to have truly unbiased analysis over a defined geographic region, we need to first consolidate all telecom data from many different operators present within the region and then analyze them. This necessitates an integration architecture that will allows to integrate, consolidate and store different type of operational telecom data such as call detail record (CDR), customer relationship management (CRM) data, network data etc into a central data warehouse (DW). Though almost all operator stores semantically same type of data but in a little different form and format. Again these operators are separate business entities so integrating data from their operational databases and data marts needs to be independent of each other and flexible in fetching data for extract, transform and loading processes.**

In this paper we shall review a few existing DW schemas used to consolidate CDR and CRM data and propose a generic model DW schema with indentified facts and dimensions for the central DW. Next we propose integration architecture to collect data and consolidate in the central DW from many different operational databases of different telecom operators. Web services (WS) has been proposed as a mode of communication between central DW and other operational databases.

Keywords—Datawarehouse; Integration Architecture; Telecom CDR; DW Integration Using Webservices;

I. INTRODUCTION

Mining telecom operational data such as Call Detail Record(CDR), Customer Relationship Management (CRM) data, subscriber master data etc have huge potential in predicting customer behavior, customer mobility and many other business applications such as market share analysis of different mobile brands[23], telecom pricing strategy analysis[10], churn prediction[25], social network analysis[26] etc. In developing and emerging economies where there are multiple big players in telecom market, collecting data from only a single telecom operator and analyzing will not give much accurate outcome. As cities in country like India has around 8 to 10 telecom service provider and there exists no single operator with more than 60% market share. So analyzing data from only a single operator will give biased and partial result. There are situations where there are social, cultural and economic biases among people regarding subscription of some particular telecom operator. For example a high quality service oriented operator will get more subscribers from wealthy and upper class communities. Again low cost operators will have more subscribers from lower middle class community. Also number of subscriber of an operator changes depending on the infrastructure present throughout the region. May be some operator has well established infrastructure for voice data and other services in a part of a city and not good infrastructure in other parts, then subscriber concentration will be more in that part of the city for that operator. Again as different social communities tend to be clustered geographically so there will be high probability that that area of the city is inhibited by people of certain community. So any analysis based on data solely collected from a single operator will have all kind of biases.

Once we integrate all data from different operator for a city or region we can have truly unbiased result that can be represented as the outcome for the whole region without any partiality. All these competing operators store business data like CDR, CRM, Subscriber master etc in OLTP (Online Transaction Processing) systems as per their own form and format. Even sometimes these operators can have their own data marts with consolidated CDR and CRM data. Though almost all telecom operator stores same type of details but there is no standards regarding data format or form to store in operational Databases or data marts. So to perform any analysis based on CDR and CRM data for a city or state in these type of countries where there are multiple big players It is of utmost important to integrate those operational (OLTP) data and in some cases consolidated data from their individual data marts into an centralized Data warehouse for Online analytical Processing (OLAP). This will allow data mining and business intelligence analysis on whole consolidated data for true un-biased and general result. In this study we will propose an integration architecture that will integrate CDR and CRM data from individual telecom operators' operational data sources or data marts into a central Integrated Data warehouse. Mobile location data collected from CDR of different telecom operators' OLTP systems along with customer details fetched from CRM databases are integrated into a central Data warehouse (DW) system with properly designed Extract-Transform-Load (ETL) processes and then aggregated with different dimensions to find various business requirements.

II.   RELATED STUDY

Due to the widespread adoption of mobile service in many developing countries, an extensive amount of mobile usage data is being created every day that can potentially power development initiatives. Important information about users is captured by telecommunications systems each time an operator enrolls a customer, as well as each time a call is made, a text message is sent, or a mobile money transaction is conducted. These details are stored in their day to day operational systems (OLTP systems) and commonly known as Call Detail Records or CDR. A typical CDR contains details like unique customer identity, Calling party cell tower location, called party cell tower location, call duration, call start time, rate plan used, roaming status etc among others. CDR data is linked to customer master table by unique customer identity and it contains other customer details like name, address, sex, marital status, age, annual income etc. All these data has been used for many different purposes like billing, promotional campaigning, churn prediction etc by the concerned operators.

While we think about using the geospatial aspect of CDR to understand social behavior or want to use mobile location using cell tower Ids from CDR to analyze social development and urban planning of a geographic region then we need whole integrated data of that area from all telecom operators [1]. Integrated CDR data of a region can be used for market share analysis, cultural studies of a group of population based on mobility and social events in a geographic regions [23],[2].

Figure 1: multidimensional data cube with dimension hierarchy
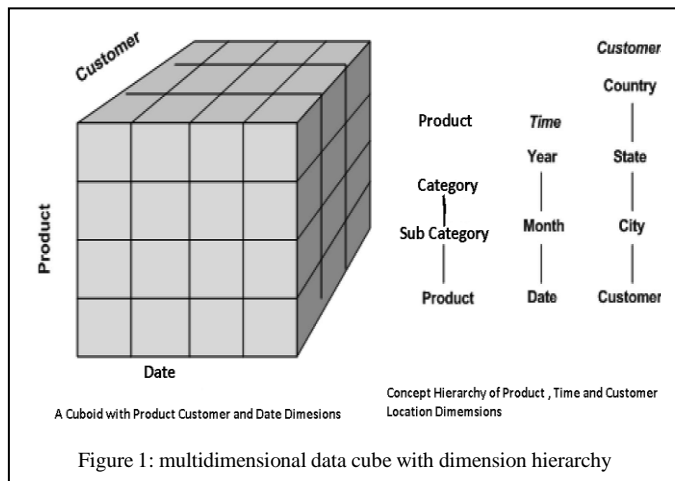
Any Integration system typically consists of wrappers and mediators. A wrapper is a software module used to access a data source, extracts data, and presents it in a predefined format, usually as a set of relational tables. A mediator collects, transforms and consolidates data received from wrappers depending on specific information need of the integration system. The specification and the design of wrappers and mediators are the core issues of an integration system [13].

A.   Data warehouse and OLAP

Inmon defines a DW as a subject-oriented, integrated, non-volatile, and time variant collection of data in support of management's decisions [3]. Again Kimball defined a DW as a replica of transaction data specifically structured and organized for query and analysis [4]. Thus, data warehousing involves the construction of a huge repository where an integrated view of data is given, which is optimized for analysis purposes.

OLAP tools are used to exploit the huge amount of information stored in Data warehouse. Codd first coined the term OLAP and presented 12 rules to evaluate any OLAP system [5]. OLAP tools conceptually model the information as multidimensional cubes. Figure.1 shows an example data cube. In the cubes, data is divided into facts, and dimensions. Facts are the central events for any concerned analysis (e.g. a sale), and dimensions used to provide contextual information to the facts (e.g. the products sold, duration of sale etc). Generally dimensions are organized into hierarchical levels. For instance, products can be grouped into product categories that can be further grouped under a brand. Generally, the facts are associated with numerical measures like sales amount, sold item count etc, and queries aggregate these measure values up to a certain level (e.g., total sales by product category by quarter), followed by either roll-up (further consolidation, e.g., to year) or drill-down (more detail, e.g., sales per month or week) operations.

OLAP tools are mainly of two types such as Multidimensional online analytical processing or MOLAP which stores data on disks in specialized Multidimensional structures. MOLAP systems typically include provisions for handling sparse arrays and apply advanced indexing and hashing to locate the data when performing queries; and Relational OLAP (ROLAP) systems which uses relational database technology for storing data, and they also employ specialized index structures, such as bit-mapped indices, to achieve good query performance.

We will employ a ROLAP system as our integrated DW as ROLAP system scales better in the number of facts they can store, and is more flexible with respect to cube redefinitions, also provide better support for frequent updates.

ROLAP implementations generally employ star or snowflake schemas, both of which store data in fact tables and dimension tables. A fact table holds one row for each fact in the cube. It has a column for each measure, containing the measure value for the particular fact, as well as a column for each dimension that contains a foreign key referencing a dimension table for the particular dimension.

Star and snowflake schemas differ in how they handle dimensions, and choosing between them largely depends on the desired properties of the system being developed. A star schema has one table for each dimension. The dimension table contains a key column, one column for each dimension level containing textual descriptions of that level's values, and one column for each level property in the dimension.

B.   Telecom Call Details Records (CDR)

A CDR is a call detail record that is generated for every user event on the telecom network and contains a broad spectrum of relevant information, including the phone numbers of both the calling and receiving parties, the tower location, the start time of a call, and its duration. CDRs are a subtype of event detail records (EDRs), which contain information about calls, SMS, value-added services, data usage, and all other services subscribers use. CDR as a term is often used to describe a phone usage data record beyond just placing and receiving calls and is used in that way in the remainder of this paper. A typical CDR data fields are shown in Figure 2.

CDRs show great promise for applied research–they have recently been used to explore human communications like Interplay between telecommunications and Face-to-Face Interactions [7], urban dynamics [6], and human mobility patterns [8],[9]. Furthermore, CDRs have been used to extract proxies of social networks. To safeguard privacy, the individual phone numbers are anonymized by the mobile phone operator and replaced with a unique, surrogate security ID.

CDR Field Name	Sample Values	Remarks
Calling Party Number	9195358XXXXX	
Called Party Number	9185358XXXXX	
Caller Cell ID	5102	
Caller Location Area Code	41708	
Callee Cell ID	2308	
Callee Location Area Code	43876	
Caller IMSI	4600013511068690	International Mobile Subscriber Identity
Served IMSI	4600013832018059	
Call Start Time	2015-10-06 14:24:32	
Call Duration	01:25:34	HH:MM:SS
Call Type		Voice , SMS or Data
Roaming Status	Y/N	
Rate Plan		
Prepaid or Post Paid	Pre/Post	
Calling Party IMEI Number	911345420081793	
Called Party IMEI Number	911845420151327	
Call Route	CellTowerId-5673	

Figure II: Example Set of CDR Fields

C. *Heterogeneous database and data warehouse integration Problems*

There are certain common general categories of issues that occur during any heterogeneous database or data warehouse integration. As Lee et al. [19] have proposed a classification scheme for the conflicts arising in heterogeneous databases:

(1) Value-to-Value conflicts: These conflicts occur when databases use different representations for the same data. This type of conflict can be further divided into data representation conflicts, data scaling conflicts, and inconsistent data.

(2) Value-to-Attribute conflicts: These conflicts occur when the same information is expressed as attribute values in one database and as an attribute name in another database.
(3) Value-to-Table conflicts: These conflicts occur when the attribute values in one database are expressed as table names in another database.
(4) Attribute-to-Attribute conflicts: These occur when semantically related data items are named differently or semantically unrelated data items are named equivalently. The former case is also called synonyms and the latter case homonyms [20].
(5) Attribute-to-Table conflicts. These conflicts occur if an attribute name of a table in a database is represented as a table name in another database.
(6) Table-to-Table conflicts. These conflicts occur when information in a set of semantically equivalent tables is represented in a different number of tables in another database.

In our integration architecture we shall resort to a global DW schema (star or snowflake depending on the concerned application) with all required facts and dimension table fields. As XML Schema has been successfully adopted in [21] for resolving the conflicts in heterogeneous database systems, we will use XML as the canonical standard to store data derived from local data cubes and use XML Schema to define the global schema. This schema shall be published in an xsd definition format and during ETL process this XSD (XML schema Definition) will need to be consulted for a one to one mapping between global integrated DW schema fields and individual operational database fields.

D. *Different CDR based DW with many differences in data form and format*

In this section we have shown a few of the existing CDR and CRM based standalone data warehouses with their fact and dimension table fields. Our integration architecture can be used to fetch data from many operational databases that stores these CDR or customer data or from standalone CDR and CRM based data warehouses of different telecom operators to integrate for a particular region. Most of the scenarios shown below there exists almost same data but with varying formats and forms. As we have discussed in sec 2.1 that some telecom operator has around 50 different relational tables to capture all call detail, billing and customer related data, again some operator already has their data marts or DWs with consolidated data. So our target would be to fetch data from either directly from relational tables of operational databases or from consolidated data marts or DWs but to maintain the mapping with our central DW schema. Also data conformity by all respect like same measuring unit, same data format , also semantically correct data needs to be fetched and transformed accordingly before integrating them into integrated central DW.

With a close look between the below two DW schema of same CDR and CRM data we see a lot of form and format and unit dissimilarity. And to overcome we will next propose a generic star schema for our central DW and also there will be a mapping guide book that will detail the semantics of each field of generic schema with proper form and format in which data has to be transformed for proper integration.
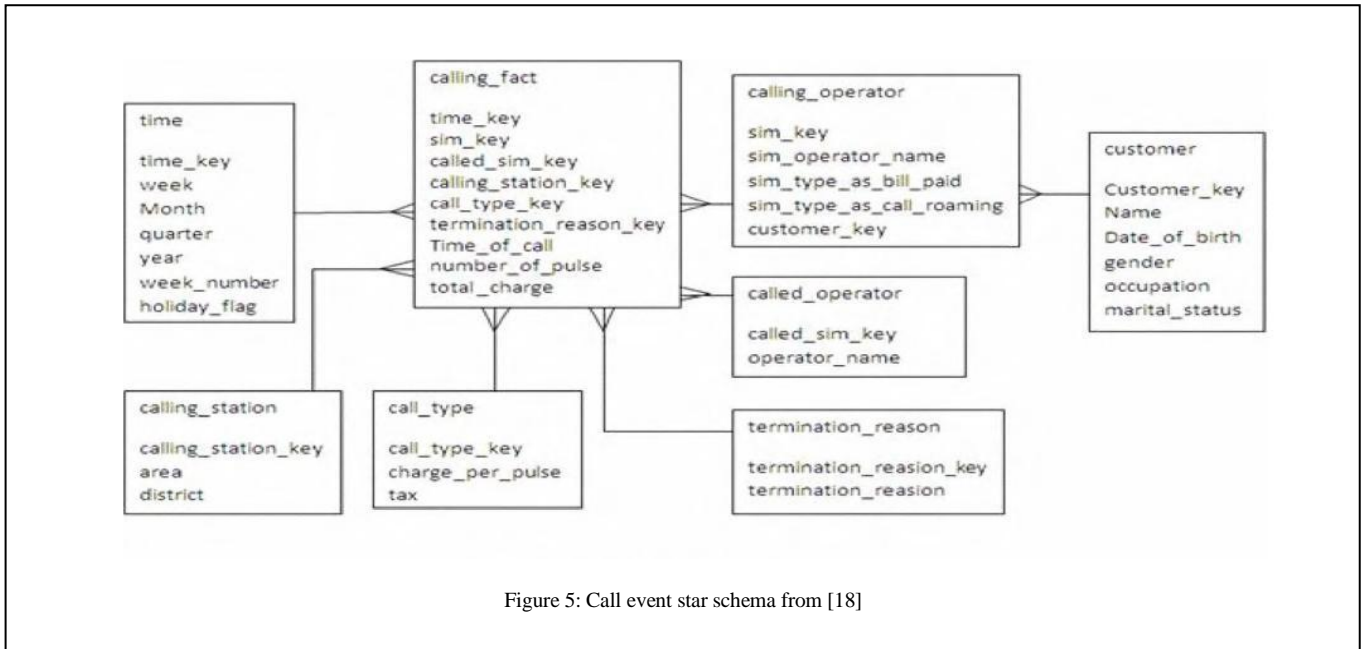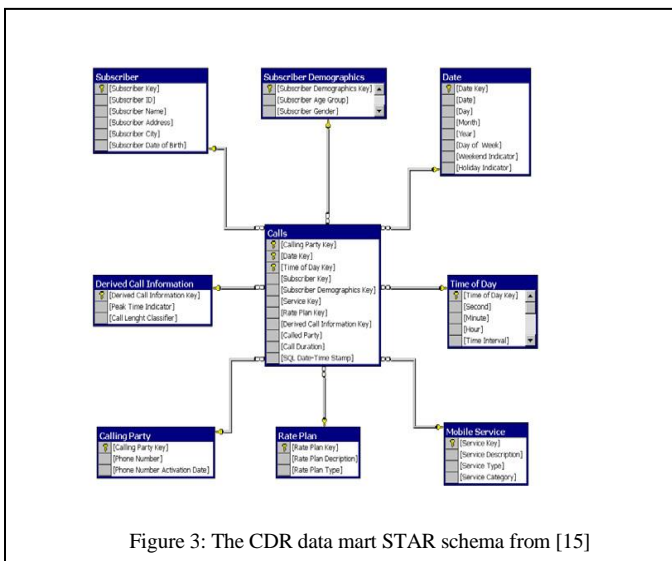
Figure 5: Call event star schema from [18]

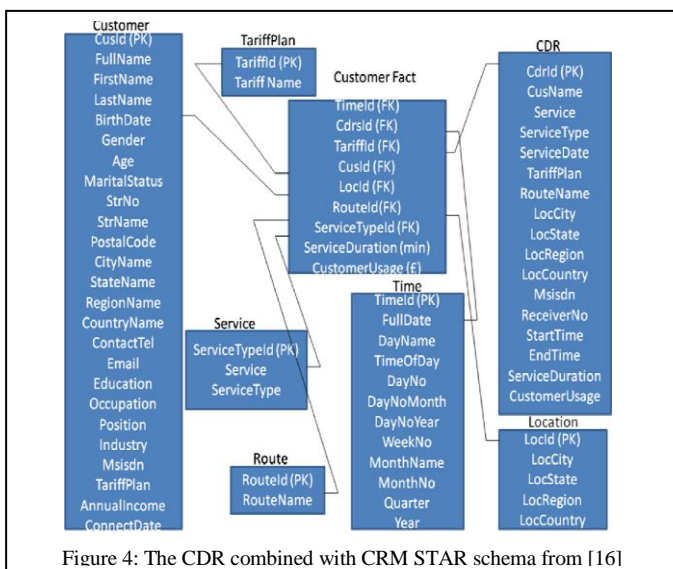Figure 3: The CDR data mart STAR schema from [15]

Figure 4: The CDR combined with CRM STAR schema from [16]

E. Related work on heterogeneous data warehouse integration

The authors in [22] considered four scenarios that have to be taken care with during integrating different data warehouses:

- Dimension hierarchies or levels with the same name, but different resources (identical dimension schemas).
- Different levels of detail for equivalent dimensions.
- Dimension levels with the same name but different meanings.
- Different dimension hierarchy names but the same meaning.
-

Our proposed system will handle all of this in a semi-automatic way. As there will be a global schema fields detailed description guide containing business meaning with semantics and syntax of the fields. And any client telecom system wrapper module has a mapping system which needs to be configured manually by the business user of each telecom system consulting the global schema guide. Once this mapping of each telecom operators' OLTP schema fields are mapped with the corresponding schema fields of the global schema transformation rules can now easily be defined. So after the mapping and defined transformation rules data can be easily fetched or extracted from the operational data sources or data marts of telecom operators and then transformed depending on the type of data and transformation rules and then loaded into the centralized DW system through a staging setup.
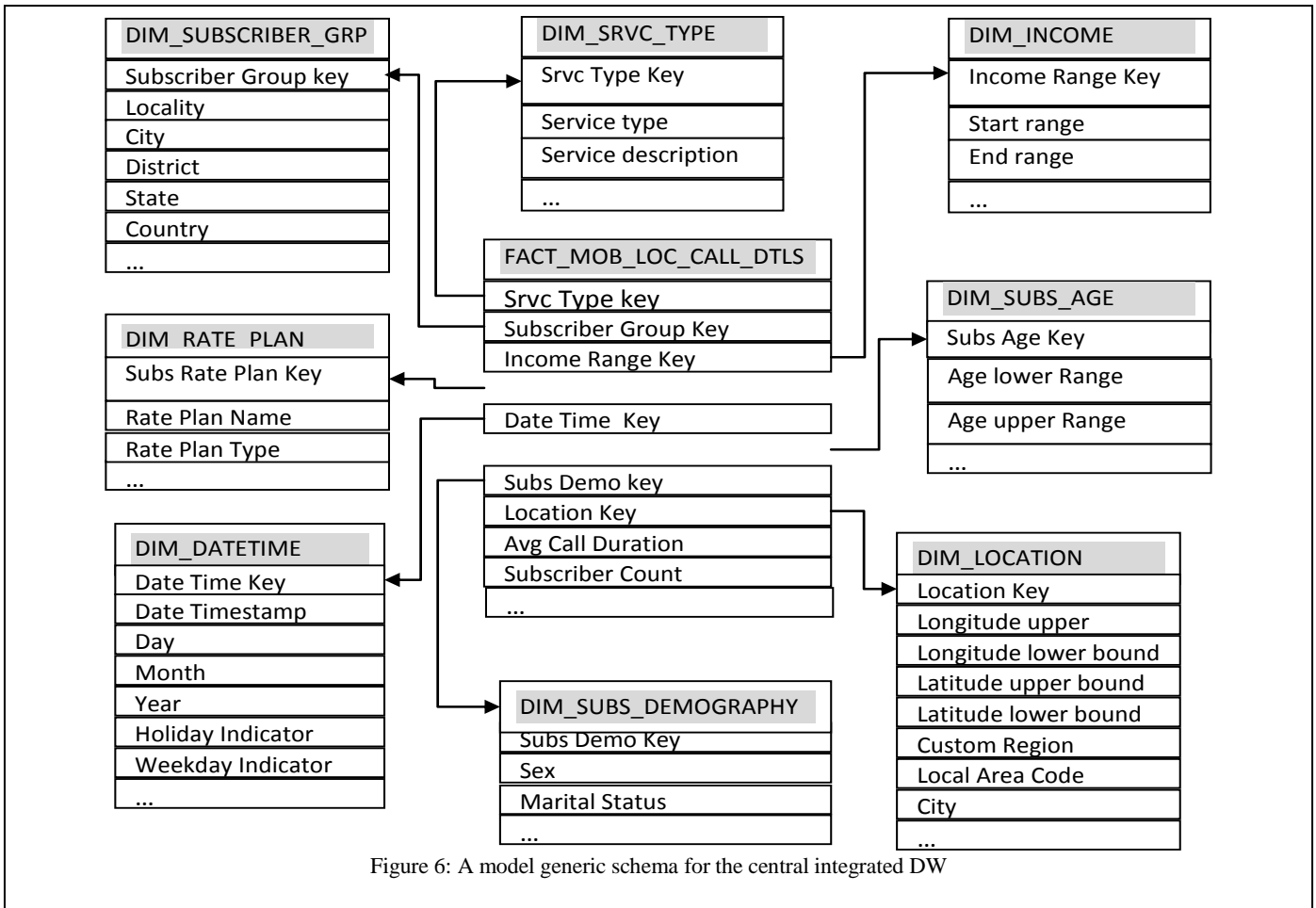
Figure 6: A model generic schema for the central integrated DW

Call Data Analysis group of Korea Telecom have made a Prototype DW System called Strategic Information Management System(SIMS), for taking strategic decision regarding customer oriented pricing in telecommunications sector by analyzing a huge volume of CDR [10]. SIMS uses oracle RDBMS and implements a star schema to facilitate multidimensional queries efficiently. Our proposed architecture will also use oracle as RDBMS for ROLAP operations. Main difference of our proposed system will be in not analyzing CDR from only a single telecom company rather building an architecture that can integrate CDRs from many telecom operators seamlessly for any general purpose analysis. Telecom Italia has their own DW project IBDA which now consists of 52 databases with many of them legacy systems [11],[13]. Telecom Italia stores their important historical data (CDR and a few other related information regarding customers etc) in a few sub systems like:

- IPDW—Interconnection Traffic Primary Data Warehouse, containing CDRs, whose main purpose is to analyze network usage patterns between TELECOM ITALIA Network Nodes and other service providers.
- CPDW—Customer Primary Data Warehouse, containing information on TELECOM ITALIA

- products and services by customer, to analyze customer data.
- TPDW—Voice Traffic Primary Data Warehouse, containing additional information on CDRs records from PSTN and ISDN switches to support various analysis tasks for Marketing, Administration, and Customer Management.

LGR Telecommunications, a specialized solutions provider to the global telecommunications industry, offering a unique business solution that taps directly into the source of each customer interaction with the network by accessing the CDRs it creates. LGR's approach intelligently captures CDR data in real-time, appends additional business information to the record, stores the data within a comprehensive Oracle data warehouse solution, and provides real-time analysis to the telecom service provider [12].

A Service Oriented Business Intelligence (SOBI) architecture to integrate data from heterogeneous data sources of a telecom organization has been developed and prototyped in [16]. The researchers have investigated the use of Service Oriented approach to BI in Telecom organizations towards providing real-time and accurate decision support about customer tariff plan in order to ensure customer satisfaction. They have used web services and SOAP/WSDL as a mean to communicate and transfer data between different components of the system. Web Services (WS) are platform independent systems which use standard network technologies to consent

Figure 7: Proposed Architecture for integrating diff. Telecom Data into a DW using WS
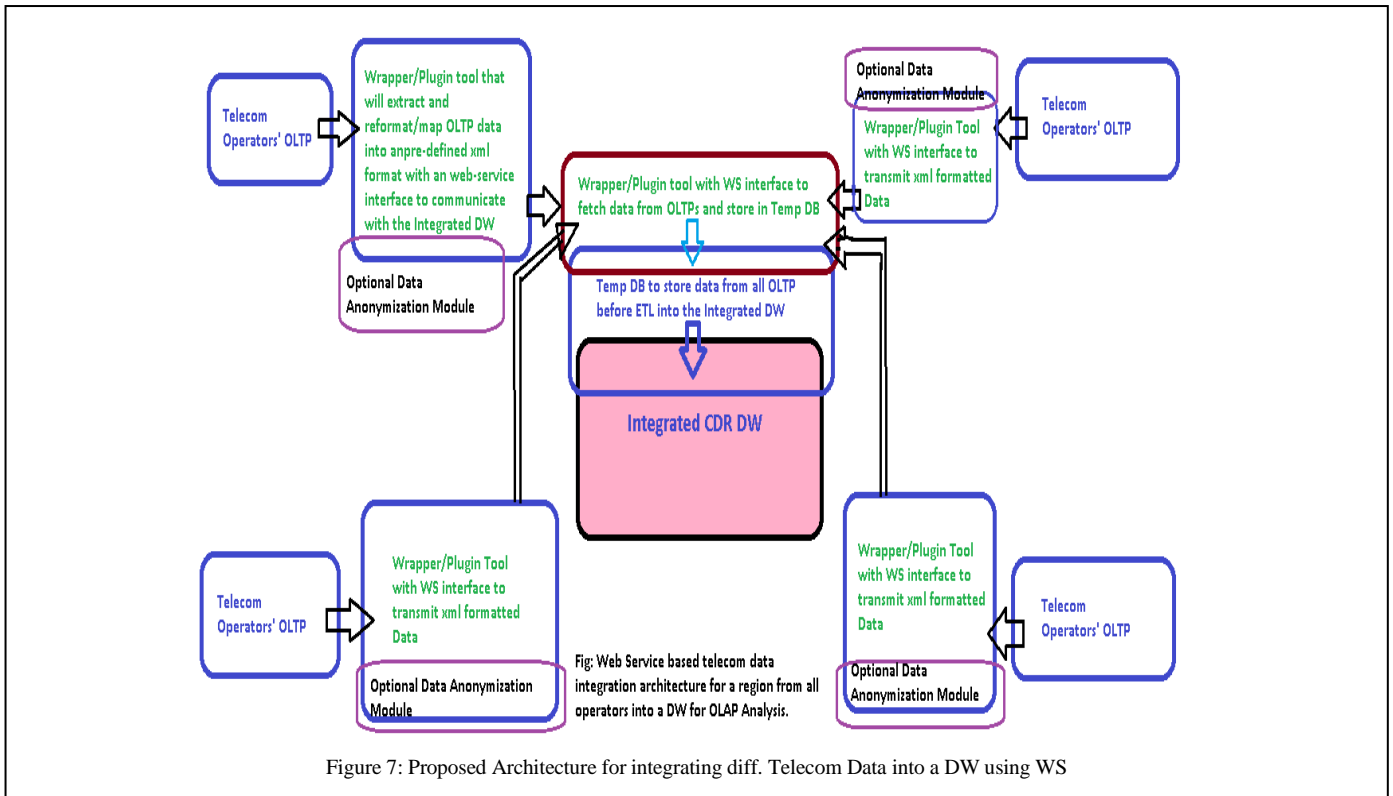
the interoperability between distributed and heterogeneous services in a transparent and flexible way [14],[17]. Our proposed architecture to integrate CDR and related operational data into a single integrated DW schema will exploit WS as a part in the wrapper and mediator module for the communication and data transfer in almost real time.

III. PROPOSED DW INTEGRATION ARCHITECTURE USING WEB SERVICES

The aim of the architecture is to integrate heterogeneous telecom operators' billing details, mobile location data and subscriber data (mainly CDR database and CRM databases). These databases contains almost same type of data for all operator but they store them differently (May be name of field is different or same data stored in different tables etc discussed in sec II.C and II.E). Our purpose here is to collect all those CDR and CRM data for a particular area (city, region) for a particular time (can be done near real time dynamically too) and integrate them into a central DW for OLAP processing to mainly understand the user mobility and urban development scenarios for the concerned area (city, region). This architecture has the following main components:

1. Integrated Central DW: It is the DW where all telecom providers' data will be aggregated and loaded in a star schema (discussed in the next section) fact and dimension tables. OLAP processing will be done on these data over the integrated data for a particular area. Global schema for the central DW will be designed and published with a detailed guide documents containing the business meaning with context

of all the fields in the DW schema. This DW star schema can be used as a base cuboid for ROLAP analysis.

2. Telecom Operators' OLTP: It is the existing heterogeneous operational OLTP (Online Transaction Processing) databases of different telecom operators. It contains CDR databases, CRM databases, customer master details etc in relational tables. Though almost all of them contains same or similar data but in various form and format. For example date is represented in different format in different systems, call duration and other measured quantities may have different units like second or minutes etc. same fields may have different names etc.

3. Wrapper Module in OLTP: This wrapper or plug-in module has multiple functionalities. It has a web service interface and a' mapping and transformation' sub module. Mapping module is used to map OLTP database fields to the central DW schema fields. This mapping will be based on some rule files which will be written by business analysts with consulting the developer's manual containing the business meaning and significant of each of the elements in the central DW schema. This module will query the operational databases as per the mapping and then apply transformation and populate the data files to be transmitted. This transformation part comes into play when there is a discrepancy (in form or format or in measuring units etc) between OLTP databases and central DW schema fields.

Once data file is populated in an pre agreed format (may be as csv or dat or insert script etc) then web service (WS) sub module will communicate that data payload to the corresponding WS interface in the OLAP server(Central DW).

4. Optional Anonymizing Module: As different telecom providers will be concerned over privacy of their subscribers, so to overcome we need to anonymize the customer identification data without hampering the temporal and spatial relationship. For the purpose these module will use hashing of the unique customer ID. And same hash function will be used for all the telecom operators.

5. Wrapper Module in OLAP: This module too has two sub modules like the OLTP wrapper module. Web service interface will communicate with other OLTP systems' wrapper module's WS interface and collect the data payload in pre agreed format. Now all collected data will be loaded into a temporary database and finally that data will be aggregated and transformed again so as to insert/update the central DW for OLAP processing. Once data is processed and loaded into the DW then temporary database can be truncated for the next round of operation.

In this architecture first the integrated DW is designed by combining the entire requirement and defining the dimensions, facts and measures. Once that is done an xml schema is published that contains a section for dimension data and a section for fact data. Tags are the name of table and elements are the name of the fields. Also a developer's guide will contain data type format and meaning of each and every field of dimension table and fact table of the central Integrated DW (OLAP).

Now the wrapper module in each of the heterogeneous OLTP system requires manual mapping from that OLTP fields to the elements of the xml schema of integrated DW. These wrapper modules will transform and aggregated data from OLTP systems and generate the XML filled with dimension and measure data and send across to the web service interface of the counter part of the central DW. There will also be a tag in the xml to indicate the data source (name of the Telecom operator) with data fetch time for dynamic update.

During the transform process each of the wrappers will do the following:

- Check for the missing value

- Convert data to the format and form specified in the mapping manual

- Data type uniformity check.

- Convert data values to same units(currency etc)

- Convert data values to same scaling.

Business user needs to do the mapping between the fields of the central DW's dimension and fact tables with the OLTP fields. So that there is none of the below problems remains:

- Mapping of same field represented by different names in OLTP and OLAP by proper understanding of business meaning of the field.

- Mapping of different field by same name (Businesswise two entirely different fields ma have same name) needs to be done to the proper counterpart.

- Some fields of table A (in OLAP) may be present as some field in some other table B in OLTP. Proper mapping needs to be done by the business user with the help of mapping manual of central OLAP.

Once all the fields of the XML schema of the integrated central DW are mapped to the fields of the telecom providers' OLTP fields then data can be queried and populated in the XML as per the schema after proper transformation while required. These xml data are sent to the central DW through web services. Once these data reaches the WS interface at the Central DW it parses the xml data and populates the Temporary Database with data from all the telecom providers' data. Now we can use wrapper module in the central DW server to aggregate and load the data into DW.

Once loading is complete Temporary DB is truncated. Also this architecture is highly dynamic as it can be run at pre-defined intervals may be daily weekly or monthly.

The whole web services communication will take place like below:

- WS client at the OLTP wrapper module will send a request after a set interval (may be once a day, once a week or once a moth etc) to the WS server present at the OLAP wrapper module to ask for permission to send data.

- Central DW WS (web service at the OLAP side) will send a confirmation (ready to accept and free to process) for data with the last access time stamp.

- The telecom providers' OLTP wrapper module will transform and populate the xml data. WS (at OLTP side wrapper module) send the payload (xml files with filled in data) to WS at central DW wrapper module.

- The WS (OLAP side) sends a ACK (acknowledgement) message to corresponding OLTP to indicate that data has been successfully received.

- If due some reason (Network issue or problem while processing data at OLAP side) then WS server at OLAP side will send a "RESEND" request to the OLTP wrapper module.

IV. CONCLUSION AND FUTURE WORK

In this study we have discussed the problems arise during analysis of data collected from a single telecom operator due to different types of social, cultural, economic and technical biases. So to overcome this partiality in telecom data analysis and prediction we proposed an architecture to integrate and

aggregate CDR, CRM and subscriber master data into an integrated central Data warehouse for analysis. We have also taken care of privacy concerns by anonymizing the customer identity Ids. With this integrated DW for a city or region any analysis done will be impartial and un-biased as it will be based on all available data throughout the region. The DW schema depicted here is a generic one and can be remodeled as per the business requirements and analysis purposes. This central integrated DW can be used to feed to OLAP tools for visual analysis, interesting pattern finding and reporting of findings. Further work includes implementing this system for specific applications and solving related challenges encountered.

REFERENCES

[1] Calabrese F., Kloeckl K., Ratti C., (2007). WIKICITY: real-time location sensitive tools for the city.

[2] Calabrese F, Pereira F, DiLorenzo G, Liu L (2010) The geography of taste: analyzing cell-phone mobility and social events. In: International Conference on Pervasive Computing.

[3] W.H. Inmon, Building the Data Warehouse. John Wiley & Sons, 2005.

[4] R. Kimball and M. Ross, The Data Warehouse Toolkit. John Wiley & Sons, 2002.

[5] E.F. Codd, S.B. Codd, and C.T. Salley, Providing OLAP to User-Analysts: An IT Mandate. Codd & Date, Inc., 1993.

[6] F. Simini, M. C. Gonzlez, A. Maritan and A. L. Barabsi (2012), A universal model for mobility and migration patterns. Nature, vol. 484, pp. 96100 (2012).

[7] Calabrese F, Smoreda Z, Blondel VD, Ratti C (2011) Interplay between telecommunications and Face-to-Face Interactions: A Study Using Mobile Phone Data. PLoS ONE 6(7): e20814. doi:10.1371/journal.pone.0020814

[8] Gonzalez M, Hidalgo C, Barabasi AL (2008) Understanding individual human mobility patterns. Nature 453: 779–782.

[9] Redrawing the map of Great Britain from a network of human interactions. PLoS One, vol. 5, e14248 (2010).

[10] Seungjae Shin, Gilju Park, Wonjun Lee, Sunmi Lee Korea Telecom R&D Group, Telecom Economic Research Lab, "Case Study: How to Make Telecom Pricing Strategy Using Data Warehouse Approach", In the Proceeding of 31st Hawaii international conference (Kohala Coast, HI, USA), Vol. 6, pp. 55-60, 1998.

[11] Stefano M. Trisolini, Maurizio Lenzerini, Daniele Nardi, "Data Warehousing and Integration in Telecom Italia", In the Proceedings of the 1999 ACM SIGMOD international conference on Management of data(Philadelphia, Pennsylvania, United States) pp. 538 539, 1999.

[12] Oracle Customer Snapshot:

http://www.oracle.com!customers/snapshots/lgrtelecommunications-db-business-benefits-casestudy.pdf.

[13] Calvanese, D., Dragone, L., Nardi, D., Rosati, R., & Trisolini, S. M. (2006). Enterprise modeling and data warehousing in Telecom Italia. Information Systems, 31(1), 1-32.

[14] Zhu, F., Turner, M., Kotsiopoulos, I., Bennett, K., Russell, M., Budgena, D., & Xu, J. (2004, July). Dynamic data integration using web services. In Web Services, 2004. Proceedings. IEEE International Conference on (pp. 262-269). IEEE.

[15] Ćamilović, D., Bečejski-Vujaklija, D., & Gospić, N. (2009). A call detail records data mart: Data modeling and OLAP analysis. Computer Science and Information Systems, 6(2), 87-110.

[16] Ishaya, T., & Folarin, M. (2012). A service oriented approach to Business Intelligence in Telecoms industry. Telematics and Informatics, 29(3), 273-285.

[17] Hummer, W., Leitner, P., & Dustdar, S. (2011, March). Ws-aggregation: distributed aggregation of web services data. In Proceedings of the 2011 ACM Symposium on Applied Computing (pp. 1590-1597). ACM.

[18] Hossain, M. M., Azim, T., Karim, M. Y., & Hoque, A. L. (2009, December). Integrated data warehousing for telecommunication industries. In Computers and Information Technology, 2009. ICCIT'09. 12th International Conference on(pp. 657-662). IEEE.

[19] C. Lee, C.J. Chen and H. Lu, An aspect of query optimization in multidatabase systems, ACM SIGMOD Record 24(3) (1995) 28–33.

[20] M.P. Reddy, B.E. Prasad, P.G. Reddy and A. Gupta, A methodology for integration of heterogeneous databases, IEEE Transactions on Knowledge and Data Engineering 6(6) (1994) 920–33.

[21] K.H. Lee et al., Conflict classification and resolution in heterogeneous information integration based on XML schema. In: B.Z. Yuan and X.F. Tang (eds), Proceedings of IEEE TENCON'02, 2002 (IEEE Computer Society, Beijing, 2002) 93–6.

[22] R.M. Bruckner, T. Wang Ling, O. Mangisengi and A.M. Tjoa, A framework for a multidimensional OLAP model using topic maps. In: C. Claramunt et al. (eds), Second

[23] International Conference on Web Information Systems Engineering (WISE'01), Vol. 2, Dec. 2001 (IEEE Computer Society, Kyoto, 2001) 109–18.

[24] Maji, G., Sen, S. " A Data Warehouse Based Analysis on CDR to Depict Market Share of Different Mobile Brands." The 12th IEEE India International Conference INDICON. 2015.

[25] Owczarczuk, Marcin. "Churn models for prepaid customers in the cellular telecommunication industry using large data marts." Expert Systems with Applications 37.6 (2010): 4710-4712.

[26] Dasgupta, Koustuv, et al. "Social ties and their relevance to churn in mobile telecom networks." Proceedings of the 11th international conference on Extending database technology: Advances in database technology. ACM, 2008.