RESEARCH ARTICLE OPEN ACCESS

An Approach to Formulate Intelligent Data Repository

A. Prema[1], A.Pethalakshmi[2]
Research Scholar[1], Associate Professor[2],
Department of Computer Science, Research and Development Centre,
Bharathiar University[1], Coimbatore,
MVM Government Arts College[2], Dindigul,
TamilNadu-India

ABSTRACT
Local Data Warehouse (Data Mart) is emerging as one of the newest corners of development in large-scale production and is elected to accomplish tactical decision making for management to produce more turnover. The term Hyper ETL (Extraction. Transformation, and Loading) is a group of three functions to extract the data from heterogeneous data sources, transforming the data into an appropriate format, and loading the data into a data repository. This paper integrates the hyper ETL, decision making methodology and fuzzy optimization technique. We have estimated decision Matrix methodology to boost the sales endorsement in data mart using fuzzy optimization technique. This incorporated approach which improves efficiency of Hyper ETL and the decision making processes for better performance in Data Mart.

***Keywords*: -** DataMart, HyperETL, Decision Matrix, Fuzzy Optimization, Decision making

I. INTRODUCTION

The ETL procedure consist of designing a target, transforming data for the target, scheduling and monitoring processes. The reason of using ETL tools is to save time and make the whole process more consistent. The ETL tools are customized to provide the functionality to meet the enterprise necessity. Hence, many of them choose to construct their own data warehouse themselves[11,12,13].Data mart conquers different troubles that result from the require to connect large numbers of decision support systems to large numbers of operational Data source systems .Many managerial decisions, however, are made with some uncertainty. Managers, for example, authorize substantial finical investments with less than complete information about product demand. As the decision taken by a manager governs the fortunes of business, right decisions will have a salutary effect while the wrong ones may prove to be disastrous, it is extremely important to choose the appropriate decision. Decision theory provides a rational approach to the managers in dealing with problems confronted with partial, imperfect or uncertain future conditions. Under conditions of uncertainty, the decision maker has knowledge about the states of nature that happen but lacks the knowledge about the probabilities the source of their occurrence.

Situations like launching a new product fall under this category. The process with insufficient data, leads to a more complex decision model and, perhaps, a less satisfactory solution. However, one uses scientific methods to exploit the available data to the fullest extents. Under conditions of uncertainty, a few decision criteria which are available could be of help to the decision maker and a choice among them is determined by the company's policy and attitude of the decision maker. In Laplace based method, the weight of each criterion and rating of each alternative are described using linguistic terms. [23]

Section 2 of this paper deals with related work done in Data Mart, Extract, Transformation and Loading. Section 3, explains an actual process of Hyper Extract, Transform and Load, Section 4 explains the proposed work, in section 5, Experimental analysis and implementation results are given, and finally, section 6 presents a conclusion of this paper.

II. RELATED WORKS

Li Jain conquered the week points of traditional Extract, Transform and Load tool's architecture and proposed a three layers architecture base on metadata. That built ETL process more flexible, multipurpose and efficient and finally they designed and implemented a new ETL tool for drilling data ware house. A systematic review method was proposed to identify, extract and analyze the main proposals on modeling conceptual ETL processes for Data Warehouse. The main proposals were identified and compared based on the features, activities and notation of ETL processes and concluded the study by reflecting on the approaches being studied and providing an update skeleton for future study [7].

Sabir Asadullaev talked about centralized Extract, Transform and Load with similar Data warehouse and Data mart, applications of data mart, data warehouse with integration bus and recommended

data warehouse architecture [8]. Over the years, data warehouse technology has been used for analysis and decision making in enterprises [4]. Different varieties of approaches for the integration of ETL tool in data warehouses have been proposed. Shaker H. Ali El-Sappagh tried to navigate through the effort done to conceptualize abbreviations for ETL, DW, DM, OLAP, on- line analytical processing, DS, ODS, and DSA[9]. A data warehouse gives a set of numeric values that are based on set of input values in the form of dimensions [6]. Data Mart can hold information which addresses both strategic and tactical information need and provides information which allows key operating functions to effectively manage performance. It unifies information from various databases into a single database. A data mart contains data from a particular business area and multiple data marts can form a data warehouse. Data marts are the cornerstones of the enterprise, and each unique knowledge data mart is maintained by the divisional or departmental group. The motives for building a data mart are specified below [14].

a) Improves end- user response time
b) Creates collective view by a group of users
c) Provides ease of creation
d) Easy access to frequently need data
e) Lower cost than implementing a full Data warehouse

Daniel Fasel demonstrates the uses of a fuzzy data ware house approach to support the fuzzy analysis of the customer performance measurement. The potential of the fuzzy data warehouse approach is illustrated using a concrete example of customer performance measurement of hearing instrument manufacture. A few for combining fuzzy concepts with the hierarchies of data ware house have been proposed. A method of summary can be guaranteed using this approach and the data ware house concepts can retain flexibility. Using a fuzzy approach in data ware house concepts improves information quality for the company. It provides broader possibilities to create indicators for customer performance measurement as in the example given of a hearing instrument manufacturer. The proposed approach does not include fuzzy linguistic concepts directly in to the hierarchical structure of dimension or into fact tables of the data ware house model and also explains how the fuzzy concepts can be aggregated over dimensions without having to redefined the fuzzy sets in every degree of granularity [17]. Visualization should provide easy understanding of the results for fuzzy queries in the fuzzy data ware house Owen Kisser et al., in "The lito project data ware houses with Literature" describe to apply the business intelligence techniques of the data warehousing and OLAP to the domain of text processing. A literary data ware-house is the conventional corpus but its data stored and organized in multidimensional stages, in order to promote efficient end user queries. This work improves the query engine, ETC process and the user interfaces. The extract, transform, load stage retains the information which are build by the data ware house. The overall idea of applying OLAP to literary data is promising. The initial custom engine is slow for production use but until more optimization is attempted, its promise is unclear [21].

A concrete ETL service framework was proposed and talked about metadata management service, metadata definition services, ETL transformation rules service, process definition service etc [3]. Two heuristic algorithms with greedy characteristics were proposed to reduce the execution cost of an ETL workflow [10]. Lunan Li has recommended to intensively manage ETL by metadata repository and makes metadata easier to understand; therefore metadata management becomes more direct, simple and centered. Numeric values of a classical data warehouse can be difficult to understand for business users, or may be interpreted incorrectly. Therefore, for a more accurate interpretation of numeric values, business users require an interpretation in meaningful non- numeric terms. However, if the transition between terms is crisp, true values cannot be measured and smooth, transition between classes cannot take place [1]. At last, definition method and related algorithms of ETL rules are designed and analyzed. A data mart contains data from a particular business area and multiple data marts can form a data warehouse [5]. ETL is an authoritative meta data based process that extracts the data from source system and loads into the data warehouse and this process improves overall data quality and report ability.

Radhakrishnan and Sreekanth proposed a web based framework model for representing the extraction of data from one or more data sources and use transformation business logic and load the data within the data warehouse. This is good starting point for gathering information in the exiting documentation for the system and also researching for ETL phase in web based scenario modeling in distributed environment which provides an effective decision results for various organization [22]. The models of the entire ETL process using UML because these structural and dynamic properties of an information system at the conceptual level are more natural than the naive approaches. It is more flexible and is used to support trading corporation, banks, financial and human resource management system of an organization at various levels. The future direction of this paper includes analyzing multimedia

information sources automating mechanisms for ETL process.

James F. Brule's "Fuzzy Systems – tutorial" demonstrates the fuzzy system is an alternative to traditional notions of set membership and logic that has its origin in ancient Greek philosophy and its applications are the leading edge of artificial intelligence and it presents the foundation of fuzzy systems with formal mathematics[18]. It is used in many applications such as information retrieval system, a navigation system for automatic cars, a predictive fuzzy logic controller for automatic operation of trains, and laboratory water level controllers for ROBOT are welders, feature definition controllers for ROBOT vision, graphics controller for automated police sketchers and more. Fuzzy systems including fuzzy logic and fuzzy set theory provide a rich and meaningful addition to standard logic. The mathematics generated by theories is consistent; a fuzzy logic may be a generalization of classic logic. Many systems may be modeled and event replicated with the help of fuzzy systems.

Lior Sapir et al., in "A methodology for the design of a fuzzy data warehouse" suggest that a data ware house is a special database used for storing business oriented information for future analysis and decision making. In business scenario, where some of the data or the business attributes are fuzzy, it may be useful to construct a ware house that can support the analysis of fuzzy data and also outlined the Kimball's methodology for the design of a data ware house can be extended to the construction of a fuzzy data ware house. A case study demonstrates the visibility of the most commonly used methodology today is Kimball's. It describes the process of translating business data and prose into a dimensional model. It has several advantages, such as users can make more intuitive and easy to understand queries in a natural language. Defining fuzzy dimensions allows the user to describe the facts with abstract human concepts which are actually more realistic [20]. The fuzzy dimensions also allow more flexible and interesting filtering of the facts. We have demonstrated that fuzzy measures used with fuzzy aggregation operators allow the user to better understand his business and data ware house measures.

D. Ashok Kumar and M. C. Loraine explained modern electronic health records are designed to capture and render vast quantities of clinical data during the health care prone. Utilization of data analysis and data mining methods in medicine and health care is sparse. Medical data is one of the heavily and categorical type data. A Dichotomous variable is type of categorical variable which is binary with categories zero and one. Binary data are the simplest form of data used for medical database in which close ended questions can be used. It is very efficient based on computational efficiency and memory capacity to represent categorical type data. Data mining technique called clustering is involved here for dichotomous medical data due to its high dimensional and data scarcity. Usually the binary data clustering is done by using 0 and 1 as numerical value. The clustering is performed after transforming the binary data into real by wiener transformation. The proposed algorithm in this paper can be usable for large medical and health binary data bases for determining the correction are the health disorders and symptoms observed [16].

Christ Sophie et al., focus that in the field of human resources there is a growing trend towards moving from activity based functions to a more strategic, business oriented role. The data mart defined on the HR information needs is the best solution to meet the objectives [15]. The main purpose of this paper is to explain how the SAS system can be used in top of SAP R/3 HR, and obtain real business benefits on a very short time. It is also based on the practical experience at the Belgian Gas and electricity provider. The structure of this paper first explained the business functions that cover shortcomings of the system. The solution to short comings is explained and business objectives for the date mart are discussed. Finally this paper explains the project approach and focuses on the specific attention points when building a data mart. It provides end to end solution and data management facilities possible to deliver quick results to the end users.

Jeremy, Andeas et al., have built powerful data marts that require minimal administration and are simple to change. This may seem like an impossible goal to anyone who has been involved in the usual complexity but there are a number of simple, practical concepts and methodologies that have been employed and tested over many years of successful data ware house implementation that are repeatable and are easy to understand [19]. For the purposes of data ware housing ETL is used to pull data from business systems into a database that is designed for analysis and reporting. Building data mart and ETL processes involves large volumes of complex business data and the easiest outcome is complexity. Lack of results are expected the easiest outcome of the more resources. It is also used to achieve powerful results in a short amount of time that is useful to users and fulfills the core requirement of effective visibility in to their complex business data.

Fuzzy union and intersection are used to take optimal solution [25].Extract, Transform and Load (ETL) is a process that involves extracting data from product source, transforming it through encoded business rules to fit business needs, and loading it into the data warehouse from where reports are generated. One can customize the ETL jobs to suit your specific business requirements. The three data base functions are combined into one tool that automates the process to pull data out of one database into another database [2]. The Testing of ETL mainly deals with how, from, when, what and where we carry in our data base. All tables except the reference table are transferred to the Data warehouse using an ETL process [8]. Many of the tables are split into smaller tables in order to expedite queries. The ETL process [15] includes designing a target, mapping sources to target, extracting data from sources, transforming data for the target, scheduling and monitoring processes, and managing the Business Intelligence environment. The ETL tools [17] were created to improve and facilitate data warehousing. ETL eliminates the step of loading the text files into intermediate storage, saving significant space and time. The ETL process consists of the following steps: 1. Initiation 2. Build reference data 3. Extract from sources 4. Validate 5. Transform 6. Load into stages tables 7. Audit reports 8. Publish 9. Archive 10. Clean up. Benefits of an ETL [16] Tool are given below:

a) To Simplify the process of migrating data
b) To Store all data transformation logic/rules as Meta data
c) To Reduce cost and effort associated with building interfaces
d) To Standardize the method of data migration
e) To Enable Users, Managers and architects to understand, review, and modify the various interfaces.

III. HYPER ETL

The steps for designing the Hyper ETL (Extract, Transform and Load) are given below [24]

1. Extract the data from operational data source**.** Data extraction is one of the three main functionalities of the ETL tools. A main consideration to assess is the product's ability to extract from a variety of various data sources.

2**.** Create tables with relevant attributes based on user requirements.

3. Transform it to fit operational needs. Generate the XML document file for the collected data.

4. Construct the Meta-Data for XML document File. This Research work implements three protocols namely Oracle Database, XML Data File and JDBC. The Protocol will be part of the url attribute of the target or source node. Every transformation will have a source and target.<source url="**xml**://localhost/etl/test.xml“>

…<target url="**jdbc**:oracle:thin:@localhost:1521:XE"

5. Eliminate the inconsistent data.
6. Split the table.
7. Assign the data.
8. Load it into the end target (ie) Pump the data into Oracle data warehouse. The loading phase is the last step of the ETL process. The information from data sources are loaded and stored in a form of tables. There are two types of tables in the database structure: Fact tables and Dimensions tables. Once the fact and dimension tables are loaded, it is time to improve the performance of the Business Intelligence data by creating Aggregates.
9. Audit reports.
10. Publish is the manual step to recap the position of the previous day’s.
11. Archive
12. Clean up (giving out data to advance its excellence)

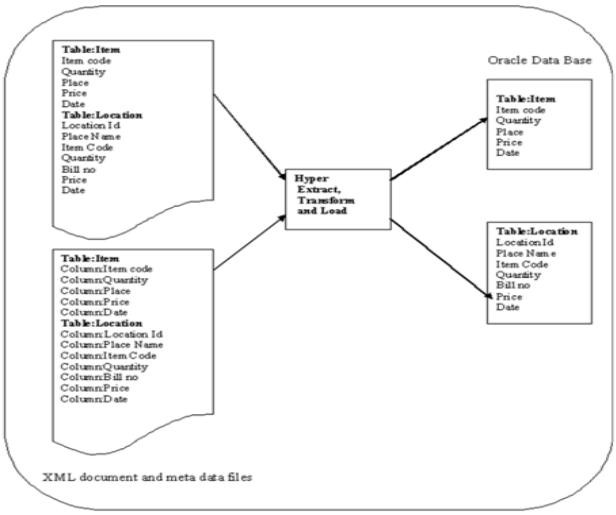
Input Tables: **Table 1.Item table**

Column name	Data Type	Nullable	Primary Key
Item code	Number(4,0)	No	1
Quantity	Number(4,0)	No	-
Place	Varchar(25)	Yes	-
Price	Number(10,2)	Yes	-
Date	Date	Yes	-

Table 2:Location Table

Column name	Data Type	Nullable	Primary Key
Location Id	Number(4,0)	No	1
Place name	Varchar(25)	Yes	-
Item code	Number(4,0)	No	-
Quantity	Number(4,0)	No	-
Bill no	Number(4,0)	No	-
Price	Number(10,2)	Yes	-
Date	Date	Yes	-

Process diagram



Output tables: Table 3: Item Table

Item code	Quantity	Place	Price	Date
M12	662	Madurai	12,30,000	12-nov-13
N14	858	Sivagangai	45,000	23-nov-13
Q11	619	Virdhunagar	1,06,658	30-nov-13
A12	1147	Theni	23,09,678	01-Dec-13
C14	325	Manamadurai	12,45,000	10-Dec-13
D14	734	Ramnad	78,090	15-Dec-13
F15	372	Paramakudi	1,11,345	06-Jan-14
H11	417	Sivakasi	7,56,098	22-Jan-14

Table 4:Location Table:

Location Id	Place name	Item code	Quantity	Bill no	Price	Date
L1	Madurai	M12	662	2345	12,30,000	12-nov-13
L2	Sivagangai	N14	858	3478	45,000	23-nov-13
L3	Virdhunagar	Q11	619	4320	1,06,658	30-nov-13
L4	Theni	A12	1147	4563	23,09,678	01-Dec-13
L5	Manamadurai	C14	325	5100	12,45,000	10-Dec-13
L6	Ramnad	D14	734	5278	78,090	15-Dec-13
L7	Paramakudi	F15	372	5366	1,11,345	06-Jan-14
L8	Sivakasi	H11	417	5490	7,56,098	22-Jan-14

IV. PROPOSED WORK

We have presented the Hyper Extract, Transform and Load tool with decision matrix methodology with fuzzy concept for producing effective

Decision to promote the sales in sales data mart. This paper, shows the quantity of various sales items to different places and applies to the decision matrix .Based on the matrix, mark the rank and depending on the rank, select the best movement of items and these items are applied to Fuzzy union and intersection technique to get an optimal solution. This technique is used to produce the maximum movement of sales on particular places. In Fuzzy Logic, intersection, union and complement are defined in terms of their membership functions. The most commonly used method of fuzzy union is to take the maximum. That is, given two fuzzy sets A and B with membership functions $\mu_A(x)$ and $\mu_B(x)$ The most commonly adopted t-norm is the minimum. That is, given two fuzzy sets A and B with membership functions $\mu_A(x)$ and $\mu_B(x)$

$$\mu_{A\cap B}(x) = \min(\mu_A(x), \mu_B(x))$$

This integrated perception progress better decision making for sales promotion in sales Data mart[26]. We took 2450 real time sales records and selected top eight places and top best eight items after preprocessing. These real time Sales data are taken into the account for evaluating this hyper ETL.

$$\mu_{A\cup B}(x) = \max(\mu_A(x), \mu_B(x))$$

Place/ Item	Madurai	Sivagangai	Virdhunagar	Theni	Manamadurai	Ramnad	Paramakudi	Sivakasi
M12	200	100	100	42	65	70	35	50
N14	300	200	100	100	47	58	23	30
Q11	200	100	100	60	25	57	42	35
A12	250	250	150	150	100	70	82	95
C14	40	80	40	40	20	30	45	30
D14	144	192	144	96	40	48	30	40
F15	160	40	30	25	20	35	30	32
H11	50	150	50	40	35	25	37	30

Prepare the decision matrix for the above items and places in order to find the rank for these items and places.

Table 5:Decision Matrix

DECISION ANALYSIS MATRIX																			
Places		Madurai		Sivagangai		Virdhunagar		Theni		Manamadurai		Ramnad		Paramakudi		Sivakasi		ITEMWISE	
Items	Weight	Rating	Score	Rating	Score	Rating	Score	Rating	Score	Rating	Score	Rating	Score	Rating	Score	Rating	Score	Total	Rank
M12(I1)	10	200	2000	100	1000	100	1000	42	420	65	650	70	700	35	350	50	500	6170	4
N14(I2)	10	300	3000	200	2000	100	1000	100	1000	47	470	58	580	23	230	30	300	8310	2
Q11(I3)	10	200	2000	100	1000	100	1000	60	600	25	250	57	570	42	420	35	350	5875	5
A12(I4)	10	250	2500	250	2500	150	1500	150	1500	100	1000	70	700	82	820	95	950	10615	1
C14(I5)	10	40	400	80	800	40	400	40	400	20	200	30	300	45	450	30	300	2980	8
D14(I6)	10	144	1440	192	1920	144	1440	96	960	40	400	48	480	30	300	40	400	6980	3
F15(I6)	10	160	1600	40	400	30	300	25	250	20	200	35	350	30	300	32	320	3432	7
H11(I7)	10	50	500	150	1500	50	500	40	400	35	350	25	250	37	370	30	300	3900	6
Total			13440		11120		7140		5530		3520		3930		3240		3420		
Rank			1		2		3		4		6		5		8		7		

The following picture shows the place wise and item wise analysis.

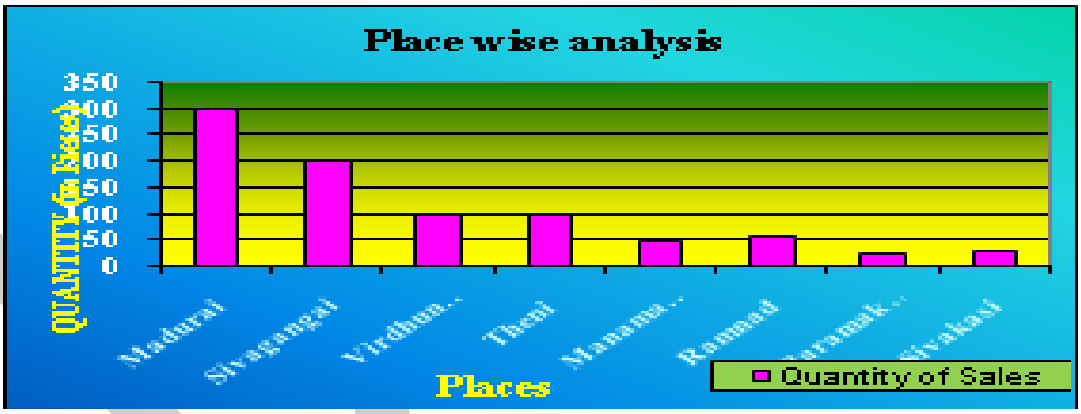


Figure 1:Placcewise Analysis

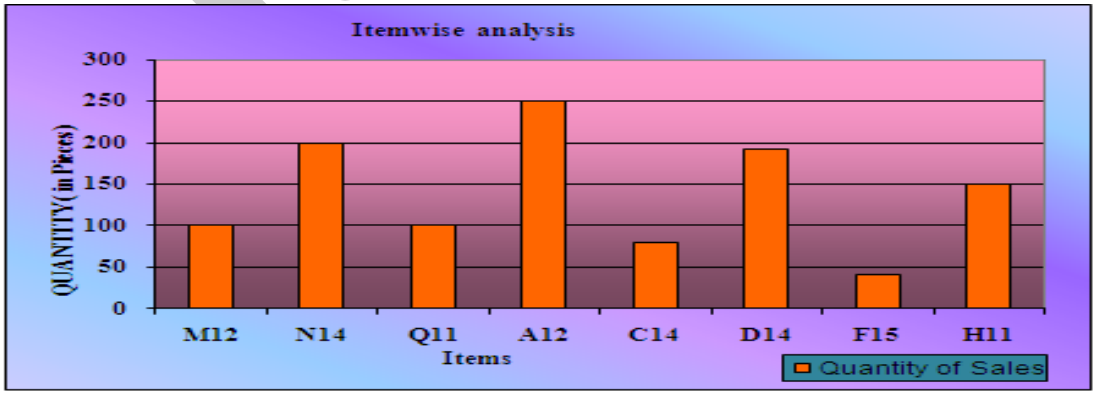


Figure 2:Itemwise Analysis

According to the decision matrix analysis, item4(A12) ,item2(N14)got first and second rank among overall items and in place Madurai and Sivagangai got first and second rank. We found the fuzzy values for these top two items and places to optimize the results.

Table 6:fuzzy values of ItemA12 & N14

Places	Item4(A12)		Item 2 (N14)	
	Quantity	Fuzzy (I4)	Quantity	Fuzzy (I2)
Madurai(P1)	250	1	300	1
Sivagangai(P2)	250	1	200	0.6667
Virdhunagar(P3)	150	0.6	100	0.3333
Theni(P4)	150	0.6	100	0.3333
Manamadurai(P5)	100	0.4	47	0.1567
Ramnad(P6)	70	0.28	58	0.1933
Paramakudi(P7)	82	0.328	23	0.0767
Sivakasi(P8)	95	0.38	30	0.1

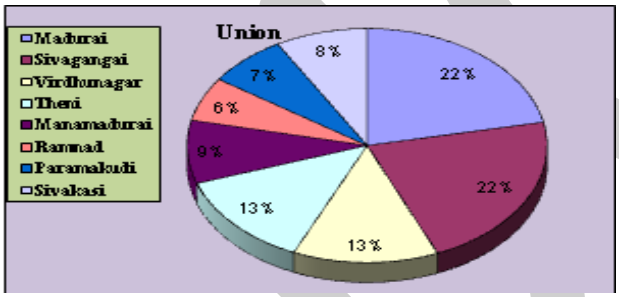


Figure 3: Fuzzy Union for items(A12 & N14)

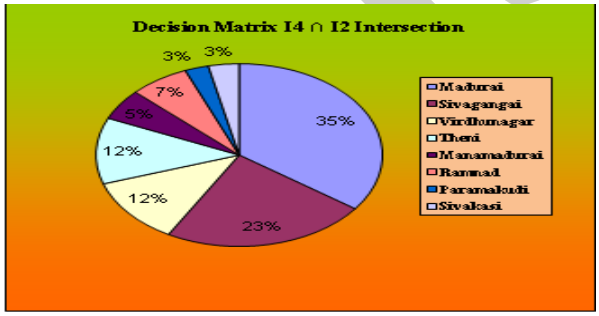


Figure 4:Fuzzy Intersection for items(A12 & N14).

Place wise Analysis:

Table 7: Fuzzy values for Places

Items	Madurai		Sivagangai	
	Quantity	Fuzzy Value	Quantity	Fuzzy Value
M12(I1)	200	0.6667	100	0.4
N14(I2)	300	1	200	0.8
Q11(I3)	200	0.6667	100	0.4
A12(I4)	250	0.8333	250	1
C14(I5)	40	0.1333	80	0.32
D14(I6)	144	0.48	192	0.768
F15(I7)	*160*	*0.5333*	*40*	*0.16*
H11(I8)	50	0.1667	150	0.6

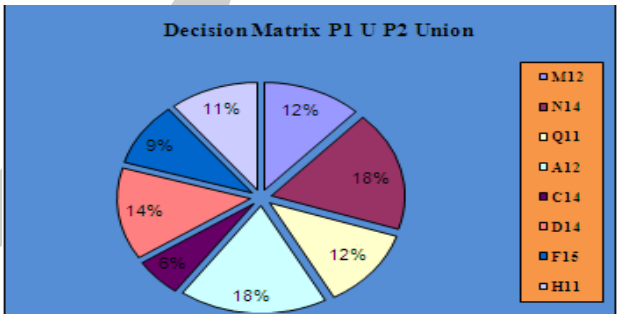


Figure 5: Fuzzy union for Places

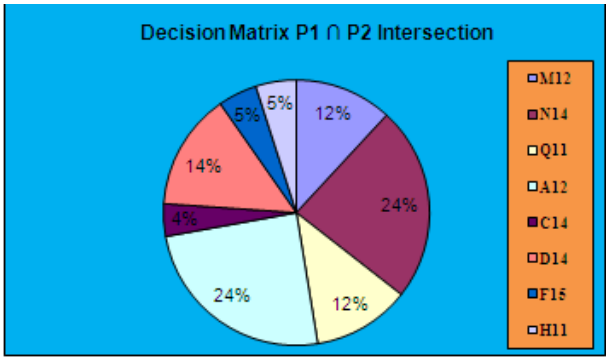


Figure 6: Fuzzy intersection for Places

Table 8:Fuzzy Union and Intersection for best items .

Places	I4 U I2	I4 ∩ I2
Madurai	1	1
Sivagangai	1	0.0033
Virdhunagar	0.6	0.002
Theni	0.6	0.002
Manamadurai	0.4	0.0013
Ramnad	0.28	0.0009
Paramakudi	0.328	0.0011
Sivakasi	0.38	0.0013

Table 9:Fuzzy Union and Intersection for best Places.

Items	P1 U P2	P1 ∩ P2
M12	0.6667	0.4
N14	1	0.8
Q11	0.6667	0.4
A12	1	0.8333
C14	0.32	0.1333
D14	0.768	0.48
F15	0.5333	0.16
H11	0.6	0.1667

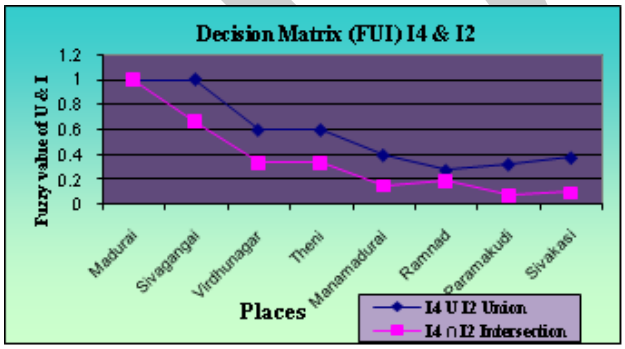


Figure 8: Fuzzy union and intersection for Items.

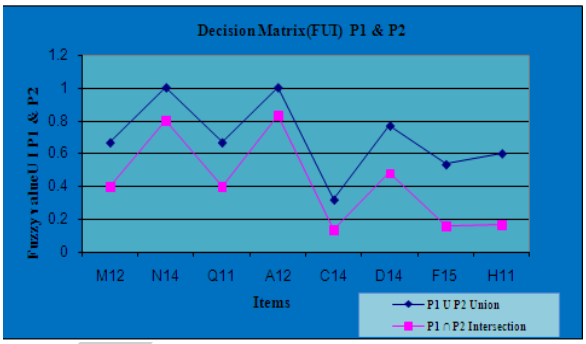


Figure 9: Fuzzy union and intersection for Places.

V. EXPERIMENTAL ANALYSIS AND IMPLEMENTATION RESULTS

We assess Hyper ETL with decision matrix methodology in sales data and then we evaluate fuzzy union and fuzzy intersection techniques to obtain an optimal solution which is used to take right decision. After this process, we tested this result with traditional ETL through some sales records. We used Intel® Xeon® E5-4600 Series Processor, 16GB DDR-III RAM to analyze this proposed Hyper ETL for sales Data mart. Hyper ETL reduces the transformation time than an existing one. This improvement become wider the aggregation method, communicate information sharply and is useful for an effective decision making. We have taken about 12 essential attributes of ETL for evaluating the performance of Hyper ETL and the calculations are given below. Depending on the attributes we have applied the rating.

Table 10:

Rating	6	5	3	1
Meaning	Very high	High	Medium	Low

Table 11:Parametrs of ETL

Parameters	Decision Matrix	
	Existing ETL	Hyper ETL
Scalability	1	6
Manageability	3	3
Disk I/ o utilization	6	5
CPU utilization	5	6
Reliability	1	6

Size (GB)	1	6
Throughput	3	6
Modifiability	3	3
Maintenance cost	6	5
Speed	1	6
Correctness	5	6
Consistency	5	5

The results are shown in the diagram.

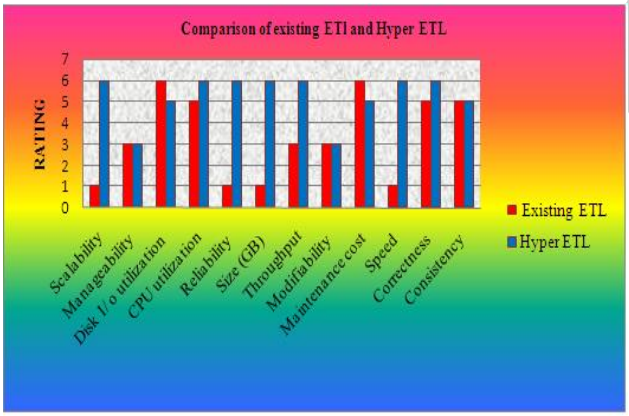


Figure 10: Comparison of existing ETL and Hyper ETL

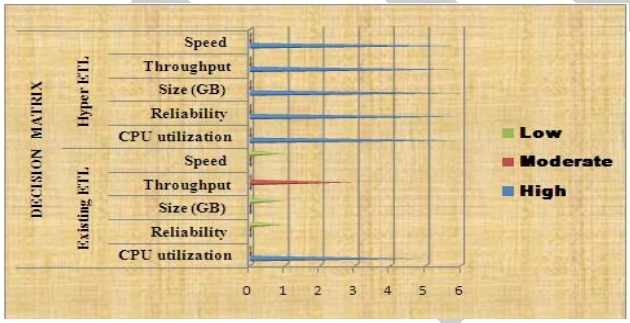


Figure 10: Essential attributes of existing ETL and Hyper ETL

VI. CONCLUSION

We have recommended an innovative model of Hyper ETL and decision matrix analysis which accomplishes an enhanced presentation of ETL Process which save time, space and improves the decision-making progression. This decision matrix method suggested the finest result to boost an efficient decision making to hike the sales promotion using fuzzy optimization technique. The objective of the paper is to find out an effective decision making and to get better performance of ETL process through attaining high Scalability, CPU utilization, Throughput, Reliability, Execution speed than an existing ETL. This Paper suggests the design of Hyper ETL with Decision Matrix method and Fuzzy optimization technique to formulate right decision making to raise the sales promotion with the intention of making efficient Data mart.

REFERENCES

[1] D.Fasel and D.Zumstein.,, A fuzzy data warehouse approach for web analytics. "A Web Science Perspective, volume 5736 of Lecture Notes in Computer Science, page 276-285. Springer, 2009.

[2] Master Data Management An Oracle White Paper September 2011.

[3]. Munoz L., Mazon, J., Trujillo, J_, "Systematic review and comparison of modeling ETL processes in data warehouse", Iberian Conference on information Systems and Technologies, June 2010.

[4]. Sabir Asadullaev, Data Warehouse Architectures – III SWG IBM EE/A 03.11.2009

[5]. Shaker H. Ali EL- Sappagh a, Abdeltawab M.Ahmed Hendawi b, Ali Hamed El Bastawissy b" A proposed model for data warehouse ETL processes"

[6]. Simitsis, A Vassiliadis, P.;Sellis, T._, "State-space optimization of ETL workflows", IEEE Transactions on Knowledge and Data Engineering, Vol 17, Issue 10, Oct 2005.

[7]. Inmon, William (2000-07-18). "Data Mart Does Not Equal Data Warehouse". DMReview.com.

[8]. Hariprasad T,"ETL testing Fundamentals" on March 29, 2012

[9]. Huamin Wang, "An ETL Services Framework Based on Metadata", 2nd International Workshop on Intelligent Systems and Applications, May 2010

[10]. W. H. Inmon." Building the Data Warehouse". Wiley Publishing, Inc., 4 edition, 2005.

[11]. Inmon, William (2000-07-18). "Data Mart Does Not Equal Data Warehouse". DMReview.com.

[12]. Jeffrey R. Bocarsly, Complex ETL Testing-A Strategic Approach

[13]. R. Kimball and M. Ross. "The Data Warehouse Toolkit." WileyPublishing, Inc., 2002.

[14]. Li Jian, Xu Bihua_, "ETL tool research and implementation based on drilling data warehouse", Seventh International Conference on Fuzzy Systems and Knowledge Discovery, Aug 2010.

[15]. Lunan Li, "A framework study of ETL processes optimization based on metadata repository" International Conference on Computer Engineering and Technology, April 2010.

[16]. D.Ashok Kumar and M.C.Loraine Charlet Annie," Decision Making on Dichotomous Medical Data using Novel Clustering approach", National conference on Signal and Image processing(NCSIP) 2012.8

[17]. Daniel Fasel, "A fuzzy datawarehouse approach for the customer performance measurement for a hearing instrument manufacturing company", Sixth International conference on fuzzy systems and knowledge discovery,2009.

[18]. J.F.Baldwin,"Fuzzy systems – logic and reasoning in fuzzy applications". London: Academic press,1981.

[19]. Jeremy, Jean King and Andreas schindler,"Understanding of Business Intelligence: ETL and Data mart Practises".

[20]. Lior Sapir and Armin Shmilovice.,"A methodology for the design of a fuzzy data warehouse"2008.

[21]. Owen kaser, Steven Keith and Daniel Lomire,"Dataware housing with literature", September 11,2006.

[22]. Radha Krishna and Sree kanth,"An Object Oriented modeling and Implementation of Web based ETL process" in IJCSNS, International Journal of Computer Science and Network Security, vol 10 no.2, February 2010.

[23]. A.Prema and A.PethaLakshmi"An Improved Enhancement of Decision – making Analysis to increase the sales promotion using Hyper ETL in Data Mart" in IJCII 2012 dec.

[24]. A.Prema and A.Petha lakshmi,"Novel approach in ETL"IEEE Xplore"Mar 2013

[25.] Kankana Chakrabarthy "Fuzzy union and intersection of fuzzy set"

[26] A.Prema.,et al,"Formulate competent decision making in datamart using fuzzy optimization technique"in (IJCSIT) International Journal of Computer Science and Information Technology vol 5(3),2014,3196-3202