Research on Consistencies of View Model in Information Resource Planning

Ming Liu[1,2,*], Shaohua Dong[1]

[1]University of Science and Technology Beijing, Beijing, China

[2]School of Mathematics and Computer Science, Tongling University, Tongling, China

* liuming-infor@163.com

Keywords: E-R diagram, Matrix mapping method, mapping relationship, view model.

Abstract: In information resource planning, existing mapping methods based on E-R diagram are difficult to go deep into the correspondence between data element of user view and specific property of logical entity, unable to form a clear expression of the whole process of information integration of the enterprise, and brings inconvenience for analysis and statistics of the subsequent data model and more in-depth analysis and calculation. Matrix mapping method is proposed to realize the consistency of view model. This method is a universal method, which improves and perfects the E-R method, can express the mapping relationship between the data element of user view and the specific entity property, ensure the consistencies of view model in information system integration.

1. Introduction

Many enterprises lack of a rational and correct perception on the importance of information resource planning (IRP) [1] in the early phase of informationization, first developed a number of information systems, such as OA, CRM. Due to the initial development of these systems are more simple, data exchange between different information systems is accomplished by data coding map, but as development and deployment of more complicated information systems, if still use this way to maintain these system of data consistency, the number and complexity of interfaces among these system will increase geometrically along with the number of new information systems increases [2]. It is impossible to adopt interfaces to implement an integrated system of multiple scattered, isolated systems.

Aiming at these difficulties, there are several solutions: Develop a new project, completely according to the function model, data model and data standards that are planned uniformly, organize system development on its own, joint development or entrusted development. Procure the matching application information systems, choose and procure application software that are consistent with the established function model, data model, and existing data standards. Transform existing information systems create a new unified data model and data standards to transform the old data structure, and modify the application accordingly. In comparison with those methods, adopting the transformation of existing information systems can be in accordance with the new unified data model to transform old data structure and data standards, and modify the application accordingly. This is not through the interface for data integration, which has advantages: less time consuming and lower cost but better effect [3]. There are still some problems to be solved, how to integrate data resources and information resources on the basis of the existing data of old information system, and effective integration with new systems. There are some problem Sneed to solve: Data correctness, ensure data elements of user view with the same semantics can be mapped to the same attribute of the same entity. Data integrity, guarantee that data elements of each user view can be found, and only one entity attribute is found that match for one element. Data non-redundancy, each non-key attribute of an entity can be found at least one matched data element of the user view, but can find only one data element of user view as a data entry [4].

2. Matrix Mapping Method

Matrix mapping method is proposed to realize the consistency of view model. This method is a universal method, which improves and perfects the E-R method. It does not depend on a field, so it has a strong generality, especially suitable for data environment with large amount of data and multiple user views. It also expresses the mapping relationship between the data elements of user view and the specific entity properties.

Definition 1: Entity

Assume that e denotes an entity, B_e denotes a set of properties of that entity, $B_\varepsilon=\{b_1,b_2,\cdots\cdots b_n\}$ in which b_i denotes the ith property element of entity, n is the number of properties. So an entity is represented as two-tuples (e,B_ε). The subject database is an entity set composed of multiple entities.

Definition 2: User view

Assume that v denotes a user view, A_v denotes a set of source data elements of user view, $A_v=\{a_1,a_2,\cdots\cdots a_m\}$ in which a_i denotes the ith data element of user view, m is the number of data elements. So, a user view is represented as two-tuples (v,A_v). The business domian is a user view set composed of multiple user views.

Definition 3: Entity incidence matrix

Given an entity set E and property set C of an information system S, then define a relation r of E as $r=\{O,I,M,C\}$ in which C denotes corresponding property set of the relation r; $o\subseteq E\times C, I\neq\varnothing$,O denotes the property set of associated entity; $I\subseteq E\times C, I\neq\varnothing$, I denotes the property set of entity which exert the relation to the associated entity; M denotes incidence matrix, $M=[m_{ij}]$.Assumed that $x_i\in I$, $i\in\{1,2,\cdots p\}$,and $y_j\in O$, $j\in\{1,2,\cdots q\}$,in which $p=|I|$, $q=|O|$, so if there is a correlation relationship between x_i and y_j , $m_{ij}=1$;Otherwise $m_{ij}=0$.M is a $p\times q$ Boolean matrix.

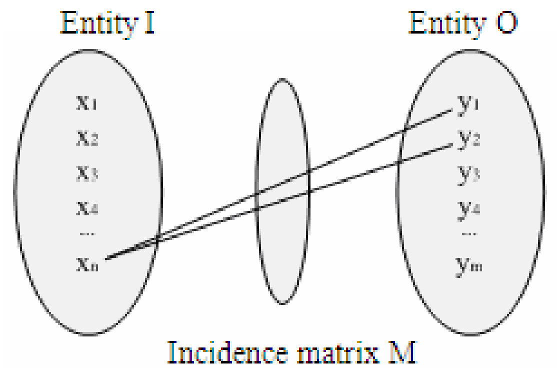

Figure 1. Entity incidence matrix M

Definition 4: Reachability relation

The Reachability relation means: given a relation $r=\{O,I,M,C\}$, if a corresponding input entity $x\in I$ can be associated with $y\in O\cup I$ through director indirect entity, then consider: from x to y is reachable and call there is a reach-relation of entities from x to y.

Lemma 1:

Assume that e denotes an entity; $\bigcup_{i=1}^{n} e_i$ is a set of entities e_i , which indicates the subject database that corresponds to business domain Y. Any user view of business domain Y is expressed as v. A_v

denotes the set of source data elements, $A_v = \{a_1, a_2, ..., a_m\}$,in which $a_i (\forall i \in \{1, 2, ...m\})$ denotes the ith data element of user view, m denotes the number of corresponding data elements.

So, for each data element in a_i can find the corresponding property in e_i .

Proof: Obtained from the second layer relationship in the process of data modeling, the business domain corresponds to the subject database, so each user view of the business domain can only respond to the logical entity in the corresponding subject database. By the completeness of correspondence between data element of user view and logical entity, we can prove the lemma 1.

Theorem 1:

For any one user view v of the business domain Y, there exists but only one entity e, which allows all attributes of that user view v can correspond to the attributes of other entities by using its attributes or its reachability relation.

Proof: By the lemma 1, for each data element of user view, there exists but only one attribute of one entity that corresponds to it. We create a new entity e in subject database, and then classify the corresponding entity properties of each data element in user view v. If one data element of the user view v can correspond to one property of an existing entity $e^{'}$ in subject database, then add primary key property of $e^{'}$ to e as its foreign key; otherwise add that data element to the entity e as its one of normal properties. Follow the steps above to perform the user view v, e is the desired result by the Definition 4, so the Theorem 1 is proved.

Definition 5: User view and entity incidence matrix

This relation is defined as $relation = \{A_v, B_e, N\}$, in which A_v denotes the data element set of user view v, B_e denotes the property set of entity e, N denotes the incidence matrix $N = [n_{kl}]$. Assumed that $a_k \in A_v$, $k \in \{1, 2, \cdots, \mathrm{s}\}$, and $b_l \in B_\varepsilon$, $l \in \{1, 2, \cdots, \mathrm{d}\}$, in which $s = |A_v|$, $d = |B_\varepsilon|$, so if there is a correlation relationship between a_k and b_l, $n_{kl} = 1$;Otherwise $n_{kl} = 0$ .N is a $s \times d$ Boolean matrix.

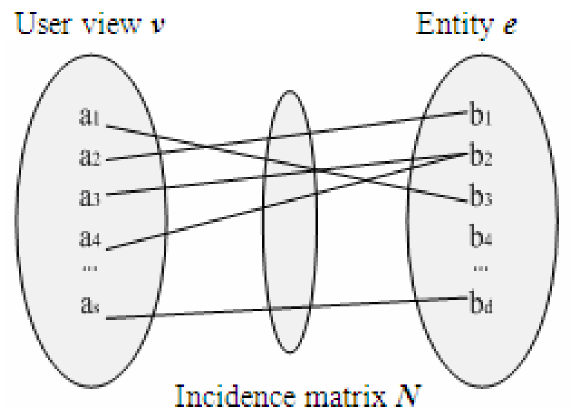

Figure 2. User view and entity incidence matrix N

Theorem 2:

Matrix R=N×M indicates the correspondence between data elements of user view and entity properties of entity set.

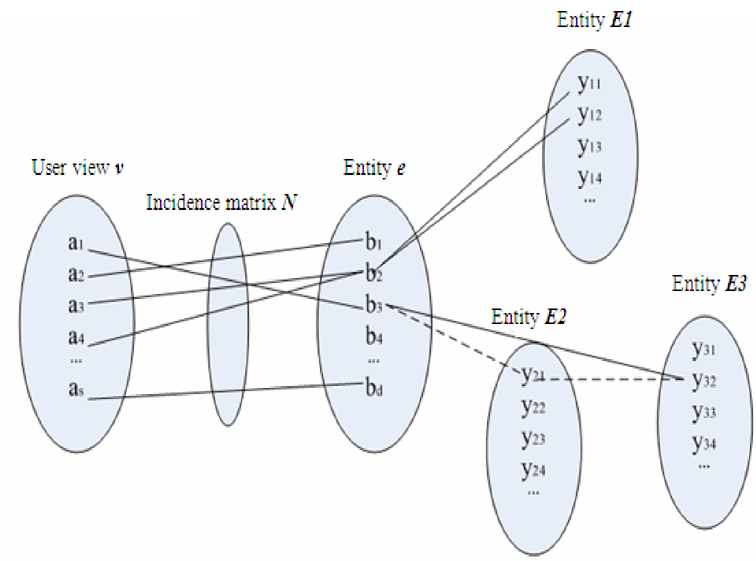

Figure 3. Calculation results of matrix R=N×M

3. Application Example

There is a user view "Type evaluation report of measuring equipment" in business domain "Type evaluation/ Prototype testing" of one enterprise. The data element of that user view comprises the following components: Apllication ID, Evaluation report ID, Evaluation unit name, Manufacturing unit name, Manufacturing unit contacts, Entrusting unit name, Date of entrustment, Environmental conditions of type evaluation, Technical superintendent name, technical superintendent title.

The corresponding entity e "Type evaluation registration of measuring equipment" of that user view has the following properties: Apllication ID, Evaluation report ID, Evaluation unit ID, Manufacturing unit ID, Entrusting unit ID, Date of entrustment, Environmental conditions of type evaluation, Technical superintendent ID, in which the "Evaluation report ID" is the primary key, the "Apllication ID", "Evaluation unit ID", "Manufacturing unit ID", "Entrusting unit ID", "Technical superintendent ID" are the foreign keys. Take (Apllication ID, Evaluation report ID, Evaluation unit ID, Manufacturing unit ID, Entrusting unit ID, Date of entrustment, Environmental conditions of type evaluation, Technical superintendent ID) as column, (Apllication ID, Evaluation report ID, Evaluation unit name, Manufacturing unit name, Manufacturing unit contacts, Entrusting unit name, Date of entrustment, Environmental conditions of type evaluation, Technical superintendent name, technical superintendent title) as row. The corresponding matrix N is organized as follows:

$$
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
$$

Figure 4. The matrix N

In addition to entity e "Type evaluation registration of measuring equipment", the existing entities "Company Profile", "Staff basic information" and "Type approval application of measuring equipment" are also be used. The entity "Company Profile" has the following properties: Unit ID, Unit name in Chinese, Unit name in English, Work telephone, Zip code, Address, Contacts name, Contacts address, Contacts telephone, Contacts position, Total assets, Remark. Take (Unit ID, Unit name in Chinese, Unit name in English, Work telephone, Zip code, Address, Contacts name, Contacts address, Contacts telephone, Contacts position, Total assets, Remark) as column, properties of entity e as row. The entity incidence matrix M from entity e to entity "Company Profile" is organized as follows:

$$\begin{pmatrix} 0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&1&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0 \end{pmatrix}$$

Figure 5. The entity incidence matrix M

By that analogy, we can also obtain the entity incidence matrix M from entity e to entity "Staff basic information", from entity e to entity "Type approval application of measuring equipment".

The correspondence between user view and each logical entity can be obtained by multiplying the matrix N with matrix M. The user view and entity incidence matrix R between user view" Type evaluation report of measuring equipment" and entity "Company Profile" is organized as follows:

$$\begin{aligned} R &= N * M \\ &= \begin{pmatrix} 1&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&0&1 \end{pmatrix} * \begin{pmatrix} 0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&1&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0 \end{pmatrix} \\ &= \begin{pmatrix} 0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&1&0&0&0&0&0 \\ 0&1&0&0&0&0&1&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0 \end{pmatrix} \end{aligned}$$

Figure 6. The user view and entity incidence matrix R

Through the R Matrix, it can be seen that the relationship between the user view of "Type evaluation report of measuring equipment" and the logical entity of "Staff basic information".

From each line, it can be seen that the name of the evaluation unit, the name of the manufacturing unit, the contact person of the manufacturing unit, and the name of the entrusting unit in the user view of the "Type evaluation report of measuring equipment" are all derived from the "Staff basic information" logical entity. Further analysis can show that, Since the 4th and 5th lines of the R matrix are exactly the same, the meaning expressed by the matrix shows that the two data elements of the manufacturing unit name and the manufacturing unit contact in the user view are mapped to the same "Staff basic information" logical entity, that is, the manufacturing unit entity. From each column, it can be seen that the unit Chinese name and unit contact name in the "Staff basic information" logical entity are used by the "Type evaluation report of measuring equipment" user view, in which the unit Chinese name is used 3 times. The name of the unit contact was used once.

Through the analysis and computing of other user views in business domain, not only can clear and definite the mapping relationship between data elements in user view and properties of logical entity, but also can calculate the usage frequency of each entity, even each property of entity. It is very important for the Information Resource Planning and the development and implementation of subsequent information systems.

4. Conclusion

The proposed method adopts matrix to solve practical problems, has great applicable value in enterprise modeling, for the research and analysis of matrix is relatively mature in Mathematic. Using matrix in modeling analysis, not only can realize clear expression of information in the information integration, but also benefit the quantification, calculation and analysis of information elements. It can be used as a basis of generalized application platform, application tools, database and data warehouse system, promotes the integrative development of systems and applications. Practice shows that by this method can not only improve the management and efficiency of information resources, but also have certain reference value for the related research of information resource planning and data warehouse technology.

Acknowledgments

This work was supported by Natural Science Foundation of Anhui Higher Education Institutions of China (Grant No. KJ2017A467, Grant No. KJ2018A474, Grant No. KJ2017A470,). The author would like to acknowledge the financial support of the National Natural Science Foundation of China (Grant No.61473035).

References

[1] Kleep A, Warmer J, Bast W.MDA explained: the practice and promise of the model driven architecture. Boston: Addison Wesley,2003.

[2] Rolstadas A, Andersen B. Enterprise Modeling Improving Global Industrial Competitiveness. Kluwer Academic Publishers,2000.

[3] Holzmann G.J. and Joshi R. Model-driven software verification. Lecture notes in computer science, 2989:76-91,2004.

[4] Skene J. and Emmerich W. Model driven performance analysis of enterprise information systems. Electronic notes in theoretical computer science, 82(6),2003.