Collaborating Semantic Link Network with Resource Space Model*

Yunpeng Xing[1,2,3], Jie Liu[1,2], Xiaoping Sun[1,2,3], and Erlin Yao[1,2,3]

[1] China Knowledge Grid Research Group, Key Lab of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, China
http://kg.ict.ac.cn
[2] Graduate School of Chinese Academy of Sciences, Beijing, China
{ypxing, lj}@kg.ict.ac.cn
[3] Hunan Knowledge Grid Lab, Hunan University of Science and Technology, Hunan, China

Abstract. The Semantic Link Network model SLN and Resource Space Model RSM are semantic models proposed separately for effectively specifying and managing versatile resources across the Internet. Collaborating the relational semantics of SLN with the classification semantics of the RSM can support richer semantic modeling in applications. This paper introduces the Resource Class Hierarchies into the SLN and investigates to cooperatively model with the RSM. The proposed approach can be used as a normalized semantic overlay of the Knowledge Grid environment.

1 Introduction

The relational data model is the milestone of data management [1, 2]. Different from the hierarchical and network models, it separates the logical data representation from physical implementation. Object-oriented databases and object-relational databases extend the application scope of the relational databases [5, 8, 10]. Several graph-based semantic models built on the traditional hierarchical and network data models have been proposed [4, 6, 7, 9, 11]. But, these data models are incapable of effectively managing heterogeneous, distributed and ocean resources in an open and dynamic Internet environment.

The Resource Space Model RSM is a semantic data model for uniformly, normally and effectively specifying and managing resources. Its theory basis is the normal forms based on orthogonal classification semantics and relevant resource operations [13, 14, 15, 18]. The Semantic Link Network model SLN is a semantic model for semantically interconnecting resources [15]. The SLN has the advantages of enabling rich semantic representation, reasoning, execution, referential search, and normalization.

To facilitate the cooperation between RSM and SLN, this paper extends the SLN by introducing a structure called Resource Class Hierarchy, which can be derived

* This work was supported by the National Basic Research Program of China (973 project No.2003CB317000) and the National Science Foundation of China (Grants 60273020 and 70271007).

H. Zhuge and G.C. Fox (Eds.): GCC 2005, LNCS 3795, pp. 750 – 755, 2005.

from the RSM. The combination of the classification semantics of RSM and the relational semantics of SLN can support richer semantic modeling and applications in the open and dynamic Internet environment [16, 19].

2 Resource Space Model and Semantic Link Network Model

A resource space is an *n*-dimensional space, denoted as $\boldsymbol{RS}(\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n)$ or just by name ***RS*** in simple [13]. Axis is defined by a set of coordinates denoted as $\boldsymbol{X}_i = \{\boldsymbol{C}_{i1}, \boldsymbol{C}_{i2}, \ldots, \boldsymbol{C}_{im}\}$. A point $\boldsymbol{p}(\boldsymbol{C}_{1,j1}, \boldsymbol{C}_{2,j2}, \ldots, \boldsymbol{C}_{n,jn})$ is determined by the coordinate values at all axes. Every point in the space can uniquely determine a resource set, where each element is called a resource entry. Point and resource entry are two fundamental operation units of RSM.

A Semantic Link Network (SLN) is a network consisting of nodes and semantic links between nodes. A node can be either a resource defined in the name space or an SLN. The Semantic Link Network can be formally represented as a directed graph *SLN*=<*Resources*, *SemanticLinks*>, where *Resources* and *SemanticLinks* represent the sets of nodes and semantic links respectively. The operations, normal form theory, characteristics, criteria and integrity constraints have been proposed in [15]. A set of semantic relationships such as *Cause-effect*, *Implication*, *Inclusion*, *Similar-to*, *Instance*, *Membership*, *Sequential* and *Reference* has been introduced in [12]. More types of semantic links can be imported according to application requirements. The SLN theory has been applied to the peer-to-peer environment to achieve efficient query routing [17].

3 The Extended Semantic Link Network

Fig.1 illustrates a part of an extended semantic link network representing a simple teaching system. The dotted circles represent *resource class hierarchies*, each of which corresponds to a resource space. The rectangles in dotted circles denote the *resource classes* corresponding to classes of resources in a resource space. The edges in the dotted circles represent the *inclusion relationships* between resource classes. The formal definitions of resource class hierarchy, resource class and inclusion relationship will be introduced in section 4. And other nodes and edges are from the generic concepts defined in the original SLN.

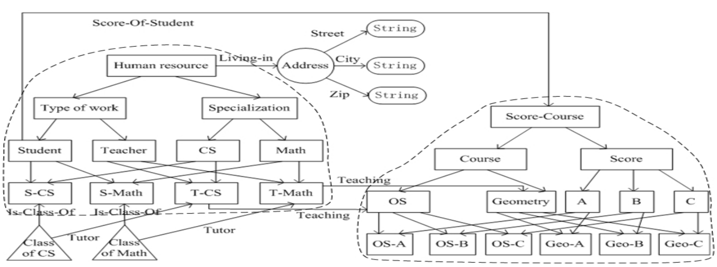

Fig. 1. A Semantic Link Network with resource class hierarchies

Definition 1. The extended Semantic Link Network is a triple $S = (VE, RE, RCH)$, where VE is a finite set of nodes in the extended Semantic Link Network which could include resources, generic classes and resource classes; RE is a finite set of triple $<v_1, v_2, re>$, where re represents the relationship between nodes v_1 and v_2 coming from VE; and, RCH is a finite set of resource class hierarchies, each of which corresponds to a resource space in the RSM.

Through incorporating the RSM and the SLN, the extended SLN has the following three advantages.

(1) The extended SLN provides an efficient mapping mechanism from RSM to SLN. The extended SLN has increased the mapping granularity from RSM to SLN. In contrast with the original SLN, a point, an axis and even a resource space in the RSM can be mapped to a node in the extended SLN. Secondly, the extended SLN provides not only resource mappings but also operation mappings from the RSM to the SLN.

(2) The extended SLN makes use of the RSM to facilitate the SLN modeling. Since quick and easy modeling is one of the salient features of the RSM, the *resource class hierarchies* in the extended SLN make the SLN modeling easier without conflict with other approaches such as Entity-Relationship model [3].

(3) The extended SLN can enhance the interoperability between SLNs. Most of the operations of the original SLN are mainly based on the structure information. The extended SLN has introduced some new RSM-based operations, which emphasize on not only structure but also semantic information of SLNs.

4 Resource Class Hierarchy

Let δ be a resource space, an axis, a point or a coordinate of the RSM, and $\boldsymbol{R}(\delta)$ be the resources that δ can contain in the RSM. Let $\boldsymbol{RS}(\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n)$ be a resource space. Its axis $\boldsymbol{X}_i = \{\boldsymbol{C}_{i1}, \boldsymbol{C}_{i2}, \ldots, \boldsymbol{C}_{im}\}$ can be extended to $\boldsymbol{X}_i^* = \{\boldsymbol{C}_{i1}, \boldsymbol{C}_{i2}, \ldots, \boldsymbol{C}_{im}, \pi_i\}$, $1 \le i \le n$. A point taking the coordinate value π_i on the axis $\boldsymbol{X}_i$ means that the axis $\boldsymbol{X}_i$ has no restriction on the point. Meanwhile, we introduce a constant γ_i at each axis $\boldsymbol{X}_i$ such that $\boldsymbol{R}(\gamma_i) = \boldsymbol{R}(\boldsymbol{C}_{i1}) \cup \boldsymbol{R}(\boldsymbol{C}_{i2}) \cup \ldots \cup \boldsymbol{R}(\boldsymbol{C}_{im})$.

Definition 2. Let $\boldsymbol{RS}(\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n)$ be a resource space and $\boldsymbol{X}_i^*$, $1 \le i \le n$, be its extended axes. The set of resources represented by $(\pi_1, \ldots, \pi_{i-1}, \gamma_i, \pi_{i+1}, \ldots, \pi_n)$ is defined as the *axis resource class* of axis $\boldsymbol{X}_i$, denoted as ac_i. Axis resource class and each relation in the Cartesian product $\boldsymbol{X}_1^* \times \boldsymbol{X}_2^* \times \ldots \times \boldsymbol{X}_n^*$ are generally called *resource classes*. Particularly, the relation $(\pi_1, \pi_2, \ldots, \pi_n)$ is called the *base resource class*, denoted as $\text{root}_{\boldsymbol{RS}}$.

For any two resource classes $c = (\boldsymbol{C}_{1,j1}, \boldsymbol{C}_{2,j2}, \ldots, \boldsymbol{C}_{n,jn})$ and $c' = (\boldsymbol{C}_{1,j1}', \boldsymbol{C}_{2,j2}', \ldots, \boldsymbol{C}_{n,jn}')$ in a resource space, if $\boldsymbol{R}(\boldsymbol{C}_{1,j1}) \subseteq \boldsymbol{R}(\boldsymbol{C}_{1,j1}')$, $\boldsymbol{R}(\boldsymbol{C}_{2,j2}) \subseteq \boldsymbol{R}(\boldsymbol{C}_{2,j2}')$, …, $\boldsymbol{R}(\boldsymbol{C}_{n,jn}) \subseteq \boldsymbol{R}(\boldsymbol{C}_{n,jn}')$ hold respectively, then the resource class c is called the subclass of the resource class c' and the resource class c' is called the superclass of the resource class c. And this inclusion relationship is denoted as $c \subseteq_c c'$.

Definition 3. Let $\boldsymbol{RS}(\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n)$ be a resource space. The *resource class hierarchy* of $\boldsymbol{RS}$ is defined as the directed graph $\boldsymbol{RSG}(CS, E)$ where

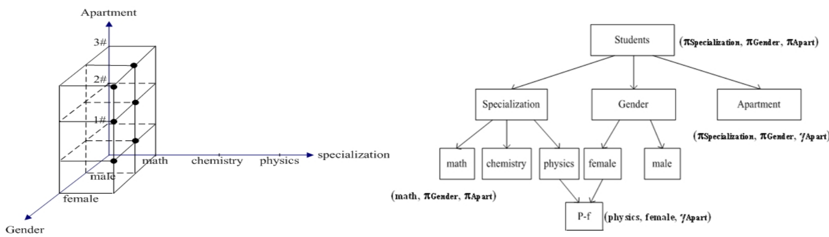

Fig. 2. A resource space and a part of its resource class hierarchy

(1) $CS = \boldsymbol{X}_1^* \times \boldsymbol{X}_2^* \times \ldots \times \boldsymbol{X}_n^* \cup \{(\pi_1, \ldots, \pi_{i-1}, \gamma_i, \pi_{i+1}, \ldots, \pi_n) \mid 1\le i\le n\}$;

(2) For any resource classes $c_1, c_2 \in CS$, if $<c_1, c_2>\in E$, then $c_2 \subseteq_c c_1$ holds;

(3) For any resource classes $c_1, c_2 \in CS$, if $c_2 \subseteq_c c_1$ holds, then there exists at least one path $\Gamma(c_1, c_{j1}, c_{j2}, \ldots, c_{jm}, c_2)$.

The resource class hierarchy of a given resource space can be viewed as a directed graph consisting of resource classes and the inclusion relationships. $\Gamma(v_1, v_2, \ldots, v_n)$ is used to denote one path from v_1 to v_n in a directed graph. Fig. 2 is the illustration of a resource space and a part of its resource class hierarchy.

5 Operations in the Extended Semantic Link Network

The extended SLN not only maps resource spaces to resource class hierarchies but also inherits all operations of the RSM. There exist four types of operations in the RSM: Join, Disjoin, Merge and Split [13]. The following is the formal definitions of these four operations on resource class hierarchies of the extended SLN.

Theorem 1. (**Rsm-Join**) Let $\boldsymbol{RS}_1(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_t, \boldsymbol{X}_{t+1}, \ldots, \boldsymbol{X}_m)$ and $\boldsymbol{RS}_2(\boldsymbol{X}_{t+1}, \ldots, \boldsymbol{X}_m, \boldsymbol{X}_{m+1}, \ldots, \boldsymbol{X}_n)$ be two resource spaces that can be joined together to form the resource space $\boldsymbol{RS}(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_t, \boldsymbol{X}_{t+1}, \ldots, \boldsymbol{X}_m, \boldsymbol{X}_{m+1}, \ldots, \boldsymbol{X}_n)$. Assume that $\boldsymbol{RSG}_1(CS_1, E_1)$ and $\boldsymbol{RSG}_2(CS_2, E_2)$ are the resource class hierarchies of $\boldsymbol{RS}_1$ and $\boldsymbol{RS}_2$ respectively. The Rsm-Join operation on $\boldsymbol{RSG}_1(CS_1, E_1)$ and $\boldsymbol{RSG}_2(CS_2, E_2)$ is defined as the directed graph $\boldsymbol{RSG}(CS, E)$ such that:

(1) $CS = \boldsymbol{X}_1^* \times \boldsymbol{X}_2^* \times \ldots \times \boldsymbol{X}_n^* \cup \{(\pi_1, \ldots, \pi_{i-1}, \gamma_i, \pi_{i+1}, \ldots, \pi_n) \mid 1\le i\le n\}$;
(2) For any two classes $c = (\boldsymbol{C}_1, \ldots, \boldsymbol{C}_t, \boldsymbol{C}_{t+1}, \ldots, \boldsymbol{C}_m, \boldsymbol{C}_{m+1}, \ldots, \boldsymbol{C}_n)$ and $c' = (\boldsymbol{C}_1', \ldots, \boldsymbol{C}_t', \boldsymbol{C}_{t+1}', \ldots, \boldsymbol{C}_m', \boldsymbol{C}_{m+1}', \ldots, \boldsymbol{C}_n')$ in CS, $<c, c'> \in E$ holds if and only if both $<(\boldsymbol{C}_1, \ldots, \boldsymbol{C}_t, \boldsymbol{C}_{t+1}, \ldots, \boldsymbol{C}_m), (\boldsymbol{C}_1', \ldots, \boldsymbol{C}_t', \boldsymbol{C}_{t+1}', \ldots, \boldsymbol{C}_m')> \in E_1$ and $<(\boldsymbol{C}_{t+1}, \ldots, \boldsymbol{C}_m, \boldsymbol{C}_{m+1}, \ldots, \boldsymbol{C}_n), (\boldsymbol{C}_{t+1}', \ldots, \boldsymbol{C}_m', \boldsymbol{C}_{m+1}', \ldots, \boldsymbol{C}_n')> \in E_2$ hold.

Then, $\boldsymbol{RSG}(CS, E)$ is the resource class hierarchy of resource space $\boldsymbol{RS}$.

Theorem 2. (**Rsm-Disjoin**) Let $\boldsymbol{RS}_1(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_t, \boldsymbol{X}_{t+1}, \ldots, \boldsymbol{X}_m)$ and $\boldsymbol{RS}_2(\boldsymbol{X}_{t+1}, \ldots, \boldsymbol{X}_m, \boldsymbol{X}_{m+1}, \ldots, \boldsymbol{X}_n)$ be two resource spaces which derive from the Disjoin operation on the resource space $\boldsymbol{RS}(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_t, \boldsymbol{X}_{t+1}, \ldots, \boldsymbol{X}_m, \boldsymbol{X}_{m+1}, \ldots, \boldsymbol{X}_n)$. Assume that $\boldsymbol{RSG}(CS, E)$ is the resource class hierarchy of $\boldsymbol{RS}$. The Rsm-Disjoin operation on $\boldsymbol{RSG}(CS, E)$ is defined as $\boldsymbol{RSG}_1(CS_1, E_1)$ and $\boldsymbol{RSG}_2(CS_2, E_2)$ (similar to $\boldsymbol{RSG}_1(CS_1, E_1)$) such that:

(1) $CS_1 = \boldsymbol{X}_1^* \times \boldsymbol{X}_2^* \times \ldots \times \boldsymbol{X}_m^* \cup \{(\pi_1, \ldots, \pi_{i-1}, \gamma_i, \pi_{i+1}, \ldots, \pi_m) \mid 1 \le i \le m\}$;
(2) For any two resource classes $c = (\boldsymbol{C}_1, \ldots, \boldsymbol{C}_t, \boldsymbol{C}_{t+1}, \ldots, \boldsymbol{C}_m)$ and $c' = (\boldsymbol{C}_1', \ldots, \boldsymbol{C}_t', \boldsymbol{C}_{t+1}', \ldots, \boldsymbol{C}_m')$ in CS_1, $<c, c'> \in E_1$ holds if and only if there exists at least a pair of resource classes $(\boldsymbol{C}_1, \ldots, \boldsymbol{C}_t, \boldsymbol{C}_{t+1}, \ldots, \boldsymbol{C}_m, \boldsymbol{C}_{m+1}, \ldots, \boldsymbol{C}_n)$ and $(\boldsymbol{C}_1', \ldots, \boldsymbol{C}_t', \boldsymbol{C}_{t+1}', \ldots, \boldsymbol{C}_m', \boldsymbol{C}_{m+1}', \ldots, \boldsymbol{C}_n')$ in $\boldsymbol{RSG}(CS, E)$ such that $<(\boldsymbol{C}_1, \ldots, \boldsymbol{C}_t, \boldsymbol{C}_{t+1}, \ldots, \boldsymbol{C}_m, \boldsymbol{C}_{m+1}, \ldots, \boldsymbol{C}_n), (\boldsymbol{C}_1', \ldots, \boldsymbol{C}_t', \boldsymbol{C}_{t+1}', \ldots, \boldsymbol{C}_m', \boldsymbol{C}_{m+1}', \ldots, \boldsymbol{C}_n')> \in E$ holds.

Then, $\boldsymbol{RSG}_1(CS_1, E_1)$ is the resource class hierarchy of resource space $\boldsymbol{RS}_1$.

Theorem 3. (**Rsm-Merge**) Let $\boldsymbol{RS}_1$ and $\boldsymbol{RS}_2$ be two resource spaces that can be merged into the resource space $\boldsymbol{RS}$ by merging $\boldsymbol{X}'$ and $\boldsymbol{X}''$ into $\boldsymbol{X}$. Assume that $\boldsymbol{RSG}_1(CS_1, E_1)$ and $\boldsymbol{RSG}_2(CS_2, E_2)$ are the resource class hierarchies of $\boldsymbol{RS}_1$ and $\boldsymbol{RS}_2$ respectively and that ac, ac' and ac'' are the axis resource classes corresponding to axes $\boldsymbol{X}$, $\boldsymbol{X}'$ and $\boldsymbol{X}''$ respectively. The Rsm-Merge operation on $\boldsymbol{RSG}_1(CS_1, E_1)$ and $\boldsymbol{RSG}_2(CS_2, E_2)$ is defined as the directed graph $\boldsymbol{RSG}(CS, E)$ such that:

(1) $CS = CS_1 \cup CS_2 \cup \{ac\} - \{ac', ac''\}$;
(2) $E = E_1 \cup E_2 \cup \{<ac, c> \mid <ac', c> \in E_1 \text{ or } <ac'', c> \in E_2\} \cup \{<c, ac> \mid <c, ac'> \in E_1 \text{ or } <c, ac''> \in E_2\} - \{<ac', c> \mid <ac', c> \in E_1\} - \{<ac'', c> \mid <ac', c> \in E_2\}$.

If $\boldsymbol{RS}$ is in 2NF, then $\boldsymbol{RSG}(CS, E)$ is the resource class hierarchy of resource space $\boldsymbol{RS}$.

Theorem 4. (**Rsm-Split**) Let $\boldsymbol{RS}_1(\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_m, \boldsymbol{X}')$ and $\mathbf{RS}_2(\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_m, \boldsymbol{X}'')$ be two resource spaces deriving from the Split operation on the resource space $\boldsymbol{RS}(\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_m, \boldsymbol{X})$ in 2NF. Assume that $\boldsymbol{RSG}(CS, E)$ is the resource class hierarchy of $\boldsymbol{RS}$ and that ac, ac' and ac'' are the axis resource classes of axes $\boldsymbol{X}$, $\boldsymbol{X}'$ and $\boldsymbol{X}''$ respectively. The Rsm-Split operation on $\boldsymbol{RSG}(CS, E)$ is defined as $\boldsymbol{RSG}_1(CS_1, E_1)$ and $\boldsymbol{RSG}_2(CS_2, E_2)$ (similar to $\boldsymbol{RSG}_1(CS_1, E_1)$) such that:

(1) $CS_1 = \boldsymbol{X}_1^* \times \boldsymbol{X}_2^* \times \ldots \times \boldsymbol{X}_m^* \times \boldsymbol{X}'^* \cup \{(\pi_1, \ldots, \pi_{i-1}, \gamma_i, \pi_{i+1}, \ldots, \pi_m, \pi) \mid 1 \le i \le m\} \cup ac'$;
(2) $E_1 = \{<c, c'> \mid c \in CS_1 \wedge c' \in CS_1 \wedge <c, c'> \in E\} \cup \{<ac', c> \mid c \in CS_1 \wedge <ac, c> \in E\} \cup \{<c, ac'> \mid c \in CS_1 \wedge <c, ac> \in E\}$.

Then, $\boldsymbol{RSG}_1(CS_1, E_1)$ is the resource class hierarchy of resource space $\boldsymbol{RS}_1$.

Theorem 5. All of the RSM-based operations on the extended SLN (Rsm-Join, Rsm-Disjoin, Rsm-Merge and Rsm-Split) keep 1NF, 2NF and 3NF of the Semantic Link Network.

6 Conclusion

The collaboration between the Semantic Link Network and the Resource Space Model is realized by extending the Semantic Link Network and defining the resource class hierarchy and the RSM-based operations on the extended SLN. The extended SLN provides an efficient mapping mechanism for the cooperation between RSM and SLN, facilitates modeling with RSM and SLN, and enhances the interoperability between SLNs by defining semantic-rich operations.

Ongoing work is to design an efficient algorithm for automatically transforming resource spaces to resource class hierarchies, to propose more operations on the extended SLN, to propose its theory on integrity constraint, and to develop tools for assisting integrity and normalization checking in future environment.

References

1. Abiteboul, S., Hull, R. and Vianu, V.: Foundations of Databases. Addison-Wesley, Reading, MA (1995).
2. Codd, E.F.: A Relational Model of Data for Large Shared Data Banks. *Communications of the ACM*, 13(6) (1970) 377-387.
3. Chen, P.: The Entity-Relationship Model□Toward a Unified View of Data. *ACM Transactions on Database Systems*, 1(1) (1976) 9-36.
4. Gyssens, M., Paredaens, J., Bussche, J. and Gucht, D.: A Graph-Oriented Object Database Model. *IEEE Transactions on Knowledge and Data Engineering*, 6(4) (1994) 572-586.
5. Kim, W.: Introduction to Object-oriented Databases. MIT Press, Cambridge, 1990.
6. Levene, M. and Loizou, G.: A Graph-based Data Model and Its Ramifications. *IEEE Transactions on Knowledge and Data Engineering*, 7(5) (1995) 809-823.
7. Levene, M. and Poulovassilis, A.: The Hypernode Model and Its Associated Query Language. In: Proceedings of Jerusalem Conference on Information Technology, Jerusalem, 520-530 1990.
8. Mok, W.Y.: A Comparative Study of Various Nested Normal Forms. *IEEE Transactions on Knowledge and Data Engineering*, 14(2) (2002) 369-385.
9. Poulovassilis, A. and Levene, M.: A Nested-graph Model for the Representation and Manipulation of Complex Objects. *ACM Transactions on Information Systems*, 12(1) (1994) 35-68.
10. Rumbaugh, J., Blaha, M., Premerlani, W., Eddy, F. and Lorensen, W.: Object-Oriented Modeling and Design. Prentice-Hall, 1991.
11. Ullman, J.D.: Principles of Database and Knowledge-Base Systems, vol. I. Computer Science Press, 1988.
12. Zhuge, H.: Active e-Document Framework ADF: Model and Platform. *Information and Management*, 41(1) (2003) 87-97.
13. Zhuge, H.: Resource Space Model, Its Design Method and Applications. *Journal of Systems and Software*, 72(1) (2004) 71-81.
14. Zhuge, H.: Resource Space Grid: Model, Method and Platform, *Concurrency and Computation: Practice and Experience*, 16(14) (2004) 1385 – 1413.
15. Zhuge, H.: The Knowledge Grid. World Scientific Publishing Co., Singapore, 2004.
16. Zhuge, H.: China's E-Science Knowledge Grid Environment, *IEEE Intelligent Systems*, 19(1) (2004) 13-17.
17. Zhuge, H., Liu, J., Feng, L., Sun, X. and He, C.: Query Routing in a Peer-to-Peer Semantic Link Network. *Computational Intelligence*, 21(2) (2005)197-216.
18. Zhuge, H. and Xing, Y.: Integrity Theory for Resource Space Model and Its Application, Keynote at WAIM2005, *Springer LNCS* 3739, pp.8-24, 2005.
19. Zhuge, H.: Semantic Grid: Scientific Issues, Infrastructure, and Methodology, *Communications of the ACM*. 48(4) (2005)117-119.