Granular Structures in Graphs

Guang Chen[1], Ning Zhong[1,2]

[1] International WIC Institute, Beijing University of Technology
Beijing, 100124, P.R. China
cg@emails.bjut.edu.cn

[2] Department of Life Science and Informatics , Maebashi Institute of Technology
Maebashi-City, 371-0816, Japan
zhong@maebashi-it.ac.jp

Abstract. The granular structures emphasize a multilevel and multiview understanding of problems. This paper focuses on a study of how to granulate a graph, and how to extract the granular structures in the graph. The granular structures can be seen as an abstract, summary or epitome of the graph. Two granular structures extraction models are proposed, one is based on the degree of the vertex, the other is based on the weight of the edge. Each model is a multilevel representation, and the models integrated together presents a multiview comprehension of the graph.

Keywords: Granular Structures, Granular Computing, Graph, Multilevel, Multiview.

1 Introduction

1.1 Purpose of the Study

Granular computing, as an emerging field of studies, aims at a systematic investigation of granular based theories and methodologies for supporting human problem solving on one hand and machine problem solving on the other [15]. Many researchers have made significant progress on concrete models and methods of granular computing [1,5,6,7,8,10,14,18,19,20,21,22].The triarchic theory of granular computing [16] offers a conceptual framework of granular computing by weaving together three powerful ideas: structured thinking, structured problem solving, and structured information processing. It emphasizes the exploitation of useful structures known as granular structures characterized by multilevel and multiview. A single hierarchical granular structure provides a multilevel understanding and representation of a problem or a system. But it typically captures one particular aspect and therefore offers one view. By constructing a family of hierarchies, it is possible to obtain multiple different views. Granular structures are a family of complementary hierarchies working together for a complete and comprehensive multiview understanding and representation. The set-theoretic approaches to granular computing [17] propose a framework to investigate granular structures used in several studies.

J.T. Yao et al. (Eds): RSKT 2011, LNCS 6954, pp. 649–658, 2011.

Many real world problems depicted and analyzed by graphs, have achieved good impacts. Representation the problem by graph has advantage of visualization and intuition. In the graph, vertices and edges connection between vertices constitute the network to indicate specific problems. The vertices represent the objects; and the edges represent the relation of objects. To all problems related to binary relations, graph can provide a model, which, in many fields of science and real life has become increasingly important role.

In this paper, we combine the two notions 'granule' and 'graph'. Problem solving with granular structures in graph have these two advantages: one is visual representation of the graph, the other is merits of multilevel and multiview of the granular structures. So we study how to granulate the graph, and how to extract the granular structures in graph. The granular structures are an abstract and summary of the graph; it can be seen as a epitome of the graph. We can achieve a well-rounded and manoeuvrable understanding of the graph by granular structures. Two models are proposed to study the extraction of the granular structures, one is based on the degree of vertex, the other is based on the weight of edge. Each model is a multilevel description of the graph, and the models present a multiview comprehension of the graph together.

1.2 Related Studies

Some researchers have considered the relation of granular computing and graph. Stell [11,12] discussed the granularity for graphs and hypergraphs, and its relation with rough set theory and mathematical morphology. His work denoted that handling spatial data at several levels of detail the granulation of graphs was an important topic, and demonstrated that there were several quite different kinds of granulation for graphs. Wong and Wu [13] used the idea of granularity and hypergraph method to investigate database scheme. The database scheme was presented by hypergraph. Chen, et al. [3] studied how to use hypergraph to present granules and granular structures, and concreted a hypergraph model for granular computing.

Many researchers have studied clustering on graphs [2,4,9]. Graph clustering consists in grouping the vertices of the networks into different clusters according to various different criteria, i.e., edge structure of the graph, vertices of the same type, etc. Commonly, there should be many edges within each cluster and relatively few between the clusters.

Previous researches explored how to construct granules or group clusters for graph in different domains. There was less discussion in how to construct granular structure for graphs, how to present the relationships and mappings among multiple levels, how to transform among different granularities, and how to solve problem under the granular structures. So our work in addition to studies grouping granules, focuses on extracting granular structure in graph, and presents relationships among the levels.

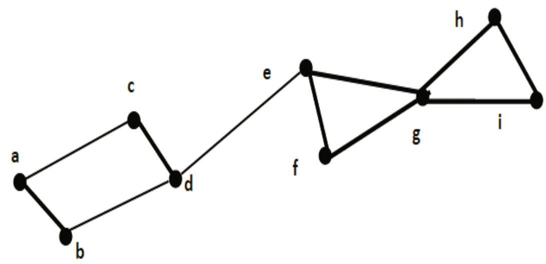

Fig. 1. A graph

2 Granules and Granular Structures in Graphs

2.1 Granules and Granular Structures

Granular computing exploits structures in terms of granules, levels, and hierarchies based on multilevel and multiview representations [15,16,18]. A granule is a group of elements which have similar characteristics and can be seen as a whole unit. A granule plays two distinct roles. It may be an element of another granule and be a part of forming the other granule. It may also consist of a family of granules and be considered to be a whole. All granules in a level may collectively show a certain structure. This is the internal structure of a granulated view. Granules in a level, although may be relatively independent, are somehow related to a certain degree. A hierarchy represents one view of a problem with multiple levels of granularity, and a hierarchical structure is a understanding of the problem in a specific perspective. So granular computing exploits multi-level and multi-view representations in problem solving with granular structure.

2.2 Granules in a Graph

Definition 1. *A* ***graph G****, denoted as $G = (V, E)$, is defined as a set of vertices V and a set of edges E among those vertices.*

In an undirected graph, each edge is an unordered pair (v, w). In a directed graph (also called a digraph in much literature), edges are ordered pairs. The vertices v and w are called the endpoints of the edge. In a weighted graph, a weight function is defined that assigns a weight on each edge. Fig.1 shows an example of graph.

Definition 2. *Let $G = (V, E)$ be a graph. Any $g \subseteq V$ is called a granule on G.*

A granule g is a subset of the universal set of vertices in graph, it is a integration of the vertices which we process together when solving problems, it is a reflection of the feature about the integration and a representation of the knowledge. A granule is the abstract of the whole, not only the set of the vertices. The power set 2^V consists of all possible granules formed from a universal set V. The standard set-inclusion relation $\subseteq$ defines a partial order on 2^V, which leads to sub-super relationship $\subseteq$ between granules.

Definition 3. *Let $g_1 \subseteq g_2 \subseteq V$, we call g_1 a* ***sub-granule*** *of g_2 and g_2 a* ***super-granule*** *of g_1.*

A granule plays two distinctive roles, it may be an element of another granule and considered to be a part forming the other granules. A granule may also be a family of granules and considered to be a whole. Under the partial order $\subseteq$, the empty set ϕ is the smallest granule and the universe V is the largest granule. When constructing a granular structure, we may consider a family G of subsets of V and an order relation on G.

2.3 Granular Structures in a Graph

Definition 4. *When all the vertices are integrated into granules respectively, a granulation process is finished, We call a granulated graph is a* ***level****.*

According to different granulation standards, we can make different granulation processes, and get different levels. Each level has a specific granularity, and may be viewed as a specific comprehension of the problem.

Definition 5. *Consider two levels l_1 and l_2, for each granule g_1 in level l_1, if there is a granule g_2 in level l_2, and $g_1 \subseteq g_2$, that is, for each granule g_1 in level l_1, we can find a sup-granule g_2 in level l_2, then we call that l_1 has a finer granularity than l_2, and l_2 has a coarser granularity than l_1.*

The finer level is a representation of the problem in a concrete view, the coarser level is a representation of the problem in an abstract view. The transformation among the levels will respond the changes from concrete to abstract, and refinement to coarseness.

Example 1. Fig.2 shows a coarser granulation of the graph in Fig.1. In this level, the graph is granulated to three granules, granule g_1 contains four vertices, granules g_2 and g_3 contains three vertices respectively. Fig.3 shows a finer granulation

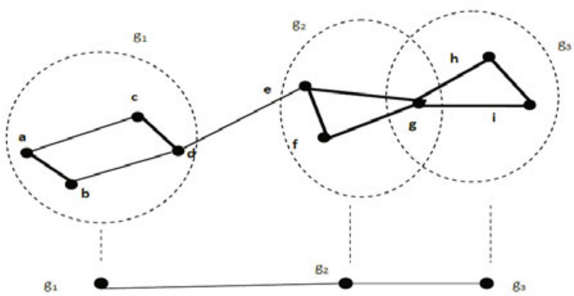

Fig. 2. Level 1: a coarser granulation of graph

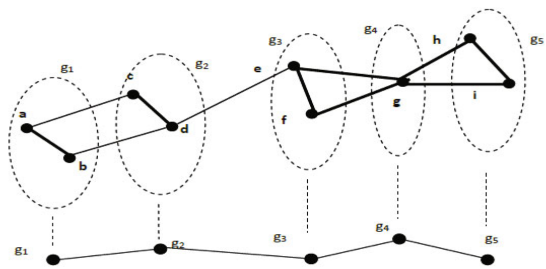

Fig. 3. Level 2: a finer granulation of graph

of the graph in Fig.1. In this level, the graph is granulated to five granules, granules g_1, g_2, g_3, g_5 contains two vertices respectively, granule g_4 contains one vertex only. So level 1 has a coarser granularity than level 2, and level 2 is finer than level 1. Through granulated graph, we can realize the distribution of vertices and the structure of graph clearly.According to Definition 4 and Fig.2, we can see that a level is essentially a covering of V.

Definition 6. *Suppose $G \subseteq 2^V$ is a nonempty family of granules of the levels. The poset $(G, \subseteq)$ is called a **granular structure**, where $\subseteq$ is the set-inclusion relation.*

By the partial order relation $\subseteq$, different levels can be integrated to a hierarchical, and the granules in different levels can be arranged into a granular structure. For each level has a special granularity and different levels present different descriptions of the graph, the granular structure offer multiple granularities and multiple views of the problem. One can switch among these granularities by the granular structure. The relation $\subseteq$ is an example of partial orders. In general, one may consider any partial order on G and the corresponding poset $(G, \preceq)$.

Definition 7. *Considering two levels l_1 and l_2 of a granular structure, A mapping f by set-inclusion relation $\subseteq$ from l_1 to l_2: $l_1 \longrightarrow l_2$ maps a granule g_1 in l_1 to granules $\{g_i | g_i \subseteq g_1\}$ in l_2. f^{-1}: $g_2 \longrightarrow g_1$ maps a granular g_2 in l_2 to a granular $\{g_j | g_2 \subseteq g_j\}$ in l_1.*

Example 2. Fig.4 shows a granular structure of the graph in fig.1 by mapping or the partial order relation $\subseteq$. For simplicity, we assume that the root is the entire universe. The structure integrates three levels of the graph together, it presents three granularities of the description and comprehension for the graph. Conceptually, this structure may be viewed as a successive top-down decomposition of the universe vertices, The root is divided into a family of granules. Alternatively, this structure may also be viewed as a successive bottom-up combination of smaller granules to form larger granules.

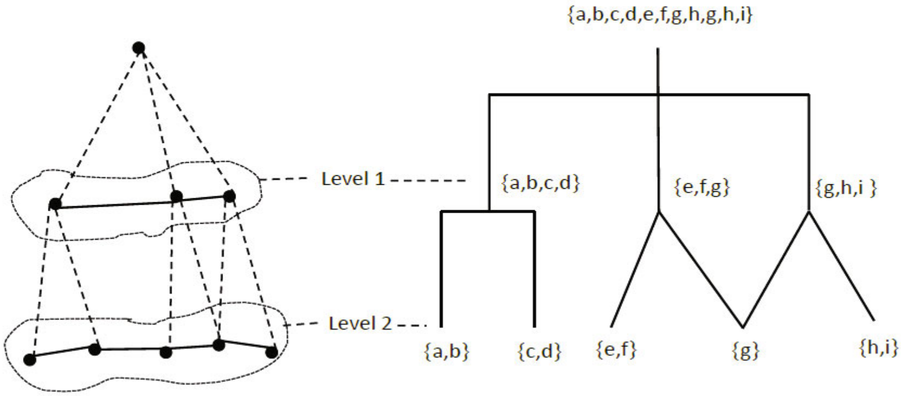

Fig. 4. A multi-level granular structure of the graph in Fig.1 by mapping or partial order relation $\subseteq$

3 Extraction of the Granular Structures in a Graph

Granular structures are an abstract and summary of the graph. It can be seen as an epitome of the graph. Through the granular structures, the comprehension of the graph is clear at a glance. The granular structures have both glancing and particular information of the graph. When we want to have a finer or coarser understanding of the graph, we only need to switch among the levels in structures.

We introduce two models to extract the granular structure in graph, of which one is based on the degree of vertex, the other is based on the weight of edge. Each granulation used different models can be seen as a special view of the graph. We can get a multi-level granular structure by these two models respectively, and they integrate together as a multi-level and multi-view representation of the graph.

3.1 Degree-Based Model

This model is based on the size of vertex degree. A indirected and non-weighted graph is used in this model. Firstly, We define the degree of vertex and the distance between two vertices for this model.

Definition 8. *Considering vertex $v \in E$, the **degree** of v is the number of the edges which connect to v, denoted as $deg(v)$.*

Definition 9. *The **distance** between vertex v_1 and vertex v_2 is the number of the edges in the Shortest Path from v_1 to v_2, denoted as $d(v_1, v_2)$.*

Definition 10. *A **center** of granule g is the vertex which degree size is largest of the vertices in this granule, denoted as $center(g)$.*

Algorithm 1.. Degree-based Algorithm

```
1. Construct the undirected non-weighted graph
2. Set thresholds
3. FOR each threshold
4.     Select the vertices which degree size exceeds the threshold
5.     FOR each vertex except centers in graph
6.         Compute its distance to the centers
7.         Group it into the granule which center has a shortest distance to it
8.     END FOR
9. END FOR
10.FOR each level except the last
11.    Mapping it to the next one
12.END FOR
```

Definition 11. *The* ***distance*** *between granule g_1 and granule g_2 is the distance between the centers of the two granules, denoted as $d(g_1, g_2)$.*

The center is a delegate of the granule, for it tights other vertices most closely.

A typical algorithm for this model is described as Algorithm 1: to select the vertices which degree size is larger than a threshold θ in vertices set V as centers of granules; for other vertices in graph, to compute their distances to the centers; to group them into the granules which center has a shortest distance to them respectively; at last the global graph is granulated. By using a vertex to represent a granule, and using edges to connect the vertices, we can get a new graph. This graph is a abstraction or summary of the original graph. It is a level of the granular structure. We can observe the graph clearly at a glance by the level. This process is shown as Fig.5. By changing the threshold, we can get several levels which have different granularities. Then we get a granular structure by integrating the levels used *mappingf*. In the structure, the more coarser granularity a level has, the more abstractly the graph is summarized, and the more quickly we understand the graph.

3.2 Weight-Based Model

This model is based on the information of the weight on the edge. For this model, we use an undirected weighted graph, a weight function is defined that assigns a weight on each edge. The weight denotes a distance between the endpoints. In many situations, the vertices near each other should be seen as an integral whole for convenience. So we consider how to granulate the vertices by distance. We define the distance of the vertices and granules for this model at first.

Definition 12. *The* ***distance*** *between vertex v_1 and vertex v_2 is the sum of the weights on the edges in the Shortest Path from v_1 to v_2, denoted as $d(v_1, v_2)$.*

Definition 13. *The* ***distance*** *between granule g_1 and granule g_2 is the average of the distance between any two vertices belongs to the two granules respectively, denoted as $d(g_1, g_2)$.*

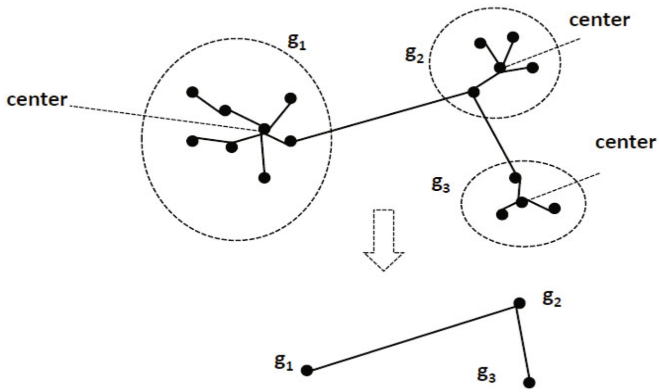

Fig. 5. Granulation in degree-based model

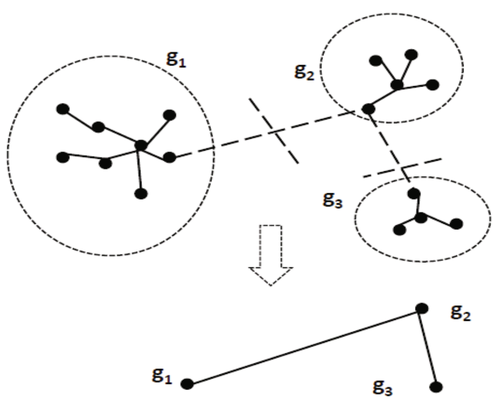

Fig. 6. Granulation in weight-based model

Algorithm 2.. Wight-based Algorithm

1. Construct the undirected weighted graph
2. Set thresholds
3. FOR each threshold
4. Delete edges which weight exceeds the threshold, Then the graph turns into a forest
5. Each sub-connected-graph of the forest is a granule
6. END FOR
7. FOR each level except the last
8. Mapping it to the next one
9. END FOR

A typical algorithm for this model is described as Algorithm 2: for an undirected weighted graph, to delete the edges which weight exceeds given threshold θ; then to out put the connected components of the resulting graph; finally each connected sub graph constructs a granule. This process is shown as Fig.6. When the given threshold θ changed, a new level will be constructed.

4 Conclusion

The prominent characteristic of granular structures is multilevel and multiview. Problem solving with granular structures in graph allows people to obtain both of the two advantages: visual representation of graph and multilevel and multiview of the granular structures. This paper defines the granules and granular structures in graph, and studies the extraction of the granular structure in graph by two typical models. The first model is based on the information of the vertex degree, the second is based on the information of weight on edge. The methods is expected to be a way more consistent with human being thinking in problem solving.

Acknowledgments. This work is partially supported by National Natural Science Foundation of China (No.60905027) and Beijing Natural Science Foundation (No.4102007).

References

1. Bargiela, A., Pedrycz, W.: Granular Computing: An Introduction. Kluwer Academic Publishers, Boston (2002)
2. Brandes, U., Gaertler, M., Wagner, D.: Experiments on Graph Clustering Algorithms. In: Di Battista, G., Zwick, U. (eds.) ESA 2003. LNCS, vol. 2832, pp. 568–579. Springer, Heidelberg (2003)
3. Chen, G., Zhong, N., Yao, Y.Y.: Hypergraph Model of Granular Computing. In: 2008 IEEE International Conference on Granular Computing, pp. 80–85. IEEE Press, New York (2008)
4. Flake, G.W., Tarjan, R.E., Tsioutsiouliklis, K.: Graph Clustering and Minimum Cut Trees. Internet Mathematics 1(1), 385–408 (2004)
5. Hobbs, J.R.: Granularity. In: 9th International Joint Conference on Artificial Intelligence, pp. 432–435 (1985)
6. Luo, J.G., Yao, Y.Y.: Granular State Space Search. In: Butz, C., Lingras, P. (eds.) Canadian AI 2011. LNCS(LNAI), vol. 6657, pp. 285–290. Springer, Heidelberg (2011)
7. Pawlak, Z.: Rough Sets. International Journal of Computer and Information Sciences 11, 341–356 (1982)
8. Polkowski, L.: A Model of Granular Computing with Applications: Granules from Rough Inclusions in Information Systems. In: IEEE International Conference on Granular Computing, pp. 9–16 (2006)
9. Schaeffer, S.E.: Graph Clustering. Computer Science Review 1(1), 27–64 (2007)
10. Skowron, A., Stepaniuk, J.: Information Granules: Towards Foundations of Granular Computing. International Journal of Intelligent Systems 16, 57–85 (2001)

11. Stell, J.G.: Granulation for graphs. In: Freksa, C., Mark, D.M. (eds.) COSIT 1999. LNCS, vol. 1661, pp. 417–432. Springer, Heidelberg (1999)
12. Stell, J.G.: Relational Granularity for Hypergraphs. In: Szczuka, M., Kryszkiewicz, M., Ramanna, S., Jensen, R., Hu, Q. (eds.) RSCTC 2010. LNCS(LNAI), vol. 6086, pp. 267–276. Springer, Heidelberg (2010)
13. Wong, S.K.M., Wu, D.: Automated Mining of Granular Database Scheme. In: IEEE International Conference on Fuzzy Systems, vol. 1, pp. 690–694 (2002)
14. Yager, R.R., Filev, D.: Operations for Granular Computing: Mixing Words with Numbers. In: IEEE International Conference on Fuzzy Systems, pp. 123–128 (1998)
15. Yao, Y.Y.: A Partition Model of Granular Computing. In: Peters, J.F., Skowron, A., Grzymała-Busse, J.W., Kostek, B.z., Świniarski, R.W., Szczuka, M.S. (eds.) Transactions on Rough Sets I. LNCS, vol. 3100, pp. 232–253. Springer, Heidelberg (2004)
16. Yao, Y.Y.: The Art of Granular Computing. In: Kryszkiewicz, M., Peters, J.F., Rybiński, H., Skowron, A. (eds.) RSEISP 2007. LNCS (LNAI), vol. 4585, pp. 101–112. Springer, Heidelberg (2007)
17. Yao, Y.Y., Miao, D.Q., Zhan, N., Xu, F.F.: Set-Theoretic Models of Granular Structures. In: Yu, J., Greco, S., Lingras, P., Wang, G., Skowron, A. (eds.) RSKT 2010. LNCS, vol. 6401, pp. 94–101. Springer, Heidelberg (2010)
18. Yao, Y.Y., Zhong, N.: Granular Computing. Wiley Encyclopedia of Computer Science and Engineering 3, 1446–1453 (2009)
19. Zadeh, L.A.: Towards a Theory of Fuzzy Information Granulation and Its Centrality in Human Reasoning and Fuzzy Logic. Fuzzy Sets and Systems 90, 111–127 (1997)
20. Zhang, L., Zhang, B.: The Quotient Space Theory of Problem Solving. Fundamenta Informatcae 59, 287–298 (2004)
21. Zhu, W., Wang, F.Y.: Relationships among three types of covering rough sets. In: 2006 IEEE International Conference on Granular Computing, pp. 43–48. IEEE Press, New York (2006)
22. Zhu, W., Wang, F.Y.: On Three Types of Covering-based Rough Sets. IEEE Transactions on Knowledge and Data Engineering 19(8), 1131–1144 (2007)