Granular Computing in Intelligent Transportation: An Exploratory Study

Seng Dewen, Cheng Xinhong, Chen Jing, Fang Xujian

School of Computer Science and Technology, Hangzhou Dianzi University, 310018 Hangzhou, China
Key Laboratory of Complex Systems Modelling and Simulation, Ministry of Education, China
Email: sengdw@163.com

Abstract: With the continuous development of the cities, the traffic situation has gradually become a topic of concern. The concept of an intelligent transportation system has been proposed and developed. In the field of intelligent transportation, the traffic data has gradually increased. People have higher demands to real time data. The traditional data processing methods and tools have become unable to meet the needs of urban transport development. In this paper we analyzed the basic theory of granular computing, the methods, technology and current situation of granular computing. Besides, we discussed the hot issues of granular computing in an intelligent transportation system. Finally, granular computing in the development of intelligent transportation fields was also discussed.

Keywords: Intelligent Transportation system, granular computing, traffic data, rough set.

1. Introduction

With the rapid development of the computers and networks, the contradiction between the rich data and poor knowledge has become increasingly prominent. In this connection the paper “Fuzzy sets and information granularity” [1] was published. In all aspects of life the researchers accepted that granular information has generated strong interest. The paper noted that the concept of information granules is found in many areas, but in different forms. In 1985 in Hobbsthesis-“granularity” [2], the concept of granularity was formally put forward and decomposition and combination were discussed, as well as how to get particles of different sizes. Since the granularity concept is proposed along with a sharp increase in the amount of information, the uncertainty of information and data is becoming more and more outstanding, and how to get useful information from
uncertainty data, has become the topic of many researchers. Thus uncertain knowledge discovery methods have also appeared and in 1965 the theory of fuzzy sets was proposed by Zadeh [27], the rough set theory – in 1982 by Pawlak [28] and the quotient space theory – in 1990 by Bo Zhang and Ling Zhang [29]. They are three basic mathematical theories for granular computing, which are efficient methods for dealing with uncertainty and knowledge discovery.

The theory of fuzzy sets is the opposite of the exact set theory, it is the basic method of depicting a fuzzy phenomenon and fuzzy concept; the concept is one of the basic forms of thinking, it reflects the essential characteristic of objective things. In the process of cognition, the common characteristic of a thing that feels is summed up, forming the concept. The fuzzy concept refers to the extension of the concept of uncertainty, thus the fuzzy set theory was formed. As an important content of the fuzzy set theory, the model of computing with words has a profound effect on the development of granular computing.

The rough set is a mathematical tool dealing with uncertainty problems, after the probability theory, fuzzy sets, evidence theory were proposed. It is characterized by incomplete and uncertain mathematical tools and can analyze the imprecise, inconsistent, incomplete and imperfect information efficiently. It can also carry on the analysis and reasoning of the data and find the implicit knowledge and reveal potential rules. In the process of dividing the rough set, the individual of the difference is often attributed to the same class, and the relationship is an indiscernible relation. The rough set theory is built on the basis of the classification mechanism. It considers that the classification is equivalent to a specific space, and the equivalence relations constitute the division of the space. It holds that knowledge is division of the data, each of which is called a concept. The main idea of the rough set theory is to approximate the description of the knowledge in the known knowledge, which is not exact or uncertain.

Along with the globalization and the information development trend, the traditional transportation technology and tools cannot already adapt to the request of the economic social development. The Intelligent Transportation Systems (ITS) are the inevitable choice of development in transport, it is a revolution in the transportation infrastructure. The integration and application of advanced information technology, communication technology, control technology, sensor technology, computer technology and system comprehensive technology, enables the relationship between people, vehicles, roads to be presented in a new way, in order to achieve the real time, accurate, efficient, safe, energy-saving goal. ITS can use the existing transportation facilities to reduce the traffic load and environmental pollution, ensure the traffic safety, improve the transportation efficiency, and therefore, the attention towards them is increased in all the countries.

The traffic data is in real time, ambiguous and uncertain, and granular computing has the ability to deal with uncertain and changing data. Therefore, it is not hard to imagine how by combination of granular computing and an intelligent transportation system, using the method and theory of granular computing to solve the problems and methods in the intelligent transportation systems.
In this paper the first section introduces the main particle model. The second section further introduces the basic composition of granular computing. The third section and fourth section respectively introduce the basic issues of granular computing and related concepts. The fifth section mainly discusses the research status quo of granular computing; the sixth section analyzes and summarizes the research progress of granular computing in the field of intelligent traffic. Finally, the whole paper is summarized and the development direction of intelligent transportation is analyzed.

2. GrC model

Granular computing is mainly divided into three models, respectively, the model of computing with words based on the fuzzy set theory, the Granular computing model based on the rough set theory and the Granular computing model based on the Quotient space theory. They are models that the people use in order to describe the ability to deal with matters of different size. The method of computing with words is using words instead of numbers and reasoning [3]. It believes that the knowledge of particles can be represented by a “word”, different levels of the word represent different levels of knowledge, how to use the language to reason, so it is necessary to build a word calculation model.

The rough set is a mathematical tool dealing with uncertainty problems, after the probability theory, fuzzy set, evidence theory have been proposed. As a new intelligent information processing technology, it is proposed by a Polish scholar Pawlak for incomplete data analysis, reasoning, study and discovery [4, 5]. In recent years, as a new soft computing method, rough set gets more and more attention. Its efficiency has been proven in many successful applications in science and engineering fields, and it is one of the focuses in the field of artificial intelligence theory and its applications.

The emphasis of the quotient space theory is different, it is not only used for a given quotient space (knowledge base) to discuss the issue of knowledge expression, but in the space of all possible suppliers to identify the most appropriate commercial space. The aim is by using different providers’ space observation (from different angles) of the same issue, to get different perspectives of the problem of the ultimate synthesis of the overall understanding of the problem. In 2003 Zhang Ling and Zhang Bo combined fuzzy quotient space theory with the concept of a fuzzy quotient space granular computing theory in order to provide new mathematical models and tools that are successfully applied to data mining and other fields (see [6-10]).

3. Composition of GrC

Granular computing basic components include a particle, a particle layer and a grain structure.
3.1. Particles

The particles, the basic elements of a granular computing model [11, 12], present a primitive grain calculation model. In life the particles are everywhere, they can be any microscopic or macroscopic object or a number of nodes definition. In the process of definition, the particle can have also mutual existence. A factor, measuring the particle size, called granularity [13], reflects the degree of the granulation space. The smaller the particle size is, the greater the corresponding particle becomes and vice versa.

3.2. Granular layer

All the particles are obtained according to an actual demand of the grain rule. Constituting a granular layer is a kind of abstract description of the problem space [13]. The particles in the same layer tend to have certain common attributes or properties. The internal structure of the grain layers refers to the domain structure of each particle in the same layer, as well as to the relationship between the particles. In the process of problem solving, the selection of an appropriate particle layer is very important for problem solving. As the particle size changes, the particle layer can also be transformed into each other. A high level of granularity larger particle layer can be divided into a plurality of the low level of the granular layer, whereas the particle size becomes smaller, the low level of the granular layer can be synthesized with a high level of the grain layer. The results between the different grain layers can be transformed into each other.

3.3. Grain structure

A grain rule corresponds to a grain layer, and different granulation criteria correspond to different layers of grain. It reflects the ability of people to see problems from different angles and different directions. Different particle layers are linked together to constitute the grain structure. The grain structure gives a description of the structure of the problem or space. It is possible to simplify the computational complexity of the problem space to some extent.

The particle, the grain layer and the grain structure are very important in the process of solving granular computing. In general, the solution space is divided into a number of particles; each particle is solved according to the specific criteria for synthesis of the particle layer solutions. Finally, the final solution of the whole solution space is synthesized according to the solution of each layer. In this way, a complex solution space is divided into a variety of small particles for calculation, and the complexity of the space is reduced. Calculating of the particle is the basis for solving the entire solution space; therefore, better granulation becomes a hot topic of current research.

Fig. 1 shows that the collection is divided according to different granulation criteria. If the particle rule is “all circular pattern” (a part of the green box in the picture), the circular pattern in the box is clustered together with the granulation criterion. The grain layer is composed of the particles according to different granulation criteria – (a collection of blue boxes and red boxes in the picture).
different grain layers are composed of a grain structure (a collection of black boxes in the picture).

![Fig. 1. Composition of GrC](image)

4. The basic problem of GrC

4.1. Granulation

Granulation is a refinement process for the solution space under central granulated guidelines. Spatial extent of the grain size is reflected by the particle size. Graining results in different reactions of the way people look at things. Granulation is the foundation of granular computing. The quality of granulation is often determined by the fact whether granular computing in the space complexity is good or bad [13].

4.2. The calculation of a particle

Granular computing, narrowed on the basis of granulated particles in the solution space, is calculated separately. At present, granular computing is the result of conversion on the basis of the grain level structure, including two aspects: First, the internal granular layer transformation, where the internal layer of particles can be transformed into each other. The other one is conversion between the grain layer, including conversion between the layers and between the structures [13]. With these two aspects the complexity of problem solving can be reduced, and the problem becomes easy to solve.

Granulation is the key in two basic problems of granular computing. So the granulation process is currently an important research problem.

5. Definition of GrC

**Definition 1 (Information sheet knowledge expression system [14]).** An information sheet knowledge expression system $E$ can be expressed as a four-tuple: $E = (S, A, V, f)$, in which $S$ is the object collection, also known as the universe; $A = C \cup R$ are the property complete works, they are called a condition attribute set and a result attribute set; $V = \bigcup_{r \in A} V_r$ is a collection of the property values, $V_r$ represents an attribute $r \in A$ property value range, that is the range of values of the attributes of $r$; $f : S \times A \rightarrow V$ is a message function that specifies the $S$ values for each attribute of the object $x$ [14].
Definition 2 (Fuzzy relation [14]). For any set of the attribute $Q \subseteq A$, indiscernible to define a binary relation (fuzzy relation)

\[
\text{IND}(Q) = \{(x, y) | (x, y) \in S^2 \land \forall b \in R(b(x) = b(y))\}.
\]

Definition 3 (Upper and lower approximations of the rough set [14]). Given the information sheet knowledge expression system $E = (S, A, V, f)$, for any collection of objects $X \subseteq S$ and properties collection $Q \subseteq A$, $X$ on the approximation set on $R$, $\overline{Q}(X)$ and the lower approximation set $\underline{Q}(X)$ are defined as follows:

\[
\begin{align*}
\overline{Q}(X) &= \bigcup \{Y_i \in S | \text{IND}(Q) \land Y_i \cap X \neq \emptyset\}, \\
\underline{Q}(X) &= \bigcup \{Y_i \in S | \text{IND}(Q) \land Y_i \subseteq X\},
\end{align*}
\]

where $S|\text{IND}(Q) = \{X | X \subseteq S \land \forall x \in X, y \in X, b \in R(b(x) = b(y))\}$ is not clearly divided in relation $R$ on $S$.

The upper and lower approximation sets can also be defined by sets of the following form:

\[
\begin{align*}
\overline{Q}(X) &= \bigcup \{x | x \in S \land [x]_R \subseteq X\}, \\
\underline{Q}(X) &= \bigcup \{x | x \in S \land [x]_R \cap X \neq \emptyset\},
\end{align*}
\]

where $[X]_R$ represents the equivalence class object $x$ in the relationship between the formation of $R$; $\overline{Q}(X)$ is expressed as a collection of all the elements of the target set $X$ in the collection $S$ for the relationship $Q$. $\overline{Q}(X)$ denotes a collection of the elements which overlap with the target set $X$ in the collection $S$ for the relationship $Q$.

Definition 4 (Rough set [14]). Given the knowledge representation system $E$, for any collection of objects $X \subseteq S$ and properties collection $R \subseteq A$, if and only if $\overline{Q}(X) = \underline{Q}(X)$, the collection $X$ is the rough set of $R$. As shown in Fig. 2, the shadow part of the graph is the object subset $X$.

![Fig. 2. Rough set](image)

Definition 5 (Precision set [14]). Precision set relative to the rough set, given the knowledge representation system $E$, for any collection of objects $X \subseteq S$ and properties collection $R \subseteq A$, if and only if $\overline{Q}(X) = \underline{Q}(X)$, the collection $X$ is the
precision set of $R$. As shown in Fig. 3, the shadow part of the graph is the object subset $X$.

![Fig. 3. Precision set](image)

**Definition 6 (Boundary set).** Given the information sheet knowledge expression system $E = (S, A, V, f)$, for any collection of objects $X \subseteq S$ and properties collection $R \subseteq A$, the boundary set $B(X)$ is defined as follows:

$$B(X) = \overline{Q}(X) - \underline{Q}(X),$$

where $\overline{Q}(X)$ is the upper approximation set of $X$ for $R$ and $\underline{Q}(X)$ is the lower approximation set of $X$ for $R$.

Fig. 4 shows a graphical representation of the approximate upper and lower sets and the boundary set.

![Fig. 4. Schematic diagram of the upper and lower approximations of the rough set](image)

In the calculation process of a rough set, the division of elements in the collection is the basis of rough set computing. We need to calculate the rough set with more accurate approximate sets in order to get more accurate results. Here we introduce the concept of roughness and membership degree.

**Definition 7 (Roughness and roughness accuracy [14]).** Given the information sheet knowledge expression system $E = (S, A, V, f)$, for any collection of the objects $X \subseteq S$ and properties collection $R \subseteq A$, the roughness and roughness accuracy are defined as follows:

$$\text{Roughness accuracy: } \alpha(X) = \frac{q(X)}{\overline{q}(X)},$$

128
Roughness: $\delta(X) = 1 - \alpha(X) = 1 - \frac{Q(X)}{\bar{Q}(X)} = \frac{B(X)}{\bar{Q}(X)}$.

According to the two formulas above given, one can see that $0 \leq \alpha(x) \leq 1$, $0 \leq \delta(x) \leq 1$, roughness is an indicator of the uncertainty of the object set $X$. The smaller the roughness is, more precise $X$ is; the greater the roughness is, more blurred $X$ is.

In Fig. 5 the shadow part is the boundary set $X$. The green box is a collection of all the objects. The blue boundary represents the upper approximate set $\bar{Q}(X)$. The red boundary represents the lower upper approximate set $\bar{Q}(\bar{X})$. The roughness accuracy of the shadow part of the graph can be calculated by the formula of roughness and roughness accuracy:

$$\alpha(X) = \frac{Q(X)}{\bar{Q}(X)} = \frac{2}{11},$$

$$\delta(X) = 1 - \alpha(X) = 1 - \frac{Q(X)}{\bar{Q}(X)} = 1 - \frac{2}{11} = \frac{9}{11}.$$  

**Definition 8 (Membership degree [30]).** Given the information sheet knowledge expression system $E = (S, A, V, f)$, for a given object set $X$, $P \in [X]_R$, $[X]_R$ is the equivalence class of $R$, and the membership degree function of $P$ is

$$\beta(P) = \frac{X \cap P}{P}.$$ 

According to the membership degree function $\beta(P)$, we can see that the membership degree of all elements of lower approximation is $\beta(q) = 1$, where $q \in Q(x)$; the membership degree of all elements of the boundary set is $0 < \beta(b) < 1$, where $b \in B(X)$.

In Fig. 6 the shadow part represents the boundary set $X$. The blue boundary region represents the upper approximation set $\bar{Q}(X)$ of the boundary set $X$. The
yellow region represents the lower approximation set $Q(X)$ of the boundary set $X$. We focus on the red area in the picture and define $P$ as the red area. According to the definition of membership, the membership degree of $P$ is:

\[
\beta(P) = \frac{|X \cap P|}{|P|} = 0.5.
\]

**Definition 9 (the $\gamma$ approximation set of $X$ [14]).** The $\gamma$ approximation set of boundary set $X$ is defined as follows:

\[
\beta_{\gamma}(X) = \{x | x \in S \land \beta(x) \geq \gamma\}.
\]

According to the formula, we can see that the $\gamma$ approximation set of $X$ is the collection of all elements for which the membership degree is greater than $\gamma$ in the collection of $S$.

In Fig. 7 the shadow part represents the object subset $X$. The yellow boundary region represents the upper approximation set $\tilde{Q}(X)$, and the blue region represents the lower approximation set $Q(X)$. The red dotted area is the 0.5 approximation set of $X$.

6. Development of GrC

Granular computing model must simulate human thinking and methods to solve complex problems. It is an efficient tool of complex problem solving, massive data mining and fuzzy information processing, when dealing with large amounts of complex information. Due to the limited cognitive ability, people tend to put large amounts of complex information. In accordance with their respective characteristics and performance it can be divided into a number of simple blocks. Every block is regarded as a particle. In fact, Grain refers to the individual through a fuzzy relation, similarity relation, neighbourhood relation or function relation of the block. This process is called information granulation. Due to the characteristics of granular computing itself, many researchers have introduced a method of granular computing to business.

Wang Guo-Yin and others first analyzed and pointed out some defects of the traditional rule acquisition methods in the paper, based on the maximum grain rule acquisition algorithm, so that granularity principles of attribute reduction are analyzed from the viewpoint of granular computing. They indicated that the essence
of the attribute reduction process is to find the decision space of approximate dividing of a huge space. But the rules in the most approximate partitioned space may not be simple rules. For this purpose, they proposed a rule-based algorithm to obtain the largest particles. The algorithm form divides the space in a hierarchical principle according to the condition and attribute domain to extract the maximum grain of the corresponding rule [15]. Li Lian et al., who first analyzed and pointed out some of the shortcomings of the traditional clustering algorithms, then combined granular computing and the principle of maximum and minimum distance to optimize the initial cluster centers. By dynamical adjusting of the upper and lower approximation and the boundary sets weighting factor they solve the problem of boundary data clustering. Using the distance between the classes and the inner distance equalization criterion function as algorithm termination conditions, they reduce the number of iterations of the clustering process and the sensitivity to noise [16].

Haiyan Li and his team pointed out that the traditional data cluster analysis made through the geometric distance, measuring the similarity, the clustering analysis method and traditional approach, described by the distance is too big in time consumption. Qualitative data are described by using information systems. The establishment of a mathematical model of qualitative data sets by granular computing based on the analysis of the variables determining the optimal variable subset of a clustering method is proposed. At that, not only the suitable cluster has an important classification structure data, but also a suitable majority of important data classification structures [17]. With the in-depth study of granular computing, the researchers analyzed granular computing and application development. Yao and Zhong [18] proposed a simple framework based on the rough set theory. They made a more detailed study of the relationship between the particle and the particle structure. Especially, they analyzed the granular computing application prospect in the field of knowledge discovery and data mining. Lin [19] summarizes the current status of the development of granular computing and points out the direction of development of granular computing, including the introduction of granular computing category development, the development of the grain structure, the concept of uncertainty of research, GrC, database and data mining, etc.

7. Development of GrC in the field of IT

Traffic congestion and environmental pollution can be seen as essentially contradictory between the people, vehicles, and roads. An intelligent transportation system is the basic approach to solve this contradiction. Leung, Faouzi and Kuriang [20] pointed out that the ITS (Intelligent Transportation System) is the future direction of development of the transport system. It is a revolution in transportation. Through advanced information technology, communication technology, control technology, sensor technology, computing technology and system integration technology and efficient integration and application, people, the interaction between the cars, the road, it is presenting a new way to achieve real time, accurate, efficient, safe, energy saving targets.
With the increase of the traffic data, the idea of granular computing provides the traffic problems with more obvious advantages. Gao and Liu [21] which granular computing is applied into the field of intelligent transportation, proposed a combined forecasting model, based on the rough set and knowledge entropy. The relative data model in prediction and a forecasting model and decision table are established between the conversion into discrete attribute values by means of continuous attribute values. The weight coefficient of the combination forecasting model is determined by an assessed rough set and knowledge entropy theory of each forecast model. The method proposed overcomes the limitations of the single prediction model and the determination of the weight coefficients is more objective. 

Path planning is one of the research aspects in the process of study of intelligent transportation. Granular computing is applied to the study of the large-scale network path by Zhang Ling. Combined with the social network layer and community structure characteristics to build the model of multiple granularity level, he has put the complex structure of large network mapping into different granularity spaces. The complexity of problem solving is reduced, the shortest path is mapped to a different size space, and the search process is set into a space from fine-grained to coarse-grained [22].

IT (Intelligent Transportation) as a wide range of terms includes urban transport, marine transport, rail and air transport and so on. Therefore, the use of granular computing knowledge to solve problems in intelligent transportation has become a significant research topic. Zhang Xiaodong et al. combine the characteristics of the rough sets and SVR and used them to predict the traffic flow. The RST (Rough Set Theory), GA (Genetic Algorithm) and SVR (Support Vector Regression) are combined to build a vessel traffic forecasting model RST-GA-SVR; the basic idea is RST model as a SVR pre-treatment system, the RST model is used to reduce the attributes of a vessel traffic flow [23]. Ming-Jun and Jian-Ye [24] applied the rough set theory to a railway traffic control command system. According to the advantages of the rough set in train operation control, the train operation adjustment model based on a rough set is established. The solving steps and the expression of the train operation adjustment of knowledge are presented.

Faced with the rapid development of IT, that generated a lot of traffic data, in order to analyze in time the traffic information data, there appeared an urgent need for a process, similar to human information processing to handle this amount of data. Li Jing et al. [25] in the study of the dynamic traffic data have introduced a new model for dynamic traffic data and a custom framework of granular computing which achieve a large number of traffic data automatic generalization and abstraction. The transportation system is strongly nonlinear, with strong randomness, large variability and uncertainty in large complex systems. How to analyze the complex systems has been a widespread concern among the researchers. Jiang Sunliang et al. [26], who introduced to the traffic flow the granular computing research in cellular automata, laid the foundation for the proposed in the future traffic system based on a cellular grain and transportation network three-tier structure. The paper will be combined with the idea of granular computing and the team will further improve the traffic flow simulation system architecture.
8. Conclusions

GrC, considered a wide methodology for looking upon the objective world, is abstraction of reality. Information granulating is a reflection of the human processing and information stored. GrC is important for the people solving the problem. The abstraction and dividing the complex problem in order to transform it into several relatively simple problems, help us to analyze and solve problems easily. But the generated data by the intelligent transport systems is fuzzy, complex and large-scale. The application of granular computing into intelligent transportation fields could solve the complex calculation problem of a large number of traffic data. Now, when the environment is becoming more and more serious, intelligent transportation is inclined to green commuting, that gives many solutions, such as using public bicycles, encouraging people to take the bus, subway and other public transport. Faced with such large amounts of data, how to solve the relationship between them, and how to make the trade-off between them in time in order to reduce the waste of the resources, and get the maximal efficiency, has become a key problem in the research of intelligent transportation.

Acknowledgments: The authors would like to thank the National Natural Science Foundation of China (No 61473108), the Public Projects of Zhejiang Province, Contract 2013C33082, Zhejiang Provincial Natural Science Foundation of China (No LY15F020038 and No LQ13F020005) and the research foundation of the Department of Education of Zhejiang Province (No Y201430884) and Hangzhou Dianzi University (No KYS105612008), which has made the present work.

References