Trend Relational Analysis and Grey-Fuzzy Clustering Method*

Zhijie Chen[1], Weizhen Chen[2,3], Qile Chen[4] and Mian-Yun Chen[2]

Abstract
In this paper, the latest researches on trend relational analysis and grey-fuzzy clustering method are presented. Grey systems and fuzzy systems are found everywhere. The researched results can open new prospects for the development and application of systems methodology to data mining.

Keywords Trend relational analysis, grey-fuzzy clustering method, grey systems, fuzzy systems.

1 Introduction

It is well known that the notion of a system is rather broad, and can be traced to antiquity. So to speak, any an object investigated, such as the motion of a macroscopic particle, or some socioeconomic phenomenon, may be qualified as a system. Systems such as social, economic, agricultural, industrial, ecological, and biological systems are usually those of great complexity. These complicated objects apparently have the following characteristic features:

- There is no physical prototype;
- The operation mechanism is not clear;
- The relationships between the inputs and the outputs are not obvious;
- The indeterminateness is very strong;
- Oft times, only a few of discrete data observed can be obtained.

Thus, in the studies of such objects, how to build a system model, how to forecast, how to make decision and how to control rely to a great extent on the use of information from objective reality. From a practical point of view, however, it is very difficult, or impossible, to get the adequate or complete information from investigated object in many situations. The phenomenon with incomplete information, therefore, is usually encountered. “Incomplete information” is the fundamental meaning of being “grey”. The name of “grey system” is chosen based on the amount of known information. Consider a “black box” stands for an object such that its internal structure is totally unknown to the investigator. Here, the word “black” represents unknown information, “white” for completely known information, and “grey” for those information which are partially known and partially unknown. Accordingly, systems with completely known information are called as white systems. Systems with completely unknown information as black systems, and the systems with partially known and partially unknown information as grey systems, respectively [1,2].

Since 1982, there has been a quick development in grey systems theory[3,4] in China, and it is also very successful in the application of the theory to many real projects, such as agriculture, society, economics, engineering, IT, data mining, management, biological

---

* This project was supported by the National Natural Science Foundation of China(69874018, 79970025), GF-PR-528.
[1] School of Economics and Management, Wuhan University of Science and Engineering, Wuhan 430073, China. Email:chen_zj@tom.com
[2] Department of Control Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China. Email:mychen@mail.hust.edu.cn. Email:chenwz_2005@126.com
[3] Department of Information and Electrical Engineering, Wuhan Polytechnic University, Wuhan 430023, China. Email:chenwzh@whpu.edu.cn
[4] Department of Computer Science, Wuhan University of Science and Engineering, Wuhan 430073, China. Email:cql@wuse.edu.cn

protection, ecology, environmental studies, etc.. After over 20 years of rapid development, the theory of grey systems consists of the following main blocks of concepts and results:

- Foundation, consisting of grey numbers, grey elements and grey relations;
- Grey systems analysis, including grey incidence analysis, grey statistics, grey clustering, etc.;
- Grey systems modeling, through the use of generation of grey numbers or functions so that hidden patterns can be found;
- Grey prediction;
- Grey decision making;
- Grey control;
- Grey process;
- Integral generating transform;
- Trend relational analysis; and
- Systems clouds.

The reader who takes an interest in this subject should refer to Kybernetes, Vol.33, No.2, 2004.(《Grey Systems Theory and Applications》,Guest Editors: Mian-Yun Chen, Sifeng Liu, and Yi Lin).

In this paper, the latest researches on trend relational analysis and grey-fuzzy clustering method are presented. The researched results here can open new prospects for the development and application of systems methodology to data mining.

2 On Grey Process

An excessively complex or complicated object, which generally shows a lack of completed model information, may be looked upon as a grey system. That is, with the aid of the grey systems approach we are able to solve the problems of the analysis and design of complicated systems, or excessively complex systems, including the data systems. Such a system might be looked upon as a data organizing framework according to which some data are considered to be relevant, others not. From a grey system's point of view, all the indeterminate or random concepts can be regarded to be grey. In order to describe an investigated object with incomplete information and to get a reasonably stable picture which can be communicated, some grey concepts are defined as follows:

- The most basic ingredient, which exists in a grey system, with incomplete information is called the grey element, denoted by $\otimes$;
- An indeterminate variable whose amplitude varies over a suitable range is called a grey variable, denoted by $X(\otimes)$;
- A function defined on the Cartesian product space $\otimes_\Sigma \times$T, denoted by $X(\otimes, t)$, is called a grey process, where $\otimes \in \otimes_\Sigma$ is a grey element, and $t \in T$ represents time;
- The observable output of an investigated object, which is a function of t, denoted by $X^{(0)}(t)$, is called a whitening function of grey process, $X^{(0)}(t) \in X(\otimes, t)$.

More often than not, only a few of the discrete data observed from the investigated object can be obtained, such as

$$X^{(0)} = \left\{X^{(0)}(k) \mid X^{(0)}(k) \geq 0, k = 1,2,\ldots,n\right\}$$

which is called the original time series corresponding to $X^{(0)}(t)$. We consider that, for an investigated object, all behavioral information is contained in $X(\otimes, t)$, and all relevant information through observation is contained in $X^{(0)}(t)$ or $X^{(0)}(k)$. Thus, $X^{(0)}(t)$ or $X^{(0)}(k)$ has provided the basis for constructing systems model[5].

3 Trend Relational Analysis for Grey Systems

Let us consider dynamic relationships between h factors that are present in an investigated object. Naturally, we can get

$$X_i^{(0)}(k), k = 1,2,\ldots,n, \qquad i = 1,2,\ldots,h.$$

Definition 3.1 *We call* $X_r^{(0)}$ *the reference factor (choose freely),* $X_c^{(0)}$ *the compared factor,* $r, c \in \{1,2,\ldots,h\}$. *Correspondingly,* $\{X_r^{(0)}(k)\}$ *is called the reference time series,* $\{X_c^{(0)}(k)\}$ *the compared time series.*

In order to express the approximateness and similarity between $\{X_r^{(0)}(k)\}$ and $\{X_c^{(0)}(k)\}$, we proceed to processing of data as follows:

Definition 3.2 *All of the following transformations are called to be mapping of quantity:*

$$M_r : X_r^{(0)} \times X_r^{(0)} \to \Delta X_{rr}^{(0)}$$

$$M_c : X_c^{(0)} \times X_c^{(0)} \to \Delta X_{cc}^{(0)}$$

$$M_{rc} : X_r^{(0)} \times X_c^{(0)} \to \Delta X_{rc}^{(0)}$$

where

$$X_r^{(0)} = \{X_r^{(0)}(1), X_r^{(0)}(2), \ldots, X_r^{(0)}(n)\}$$

$$X_c^{(0)} = \{X_c^{(0)}(1), X_c^{(0)}(2), \ldots, X_c^{(0)}(n)\}$$

$$\Delta X_{rr}^{(0)} = \{\Delta X_{rr}^{(0)}(2), \Delta X_{rr}^{(0)}(3), \ldots, \Delta X_{rr}^{(0)}(n)\}$$

$$\Delta X_{cc}^{(0)} = \{\Delta X_{cc}^{(0)}(2), \Delta X_{cc}^{(0)}(3), \ldots, \Delta X_{cc}^{(0)}(n)\}$$

$$\Delta X_{rc}^{(0)} = \{\Delta X_{rc}^{(0)}(1), \Delta X_{rc}^{(0)}(2), \ldots, \Delta X_{rc}^{(0)}(n)\}$$

$$\Delta X_{rr}^{(0)}(k) = X_r^{(0)}(k) - X_r^{(0)}(k-1)$$

$$\Delta X_{cc}^{(0)}(k) = X_c^{(0)}(k) - X_c^{(0)}(k-1)$$

$$\Delta X_{rc}^{(0)}(k) = X_r^{(0)}(k) - X_c^{(0)}(k).$$

Definition 3.3 *Let* $\underline{Mrc}$ *be a mapping. If*

$$\underline{Mrc} : \Delta X_{rc}^{(0)} \times \Delta X_{rr}^{(0)} \times \Delta X_{cc}^{(0)} \to \xi(k) \in [0,1],$$

then we call $\xi(k)$ *the trend relational function of both* $\{X_r^{(0)}(k)\}$ *and* $\{X_c^{(0)}(k)\}$.

Theorem 3.1 *Based on* $\{X_r^{(0)}(k)\}$ *and* $\{X_c^{(0)}(k)\}$, *if*

$$\xi_{rc}(k) = \xi_{rc}\left(\{X_r^{(0)}(k)\}, \{X_c^{(0)}(k)\}\right)$$
$$= \left[1 + \beta \,|\, \Delta X_{rc}^{(0)}(k) + \Delta X_{rc}^{(0)}(k-1) \,|\, + \gamma \,|\, \Delta X_{rr}^{(0)}(k) - \Delta X_{cc}^{(0)}(k) \,|\, \right]^{-1}$$
$$k = 2,3,\ldots,n,\ \beta, \gamma \in [0,1],$$

then $\xi_{rc}(k)$ *is a kind of trend relational functions.*

Proof. The minimal amount of information about $\{X_r^{(0)}(k)\}$ and $\{X_c^{(0)}(k)\}$ is involved in $\xi_{rc}(k)$. Thus, $\xi_{rc}(k)$ can express the dynamic relationship between $\{X_r^{(0)}(k)\}$ and $\{X_c^{(0)}(k)\}$ sufficiently, and meet the conditions in Definition 3.3. $|\Delta X_{rr}^{(0)}(k) - \Delta X_{cc}^{(0)}(k)| \Rightarrow$ similarity; $|\Delta X_{rc}^{(0)}(k) + \Delta X_{rc}^{(0)}(k-1)| \Rightarrow$ approximateness on certain similarity; $\{X_r^{(0)}(k)\} = \{X_c^{(0)}(k)\} \Rightarrow \xi_{rc}(k) = 1$ [6-11].

Corollary 3.1 *If* $\xi_{rc}(k)$ *=const,* $k = 2,3,\ldots,n$, *then* $\{X_c^{(0)}(k)\}$ *is completely similar to* $\{X_r^{(0)}(k)\}$.

Definition 3.4 $\xi_{rc}(k)$ *is a trend relational function between* $\{X_r^{(0)}(k)\}$ *and* $\{X_c^{(0)}(k)\}$, *then*