

Analysis of Time-Series Data Using the Rough Set

Yoshiyuki Matsumoto and Junzo Watada

Abstract Rough set theory was proposed by Z. Pawlak in 1982. This theory has high capability to mine knowledge based on decision rules from a database, a web base, a set and so on. The decision rule is widely used for data analysis as well. In this paper the decision rule is employed to reason for an unknown object. That is, the rough set theory is applied to analysis of economic time series data. An example shown in the paper indicates how to acquire knowledge from time series data. At the end we suggest its application to predictions.

1 Introduction

As changes in economic time-series data influence on the profits of a corporation, analyzes of such changes are widely pursued. Especially, technical and fundamental analyzes are employed as a method to analyze stock prices and dealing rates. The technical analysis is to analyze stock prices based on the time-series changes of stock prices through graphical expression of market prices. On the other hand, the fundamental analysis is to analyze stock prices based on various indices of corporate achievements and economical environments. As well, a chaotic method is also employed in forecasting of a stock [1].

The objective of this paper is to acquire knowledge from economical time-series data and forecast its change in terms of rough set theory [2, 3]. At the end, we will analyze real time-series values of TOPIX (Stock Market Index for the Tokyo Stock Exchange) and show what kind of knowledge acquisition and forecast will be done.

Y. Matsumoto (✉)

Shimonoseki City University, 2-1-1, Daigaku-Cho, Shimonoseki, Japan
e-mail: matsumoto@shimonoseki-cu.ac.jp

J. Watada

Waseda University, 2-7 Hibikino, Wakamatsu, Kitakyushu, Japan
e-mail: junzow@osb.att.ne.jp

2 Rough Set Theory

A rough set has been proposed by Z. Pawlak in 1982 [2] and is employed to analyze various applications widely [4]. It is possible to roughly express elements in a set of considered objects according to the recognizable scale. The rough set theory denotes such rough representation as approximation. This is a method of knowledge acquisition. There are two kinds of approximations: one is an upper approximation to take an element of a rough set into consideration from possibility points of view and the other is a lower approximation to take an element of a rough set from viewpoints of necessity. The visual illustration of upper and lower approximations is shown in Fig. 1.

It is named “reduction” to obtain a subset of minimal number of features that equivalently discriminate objects with all plural features that characterize some set. General speaking, there can exist plural reductions.

It is possible to decide a decision table, if features of a set can be divided into two subsets of condition features and decision features, respectively. The decision table can be understood as decision rules that correspond to a value of conditional feature to a value of decision feature. For instance, a decision table shown in Table 1 illustrates a decision rule for sample x1.

$$\text{If } a = 1 \text{ and } b = 1 \text{ and } c = 1 \text{ then } d = 1$$

This decision table has 3 conditional features and 5 samples. It is possible to derive 5 rules with 3 conditions. But the decision rules have redundancy in the conditional portion. By employing a reduction method in the rough set theory, it is possible to derive the minimal rules required for expressing the same decision rules.

Fig. 1 Upper and lower approximations

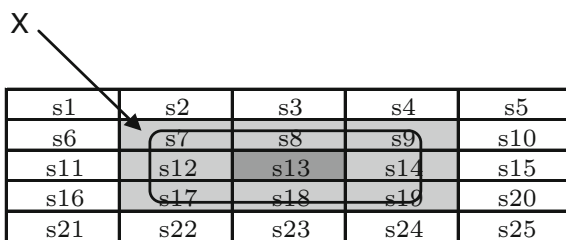


Table 1 Decision table

Sample	Conditional feature			Decision feature
	a	b	c	d
x1	1	1	1	1
x2	2	1	2	2
x3	2	1	1	1
x4	2	2	1	2
x5	1	1	2	1

In the case of a decision table shown in Table 1, rules illustrate the decision feature $d = 1$ as follows:

$$\text{If } a = 1 \text{ then } d = 1$$

$$\text{If } b = 1 \text{ and } c = 1 \text{ then } d = 1$$

As the same, the rule that illustrates the decision feature $d = 2$ can be written as follows:

$$\text{If } b = 2 \text{ then } d = 2$$

$$\text{If } a = 2 \text{ and } c = 2 \text{ then } d = 2$$

3 Determination of Decision Rules

It is required to build up a decision matrix in extracting decision rules from a decision table. For instance, the decision matrix for decision class $d = 1$ in Table 1 results in Table 2. This decision matrix is obtained using the lower approximation of decision class $d = 1$ and discriminate object $d = 2$.

The decision matrix is a table that describes feature value between samples. For example, on the case of $x1$ and $x2$, as $a = 1$ and $c = 1$ are deferent from $x1$, this value is denoted in the table. Therefore, the table explains that $a = 1$ or $c = 1$ can discriminate between $x1$ and $x2$. In the same way, $x4$ can be discriminated from $x1$ using $a = 1$ or $b = 1$.

$$\begin{aligned} x1: & (a1 \text{ or } c1) \text{ and } (a1 \text{ or } b1) \\ & = (a1) \text{ or } (a1 \text{ and } b1) \text{ or } (a1 \text{ and } c1) \text{ or } (b1 \text{ and } c1) \\ & = (a1) \text{ or } (b1 \text{ and } c1) \end{aligned}$$

In the same way, $x3$ and $x5$ can be described as follows:

$$\begin{aligned} x3: & (c1) \text{ and } (b1) = (b1 \text{ and } c1) \\ x5: & (a1) \text{ and } (a1 \text{ or } b1 \text{ or } c2) \\ & = (a1) \text{ or } (a1 \text{ and } b1) \text{ or } (a1 \text{ and } c2) \\ & = (a1) \end{aligned}$$

Table 2 Decision matrix

$A_*(d = 1) \setminus d = 2$	x2	x4
x1	a1, c1	a1, b1
x3	c1	b1
x5	a1	a1, b1, c2

As the feature results in $d = 1$ to discriminate x_1, x_3 and x_5 , we have a decision rule that x_1, x_3 or x_5 result in $d = 1$.

$$(a_1) \text{ or } (b_1 \text{ and } c_1) \text{ or } (b_1 \text{ and } c_1) \text{ or } (a_1) \\ = (a_1) \text{ or } (b_1 \text{ and } c_1)$$

On this case as shown in the previous section, the decision rule can be obtained as follows:

$$\text{If } a = 1 \text{ then } d = 1$$

$$\text{If } b = 1 \text{ and } c = 1 \text{ then } d = 1$$

It is possible to derive decision rules for decision class = 2 in the same way.

4 Analysis of Decision Rules

Only decision rules that are obtained rough set theory and have high C.I. are employed in reasoning. C.I. is an abbreviation of Covering Index that is a rate of objects that can sufficiently reach the same decision feature by the rule out of the whole objects [5].

Generally speaking, decision rules with high C.I. are highly reliable and results in good reasoning. In real situations, the number of obtained decision rules is often more than several hundreds. In these cases, reasoning does not employ almost all decision rules. That is, reasoning scattered almost decision rules.

It is necessary to make decision rules effective so as to combine decision rules by means of decision rule analysis [4]. Decision rule analysis enables us to obtain new combined decision rules by means that premises of decision rules are decomposed and given some points depending on their C.I. value. This method enables us to take all decision rules into consideration even if rules have a low C.I. value. In this paper, decision rules are combined and applied to forecasting.

Let us explain the detail of decision rule analysis. The decision rule analysis determines rules by calculating their column scores. The column score can be calculated in the following:

Let us consider the following three rules.

IF $a = 1$ and $b = 1$ then $d = 1$	(C.I. = 0.4)
IF $b = 2$ then $d = 1$	(C.I. = 0.3)
IF $a = 2$ and $b = 2$ and $c = 1$ then $d = 1$	(C.I. = 0.6)

The column score can be obtained using combination table as shown in Table 3. The combination table is an $n \times n$ matrix consisting of all features. The element of the combination table is a score of combination of two features.

Table 3 Combination table

	a = 1	a = 2	b = 1	b = 2	c = 1	c = 2	Column score
a = 1			0.2				0.2
a = 2				0.1	0.1		0.2
b = 1	0.2						0.2
b = 2		0.1		0.3	0.1		0.5
c = 1		0.1		0.1			0.2
c = 2							

For example, the first rule has a = 1 and b = 1 as its premises. On this case, the vertical column has a = 1 and the horizontal row has b = 1, and the vertical column has b = 1 and horizontal row has a = 1. We describe two scores in these elements. The score value is one or C.I. value divided by the written score value.

On this case, two elements have each score value.

$$0.4/2 = 0.2$$

On the case of the second rule, as the premise has one feature, the column and row are written 0.3 for b = 2.

On the case of the third rule, as the premise has 3 features, 6 elements (3C2 = 6) should be written scores. The written score is

$$0.6/6 = 0.1.$$

The column score is the total value of scores in each column. For example, on the case of a = 2 we obtain

$$0.1 + 0.1 = 0.2.$$

This calculation results in Table 3. Using this combination we can derive a decision table. For example, on the case of column b = 2, since there is a score in a = 2, b = 2 and c = 1, the rule of this column results in as follows:

$$\text{IF } a = 2 \text{ and } b = 2 \text{ and } c = 1 \text{ then } d = 1.$$

Usually, scores under the some threshold are not accepted. For instance, when the threshold is 0.2, the rule is written in the following:

$$\text{IF } b = 2 \text{ then } d = 1.$$

5 A Rough Set Approach to Analyzing Time-Series Data

In this paper, a rough set is applied to time-series data employing the focal time-series data and changes of related data that influence on the focal data.

Table 4 Only one time-series data

No	Conditional feature			Decision feature
	1 period previous	2 period previous	3 period previous	Present period
1	+	-	-	-
2	+	-	+	+
3	+	-	-	-
4	+	+	-	+
5	-	-	+	-

Table 5 Including related data

No.	Conditional feature						Decision feature
	Target data			Related data			Present period
	1PP	2PP	3PP	1PP	2PP	3PP	
1	+	-	-	-	-	-	-
2	+	+	-	+	+	+	+
3	+	-	-	+	-	-	-
4	+	-	+	-	-	+	+
5	-	+	-	-	+	+	-

PP: Period previous

General speaking, data treated in a rough set are categorical. In this paper, the change of the value is calculated from its single period previous value and two categories: plus and minus are defined by its going up or down changes, respectively. Such categorical data are analyzed by a rough set.

For instance, when the information of three past periods is analyzed, let us select going up or down movements from first to third periods for a conditional feature and the present change for a decision feature. That is, the present change is decided using the increasing and decreasing movement in the three past periods as shown in Table 4.

When employing other time-series data that may influence on the decision feature, such time-series data is additionally taken as a conditional feature as well and the present movement is decided depending on these features as shown in Table 5.

6 Analysis of TOPIX

The method described above is employed to analyze TOPIX time-series data. Dollar-Yen exchange rates, NY Dow-Johns Industrial Average of 30 stocks (DJIA) and NASDAQ Index are employed as a related time-series data. Let us forecast the changes of TOPIX based on the knowledge acquisition from these changes. The data employed is monthly values from 1995 to 2003. The first half 50 samples are

employed for knowledge acquisition and the latter half 50 samples are employed for verifying the model.

Increasing and decreasing movements in 6 periods (half a year) are employed in the knowledge acquisition. That is, these changes of increasing and decreasing movements from the first to sixth periods are taken for a conditional feature, the change of the present period is taken for a decision feature. Analysis was done for four combinations of the above-mentioned data as (1) TOPIX, (2) TOPIX and Dollar-yen exchange rates, (3) TOPIX and NY Dow-Johns Industrial Average, and (4) TOPIX and NASDAQ index. In the 1st case TOPIX is calculated changes from the first period to sixth period, and in other cases the other data as well as TOPIX are calculated these changes and taken for conditional features (Figs. 2, 3, 4, 5, Table 8).

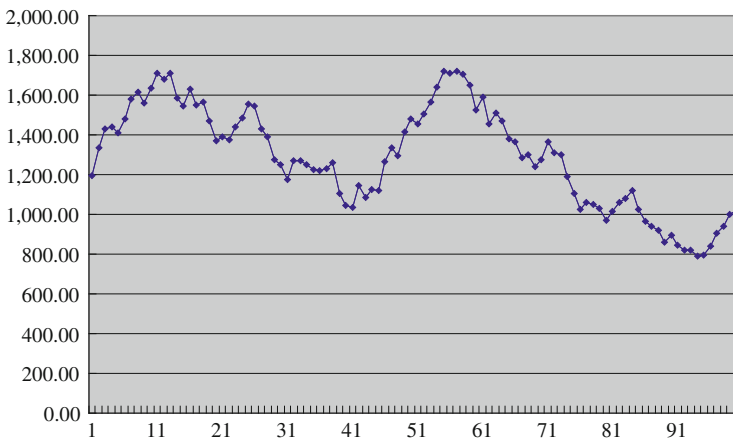


Fig. 2 TOPIX

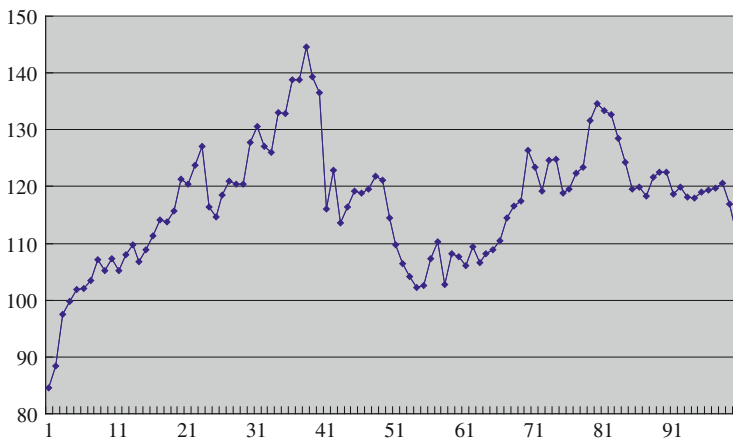


Fig. 3 Dollar-yen exchange rates

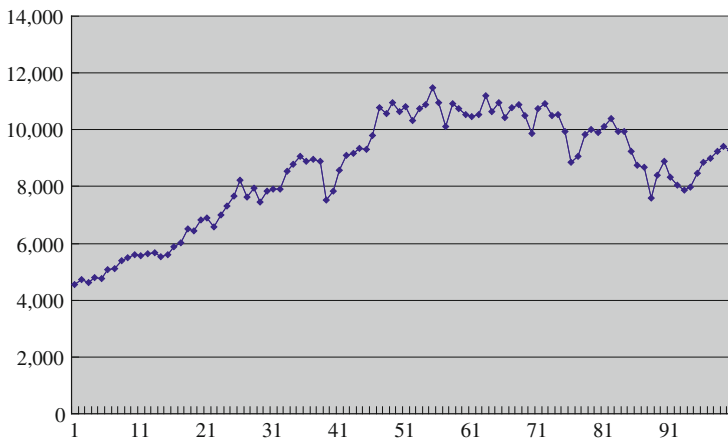


Fig. 4 NY Dow-Johns industrial average

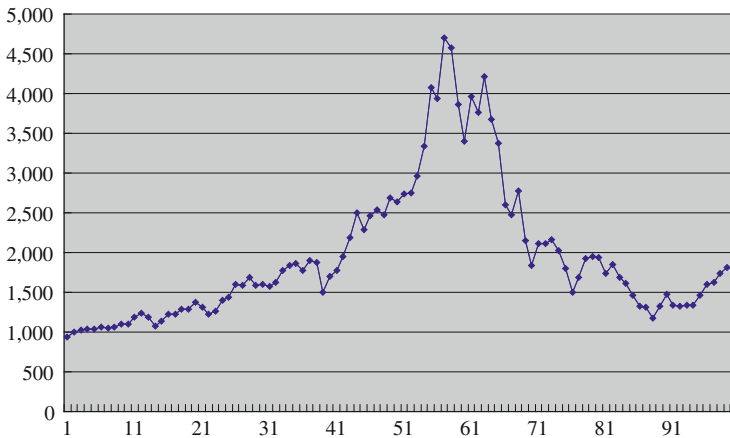


Fig. 5 NASDAQ index

Table 6 Forecast results

Using only TOPIX	52.85 %
Using TOPIX and Dollar-yen	53.80 %
Using TOPIX and DJIA	66.90 %
Using TOPIX and NASDAQ	62.50 %

Table 7 Forecast results
(Decision rule analysis method)

Using only TOPIX	44.40 %
Using TOPIX and Dollar-yen	51.30 %
Using TOPIX and DJIA	54.50 %
Using TOPIX and NASDAQ	53.30 %

Table 8 Conform rate

Using only top 3 rules	32.3 %
Using decision rule analysis method	77.3 %

7 Results

Table 6 illustrates forecasted results based on these rules. Using three top rules in a C.I. value, the last half 50 values are forecasted.

Regarding C.I. values of obtained rules, the rule obtained using related data is better than the one obtained only from TOPIX. This result shows that related data could acquire better rules that cover wider range. It was the rule of (-) movement based on TOPIX and Dollar-Yen Exchange Rates that showed the best C.I. value. It can cover 40 % of the whole range.

Regarding the forecasted results using rules obtained, it is better using related data than using only TOPIX times-series data.

Considering the result of all increasing and decreasing movements, the NY Dow-Johns Industrial Average is the best effect in forecasting among all combinations.

Table 7 shows the result obtained by forecasting using the decision rules acquainted from the decision rule analysis. It is frequent that the forecasting precision becomes worse than the result using the 3 rules of the highest C.I. values. Since decision rules with low a C.I. value are employed in forecasting, the forecasting precision should be worsened. Nevertheless, the number of objects that fit to obtained decision rules is larger on the case of the decision rule analysis. That is, even though the forecasting precision is worsened, the number of forecastable objects increases.

On the case where we use three rules with higher C.I. values, there are one third less objects fitting to rules than the number of the total 50 objects. On the other hand, it is about 80 % our of the whole objects that fit to rules obtained by decision rule analysis.

8 Concluding Remarks

In this paper, we proposed a method based on a rough set to analyze time-series data. As its application we analyzed TOPIX time-series data and forecasted future changes. As data related to TOPIX, Dollar-Yen Exchange Rate, NY Dow-Johns Industrial Average of 30 stocks and NASDAQ index are employed. For these data, decision rules are acquainted in terms of a rough set theory. Employing rules with higher C.I. values, the related data could obtain better results than TOPIX without any related data. The combination of TOPIX with NY Dow-Johns Industrial Average resulted totally in the highest precise forecasting.

Also, we forecast using rules obtained by decision rule analysis. Even if the forecasting precision was worse than on the case of using three rules with highest C.I. values, the number of objects that fit to rules is more than on the case of using C.I. values. Therefore, it is effective when we forecast data that are not fit to rules with high C.I. values.

On the other words, if we forecast using rules with higher C.I. values when objects are fit to such rules and using rules obtained by decision rule analysis for the other case, the forecasting can be compensated mutually.

In this application, we employed two categories of increasing and decreasing movements of the time-series data. If we will categorize more dementedly into several ones, it may be possible to obtain more knowledge. It should be also examined to obtain decision rules that cover whole states.

References

1. Matsumoto, Y., Watada, J.: Improvement of Chaotic Short-term Forecasting on Fuzzy Reasoning and Tuning on Genetic Algorithm. *J. Jpn Soc. Fuzzy Theor. Intell. Inform.* **16**, 44–52 (2004)
2. Pawlak, Z.: Rough Sets. *Int. J. Comput. Inf. Sci.* **11**, 341–356 (1982)
3. Tsumoto, S.: Rough sets: past, present and future. *J. Jpn Soc. Fuzzy Theor Syst.* **13**, 552–561 (2001)
4. Mori, N., Tanaka, H., Inoue, K.: Rough sets and Kansei: knowledge acquisition and reasoning from Kansei data. Kaibundo, Tokyo (2004)
5. Tanaka, H., Tsumoto, S.: Rough sets and expert system. *Math. Sci.* **378**, 76–83 (1994)
6. Watada, J., Li, H.: A rough set approach to building association rules and its applications. In: 3rd International Conference on Artificial Intelligence in Engineering and Technology, Kota Kinabaru, Malaysia, 22–24 Nov 2006