

## Rule Discovery and Matching in Stock Databases

You-min Ha

Department of Computer Science  
Yonsei University, Korea  
ymha@cs.yonsei.ac.kr

Sanghyun Park

Department of Computer Science  
Yonsei University, Korea  
sanghyun@cs.yonsei.ac.kr

Sang-Wook Kim

College of Information and Communications  
Hanyang University, Korea  
wook@hanyang.ac.kr

Jung-Im Won

College of Information and Communications  
Hanyang University, Korea  
jiwon@hanyang.ac.kr

Jee-Hee Yoon

Division of Information Engineering and Telecommunications  
Hallym University, Korea  
jhyoon@hallym.ac.kr

### Abstract

*This paper addresses an approach that recommends investment types to stock investors by discovering useful rules from past changing patterns of stock prices in databases. First, we define a new rule model for recommending stock investment types. For a frequent pattern of stock prices, if its subsequent stock prices are matched to a condition of an investor, the model recommends a corresponding investment type for this stock. The frequent pattern is regarded as a rule head, and the subsequent part a rule body. We observed that the conditions on rule bodies are quite different depending on dispositions of investors while rule heads are independent of characteristics of investors in most cases. With this observation, we propose a new method that discovers and stores only the rule heads rather than the whole rules in a rule discovery process. This allows investors to impose various conditions on rule bodies flexibly, and also improves the performance of a rule discovery process by reducing the number of rules to be discovered. For efficient discovery and matching of rules, we propose methods for discovering frequent patterns, constructing a frequent pattern base, and its indexing. We also suggest a method that finds the rules matched to a query from a frequent pattern base, and a method that recommends an investment type by using the rules. Finally, we verify the effectiveness and the efficiency of our approach through extensive experiments with real-life stock data.*

### 1. Introduction

Around us, there are a variety of objects such as stock prices, temperature values, and money exchange rates whose values change as time goes by. The list of changing values sampled at a fixed time interval is called *time-series data* [2, 13, 14, 16]. In many applications, an element value in time-series data is affected by its preceding values accumulated[7]. Thus, by analyzing past element values in time-series data, we can find the regularities and also build their model, thereby predicting the values to appear in the future.

Stock price sequences are a typical example of time-series data[3, 10]. Since the goal of stock investors is to maximize their earnings, it would help them achieve successful investments to recommend proper buying and selling points via analysis of the stock price sequences.

Time-series analysis[6] has been a well-known method for predicting stock prices in the future. It is classified into two categories: time domain analysis and frequency domain analysis. *Time domain analysis* is based on the regression model, which assumes that a current value is determined by the regression of its preceding values[6]. *Frequency domain analysis* is primarily used in analyzing stationary time-series data for predicting macroscopic tendencies of months, seasons, or years. However, they cannot reflect the conditions for investments, which could be dynamically changed by investors, and also have a problem of not being appropriate for short-term predictions.

In a machine learning perspective, there have been some