

# Storage Structures for Efficient Query Processing in a Stock Recommendation System

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

Seung-Hwan Lim

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

## Abstract

*Rule discovery is an operation that uncovers useful rules from a given database. By using the rule discovery process in a stock database, we can recommend buying and selling points to stock investors. In this paper, we discuss storage structures for efficient processing of queries in a system that recommends stock investment types. First, we propose five storage structures for efficient recommending of stock investments. Next, we discuss their characteristics, advantages, and disadvantages. Then, we verify their performances by extensive experiments with real-life stock data. The results show that the histogram-based structure performs best in query processing and improves the performance of other ones in orders of magnitude.*

## 1. Introduction

Time-series data is a list of changing values sampled at fixed time intervals [7]. It reflects the status changes of objects in nature and economy as time passes. In many applications, an element value in time-series data is significantly affected by its preceding values accumulated so far [4]. Thus, by analyzing past element values in time-series data, we can find the regularities and also form their model, thereby predicting the values to appear in the near future.

Stock price sequences are a typical example of time-series data [1, 3, 5]. Since the goal of stock investors is to earn high return, it would help investors achieve successful stock investments to recommend proper buying and selling points via analysis of the stock price sequences. Each

stock investor has his/her own conditions for buying or selling stocks. To meet requirements of a variety of stock investors, it would be so useful to develop a system that automatically recommends stock items whose price changing patterns come to satisfy the conditions required by individual investors.

In our previous work [6], we developed a system that recommends investment types to stock investors by discovering useful rules from past changing patterns of stock prices in a database. We proposed a new method that discovers and stores only the rule heads rather than the whole rules in a rule discovery process [6]. 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 proposed methods for discovering frequent patterns, constructing a frequent pattern base, and indexing those patterns. We also suggested a method that efficiently finds the rules matched to a query from a frequent pattern base, and proposed a method that recommends an investment type by using the rules.

In our stock investment system, there are a large number of stock investors, who issue queries on multiple stock items of interest. When a query is requested to process, stock prices related to the query are accessed from disk for recommending an investment type. This incurs a number of random disk accesses, thereby causing degradation of system performance. In this paper, we propose a variety of storage structures that reduce disk accesses and CPU computations, and then discuss their advantages and disadvantages. We also evaluate their performances via extensive experiments.