

A Framework for Stock Selection Using Association Rules on Combined Cash Flow and Accrual Financial Indicators

Amontep Wijitcharoen[□], Bunthit Watanapa[†], and Praisan Padungweang[‡]

School of Information Technology
King Mongkut's University of Technology Thonburi
Bangkok 10140 Thailand

Email: [□]amontep.wj@mail.kmutt.ac.th, [†]bunthit@sit.kmutt.ac.th, [‡]praisan.pad@sit.kmutt.ac.th,

Abstract—Traditionally, decision makers use accrual-based financial indicators, which are dependent on historical information, to evaluate stocks. This passive information can provide only partial intelligence to the decision makers. Additional cash-based insight would be needed for more complete information to identify any stock that could beat the stock market trend. In this study, we propose a systematic framework for financial information mining that actively considers the interaction between traditional financial ratios and cash flow signaling in identifying potential outperforming stocks in the coming period. Association detection method can help to ensure the mined rules be reliable and practical. To demonstrate the performance of the mining rules, we plan to collect the financial statement of selected stocks in the Service Industry from the Stock Exchange of Thailand during 2014 to 2016 for inducing a set of rules. The accuracy test of the model may then be performed based the latest actual outcome in 2017.

Keywords— Association rule, stock selection, cash flow, financial analysis

I. INTRODUCTION

Suitable stock selection plays an important role in financial portfolio construction. Selecting the right mixture of stocks at the right time could generate a satisfactory return to the investor. Over the recent decades, financial analysis is becoming a popular tool for analyzing stocks using financial indicators such as Return on Equity (ROE), Return on Assets (ROA), Price-to-Earnings ratio (P/E), Earnings per Share (EPS), and so on [1]. The drawback of using such accrual financial indicators only is the inability to detect cash-based signaling which complements to provide a fuller picture of a company's operation performance [2]. For example, when a company sold a product, in accrual accounting, it could mean the company already take a profit even though the customer still does not pay the money. As a complementary technique, cash flow (CF) has been used to give insight into the liquidity of the company in a given period of time [3,4]. In this study, we propose a framework to combine the traditional financial ratios and the cash flow measures to identify effective rules that can effectively suggest stocks in a selected industry that could generate a satisfactory return rate.

II. METHOD

The proposed framework is shown in Fig.1. It breaks down the relevant tasks into four sequential steps. First, a decision maker (DM) has to extract and transform all the financial indicators and cash flow into usable forms of codes or

measures for the rules mining purpose. For example, the price of stock would be transformed as % change in price or the return of stock. The available traditional financial indicators used in this study are as listed in Table I. They can be categorized into five performance measures of profitability, liquidity, leveraging, asset utilization and growth. Cash flow manipulated in this study are operating cash flow, investing cash flow and free cash flow. We shall exclude those stocks with incomplete data for ensuring the validity of the mining process.

Second, both financial ratios and cash flow are absolute numerical value for an individual firm. Thus, prior to the mining process, these measures must be encoded and normalized for compatibility according to the defined schemes shown in Table II. The encoding is required for meaningful interpretation of data, e.g. for the dependent data, the outperforming stock is coded as '1' if its return is larger than the % change of the market index, otherwise '0'. The normalization is to standardize the cash flow data so that the difference in size and scale of the companies is mitigated, e.g. the operating CF is considered in terms of % of revenue.

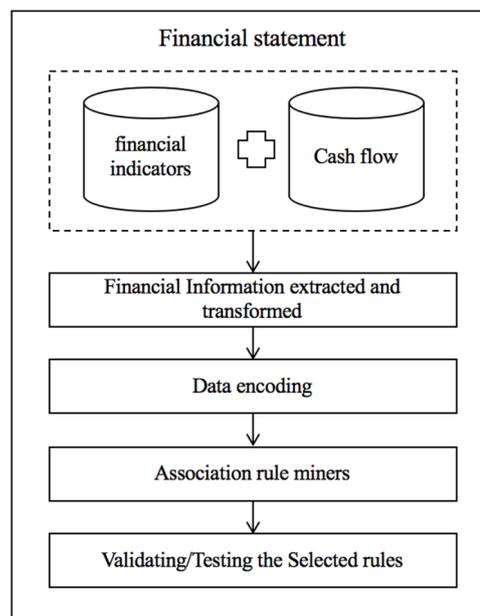


Fig.1 The proposed framework for stock portfolio selection.

Third, we use the coded values obtained from the previous step to identify the combinations of financial ratios and cash

flows which are associated with the outperformance of stocks. The desired level of reliability of the discovered rules should be heuristically identified by investors via the tuned values of Support, Confidence, and Lift parameters. [5].

Last, the validation of the rule should be performed by testing the usability of the obtained rule(s). A metric of accuracy when applying rules to the real and unseen set of the financial information should be reported.

TABLE I. LIST OF TRADITIONAL FINANCIAL RATIOS USED

Performance measures	Financial Ratios
Asset utilization	Receivable turnover
	Asset turnover
Liquidity	Current Ratio
	Quick Ratio
	Debt to equity ratio
Leverage	Leverage ratio
	Solvency Ratio-I
Profitability	Return on equity
	Return on assets
	Net profit margin
	Earnings per Share
Growth	Revenue growth rate

TABLE II. BINARY CODING INSTRUCTION FOR ALL VARIABLES: 1 INDICATES THE PRESENCE OF A CHARACTERISTIC AND 0 INDICATES THE ABSENCE.

Variables		Coding	
Dependent variable:			
1	Outperforming stock	If the stock's return at the end of the year greater than the percent change of Market index. coded as 1	If not: 0
Independent variables:			
Financial indicators			
1	Receivable turnover (RT)	If the RT greater than the average of the sector. coded as 1	If not: 0
2	Asset turnover (AT)	If the AT greater than the average of the sector. coded as 1	If not: 0
3	Current Ratio (CR)	If the CR greater than the average of the sector. coded as 1	If not: 0
4	Quick Ratio (QR)	If the QR greater than the average of the sector. coded as 1	If not: 0
5	Debt to equity ratio (DE)	If the DE is within the acceptable level (< +1 S.D.): coded as 1	If not: 0
6	Leverage ratio (LR)	If the LR greater than the average of the sector. coded as 1	If not: 0
7	Return on equity (ROE)	If the ROE greater than the average of the sector. coded as 1	If not: 0
8	Return on assets (ROA)	If the ROA greater than the average of the sector. coded as 1	If not: 0
9	Net profit margin (NP)	If the NP greater than the average of the sector. coded as 1	If not: 0
10	Price Earning (P/E)	If the P/E greater than the average of the sector. coded as 1	If not: 0
Cash flows			
1	Operating cash flow/Revenue (OCF/R)	If the OCF/R greater than the average of the sector. coded as 1	If not: 0

2	Investing cash flow/revenue (IFC/R)	If the ICF/R lower than the average of the sector. coded as 1	If not: 0
3	Free cash flow/Revenue (FCF/R)	If the FCF/R greater than the average of the sector. coded as 1	If not: 0

III. AN INSIGHT INTO EXPERIMENTAL DESIGN

For validating the model in this framework, the data set could be collected from the public data source such as Yahoo Finance, Morningstar, or The Stock Exchange of Thailand (SET). We plan to use the historical financial statements of the Service industry of SET during 2014 to 2016 for mining rules and will validate the obtained set of rules by selecting stocks in 2017.

The Service industry is one of the big three industry groups in the Thai stock market, consisting of five sectors namely Commerce, Health, Media and Publish, Tourism and Transportation. The comparative accuracy of the proposed system could be calculated from the number of the selected stocks that have an annual return greater than that of the SET.

IV. CONCLUSION AND FUTURE WORK

The integration of cash flow and the financial indicators can hypothetically fill the gap of using only accrual-based analytic information to select stocks, in order to outperform the average return of the stock market. A novel framework is proposed to mine these aforementioned information using Association rule method, so that investors may gain insight into the key features of a company's performance in a particular industry and generate effective rules for selecting stocks.

For future work, we plan to set up an experiment for discovering rules using Apriori algorithm using real data set acquired from the Service industry in SET. The concept of Apriori that prunes infrequent itemsets regarding the Support enables efficient mining and is extensible for dealing with Confidence and Lift [6]. The validation of the model in terms of rules accuracy will then be performed and the results will be disseminated accordingly.

REFERENCES

- [1] Beaver, W. H. Financial ratios as predictors of failure. *Journal of Accounting Research*, 71-111, 1966.
- [2] Barua, Suborna, and Anup Kumar Saha. "Traditional Ratios vs. Cash Flow based Ratios: Which One is Better Performance Indicator?." *Advances in Economics and Business* 3.6 (2015): 232-251.
- [3] Sloan, R., 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71, 289-315.
- [4] Lee, T. 'Laker Airways - The Cash Flow Truth', *Accountancy*, 115-116, 1982.
- [5] Agrawal, A., Mannila, H., Srikant, R., Toivonen, H. and Verkamo, A. Fast Discovery of Association Rules. In: Fayyad et al; 1996. p. 307-328.
- [6] Mandave, P., Mane, M. and Patil, S., 2013. Data mining using Association rule based on APRIORI algorithm and improved approach with illustration. *International Journal of Latest Trends in Engineering and Technology (IJLTET)*, ISSN.