

**Portfolio Optimization Using Fundamental Indicators Based
on Multi-Objective EA**

António Daniel Ventura Silva

Dissertação para obtenção do Grau de Mestre em
Engenharia e Gestão Industrial

Orientador: Prof. Rui Fuentecilla Maia Ferreira Neves

Co-orientador: Prof. Nuno Cavaco Gomes Horta

Júri

Presidente: Prof. José Rui de Matos Figueira

Orientador: Prof. Rui Fuentecilla Maia Ferreira Neves

Vogal: Prof. João Carlos da Cruz Lourenço

Março 2014

Resumo

Este trabalho apresenta uma nova abordagem para a composição de uma carteira de investimentos em acções. Será utilizada uma abordagem fundamental usando indicadores financeiros e técnicos com algoritmos evolutivos para otimizar dois objectivos (retorno e risco). O algoritmo escolhe uma a composição de uma carteira de acções. São usados dois cromossomas diferentes para representar os modelos de investimento com restrições reais, equivalentes as enfrentadas pelos gestores de fundos de investimento como hedge funds e fundos de pensão. Para validar a presente solução são estudados dois casos para o S&P 500 para o período de Junho de 2010 até o final de 2012. As simulações demonstram que a selecção de acções com base em indicadores financeiros é uma solução que pode ser usado para escolher as melhores empresas em termos operacionais com retornos acima da média do mercado e com baixas volatilidade. Neste estudo o otimizador encontra acções com alto retorno sobre o investimento feito escolhendo empresas com elevadas taxas de crescimento do lucro líquido e margem de lucro. Para obter acções com alto potencial de valorização, é necessário escolher as empresas com baixa ou média capitalização de mercado, baixo PER, com altas taxas de crescimento da receita e alta alavancagem operacional.

Palavras-chave: Múltiplos objectivos algoritmos evolutivos, Análise fundamental, análise técnica, demonstrações de resultados, rácios financeiros, Index S&P 500.

Abstract

This work presents a new approach to portfolio composition in the stock market. It incorporates a fundamental approach using financial ratios and technical indicators with a Multi-Objective Evolutionary Algorithms to choose the portfolio composition with two objectives the return and the risk. Two different chromosomes are used for representing different investment models with real constraints equivalents to the ones faced by managers of mutual funds, hedge funds, and pension funds. To validate the present solution two case studies are presented for the SP&500 for the period June 2010 until end of 2012. The simulations demonstrates that stock selection based on financial ratios is a combination that can be used to choose the best companies in operational terms, obtaining returns above the market average with low variances in their returns. In this case the optimizer found stocks with high return on investment in a conjunction with high rate of growth of the net income and a high profit margin. To obtain stocks with high valuation potential it is necessary to choose companies with a lower or average market capitalization, low PER, high rates of revenue growth and high operating leverage.

Keywords: Multi objective Evolutionary Algorithms, Fundamental Analysis, Technical Analysis, Financial Statements, Financial Ratios, S&P 500 Index.

Acknowledgements

I would like to express my deep and sincere gratitude to my supervisors, professor Rui Neves and professor Nuno Horta, for their guidance, constructive criticism and support in the development of this thesis. I need to said that wall support of my mother was the most important factor that give me the motivation in developing this thesis.

Table of Contents

| | |
|--|------|
| Resumo..... | I |
| Abstract..... | III |
| Acknowledgements..... | V |
| Table of Contents..... | VII |
| List of Figures..... | XI |
| List of Tables..... | XIII |
| 1. Introduction..... | 1 |
| 1.1 Motivation..... | 2 |
| 1.2 Approach..... | 2 |
| 1.3 Problems description..... | 2 |
| 1.4 Related Publications..... | 3 |
| 1.5 Document Structure..... | 3 |
| 2. Literature Review..... | 5 |
| 2.1 Value investing..... | 5 |
| 2.1.1 Fundamental analysis..... | 5 |
| 2.1.2 Financial Statements..... | 6 |
| 2.1.3 Fundamental indicators to use in ratio analysis..... | 8 |
| 2.1.4 Economic indicators..... | 10 |
| 2.1.5 Categories of Companies..... | 12 |
| 2.1.6 Warren Buffett method..... | 13 |
| 2.1.7 Approaches to Equity Valuation..... | 14 |
| 2.2 Technical analysis..... | 16 |
| 2.2.1 Trends in the market..... | 16 |
| 2.2.2 Strategies to invest..... | 17 |
| 2.2.3 Types of Market..... | 17 |
| 2.2.4 Designing a strategy to invest..... | 19 |
| 2.3 Introduction to Multi-objective Optimization..... | 20 |
| 2.3.1 Multi-objective Description..... | 21 |
| 2.4 Evolutionary Computation..... | 22 |
| 2.4.1 Memory module..... | 23 |
| 2.4.2 Selection module..... | 23 |
| 2.4.3 Reproduction model..... | 23 |
| 2.5 Evolutionary algorithms..... | 23 |
| 2.5.1 Representation..... | 24 |
| 2.5.2 Population..... | 24 |
| 2.5.3 Quality indicator..... | 24 |
| 2.5.4 Selection..... | 25 |
| 2.5.5 Variation or Reproduction operations..... | 25 |
| 2.5.6 Algorithm design issues..... | 25 |

| | | |
|-------|---|----|
| 2.5.7 | State of the Art Multi-Objective Evolutionary Algorithms..... | 26 |
| 2.6 | Portfolio Optimization | 27 |
| 2.6.1 | Diversification and Risk of Portfolios..... | 27 |
| 2.6.2 | Mean variance – Model..... | 28 |
| 2.6.3 | The mean-variance cardinality constrained portfolio optimization model (MVCCPO) ... | 28 |
| 2.7 | State of the Art of Portfolio Optimization | 29 |
| 2.8 | Conclusion | 32 |
| 3. | System Architecture | 35 |
| 3.1 | Algorithm architecture | 35 |
| 3.1.1 | Architecture of the Evolutionary Algorithms | 36 |
| 3.2 | Obtaining the Financial data..... | 38 |
| 3.3 | Financial Statement analysis | 39 |
| 3.3.1 | The income Statement | 39 |
| 3.3.2 | Balance Sheet Statement | 40 |
| 3.3.3 | Cash flow statement | 41 |
| 3.4 | Financial ratios | 42 |
| 3.4.1 | Debt ratio | 42 |
| 3.4.2 | Return on Equity..... | 43 |
| 3.4.3 | Profit margin..... | 44 |
| 3.4.4 | Price earnings ratio | 45 |
| 3.4.5 | Revenue growth rate (RG)..... | 46 |
| 3.4.6 | Rate of change in common stock outstanding..... | 47 |
| 3.4.7 | Net Income Growth rate | 48 |
| 3.5 | Technical analysis | 49 |
| 3.5.1 | Moving average | 49 |
| 3.6 | Conclusions | 50 |
| 4. | Multi-Objective optimization | 51 |
| 4.1 | Chromosome structure..... | 51 |
| 4.1.1 | First Model of the Chromosome..... | 51 |
| 4.1.2 | Second Model of Chromosome | 52 |
| 4.2 | Reproduction Process or Evolution Process | 52 |
| 4.2.1 | Crossover | 53 |
| 4.2.2 | Mutation | 55 |
| 4.3 | External file | 56 |
| 4.4 | Parameters of the trading system | 57 |
| 4.4.1 | Stop Loss..... | 57 |
| 4.4.2 | Take Profit | 57 |
| 4.4.3 | Position Size | 57 |
| 4.4.4 | Trigger Signal by SMA | 58 |
| 4.4.5 | Global Value input for portfolio..... | 58 |
| 4.4.6 | Stock Number in the Portfolio..... | 59 |
| 4.5 | Constraints..... | 59 |

| | | |
|-------|---|----|
| 4.6 | Simulator | 60 |
| 4.7 | Measures of Portfolio performance | 61 |
| 4.7.1 | Calculation of Return of Investment | 62 |
| 4.7.2 | Variance as measure of risk | 62 |
| 4.7.3 | Sharpe ratio..... | 63 |
| 4.7.4 | Pareto Dominance Concept | 64 |
| 4.8 | Conclusion | 65 |
| 5. | Results..... | 67 |
| 5.1 | Simulations and Results of the first model..... | 67 |
| 5.1.1 | Single objective using ROI..... | 67 |
| 5.1.2 | Multi-objective using the Pareto Dominance Concept..... | 68 |
| 5.1.3 | Multi-objective using the Pareto Dominance Concept with two years of training | 71 |
| 5.2 | Simulations and Results of the Second Model..... | 72 |
| 5.2.1 | Single objective using ROI..... | 72 |
| 5.2.2 | Single-objective using the Sharpe ratio as objective | 75 |
| 5.2.3 | Multi-objective using the Pareto Dominance Concept..... | 77 |
| 5.3 | Conclusions | 79 |
| 6. | Conclusions and Future Work..... | 81 |
| 6.1 | Conclusion | 81 |
| 6.2 | Future Work..... | 81 |
| 7. | References | |
| | Appendix A-Results Table | |
| | Appendix B-Results Table | |

List of Figures

| | |
|--|----|
| FIGURE 1 - MARKET WITH A STABLE AND QUIET VOLATILITY..... | 18 |
| FIGURE 2 - MARKET STABLE AND VOLATILE..... | 18 |
| FIGURE 3 -TRENDING AND QUIET | 19 |
| FIGURE 4 - TRENDING WITH HIGH VOLATILITY | 19 |
| FIGURE 5 - PARETO FRONT (NOILUBLAO & BUREERAT, 2009) | 22 |
| FIGURE 6 - GENERAL MODEL OF SEARCH HEURISTIC..... | 22 |
| FIGURE 7 - FLOW CHART OF A EA..... | 24 |
| FIGURE 8 - EFFICIENT FRONTIER..... | 27 |
| FIGURE 9 - SYSTEM ARCHITECTURE..... | 36 |
| FIGURE 10 - FLOW CHART OF ALGORITHM ARCHITECTURE..... | 37 |
| FIGURE 11 - FLOWCHART OF THE PROGRAM FOR GETTING THE DATA | 38 |
| FIGURE 12 - ANNUAL INCOME STATEMENT OF MICROSOFT | 39 |
| FIGURE 13 - ANNUAL BALANCE SHEET STATEMENT OF MICROSOFT | 40 |
| FIGURE 14 - ANNUAL CASH FLOW STATEMENT OF MICROSOFT | 41 |
| FIGURE 15 - COMPOUND RETURNS OF COMPANIES OF S&P500..... | 43 |
| FIGURE 16 - COMPARING PROFIT MARGIN AND REVENUE OF APPLE INC..... | 44 |
| FIGURE 17 - EPS OF THE MARKET S&P 500 VS INDEX PRICE | 46 |
| FIGURE 18 - REVENUES VS STOCK PRICE OF AMAZON..... | 47 |
| FIGURE 19 - COMMON SHARES OUTSTANDING OF MICROSOFT | 48 |
| FIGURE 20 - STOCK PRICE AND NET INCOME OF WAL-MART STORES..... | 49 |
| FIGURE 21 - TREND OF SP&500 WITH TWO SMA OF 200 AND 150 DAYS..... | 50 |
| FIGURE 22 - REPRESENTATION OF THE CHROMOSSOME..... | 52 |
| FIGURE 23 - CHOMOSSOME THAT REPRESENT THE FIRST MODEL..... | 52 |
| FIGURE 24 - CHROMOSSOME REPRESENTING THE SECOND MODEL..... | 52 |
| FIGURE 25 - CROSSOVER METHOD USING TWO CROMOSSOMES..... | 53 |
| FIGURE 26 - CROSSOVER METHOD WITH FOUR PARENTS A ONE CHILD | 54 |
| FIGURE 27 - MUTATION OPERATION WITH FOUR GENES MUTATED..... | 55 |
| FIGURE 28 - PROCESS DIAGRAM OF THE EXTERNAL FILE | 56 |
| FIGURE 29 - FLOW CHART FOR EVALUATE THE STOCKS | 59 |
| FIGURE 30 - FLOW CHART OF THE SIMULATOR OF TRADING | 61 |
| FIGURE 31 - DENSITY FUNCTION OF ABSOLUTE RETURNS OF APPLE AND INDEX SP&500... | 63 |
| FIGURE 32 - TREND OF 10 YEAR BONDS OF USA | 64 |
| FIGURE 33 - SINGLE OBJECTIVE RESULTS OBTAINED IN TRAINING USING ROI AS OBJECTIVE | 67 |
| FIGURE 34 - RESULTS OF REAL TEST IN SINGLE OBJECTIVE | 68 |
| FIGURE 35 - MO RESULTS OBTAINED IN TRAINING USING THE PARETO DOMINANCE | 69 |
| FIGURE 36 - RESULTS OBTAINED IN REAL TEST | 69 |
| FIGURE 37 - AVERAGE RETURN OF THE MOEA | 70 |
| FIGURE 38 - PARETO CURVE OBTAINED IN TRAINING USING THE DOMINANCE OF PARETO | 71 |
| FIGURE 39 - PARETO CURVE OBTAINED IN REAL TEST | 72 |
| FIGURE 40 - RESULTS OBTAINED IN TRAINING USING ROI AS OBJECTIVE..... | 73 |
| FIGURE 41 - RESULTS OF REAL TEST | 73 |
| FIGURE 42 - SHARPE RATIO TRAINING RESULTS..... | 75 |
| FIGURE 43 - RESULTS OBTAINED IN THE REAL TEST FOR THE SHARPE RATIO OBJECTIVE | 76 |
| FIGURE 44 -GRAPHIC 3D OF REAL TEST FOR THE SHARPE RATIO OBJECTIVE | 76 |
| FIGURE 45 - PARETO CURVE OBTAINED IN TRAINING USING THE DOMINANCE OF PARETO | 77 |
| FIGURE 46 - RESULTS OBTAINED IN THE REAL TEST | 78 |

List of Tables

| | |
|--|----|
| TABLE 1 - PROFITABILITY RATIOS..... | 8 |
| TABLE 2 - LIQUIDITY RATIOS..... | 9 |
| TABLE 3 - LEVERAGE RATIOS..... | 9 |
| TABLE 4 - EFFICIENCY RATIO..... | 10 |
| TABLE 5 - MARKET VALUE RATIOS..... | 10 |
| TABLE 6 - OVERVIEW APPROACHES TO PORTFOLIO OPTIMIZATION..... | 32 |
| TABLE 7 - DEBT - RATIOS..... | 42 |
| TABLE 8 - AVERAGE 5 YEARS OF ROE..... | 44 |
| TABLE 9 - RESULTS OF THE REAL TEST..... | 70 |
| TABLE 10 - GENES OF THE CHROMOSOMES..... | 71 |
| TABLE 11 - RESULTS OF THE REAL TEST..... | 74 |
| TABLE 12 - GENES OF THE CHROMOSOMES..... | 74 |
| TABLE 13 - RESULTS OF THE REAL TEST..... | 76 |
| TABLE 14 - GENES OF THE CHROMOSOMES..... | 77 |
| TABLE 15 - RESULTS OF THE REAL TEST..... | 78 |
| TABLE 16 - GENES OF THE CHROMOSOMES..... | 78 |
| TABLE 17 - TRADES OF CHROMOSOME 36..... | 87 |
| TABLE 18 - TRADES OF CHROMOSOME 51..... | 88 |

List of Acronyms and Abbreviations

Optimization and Computer Engineering Related

- ACO - Ant Colony Optimization
- AI – Artificial Intelligence
- ATI –Automatic Trading Systems
- EA – Evolutionary Algorithm
- GA – Genetic Algorithm
- MO – Multi-objective Optimization
- SA – Simulated Annealing
- PS – Pareto Set
- PSO - Particle Swarm Optimization
- TS – Tabu Search

Investment Related

- ACP – Average Collection Period
- APP –Average Payment Period
- B&H – Buy-and-hold
- CR – Current Ratio
- CAPM –Capital Asset Price Model
- DCF – Discount Cash Flow
- DER – Debt Equity Ratio
- DY – Dividend Yield
- EBIT- Earnings Before Interests and Taxes
- EMH – Efficient Market Hypothesis
- EPS – Earnings Per Share
- FA – Fundamental Analysis
- GDP – Gross Domestic Product
- HFT – High Frequency Trading
- HSI – Hang Seng Index
- MA – Moving Average
- MC – Market Capitalization
- NYSE – New York Stock Exchange
- PBV – Price Book Value
- PCF – Price Cash Flow
- PER – Price Earnings Ratio
- POR – Pay Out Ratio

- QR – Quick Ratio
- ROA – Return On Assets
- ROE – Return On Equity
- S&P 500 – Standard and Poor’s 500 stock index
- TA –Technical Analysis

1. Introduction

The stock exchange is the physical place where equity shares of companies listed in that exchange trade using manual or advanced automatic trading system (ATS). The largest and more known exchange is the New York Stock Exchange (NYSE) where are registered and traded thousands of USA companies. The evolution of the internet and the globalization allow trading stocks through internet using a broker, with a commission and low slippage¹. The possibility to know the transaction price of the stocks and volumes in real time² is fundamental to trade any common stock of the NYSE or in other exchanges, because allow the trader to perform market analysis and know the execution price of his orders.

The advances in computer science, like the increase of the power of processing of the CPU, and the memory capacities and hardware, led to large increase of the available data on financial markets. This allows the researchers to apply a new set of tools, the evolutionary algorithms (EA), artificial intelligent (AI), Fuzzy logic and machine learning, to solve a variety of problems, such as portfolio optimization and prediction of price movements of the financial instruments (Kaucic, 2012).

Investors use these applications to create trading programs to execute trades without human intervention, in different time periods, that can be annual, weekly (swing trading), daily (day trading) and less than one hour (high-Frequency trading). There are other types of programs developed to explore arbitrage opportunities and for portfolio selection. These systems are back tested in historical or current market to check the possible performance that can achieve, and improve the trading system before put in use. The main ideas to invest in the stock market are technical analysis, fundamental analysis and the random walk theory. Technical analysis studies the market patterns and demand and supply of stocks shares. Fundamental analysis studies a stock, from economic and financial point, where tries to define if a company is undervalued in the market in relation to intrinsic value, calculated by the projections of future earnings of the company. Random walk theory defends that the market discounts all future developments so that the investor cannot expect to outperform the market (Malkiel, 1999).

In the USA stock market (S&P 500, Dow Jones, Nasdaq 100), participants annually transact trillions of dollars, they are short-term traders, scalpers, speculators, companies doing hedge to reduce their risk, and the value investors (Faith, 2007). Short-term traders invest in a period lower than a week, based their investment decisions in technical analysis. Scalpers are traders that use technical analysis for win a little spread between bids and asks, in trades with a period of time inferior 5 minutes. Speculators use time frames higher than a week, normally months, to trade using technical analysis and fundamental analysis. Value investors use fundamental analysis to evaluate a company and buy a share below its intrinsic value.

¹ The difference of expected price of the trade and real executed price

² With a low delay of communication time

This thesis relies mostly on the value investor ideas from Warren Buffett, famous investor who underlines that a diversified portfolio of common stocks, each of them purchased below its intrinsic value, will generate a return above the general market.

To constitute a portfolio a manager needs to take a set of decisions to solve the portfolio optimization problem, distribute his capital in a number of assets to maximize the expected return and minimize the risk. To solve this problem it is necessary to consider the correlation of the assets, budget constraints and the investor's preference. This type of problems are complex and practically impossible to solve by exact methods, and researchers use heuristics like local search (LS), tabu search (TS), simulated annealing (SA), particle swarm optimization (PSA), ant colony optimization (ACO), and evolutionary algorithms (EA) to solve them.

EA are based in the process of biologic evolution, the solutions of the problem are encoded in chromosomes, and tested using a fitness function. The previous best solutions are used to generate new solutions, by repeating this process the algorithm is capable of reach close to the global optimal solution of the problem.

1.1 Motivation

The competition among investors is high and many techniques used to invest in financial instruments do not produce outstanding results. This results in a motivation and the challenge to implement an uncommon technique that generates better results than the stock market Index S&P 500. Another reason is to test the efficient market hypothesis theory (stock price reflects all information available and this implies that it is impossible to beat the index with an active management approach³), and try to find and explore some inefficiency in the market valuations.

One important aspect of any investment system is its adjustment where the performance obtained in training and real test are similar. Financial markets are uncertain and dynamic, change very quickly, so the goal is to develop a robust system to invest that can outperform the index, and a strategy of buy and hold, in the most adverse environments that a stock market can offer.

1.2 Approach

The selection of the best companies to invest is probably the most important factor to the success of any investor of common stocks. In this research it will be investigated the effectiveness of computer intelligent systems to invest, using the value investing approach. EA will be used to select a business to invest that have the principles of Warren Buffet, using fundamental analysis of the financial statements (Income Statement, Balance Sheet, Cash flow Statement), and ratio analysis.

1.3 Problems description

The use of a fundamental approach based in financial statements analysis to select the stocks that will have in the future a better performance in the market, is a difficult task because most of the fundamental information are discounted in the valuation done by the market.

³ Portfolio management strategy where the goal is to outperform an investment benchmark

Using the EA in a training environment for posterior use in real market situation, led to the most difficult problem to deal, the adjustment of the solutions, namely the performance that the system will achieve in others markets and conditions.

It is expected that a robust investing system will be independent of the conditions of the markets and it will have a low variance of the different performances. To guarantee more adjustment is essential to use a large sample of data in the training, and avoid the over-fitting⁴ effect.

Other problem to deal, when using some EA as search mechanisms, is the classical selection process that can lead to the algorithm get trapped in a local optimum and do not reach the global optimum, which causes a great loss of performance in the trading system.

1.4 Related Publications

During the development of this work, the following paper has been accepted for publication which demonstrated the effort under development: Silva, A., Neves, R., Horta, N. "Portfolio Optimization Using Fundamental Indicators Based on Multi-Objective EA", CIFEr 2014.

1.5 Document Structure

This thesis is structured as presented next:

- Chapter 2 presents the necessary theory and methodologies used to perform this investigation.
- Chapter 3 presents the architecture of the system, and the methods used.
- Chapter 4 presents the methodologies used in the EA and the developed algorithms.
- Chapter 5 summarizes the experiences done and the results obtained.
- Chapter 6 is the conclusions of the thesis and the improvement to done in the future work.
- The appendix are the recorded investments executed of two strategies selected

⁴ A modeling error which occurs when a function is too closely fit to a limited set of data points

2. Literature Review

This chapter presents background information and reviews the existing literature that is relevant to the development of this project. In the first part of the chapter is presented a brief description of the existing approaches to invest, in section 2.1 will be described in depth the “*Value Investing*” and the instruments and methods used in it. In section 2.2 is presented the technical analysis, the strategies to use and the tools to analyze the market. A formal definition of multi-objective optimization problems and its concepts are given in section 2.3. A description and implantation of Evolutionary Algorithms are presented in sections 2.4 and 2.5 before the review of the existing literature of the MOEA’s for portfolio management in sections 2.6 and 2.7.

2.1 Value investing

It is a comprehensive investment philosophy to perform in-depth fundamental analysis to limit risk, resist the crowd psychology⁵ and achieve long-term investment results. It is the practice of purchasing securities or assets for less than its intrinsic value (Graham & Dodd, 1934).

Value investors seek stocks of companies that trade below their estimations of intrinsic value in the market, they act based on the theory that markets overreacts to good and bad news resulting in stock price movements that do not correspond with the company’s long-term fundamentals. It is present next the analysis of financial statements, the use of economic indicators, the fundamental ratios, the method of investment of Warren Buffett, and the approaches to evaluate the stock price.

2.1.1 Fundamental analysis

A company’s stock can be undervalue because of a disappointing earnings report, a restructuring of the company, a lawsuit, if the company fails its strategic or financial objectives, bear market tendency of the general market that provoke short-term trend in the price of the stock.

Fundamental analysis entails rigorous evaluation of the company’s fundamentals as industry growth, the capacity of the competitors, the future global economic factors for its goods and service, and macroeconomic conditions that affect the company.

The work of fundamental analyst for evaluate a security entails the study of the financial statements, and industrial analysis, and ratio analysis. The analyst needs to be always looking to find information about corporate actions, like restructuring, spin-offs⁶, merger and acquisitions, to get a better estimation of the value.

⁵ Members of the crowd often adapt and act to the expectations of the surrounding culture and modify individual traits in order to identify with the crowd

⁶The creation of an independent company, through the sale of part of the company

2.1.2 Financial Statements

Financial statements, are reports issued by companies to demonstrate their financial performance, normally trimestral or annually. The objective of the financial statements is to provide information about the financial position, performance and changes in financial position of the company. These reports are composed by an income statement, balance sheet and cash flow statement and are used by the value investors to do quantitative analysis on the company.

a) **Income Statement**

Income statement measures the performance of the company for a specific period of time, shows how much money the company operations generated (Revenues/sales), the operating expenses occurred during the period, the operating profit (EBIT), the financial costs (interest and taxes), and the net income (Hillier, Ross, Randolph, Jeffrey, & Jordan, 2010).

Revenue or sales represents the amount of money a company earns through the sale of a good or service during a specific time period, sometimes managers break down revenue by business segment or geography. The best way for the company to improve its profitability is by increasing its Revenues (Buffett & Clark, 2008).

- **Competitive advantage in the Income statement**

A company has a competitive advantage when have some power in the market that others rivals not have, normally this are monopolistic or oligopoly companies. They can price its products much higher than their marginal costs, obtaining higher operating margins (Pindyck & Rubinfeld, 2006). In a case of increase of costs of production it can be transferred to the final price for the consumer, if this is not possible the company does not have the competitive advantage and the increase in costs will generate lower profit margins.

Companies with gross profit margin above 40% are more likely to have a competitive advantage and below are probably in a highly competitive industries (Buffett & Clark, 2008).

In case of a competitive advantage is the result of some technological advancement, there is always a threat that a newer technology will replace it. Companies that its business is in the area of technology and spends a lot in Research & Development, like 30% of its gross profit or more, means that probably are inserted in very competitive industry and maybe not a good company to invest (Buffett & Clark, 2008).

- **Interest Coverage**

Interest expenses represent the money out of the business to pay the interests of the debts that the company has in the balance sheet, it is a financial cost that depends of the level of leverage of the company. If the company has a low *Times Interest Earned ratio (TIE)* is highly riskier to invest, because an unexpected event that reduces the *EBIT* of the company can put it in position of not fulfil the payment of its obligations.

$$TIE = \frac{EBIT}{Annual\ Interest\ Expense} \quad (1)$$

b) Balance sheet

Balance Sheet is the instantaneous reports of the financial condition of the company on a particular date, are generated normally in the end of a trimester, or year. This report states what the company owns and how it is financed, it is divided into three parts, *assets*, *liabilities* and *total equity* with an accounting relationship given by equation 2 (Hillier, Ross, Randolph, Jeffrey, & Jordan, 2010) (Robinson, Greuning, Henry, & Broihahn, 2009).

$$\text{Assets} = \text{Total Equity} + \text{Liabilities} \quad (2)$$

Using the report format under U.S. generally accepted accounting principles (GAAP) the assets are registered in the *balance sheet* in order of liquidity, in the top are the assets more liquid (the current assets), in the bottom the less liquid assets.

The part of *assets* are divided in *current assets* (assets that can be converted in cash in a short period of time), and *not current assets* that cannot be converted in money so quickly.

The firsts assets to appear are the *cash and short-term investments*, if exist a record of accumulation of money year after years, it is a good signal for the investor, in terms of profitability and stability of the company operations (Buffett & Clark, 2008).

Liabilities are divided in two parts, the *short term liabilities*, money that company owed and need to pay in one year or less time, this include *account payables*, *accrued expensive*, and *short term debt*, the other part is the *Long term liabilities* that come due in more than one year.

In industries that have a fiercely competitive environment, the companies need to constantly upgrade their manufacturing facilities to stay competitive. The acquisition of assets creates ongoing expenses that are recorded in *property/plant/equipment* and according with U.S. GAAP is reported in the statement at the cost of acquisition less accumulated depreciation.

The recent year's assets acquisitions, demonstrates the necessity of upgrade of the operations of the company and this helps to differentiate the companies with competitive advantage from the others.

Companies with higher profits margin have a ratio of *property/plant/equipment* to *total debt* higher than the others, meaning lower operations leverage. The high value of fixed assets represent a barrier to competitors to enter in is sector, because need to do a great investment (Buffett & Clark, 2008).

Companies with durable competitive advantage require little or no long-term debt to maintain their operations, it is necessary analyze more than one year of the financial statements to see if the operations work with low levels of debt. Can occurred the case of the company generate so much profits that decide to take advantage of tax shield generated by the use of a higher level of debt, in this cases is normal that net earnings can pay off long-term debt in three or four.

Common stocks represent ownership of the company, gives the right to elect a board of directors and to receive dividends. Preferred stock is other class of equity that does not give the right to vote, but has the right to receive a dividend before common stocks dividends (Hillier, Ross, Randolph, Jeffrey, & Jordan, 2010).

Retained earnings, are profits that the company retained to invest in the operations of the company, or in others investments considered profitable. The retention ratio defines the percentage of profits that are not distributed as dividends and are retained (Hillier, Ross, Randolph, Jeffrey, & Jordan, 2010).

c) Cash Flow Statement

Since the accrual Method allows credit sales to be booked as revenue in the income statement, it is necessary to keep separate track of the actual cash that flows in and out of business, this is shown in the cash flow statement.

This statement reports only how much money entering (cash inflow) from on-going operations and external investment sources, and how much money goes out of the company (cash outflow) to pay the business activities, investments, and financial expenses, during the period.

Companies with durable competitive advantage working in their favor generate a great quantity of cash inflow from operations that can be applied by the managers to buy back the company shares reducing the number of outstanding shares. These actions are taken because the managers prefer finance the company with debt, because is more cheap, second are trying to manipulate the price of stock in the market, or the manager board thinks the company is undervalue in the market and by repurchasing their own shares increase the wealth of the shareholders. This can be noted in the statement in the section of *cash from investing activities*, by analyze the reports of lasts years (Buffett & Clark, 2008) (Hillier, Ross, Randolph, Jeffrey, & Jordan, 2010).

2.1.3 Fundamental indicators to use in ratio analysis

Ratio analysis is used to conduct quantitative analysis of the information of financial statements, it uses the ratios calculated from current year and compared with ratios of previous periods to check the performance of the company and to select the best company in an industry. Next in the tables are some of financial ratios used by investors.

| Indicator | Profitability Ratios | Description |
|-------------------------------|---|---|
| Return on Equity (ROE) | $\frac{\text{Net income}}{\text{Shareholder's Equity}}$ | ROE measures corporation's profitability, by revealing how much profit a company generates with the money shareholders have invested. |
| Return on Assets (ROA) | $\frac{\text{Net income}}{\text{Total Assets}}$ | ROA measures efficiency of management is using the assets to generate earnings. |

Table 1 - Profitability Ratios

| Indicator | Liquidity Ratios | Description |
|---------------------------|---|---|
| Current Ratio (CR) | $\frac{\text{Current Assets}}{\text{Current Liabilities}}$ | Measures the company's ability to pay short term obligations. |
| Quick Ratio (QR) | $\frac{\text{Current Assets} - \text{Inventories}}{\text{Current Liabilities}}$ | Measures the company's ability to pay short term obligations with its most liquid assets. |

Table 2 - Liquidity Ratios

| Indicator | Leverage Ratios | Description |
|---|---|---|
| Debt Equity Ratio (DER) | $\frac{\text{Total Liabilities}}{\text{Shareholders Equity}}$ | Measures of financial leverage. |
| Debt Ratio (DR) | $\frac{\text{Total Liabilities}}{\text{Total Assets}}$ | Indicates the leverage of the company along with the potential risks the company faces in terms of its debt-load. |
| Times Interest Earned Ratio (TIER) | $\frac{\text{EBIT}}{\text{Annual Interest Expense}}$ | A metric used to measure a company's ability to meet its debt obligations. |

Table 3 - Leverage Ratios

| Indicator | Efficiency Ratios | Description |
|--|---|---|
| Average Payment Period (APP) | $\frac{\text{Account Payable}}{\text{Sales}} \times 12$ | Average period taken by the company to pay to its creditors. |
| Average Collection Period (ACP) | $\frac{\text{Account Receivables}}{\text{Sales}} \times 12$ | Average period for the company receive its payments from its clients. |

Table 4 - Efficiency Ratio

| Indicator | Market Value ratios | Description |
|---------------------------------|---|---|
| Earnings Per Share (EPS) | $\frac{\text{Net income} - \text{Dividends on preferred stock}}{\text{Average Outstanding Shares}}$ | The portion of a company's profit, by each outstanding share of common stock. |
| Payout Ratio (POR) | $\frac{\text{Dividends per Share}}{\text{Earnings per Share}}$ | This ratio shows the percentage of the earnings that are distributed to shareholders as dividends. |
| Dividend Yield (DY) | $\frac{\text{Annual Dividends per Share}}{\text{Share Price}}$ | A financial ratio that shows how much a company pays out in dividends each year, relative to its share price. |

Table 5 - Market Value Ratios

2.1.4 Economic indicators

Economic indicators are separated in macroeconomic indicators to effectuate macroeconomic analysis to forecast the cycle of the economy, and the industry indicators to analyze the industries.

I. Macroeconomic indicators

Macroeconomic factors have a huge impact on the general market tendency and influence the economic fundamentals of the companies. It is important to knowing how to interpret and analyse the indicators presented next.

- **The Gross domestic product (GDP)** is a measure of the economy's total production of goods and services. Growing in GDP indicates an expansion of economy with the opportunity for companies to increase their sales and services.
- **Balance of Trade** is the difference between a country's imports and its exports. A positive balance, attract more capital (external and internal) for investments. Companies in these countries are more competitive internationally, and are more capable of increasing their operations and sales.
- **Inflation Rate** is the rate at which the general level of prices rises. High rates are associated with economies with demand for goods and services where are outstripping productive capacity, which leads to upward pressure on prices. (Bodie, Kane, & Marcus, 2010)
- **Sentiment of Consumers and Producers** have important impact in performance of economy, a optimism sentiment in relation to the future causes a behaviour to consume more products and services in the consumers and the producers are willing take more risks trough investments and increases of production.
- **Interest Rate**, are the cost of money, high rates decreases the present value of future cash flows, meaning that companies and investors will see decreasing the attractiveness of its investments.
- **Monetary expansion**, depending if it is anticipated by the financial markets, can led to an increase of stock price, because that expansion means lower interest rates for some time and higher output of the economy (Blanchard, 2006).

II. Industrial indicators

The performance of any company depends in great part from the economic future of his industry, is used industrial indicators for perform sector analysis to have a clear picture of the actual state and future of the sector (Bodie, Kane, & Marcus, 2010).

- **The number of customers** is important to evaluate the risk of the company, if total revenue was achieved by few customers or some millions is an important factor to evaluate the risk. A loss of one client when the enterprise has few clients or large portion of the revenue is generated by a couple of them is a situation more risky than when there are a large number of clients, and each of them without meaningful contribution for the total revenue .
- **Market share** gives some important information about the company and the business. A company that has more than 70% of the market suggests that probably has a competitive advantage like a barrier to enter in the market, or economies of scale.
- **Industrial growth** is a reference for estimation the growth rate of any company that composes the group. In an industry growing fast, companies will follow this growth and the best competitors can have higher growth rate than the industry.

- **The number of competitors** gives information about the competitive environment of the industry, when exist lower number of barriers to entry and a large number of competing firms creates a difficult operating environment for generating profits. With a low competition environment in the sector the companies can use the ability of pricing power and increase its profits or pass increase of production cost to the clients.

2.1.5 Categories of Companies

Lynch and Rothchild define six general categories of companies depending of the growth rate of the company, capitalization size and his economic behavior (Lynch & Rothchild, 1989).

1. **Slow Growers** are large and aging companies that are expected to grow slightly faster than the GDP, with time every fast-growing industry becomes a slow growth industry and most of the companies in the sector lose momentum too and became slow growers. The best strategy to apply for investing in this case is purchase the shares with the objective to win a dividend, the aspects that investor needed to consider is the payout ratio percentage (more low better), and see in the financial statements if the company has a good record of pay dividend and check its growth rate.
2. **The Stalwarts** are big companies (Multibillion-dollars), where its earnings growth is faster than the slow growers. The stock of this type of company purchased at fair price probably will generate 30 to 50% gains in a good year. During recessions and hard times (bear market), normally these type of company tend to perform better than the general market, this happen because they have a durable competitive advantage that in recession times allows the company to increase its earnings.
3. **The fast Growers** are small companies, new enterprises that grow more than 20% a year, they do not need necessary belong to a fast growing industries to have high rates of growth, it can expand in a slow-growing industry by taking market share of its competitors. The share price of these companies represent the investments with most potential for valorization, but they have too a higher risk of down fold and a higher volatility in the stock prices than the other types of enterprises.
4. **The cyclical** companies are inside industries that expand and contract with a high correlation with the GDP growth. The sales and earnings growth faster than in other type of companies in periods of economic expansion, but in a scenario of economic recession those companies can go to bankrupt or need pass many years to the industry recover.
5. **Turnarounds** are companies that pass hard times caused by a bad new about the future of the business, poor financial condition that can mean bankruptcy or a scandal of corruption that affect the price of the stock. When the situation or factor that caused the devaluation of the share price is solved, it is probable a quick recovery of the price.
6. **Assets Plays** is any company that holds a valuable asset that worth more than what is recorded in the balance sheet, or it is unknown the existence of it by the market in its valuation of the stock.

2.1.6 Warren Buffett method

The Warren Buffett method consists in investing exceptional companies, with a durable competitive advantage, normally an economic monopoly or oligopoly, where the increase of the price of its products, leads to growth of operating profit. The most important factor for him is the durability of the competitive advantage, if it last for a long period of time the tendency to the business and stock price is to increase in the long run.

Warren Buffett invests only in business inside what he defines by "*circle of competence*", those are companies inside industries that an investor understand and had a profound knowledge (Hagtrom, 1994). The circle can be amplified and improved with research and experience, this mean that an investor can increase the number of hypotheses to invest and diversifies his portfolio.

He looks for companies with a consistent earnings history and with favourable long-term prospects. The business need to have a good return of equity ratio, low debt, high profit margins and be managed by honest people for him to invest. After finding a company with these characteristic he calculates the intrinsic value of the shares by discounting the present value of his estimations of future cash flows of the investment, and only purchases a share below its calculations more a margin of safety.

The characteristics of companies are divided in three parts, business, management, and financial.

1. Business characteristics:

- The business needs to be simple and understandable for the estimation of future cash flows has a high degree of confident.
- Business with a consistent operating history, give the guarantee that is competitive advantage is durable.
- Sell a unique product or service.
- Companies with bargaining power in its purchases (occurs when a company that purchases a large fraction of an industry's output can demand price concessions) or seller power (company has a monolithic control by on type of product and can demand higher prices). (Bodie, Kane, & Marcus, 2010)
- Company with low need of innovation and new products (meaning a research and development costs lower).

2. Management characteristics:

- Management need to be honest and rational to allocate the capital for provide returns above the cost of capital.
- Management candid with the shareholders by disclosing all the information relevant to them.

3. Financial tenets:

- Low debt in is balance sheet.
- Consistent growth in earnings.
- High ROE, the companies can give a higher return for each dollar invested.

The method of investing of Warren Buffett defends to hold the investment as much time as possible, but sometimes it sells a winning position due to some motives. The first reason to sell is when found a better investment, and need capital for that investment. Second, when is almost certain the company will lose its durable competitive advantage. The last reason is when the stock price reaches value that is far beyond the estimated theoretical value considered the future of the business and the economy (Buffett & Clark, 2008).

2.1.7 Approaches to Equity Valuation

In this section is presented how evaluating companies based in two groups of general approaches, the discount cash flow techniques, where the value of the stock is estimated based in the present value of the future cash flows, and the relative valuation techniques, that measure the level of overvaluation or undervaluation of a stock based upon in its current price relative to earnings, cash flow, book value and sales.

a) Discount cash flow model (DCF).

DCF valuation models recognize that common stock represents the ownership of a business and the value must be related to the future returns of owners.

Blanchard defines intrinsic value of stocks as the present value (PV) of the future expected cash flows, by estimating the total cash flows that are likely to occur in the life of the business and discounting it at an appropriate rate (Blanchard, 2006).

$$PV = \sum_{i=1}^n \frac{Cash\ Flow_i}{(1 + R)^i} \quad (3)$$

R-Discount rate
i-Year to discount

Financial analysts have developed different versions of the DCF, the *Dividend discount model* (DDM), the *Free cash flow discount model*, and the *Residual income model* (Hillier, Ross, Randolph, Jeffrey, & Jordan, 2010).

Next is present the three cases using the DDM to evaluate equity, depending on the rate of growth of dividends.

I. Case with a zero growth rate

In a case that the company pays a constant dividend and the investor hold the stock forever, the value of equity is given by the equation 4.

$$P_0 = \sum_{i=1}^n \frac{Div_i}{(1 + R)^i} = \frac{Div_1}{R} \quad (4)$$

R-Discount rate
 Div_i -Dividend year i

II. Case with a constant growth rate

Considering that the company will keep a dividend growth rate constant during his life, the value of the equity is determined by the equation 5.

$$P_0 = \frac{Div_1}{R - g} \quad (5)$$

g-Growth rate of dividends

III. Case with differential growth rate

This case is the more complex, and is this that occur in real world, for determine the value of a share the financial analyst need to do a prevision of the future growth rates of dividends applied to the equation 6. More correct are the previsions of the analyst better can he take advantage of the inefficiencies in market, due to valuation errors from other investors.

$$P_0 = \frac{Div_1}{1 + R} + \frac{Div_1 \times (1 + g_1)}{(1 + R)^2} + \dots + \frac{Div_{n-1} \times (1 + g_n)}{1 + R^n} \quad (6)$$

b) Relative Valuation Techniques

This techniques estimate the economic value of a company by comparing it with similar companies and past values. The relative ratios compare the stock price to the financial and economic information of the company.

It will be discussed the PER, Price to cash flow (P/CF), Price to book value (P/BV) and Price to sales (P/sales).

I. Price earnings ratio

A valuation ratio of a company's current share price compared to its per-share earnings. It is a way to estimate the value by determining how much dollars is necessary to pay for one of the earnings.

$$PER = \frac{Share Price}{EPS} \quad (7)$$

II. The price/Cash Flow

It is a ratio similar to PER, measures the firm's financial health, the definition uses the cash flow to calculate the ratio, because of that this ratio do not have the effect of depreciation and non-cash factors, and represent the price to pay for the money entering in the company.

$$\frac{P}{CF} = \frac{Share Price}{Cash flow per share} \quad (8)$$

III. Price/Book Value Ratio

It measures the valuation done by the market to the assets of the company, the premium pay over the book value represent the valuation on the capacity to growth and the ROE.

$$\frac{P}{B} = \frac{\text{Share Price}}{\text{Total Assets} - \text{Intangible Assets} - \text{Liabilities}} \quad (9)$$

IV. Price/sales Ratio

It evaluates a company in terms of the sales, is important ratio because the sales growth is a request to the company's growth. This ratio is used by investors because is difficult to manipulate the sales in the statement, and give a more exact relative evaluation.

$$\frac{P}{S} = \frac{\text{Share Price}}{\text{Sales}} \quad (10)$$

2.2 Technical analysis

Technical analysis, is a method of evaluating securities by analysing the market activity, the past prices and transactions volume. The technical analyst uses a set of tools like indicators and chart patterns to determine the possible future movement of the market. The balance of supply and demand is the factor that changes the price and volume of the securities.

This theory ignores totally the fundamental analysis and the intrinsic value of the securities; it is based in the philosophy that the change in the price and volume discounts the fundamental factors and information available (Achelis, 2000).

Another important belief is that history tends to repeat itself in terms of price movements. This is justified by the human mentality, the investors and traders, tends to have the same reactions and actions in similar conditions of the market. The Chartist technicians analyses the actual chart pattern that the market is forming and compared with past patterns to predict the future market movement. The DOW THEORY is other form of technical analysis that define that the markets as three trend movements, and each trend is composed by a phase of accumulation of stocks, absorption phase and distribution phase. The trend for be real need to be confirmed by increase of volume and they finish with a definitive signal of end.

2.2.1 Trends in the market

The prices fluctuate constantly in the market, but sometimes tend to describe oscillations with a tendency corresponding to a bull market in case of the tendency for price growth, when the trend decrease in price is called bear market. The Dow Theory describes the market as three forces simultaneously affect the stock price, described below.

a) Primary market trend

The main tendency or primary trend is the price tendency with higher timeframe, usually lasting several months or years. Defined by economic fundamentals of the company and economy, it starts when the market recognizes the fundamentals and moves in concordance with them, creating a price trend.

b) **Secondary market trend**

The secondary trend corresponds to market fluctuations against the primary trend with a lower period of time. These trends are also often called price correction when the market evolves too quickly relative to their fundamentals and performs a secondary tendency to correct the price.

c) **Tertiary or minor trends**

Minor trends are daily fluctuations of little importance, normally called noise.

2.2.2 Strategies to invest

Technical traders define two principal strategies to invest the trend following and band trading. Next will be explained these strategies.

a) **Trend Following**

This approach is based that sometimes the stock price moves in one direction for a long period of time, generating a large price movement (Tharp, 2007).

Trend followers enter long in the market when the price break a reference value and enter short with the inverse conditions. The strategy to be successful needs to have a good exit rules to limit its losses and the trader need to be disciplined to stay with the movement for a long period of time to generate the expected return.

b) **Band Trading**

Band trading strategies are used when the investor assume that the market is in a range bound, the trading concept is the opposite of trend following, where the band traders buy at prices close of reference low and sell at a reference high (Faith, 2007).

2.2.3 Types of Market

There are different types of volatility in the market, they can be very dynamic and changing sometimes very quickly, Faith defines four types of volatility in any financial instrumented traded (Faith, 2007).

a) **In a range stable and quiet**

Prices tend to stay within a relatively small range with little movement up or down outside the range, as can be seen in Figure 1. The strategy for explore this market is the band trading, but sometime the volatility is so low that is better to stay out of the market, waiting for a break of the range and using a trend following strategy.

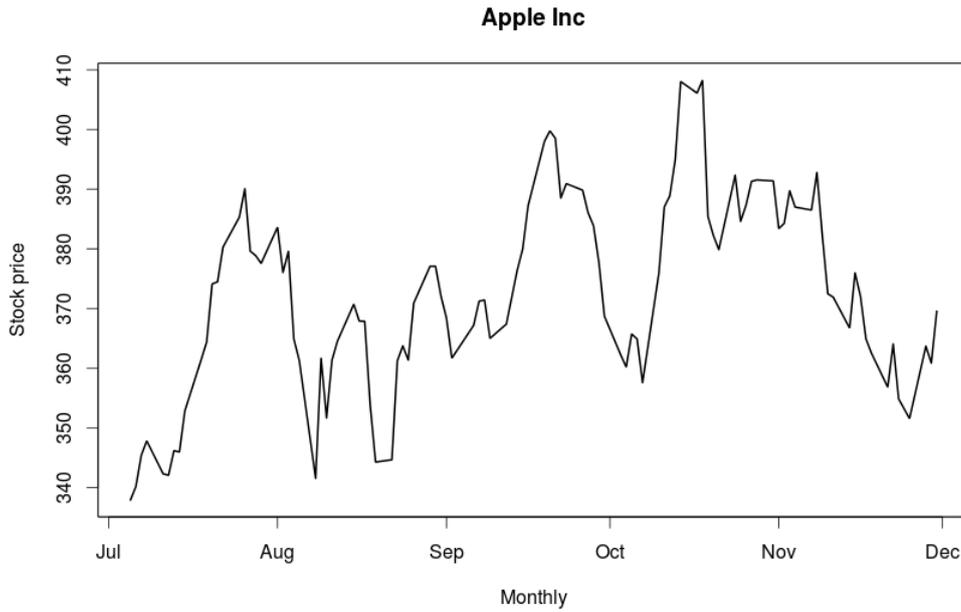


Figure 1 - Market with a stable and quiet volatility.

b) In a range stable and volatile

Prices tend to stay within a big range that are formed in time horizon of weeks, where inside the range have great volatility, as demonstrated in Figure 2. Band trading is the correct strategy to apply when this type of market is identified.

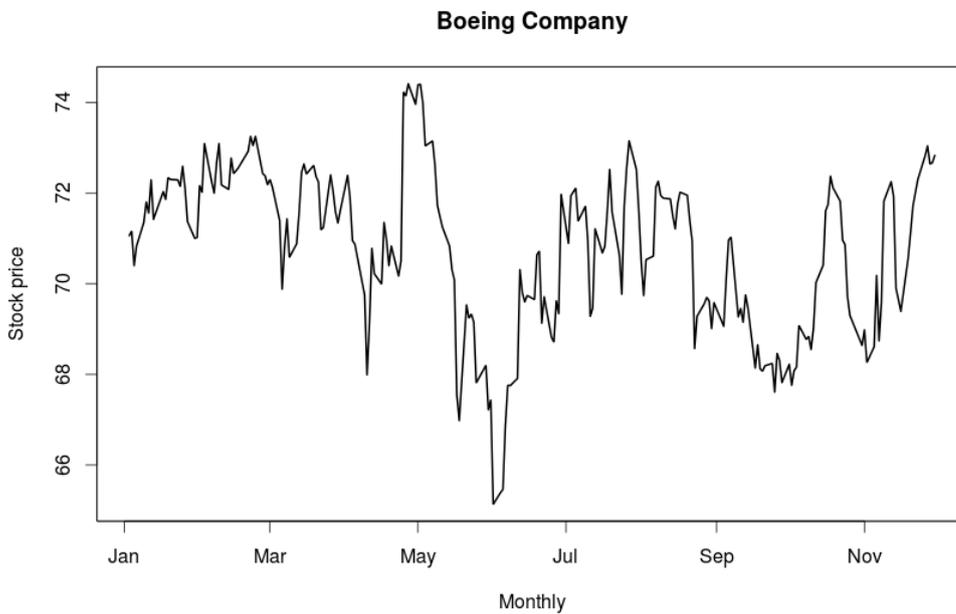


Figure 2 - Market stable and volatile.

c) Trending and quiet

Prices move in one direction, with retracements with low volatility in opposite direction, Figure 3. The trend following is the appropriate strategy to use in this case.

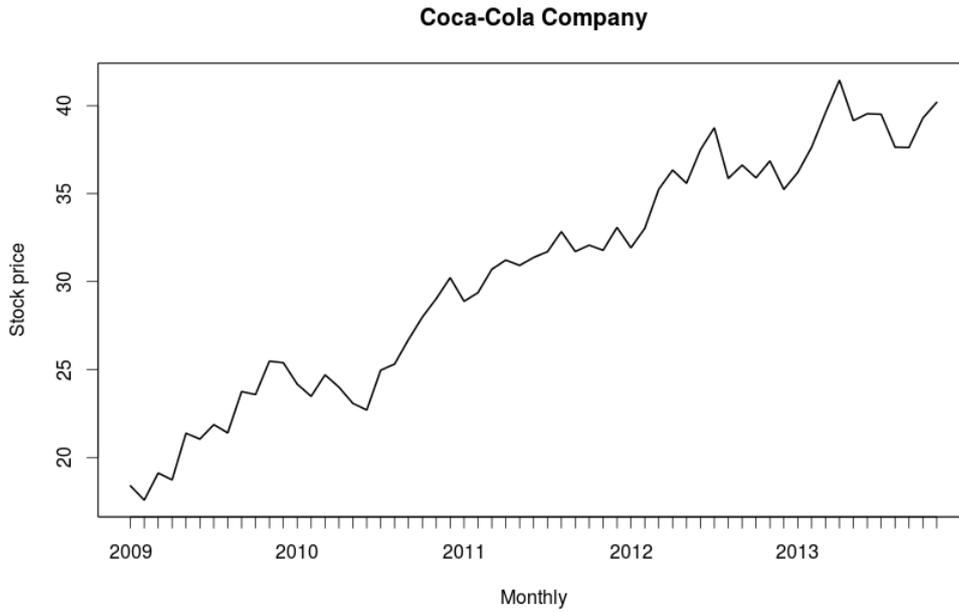


Figure 3 -Trending and quiet

d) Trending and volatile

Trending and volatile happen when there are large changes in price in one direction, with occasional significant short-term reversals of direction, Figure 4. This case is similar to the before, but is need to adapt the trend following strategy to the higher volatility of the correction.



Figure 4 - Trending with high volatility

2.2.4 Designing a strategy to invest

After developing the approach to invest there is the need to create a system to implement the strategy. The building blocks of any strategy are the trading costs, markets to invest, market timing, protection capital and exit of a wining position.

a) **Trading Costs**

Trading costs are expenses incurred when buying or selling securities, they include the commissions and spreads charged by the brokers for doing the transaction for its clients.

The costs of doing business are important because they reduce the net return of the investments and need to be considered in the investment decisions. High transaction costs and high frequency of trading mean a great percentage of money paid in commissions, this can be the difference between being profitable or not in the long run. Investor should avoid if possible high commissions and stocks with low liquidity, because these are the stocks with higher spreads.

b) **Select Markets to invest**

The first decisions of any investor is to select the markets to trade, considering the available budget, possible cost in transactions, the knowledge it has about the different products, the access to information and its experience.

c) **Entry or Market Timing**

This is the exact price and the market conditions that need to be present for entering in position of a common stock, the objective is define a set of rules that give the entry signal to improve the timing of buying and increase the reliability of the system (Tharp, 2007).

For *Peter Lynch* the right time to enter is during collapses, free-falls, and correction of the retracement of the price, due to lowest price of the securities (Lynch & Rothchild, 1989).

d) **Stops Loss and Protection of capital**

Stop loss are predetermined policies that reduce a portfolio exposure, it is a component of the system to get out of a losing position, not allowing one or more investments to continue to losing money and protect the remaining capital available to continue the investing activity.

e) **Exit of a winning position**

In any system is important to define the exit conditions of a winning position, for profits maximization, to perform this are incorporated a set of rules in the system, for selling a position. Technical traders normally sell its positions, when one or more of the next conditions are achieved:

- Sell signal given by a technical indicator
- A retracement of the market in the opposite direction
- When an amount of time pass from the time of entry
- A type of Chart Pattern done by the market
- When a predefined amount of profit is achieved

2.3 Introduction to Multi-objective Optimization

Multi-Objective optimization is a method used to solve real-world problems that involve simultaneous of several incommensurable and often competing objectives. Normally there is not only a single optimal solution, but a set of optimal solutions.

Multi-objective Optimization (MO) is the problem of finding the best solutions to optimizing two or more objectives that are in conflict, subjected to certain constraints (Metaxiots & Liagkouras, 2012).

The objective function (OF) is an equation to be optimized (maximized or minimized), composed by variables with need to respect certain constraints. Mathematical algorithms are used to optimize the OF, in single objective optimization there is only one OF to optimize, and in multi-objective there are two or more. The nomenclature used to these functions is $f_i(x)$.

The search space Ω is composed by n parameters, the decision variables. The solutions found for the optimization of the OF are called vector solution or decision vector, they are composed by a set of values each one correspond to one decision variable. The process of optimization corresponds to varying the value of the decision variables to search in the space the optimal solutions to the OF.

$$x = (x_1, x_2, x_3, \dots, x_n) \quad (11)$$

2.3.1 Multi-objective Description

The multi-objective optimization problem can be described as for each point in the space (Ω) defined by n decision variables have a vector function $F(x)$ composed by m functions $f_i(x)$. The goal is to find all the solution vectors that solve the problem, considering the restrictions imposed. The mathematical formulation to the problem is described in equations 12:

$$\begin{aligned} F(x) &= [f_1(x), \dots, f_m(x)], \quad m \text{ number of OF} \\ g_i(x) &\leq b_i, \text{ for } i = 1, \dots, k \\ x_i &\geq 0, \text{ for } i = 1, \dots, n \end{aligned} \quad (12)$$

k the number of constraints

The concept of Pareto dominance defines a vector x_1 dominates another x_2 for a maximization problem, if for each component of the vector $F_1(x)$ is bigger or equal at each component of $F_2(x)$, but in at least one component is greater (Zitzler, Laumanns, & Bleuler, 2004).

$$\begin{aligned} x_1 &> x_2 \text{ if,} \\ f_i(x_1) &> f_i(x_2), \text{ for at least one } i \text{ of } \exists i \in \{1, \dots, n\} \\ &\text{and for the remaining components of } f_i \\ f_i(x_1) &\geq f_i(x_2), \text{ for } i \in \{1, \dots, n\} \end{aligned} \quad (13)$$

A vector solution is called *Pareto Optimal* if there is no other vector in the solution space that improve any $f_i(x)$ without deteriorating at least other $f_j(x)$.

From this definition it is constituted the *Pareto set* that represent the objective vectors not dominated with different trade-offs between the objectives.

The Pareto Front (PF) or Pareto Set (PS) is composed by the set of efficient solutions non-dominated. The Figure 5 represents a multi-objective optimization problem with two OF to minimize, where the solutions to the problem are in feasible region.

$$\begin{aligned}
 PS^* &= x_i^* \subseteq \Omega \\
 F(x_i^*) &\subseteq F(\Omega) \\
 \nexists x_m \in \Omega: x_m > x_i, \forall x_i \in P^*
 \end{aligned}
 \tag{14}$$

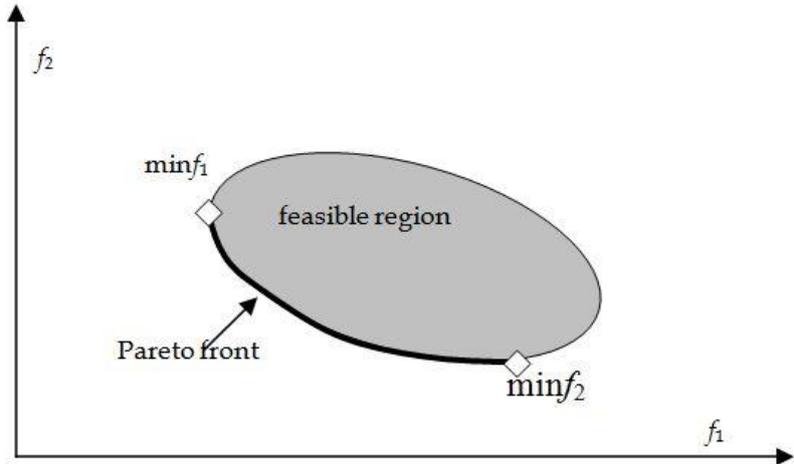


Figure 5 - Pareto front (Noilublao & Bureerat, 2009)

2.4 Evolutionary Computation

It is used by researchers a number of stochastic search strategies to solve the multi-objective optimization problem, the Evolutionary Algorithms, Genetic Algorithms, Genetic Programming, simulated annealing and ant colony optimization. Those heuristics do not guarantee the identification of optimal Pareto front, but find a good approximation to it.

A search heuristic is composed by Working Memory, Selection Module, and Reproduction Model, as represented in Figure 6.

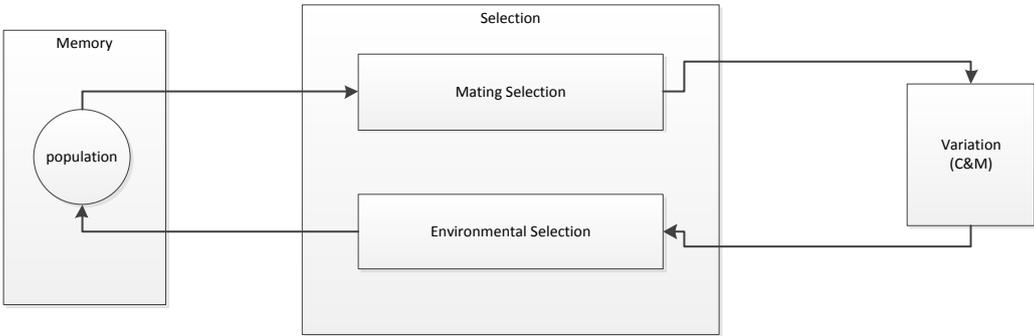


Figure 6 - General model of search heuristic

2.4.1 Memory module

This module contains the currently considered population, each time a reproduction is made the new population found is recorded in the memory by replacing the previous one.

2.4.2 Selection module

The selection model are composed by two processes the mating and environmental selection. The mating process picks the most promising individuals from the population and sends these individuals to the reproduction module. After the reproduction process creates new solutions the environment selection determines which of the new solutions are stored in the memory as the new population.

2.4.3 Reproduction model

The variation or reproduction model uses a set of solutions for systematically or randomly genetically change these solutions to generate a new set of potential better solutions.

2.5 Evolutionary algorithms

EA are optimization computer algorithms inspired by the biological model of evolution that improve automatically through experience. They use a mechanism that simulates the process of evolution to search in the space of solutions the global optimum of the problem. Using a method to evaluate the quality of randomly generated solutions they progress towards the optimum by using a fittest measure to choose the best candidates to reproduction.

The research done in the field proves that it is a heuristic capable to solve the optimization problem and find the pareto set, because EA are capable of processing a set of solutions in parallel and finding a good approximation in a single run (Zitzler & Thiele, 1998).

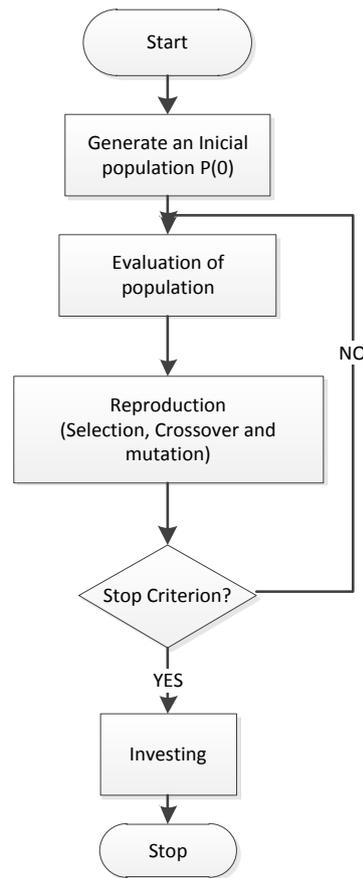


Figure 7 - Flow chart of a EA

2.5.1 Representation

EA solutions are represented as a chromosome that is a collection of *genes*, each gene represent one decision variable. The chromosome is an encoding of the solution to the problem, and is called genotype.

2.5.2 Population

It is called population the set of solutions that the algorithm finds at each iteration, normally in the first one is used a random population, or can be used a population from a previous simulation to initiate the process.

2.5.3 Quality indicator

The quality indicator or fitness function is designed to gives a quantitative measure of the quality of the chromosome in solving the problem. The selection module uses the evaluation done by the fitness function to select the chromosomes for reproduction. At each iteration it is expected that the average fitness of the population improves meaning that the algorithm is learning.

In single-objective optimization the objective function and fitness function are often equals, but for multi-objective it is used different strategies to do the fitness, they are aggregation-based, criterion-based and Pareto-based fitness assignment strategies (Zitzler, Laumanns, & Bleuler, 2004).

2.5.4 Selection

After the fitness is performed the algorithm needs to decide what are the individuals obtained in the reproduction that compose the next population, this is done by a mechanism where the best individuals from the population are selected to be used for reproduction in the next iteration. This process of selection and reproduction steers the search in direction of the non-dominated front at each selection. The common selection operators are the Fitness-Proportionate Selection, Truncation Selection, Ranking Selection, Tournament selection (Hassan, 2010).

2.5.5 Variation or Reproduction operations

Variation operations are realized in the individuals that are selected for reproduction to generate the next population, by the processes of mutation, and crossover.

a) **Crossover**

Recombination or crossover operation uses a certain set of solutions to create a predefined number of new solutions called children, by recombining parts (genes) of each solution (Zitzler, Laumanns, & Bleuler, 2004).

b) **Mutation**

Mutation is the operation that selects some individuals of the population to reproduction randomly, consists in modifying these individuals by changing some genes according with a given mutation rate.

2.5.6 Algorithm design issues

According to (Zitzler, Laumanns, & Bleuler, 2004) in MO the goal is to achieve the set of solutions that are as close as possible to the global Pareto-optimal front, by minimizing the distance of the solutions generated to the pareto front. This is obtained by selecting the individuals that are non-dominated for them to reproduce. The algorithms need to maximize the diversity of the population obtained to cover the whole pareto-front, this is done by avoid populations that contain too much identical solutions. Next is explaining the considerations to take in the design of a Multi-Objective Evolutionary Algorithm (MOEA).

a) **Diversity Preservation**

For a MOEA preserve the diversity within the current Pareto set approximation, it is need to incorporate density information in the algorithm to be used in the selection process. This implies that in the selection module the chances of individual's being selected for reproduction decreases, if the density of individuals in its neighborhood is greater. For doing the diversity preservation there are three techniques the Kernel methods, Nearest neighbor techniques, and Histograms, those techniques are described in (Zitzler, Laumanns, & Bleuler, 2004).

b) **Elitism**

Elitism addresses the problem of losing good solutions during the optimization process, due to random effects. For solving this problem in the new populations the best elements of the old population are preserved (Zitzler, Laumanns, & Bleuler, 2004).

c) **Limit behavior**

Limit behavior of the MOEA is what the algorithms can achieve in terms of performance, when they have unlimited resources, infinite time to continue the evolution and space in memory to store all the solutions.

d) **Global convergence**

Global convergence for MOEA is when the Pareto front approximation $PF^{*(t)}$ achieved by the algorithm is practical identical to the true Pareto front when the number of generation's t goes to infinity. In real implementations the algorithms have to deal with limited resources, and should guarantee the convergent approximation to the true Pareto front $PF^{*(t)} \subseteq PF^*$ (Zitzler, Laumanns, & Bleuler, 2004).

2.5.7 State of the Art Multi-Objective Evolutionary Algorithms

MOEAs are distinguished from standard EAs by employing the Pareto dominance concept in the fitness evaluation to allow the comparison between individuals based on multiple conflicting objectives.

The first implementation of a multi-objective evolutionary algorithms, date from 1984 with the algorithm VEGA an simple genetic algorithm with a modified selection mechanism , developed by Shaffer, from then many new algorithms have been developed, Srinivas and Deb developed NSGA in 1989 that ranked the population with the pareto dominance and uses a dummy fitness value proportional to the population size; MOGA uses a scheme in which each is given a rank proportional to the number of individuals, created in 1993 by Fonseca and Fleming; Horn, Nafpliotis and Golberg developed NPGA in 1994, this algorithm uses a tournament selection scheme based on pareto dominance ; SPEA uses an external archive for saving the non-dominated solutions found in each iteration, was introduced by Zitzler and Thiele in 1999; PAES developed by Knowles and Corne in 1999 uses an evolutionary strategy with one parent produce on child by mutation; PESA uses hyper-grid scheme technique for selection and diversity maintenance, developed in 2000 by Corne, Knowles and Oates. Deb, Pratab, Alagarwall and Meyarivan proposes the NSGA-II, that is main characteristic is the computational implementation more easily is that NSGA ; Erickson, Mayer and Horn they have created NPGA –II in 2001 it uses a Pareto approach with tournament selection scheme; SPEA- II introduced by Zitzler, Laumanns and Thiele in 2001 is an improved version of SPEA with an improved fitness assignment; Corne, Jerram, Knowles and Oates have develop PESA-II in 2001 uses an selection method where the selective fitness is done to the objective space and not to the individuals; in 2012 Liang, Jane You, Han, Li, proposed the DS-MOEA it uses three alternative operations to perform the learning of the population(evolution in the objective space, evolution in the solution space and self-evolution); and a new algorithm call PSS uses a method to select the non-dominated solutions using an global valuation done by fuzzy measures with the user's degree of consideration objective and partial objective evaluation, it was created by Jong Kim, Ji Han, Ye Kim, Seung Choi, Eun Kim in 2012.

2.6 Portfolio Optimization

A portfolio can be defined as a set of selected investments done by its manager. He has at their disposal an amount of capital and a range of assets, to maximize the portfolio return and minimize the risk using asset allocation⁷. The efficient frontier in Figure 8 represents the portfolios with highest expected return for any risk level.

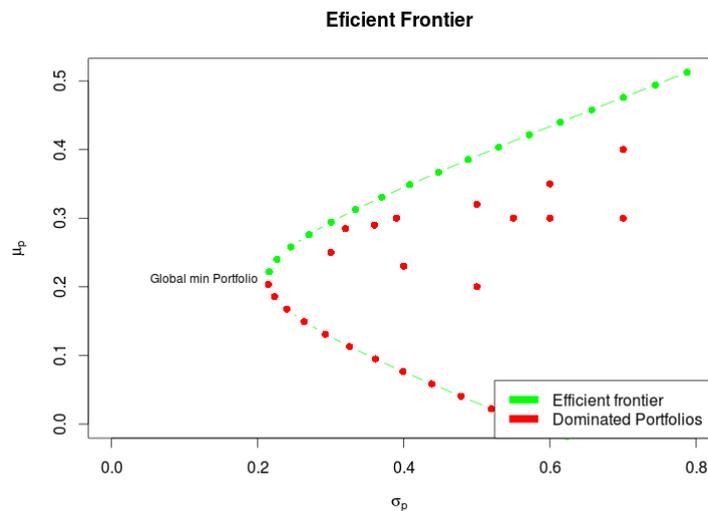


Figure 8 - Efficient frontier

2.6.1 Diversification and Risk of Portfolios

The arbitrage pricing theory using one factor model of equation 15 is a tool that allows describing and quantifying the factors that affect the rate of return of a security.

$$R = \bar{R} + \beta F + \varepsilon \quad (15)$$

The risk of a stock is the unanticipated part of the return ($\beta F + \varepsilon$), can be separated in two components, the systematic risk (βF) or market risk due the macroeconomic environment that affects the general market, and unsystematic risk (ε) that affect only the company or the industry where it is insert.

This concept applied to the portfolios theory, defines that the return of a portfolio is given by the weight sum of individual returns, equation 16. It can be noted the diversification effect in the risk reduction of the portfolio in the equation, because the unsystematic risks are independent, the increase of securities in the portfolio decreases the value of the sum of weights of unsystematic risks.

⁷ Investment strategy that attempts to balance the risk versus return by adjusting the percentage in each asset selected to incorporate the portfolio.

$$R_p = w_1 \times (\bar{R}_1 + \beta F_1 + \varepsilon_1) + \dots + w_n \times (\bar{R}_n + \beta F_{2n} + \varepsilon_n) \quad (16)$$

This theory is applied to the portfolios when is used some diversification to reduce the weight average of the unsystematic risk by diversification.

2.6.2 Mean variance – Model

The Classical Mean-Variance portfolio optimization model introduced by Markowitz aims to select a set of assets to invest from the space of available assets and determine the fractions w_i of the budget to be invested in each asset selected. In this model the objective is to minimize the risk for each level of expected return of the portfolio.

The Markowitz portfolio optimization can be stated mathematically as follow (Markowitz, 1952):

$$\text{Min } \rho(w) = \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i w_j \quad (17)$$

σ_{ij} – Covariance of asset i with the asset j

The portfolio (P) is a set of real valued weights (w_i) of the stocks selected from the n available assets in the market.

$$\sum_{i=0}^n w_i = 1, \quad 0 \leq w_i \leq 1, \quad i = 1, \dots, n \quad (18)$$

The return of each asset is given by:

$$E_i = w_i \times r_i \quad (19)$$

r_i – valorization of asset i

The expected return of the portfolio is given by the sum of the expected returns of the individual assets:

$$r_p = \sum_{i=1}^n w_i \times r_i \quad (20)$$

2.6.3 The mean-variance cardinality constrained portfolio optimization model (MVCCPO)

In the real world portfolio managers need to deal with a number of realistic constraints that arise from normal business practices and industry regulations (Skolpsdungket, Dahal, & Harnpornchai, 2007).

The Model MVCCPO is an expanded model of the Markowitz, where it is added two constraints, the cardinality constraint and the quantity constraint.

Cardinality constraint impose a limit on the number of assets to be selected, it is used by managers to restrict the number of assets in the portfolio to a number that is possible following the economics factors of each asset, and control the transaction costs.

Quantity constraint or buy-in threshold restricts the weight of each asset in the portfolio between lower and upper bounds. The low limit is used to avoid small holdings that do not have almost any effect in the performance of the portfolio, and to prevent to pay higher costs of transaction. The higher limit is used to prevent the excessive exposure of the portfolio to a single asset (Chang, Meade, & Sharaiha, 2000).

Mathematical formulation of MVCCP Model:

$$\text{Min } \rho(x) = \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} x_i x_j \quad (21)$$

$$\text{Max } r_p(x) = \sum_{i=1}^n r_i x_i \quad (22)$$

$$\sum_{i=0}^n x_i = 1, \quad 0 \leq x_i \leq 1, \quad i = 1, \dots, n \quad (23)$$

$$\sum_{i=0}^n \delta_i \leq K \quad (24)$$

$$l_i \delta_i \leq x_i \leq u_i \delta_i, \quad i = 1, \dots, n \quad (25)$$

$$\delta_i \in \{0,1\}, \quad i = 1, \dots, n$$

In this model the equations 21 and 22 are the objectives functions to optimize by the MOEA, in the equation 23 is imposed a restriction that the total amount invested need to be equal to the capital available, equation 24 represent the Cardinality constraint and the equation 25 limits the weight of each investment done ($\delta_i = 1$) to an inferior limit l_i and upper limit u_i .

2.7 State of the Art of Portfolio Optimization

Some real-world problems involve simultaneous optimization of several incommensurable and often competing objectives. Normally there isn't only a single optimal solution, but a set of optimal solutions.

Multi-objective Optimization (MO) is a method to solve the problem of finding the best solutions when optimizing two or more objectives that are in conflict with each other, subjected to certain constraints (Fonseca & Fleming, 1993).

There are many measures of risk and return to evaluate the performance of a portfolio, and these measures can be used as the objectives to be optimized by the EA. The most popular in portfolio management are the Compound annual growth rate CAGR%, Managed Account Report (MAR) ratio,

Sharpe Ratio, Value at Risk (VaR), conditional value-at-risk (CVar), the Mean (Portfolio Expected Return), and the variance (Metaxiots & Liagkouras, 2012).

Tettamanzi & Loraschi in 1995 describe a Multi-Objective Evolutionary Algorithm (MOEA) using the Markowitz model, but the measure of risk used is the lower partial moments or downside risk introduced by Harlow (Harlow, 1991). This objective takes into account the down-side part of the distribution of returns. The research proves that downside risk make the use of quadratic optimization techniques impossible, because the shape of the objective function is non-convex.

Cesarone, Scozzari, & Tardellain 2009 used an algorithm based approach that starts from a pair of assets in the portfolio and tries to add one each time for the MVCCPO model. The simulations proved this model of investment has better performance than the Markowitz classical model.

In the work done by Chang, Meade, & Sharaiha in 2000 three heuristics TS, GA and SA are used to solve the portfolio optimization problem. They also study the problem of finding the efficient frontier using the MVCCPO model. The results prove the existence of cardinality constraint affects in the shape of the efficient frontier, causing discontinuities in the curve.

In the work Lin & Wang in 2001, NSGA-II is used in the Markowitz's Model with constraints of fixed cost and minimum lots. The results show the efficiency of the GA is undermined without the fitness scaling, and the transaction costs dislocates the pareto curve in the vertical axis.

Schaerf in 2002 uses MVCCPO model to compare and combine different neighborhood relations in the Pareto front, with local search strategies to find it.

Schyns & Crama in 2003 describes the application of simulated annealing (SA) for solution of the Classic Markowitz model with more realistic constraints, the quantity constraint, cardinality constraint, turnover constraints and trading constraints. The advantage of using SA over other heuristic methods is the ability to avoid getting trapped in optimal local points, its flexibility and ability to approach global optimality. The important conclusions of this paper are the introduction of trading constraints is difficult to handle, and there is a trade-off between the quality of the solutions and the time of the simulations to find them.

Lozano & Armañanzas in 2005 uses the heuristics greedy search (GS), SA, and the ACO. For the simulations they used data from five different market indexes. They varied the number of assets (K) in the portfolio for each simulation. The simulations show, that fewer assets in a portfolio can represent a higher expected return but it is obtained with a higher variance of the return. They concluded that ACO is a better heurist than the others to obtain portfolios solutions with higher risk more close to the true pareto front, and the SA fits better for lower risk values.

Clack & Patel in 2007 compared the performance of a standard EA against an Age-Layered Population Structure EA (ALPS EA). They use in the portfolio a basket of 82 stocks of the 100 available. The simulations performed showed that ALPS EA reduces the premature convergence, providing better fittest solutions than the EA.

Ghang, Yang, & Chang in 2009, tested different risk measures in substitution of the mean-variance one of them is the variance with skewness, developed based on the theory that portfolio return may not be a symmetrical distribution, this means that the distribution of return of individual assets tend to exhibit a higher probability of extreme values, like it has been suggested first by Samuelson in 1958.

The results show that MOEO is capable of finding a wider spread of solutions than the others algorithms, and is capable of competing with NSGA-II, SPEA 2 and PAES in portfolio optimization problems.

Hirabayashi, Aranha, Hitoshi in 2009 proposed a GA to generate trading rules based in technical indicators (RSI, MA, percent difference from moving Average). The algorithm after entering in a position, will exit it based on the following genes stop loss or take profit optimized. They used this system to trade in the forex market (FX).

Golmakani & Fazel in 2011, used an extended model of Markowitz, with four constraints (minimum transaction lots, sector capitalization lots, cardinality, and quantity constraints). The authors proposed a heuristic called CBIPSO (combination of binary PSO with improved PSO) to solve the portfolio optimization problem of Markowitz.

They compared their heuristic against the GA proposed by Soleimani in 2007. In the simulation they tested different portfolio sizes and expected returns, and they conclude that the CBIPSO outperforms Genetic Algorithms (GA), can achieve better solutions in less amount of time.

Hassan & Clack in 2009 tested the combinations of two techniques mating restriction and diversity enhancement in the algorithm SPEA2, to improve the robustness and the diversity of the solutions. To evaluate the quality of the solutions they used the sharpe ratio.

Casanova in 2010 used a Learning Classifier System (LCS) in a dynamic learning system to select the stocks to invest based in technical and intuition analysis, the revaluation period RP, Average Revaluation Period(ARP), RSI, MA, DMA are the indicators used for ranking the best stocks for trade, considering the genes parameters (Days; Minimum value selection of the parameter, Variation allowed of the best stock, type of price) for each indicators; with a system for tactical asset allocation call Tradinnova-LCS, simulates the intelligent behaviour of an investor in a continuous market to form the portfolio. The system tested outperforms all the investment funds analysed by the INVERCO in the periods of simulation.

Gorgulho, Neves, & Horta in 2011 implemented an expert technical trading system, describing the system architecture and the investment simulator, and used GA to found the solutions. They tested the system against B&H strategy, and random selection, to prove the superiority of the GA system based in technical signals.

The approach of Kaucic presented in 2012 is a trading system based in technical analysis, where an investment module is used to manage a portfolio with long and short positions to generate the so-called long-plus-short portfolio. A technical module is used for detecting overbought/oversold conditions and short-term changes in relative value in contrast to long-term through a learning mechanism using EA that manages the information derived from the technical indicators incorporated.

Pandari, Azar, & Shavazi in 2012 developed a MOEA model with six objectives to optimize, and tested it against the classical model of Markowitz. The conclusion that they arrive, is that their model use less risk, due to the higher number of objectives optimized by the algorithm.

In Table 6 is presented a summary of the different solutions related to the optimization of portfolios using several parameters to describe their main characteristics.

| Reference | Period of Simulation | Algorithms Utilized | Markets tested | Fitness functions | Constraints | Portfolio analysis | Results Obtained |
|--------------------------------------|-------------------------------------|---|--|---|---|---|--|
| Chang, Meade, & Sharaiha | Mar 1992 to Sep of 1997 | GA TS SA | Hang Seng DAX, FTSE S&P, Nikkei | Mean, Variance | minimum lots | Markowitz's Model | Best results obtained with the GA Heuristic |
| Lin & Wang | Mar 1992 to Sep of 1997 | GA based on NSGA-II and Genocop | Hang Seng index | Mean, Variance | fixed transaction costs minimum lots | Markowitz's Model | The proposal GA solve the portfolio selection problem efficiently |
| Schaerf | NA | TS SA LS | Hang Seng, DAX,FTSE,S &P, Nikkei | Average percentage loss | cardinality quantity | Markowitz's Model | The Tabu search is the heuristic that achieves the best Pareto Curve. |
| Schyns & Crama | 6 Jan of 1988 to 9 Apr 1997 | SA | 151 US Stocks | Mean, Variance | Floor ceiling, turnover, trading and quantity | Markowitz's Model with | The SA is able to handle more classes of constrains than other heuristics |
| Lozano & Armañanzas, 2005 | Mar 1992 to Sep 1997 | GS SA ACO | Hang Seng DAX ,FTSE S&P ,Nikkei | Mean, Variance | cardinality quantity | Markowitz's Model | They obtained a portfolio with a return of 3 and risk 0.1 |
| Clack & Patel | 31 of May of 1999 to 31 of Dec 2005 | ALPS ¹ system incorporated in GP | FTSE 100 | Sharpe Ratio | | Non-Linear Model | The ALPS GP obtained a return of 50%,and the standard GP a return of 33% |
| (Ghang, Yang, & Chang | From Jan2004 to Dec 2006 | GA | Hang SENG FTSE S&P | Mean, variance, Semi-variance, Variance with Skewness | | Markowitz's Model | The higher return obtained for S&P was 0.0023 with a risk 0.0008 |
| Chen, Weng, & Li | Mar1992 to Sep 1997 | MOEO algorithm | Hang Seng DAX ,FTSE S&P ,Nikkei | Mean, Variance | cardinality quantity | Markowitz's Model | Best performance obtained with a MOEO with a Return of 0.00859 and risk 0.000417 |
| Anagnostopoulos & Mamanis | Mar1992 to Sep 1997 | NPGA2,NSGA-II PESA, SPEA2, e-MOEA | Hang Seng DAX, FTSE, S&P, Nikkei | Mean, Variance | cardinality quantity | Markowitz's Model | The SPEA2 is superiority than others MOEA |
| Casanova | 2005 to 2009 | LCS model | IBEX 35 | ROI | | Technical Analysis | 15,3% |
| Gorgulho, Neves, & Horta | 06 Jan 2003 to 06 Jan 2009 | GA | DJI | ROI | | Technical trading system | 60% |
| Kaucic | 25 Jan2006 to 19 Jul 2011 | EA | DJI | Information Ratio, Omega, Sortino ratio | | Technical trading system | 70% 125% 119% |
| Pandari, Azar, & Shavazi | Mar2002 to Mar2008 | MOEA | Tehran Stock | Cumulative Return, Mean Return | | GA model using Sharpe Ratio, Markowitz's Model | 600% 350% |

Table 6 - Overview approaches to portfolio optimization

2.8 Conclusion

In this chapter was presented major investment approaches, in particular the value investing, where is explained their components, such as financial analysis performed using financial statements of the companies, the economic indicators that influence the business sector, and macroeconomic indicators to take into account in forecast the economic cycle, for investors have a better perspective of the future and profitability of investments.

It has explained the method of investment of Warren Buffet, and the DCF methods used to evaluate companies depending on the cash flow generated and growth rates. It was presented the different types of markets, and the necessary components to build a complete trading system.

The optimizing concepts of multiple objectives, such as Pareto curve, and the Pareto set, were discussed in this work in a way to be applied in the development of the MOEA. It was made a brief introduction to evolutionary algorithms and how to implement them. It is explained the most important considerations to have in implementation and possible problems of the EA.

After a description of the State of the art of multi-objective EA, are introduced the Mean-Variance Model, and the mean-Variance cardinality constrained portfolio optimization model. It is presented the state of art of the investigation than with portfolio optimizations, the models of investment, approaches to manage the portfolios used, and the objectives to optimize.

The active strategies of investment based mainly in technical analysis, demonstrate that exist a number of possible applications of intelligent computing applied to investments in stocks, and can achieve good investment returns.

3. System Architecture

The system developed uses the value investing approach with a technical trading system to invest in the stock market. For the system work is necessary to obtain the financial market data to perform the trainings and real test simulations.

Due to the wide availability of information and data in the U.S. markets, it is essential to focus on the really important information. The financial information provided, as well as the economic analysis of the sector, and the management of the company, are the most relevant data information to the analysts. In this work it is used the annual financial information available, and the adjusted closing prices of the companies of S&P 500.

This chapter presents the architecture of the system, first in the section 3.1 is explained the algorithm architecture, section 3.2 describe how it is obtained the financial statements and quotations of listed companies in the S&P 500, in section 3.3 is explained the financial statements and how interpret them, then section 3.4 presents the financial ratios used in the chromosomes to evaluate the companies, and the section 3.5 describes the technical analysis used.

3.1 Algorithm architecture

The system architecture is presented in Figure 11 constituted by three main modules, the Investor Simulator, the Optimizer and the Data. The main blocks of the architecture are the Data block that is accessed by the investment simulator to test the strategies, and the optimization block to implement the MOEA. The Optimization Block uses the results Return of Investment (ROI) and Variance, obtained in the investor block to calculate the fitness function and evaluate the strategies. It also selects them for reproduction and applies the methods of crossover and mutation to create news chromosomes.

The Data Block uses the financial statements and stock quotes to calculate the ratios used by the investor simulator. The Investor Simulator Block tests the strategies obtained in the optimization block, depending on the investment model and using the inputs given by the user, for getting the data for testing. It evaluates in each day the stocks, and calculates the ROI of the portfolio from the beginning of the simulation. With the recorded values for each day of the ROI, it is then calculated the monthly variance of the return.

Each strategy is tested in the simulator using the retrieved time series, with a window that slides one day at each increment. The selected financial ratios are evaluated, and the computations are performed for doing the trading decisions of the day. After finishing the time frame of each period of training, it is recorded the results, and performed the population reproduction. This process is repeated to achieve the number of iteration that signals the end of the training. After the conclusion of the training of the population the external file with the best solutions is used in the real test.

The system uses the following parameters, the period of training, and the period of real test, also transactions costs are included at to 2% of the stock value.

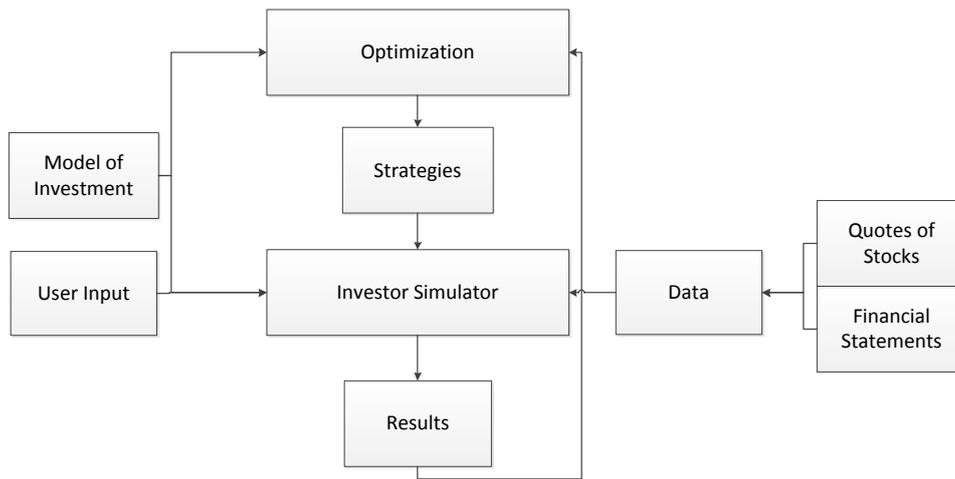


Figure 9 - System architecture

3.1.1 Architecture of the Evolutionary Algorithms

It is presented next a flow chart with the architecture of the algorithm in Figure 10, where the system required some specification to work, the period of training, real test period, commission percentage to charge in each transaction performed and the number of iterations to perform the training of the population. Next is accessed the data to the algorithm to work in the periods defined, before the start of training is created a random population. For test the chromosomes is created time series with all the stocks prices, where are executed the tasks to manage the portfolio, as calculate the financial ratios, and trading decisions each time the window advance on day. In the end each iteration is recorded the results (Return and Variance) of each chromosomes for the reproduction and selection process. When number of the iterations arrive to the maximum number the training is finish, and the algorithm perform a real simulation of the population in the external file (the best chromosomes obtained in the training) for the period defined in the specifications. After tested all chromosomes are recorded the results obtained.

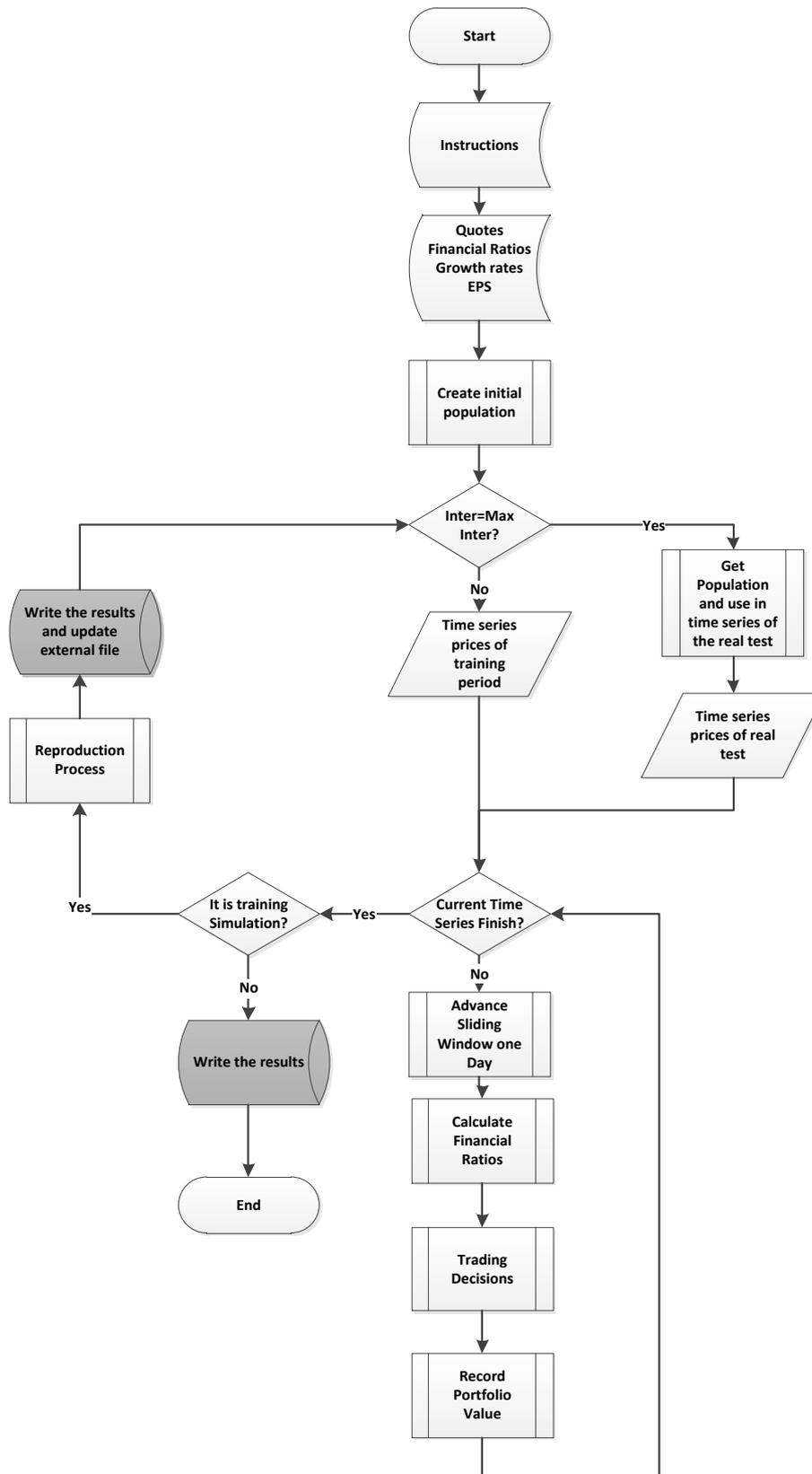


Figure 10 - Flow chart of algorithm Architecture

3.2 Obtaining the Financial data

To perform the necessary studies in this work, it was necessary to obtain the financial statements of the companies and their quotations. It was chosen the Yahoo finance Database to extract the necessary data, because allows the download of the Excel files with the markets information of each company and have on their site quarterly and annual financial statements of each company. To get the data, a program was developed in C + + in eclipse in the Linux environment to download of the quotes of the companies selected, and the respective financial statements. The flow charts in Figure 11 describe the work of the program mentioned before.

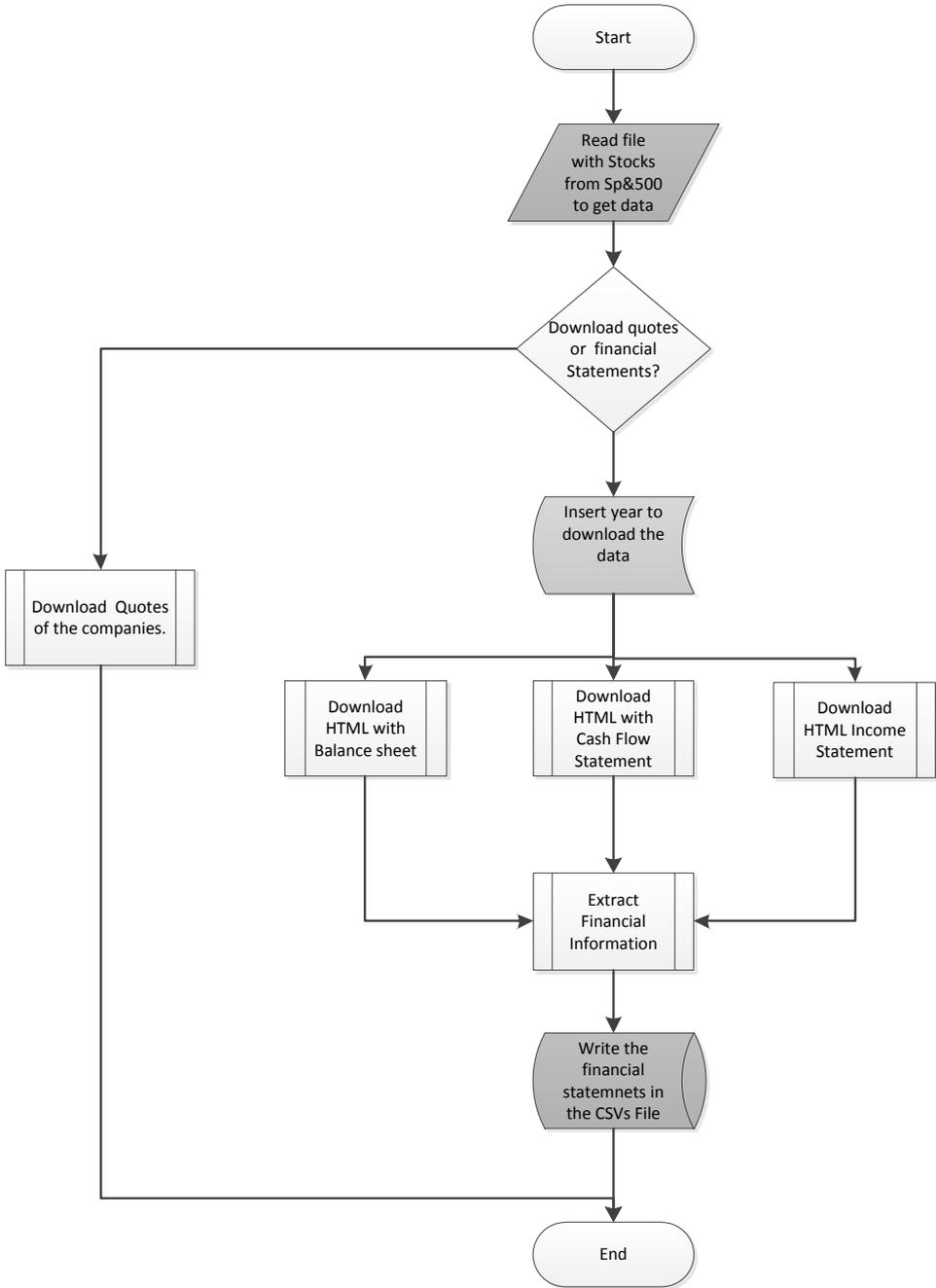


Figure 11 - Flowchart of the program for getting the data

3.3 Financial Statement analysis

In chapter 2, where presented some techniques of valuation of Stocks, but the input variables of the models are estimates of futures dividends and earnings, for the investor this is based on economic projections that can be very uncertain, the only correct information available is the financial accounting, to estimate the intrinsic value of stocks.

3.3.1 The income Statement

The income statement is a demonstration of the profitability of the company, over a period of time. It is reported the revenues obtained from the operations, the expenses incurred, and the net earnings.

The expenses are separated in four parts, the first is the cost of goods sold, this are the direct cost of generating the products, second are general and administrative costs, that are indirect costs related to the production, the third is the interest expenses of the debt, recorded as financial cost, and in the final is the taxes charged by the government.

The Figure 12 shows the income statement of Microsoft, where in the top are the revenues from operations, next are the costs, by subtracting the costs to revenue is obtained the *operating income or loss*, or *EBIT*⁸. To obtain the *Net Income* in the last row it is extract the interest expenses and taxes paid to the *EBIT*.

| In Millions of USD (except for per share items) | 12 months ending 2013-06-30 |
|---|-----------------------------|
| Revenue | 77,849.00 |
| Other Revenue, Total | - |
| Total Revenue | 77,849.00 |
| Cost of Revenue, Total | 20,249.00 |
| Gross Profit | 57,600.00 |
| Selling/General/Admin. Expenses, Total | 20,425.00 |
| Research & Development | 10,411.00 |
| Depreciation/Amortization | - |
| Interest Expense(Income) - Net Operating | - |
| Unusual Expense (Income) | 208.00 |
| Other Operating Expenses, Total | - |
| Total Operating Expense | 51,293.00 |
| Operating Income | 26,556.00 |
| Interest Income(Expense), Net Non-Operating | - |
| Gain (Loss) on Sale of Assets | - |
| Other, Net | 194.00 |
| Income Before Tax | 27,052.00 |
| Income After Tax | 21,863.00 |
| Minority Interest | - |
| Equity In Affiliates | - |
| Net Income Before Extra. Items | 21,863.00 |
| Accounting Change | - |
| Discontinued Operations | - |
| Extraordinary Item | - |
| Net Income | 21,863.00 |

Figure 12 - Annual Income statement of Microsoft

⁸ This is a measure of profitability of the firms operations

3.3.2 Balance Sheet Statement

The balance sheets give the financial condition of the company at particular moment, where it is list all the assets and debts at the time of the report occurred.

In the first part of the report are listed the assets from the more liquid to the less, they are divided in current assets, and long term assets. Next are the part of liabilities that is arranged similiary the short – term debts comes first, and the long term debts after.

The difference between total assets and liabilities is the stockholders' equity, this are the book value of the company. In the last line of the balance sheet are the total common shares outstanding, representing the number of shares in the market of the company. In Figure 13 are the exemple of annual balance sheet statement, where can be seen the assets and liabilites hold by the microsoft.

| In Millions of USD (except for per share items) | As of 2013-06-30 |
|---|------------------|
| Cash & Equivalents | 3,804.00 |
| Short Term Investments | 72,971.00 |
| Cash and Short Term Investments | 76,775.00 |
| Accounts Receivable - Trade, Net | 17,486.00 |
| Receivables - Other | - |
| Total Receivables, Net | 17,486.00 |
| Total Inventory | 1,938.00 |
| Prepaid Expenses | - |
| Other Current Assets, Total | 5,267.00 |
| Total Current Assets | 101,466.00 |
| Property/Plant/Equipment, Total - Gross | 22,504.00 |
| Accumulated Depreciation, Total | -12,513.00 |
| Goodwill, Net | 14,655.00 |
| Intangibles, Net | 3,083.00 |
| Long Term Investments | 10,844.00 |
| Other Long Term Assets, Total | 2,392.00 |
| Total Assets | 142,431.00 |
| Accounts Payable | 4,828.00 |
| Accrued Expenses | 4,117.00 |
| Notes Payable/Short Term Debt | 0.00 |
| Current Port. of LT Debt/Capital Leases | 2,999.00 |
| Other Current liabilities, Total | 25,473.00 |
| Total Current Liabilities | 37,417.00 |
| Long Term Debt | 12,601.00 |
| Capital Lease Obligations | - |
| Total Long Term Debt | 12,601.00 |
| Total Debt | 15,600.00 |
| Deferred Income Tax | 1,709.00 |
| Minority Interest | - |
| Other Liabilities, Total | 11,760.00 |
| Total Liabilities | 63,487.00 |
| Redeemable Preferred Stock, Total | - |
| Preferred Stock - Non Redeemable, Net | - |
| Common Stock, Total | 67,306.00 |
| Additional Paid-In Capital | - |
| Retained Earnings (Accumulated Deficit) | 9,895.00 |
| Treasury Stock - Common | - |
| Other Equity, Total | -117.00 |
| Total Equity | 78,944.00 |
| Total Liabilities & Shareholders' Equity | 142,431.00 |
| Shares Outs - Common Stock Primary Issue | - |
| Total Common Shares Outstanding | 8,328.00 |

Figure 13 - Annual Balance Sheet statement of Microsoft

3.3.3 Cash flow statement

The cash flow statement follows the cash transactions and recorded them. This provides important evidence of the wealth of the company, by revealing the sources of cash flow. The first line represents the net income from the report of income statement for the same period.

The cash flow statement separates cash from operations, financial and investing activities, this allow an investor to see if a good report of income statement, is due to the operations of the company or to a single event that will not repeat, like cash flow from a sell of fixed assent. The section of financial activities list the cash flow realized or spends with issue or acquired debt and equity.

| In Millions of USD (except for per share Items) | 12 months ending 2013-06-30 |
|---|-----------------------------|
| Net Income/Starting Line | 21,863.00 |
| Depreciation/Depletion | 3,755.00 |
| Amortization | - |
| Deferred Taxes | -19.00 |
| Non-Cash Items | 4,609.00 |
| Changes in Working Capital | -1,375.00 |
| Cash from Operating Activities | 28,833.00 |
| Capital Expenditures | -4,257.00 |
| Other Investing Cash Flow Items, Total | -19,554.00 |
| Cash from Investing Activities | -23,811.00 |
| Financing Cash Flow Items | 199.00 |
| Total Cash Dividends Paid | -7,455.00 |
| Issuance (Retirement) of Stock, Net | -4,429.00 |
| Issuance (Retirement) of Debt, Net | 3,537.00 |
| Cash from Financing Activities | -8,148.00 |
| Foreign Exchange Effects | -8.00 |
| Net Change in Cash | -3,134.00 |
| Cash Interest Paid, Supplemental | 371.00 |
| Cash Taxes Paid, Supplemental | 3,900.00 |

Figure 14 - Annual Cash Flow Statement of Microsoft

3.4 Financial ratios

From the financial statements information is calculated the financial ratios that are quantitative measures that allow analyzing a company in terms of profitability, liquidity, debt and growth. They are used to compare companies inside the same industry and to draw conclusions about the best companies to investment. Then will present the ratios used in this work as well as its desired function for select a company.

3.4.1 Debt ratio

An investor can find what appear to be good company to invest because in income statements is verified good results, including *net income* increasing with a high *Profit Margin*, and the company are traded in the market with a low PER ratio. This analyse draw the conclusion that is a profitable company that are undervalued. This is the problem of leverage, makes a company appear more profitable than in reality it is, and that can have competitive advantage.

To filter this type of company is used the *Debt Ratio*, companies that have high debt compared to its assets are at greater risk of going bankrupt in the occurrence of an adverse economic cycle, or by suffer a reduce in their profits by increasing interest rates of their debts. Usually companies with low *Debt Ratio* are in extremely competitive industries with constant need for innovation and updates of its products and production processes, which are carried out using external financing.

$$Debt\ ratio = \frac{Total\ Debt}{Total\ Assets} \quad (26)$$

By analyzing the *Debt-ratios* of some companies in table 7 can be verified differences in the capital structures from one company to another. The companies valorization with different debt ratios in Figure 15, demonstrates that it is need to take more considerations than the capital structure to predict the companies with higher potential of return.

| Debt Ratio | | | |
|------------|------|------|------|
| Stock | 2011 | 2012 | 2013 |
| JNJ | 0,47 | 0,50 | 0,45 |
| KO | 0,62 | 0,60 | 0,57 |
| MSFT | 0,45 | 0,45 | 0,47 |
| AAPL | 0,40 | 0,33 | 0,34 |
| ADI | 0,26 | 0,26 | 0,28 |
| ADP | 0,81 | 0,80 | 0,82 |
| AMZN | 0,63 | 0,69 | 0,74 |

Table 7 - Debt - Ratios

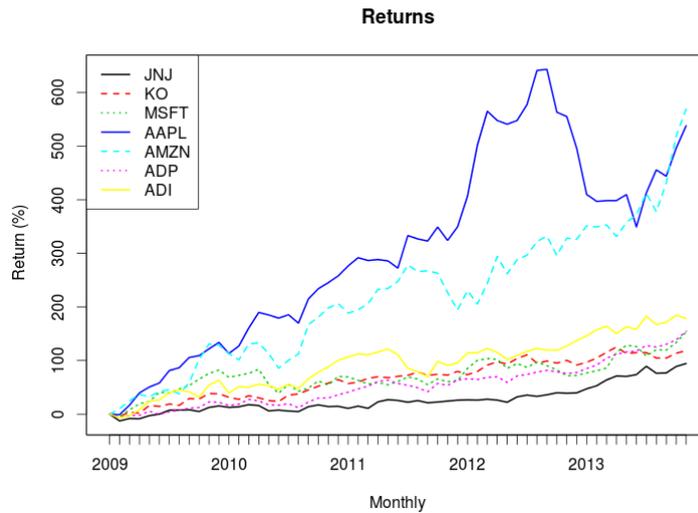


Figure 15 - Compound returns of companies of S&P500

3.4.2 Return on Equity

Return on Equity is an accounting measure of the performance of the company in generating profits. This ratio is used by investors to select companies that maximize the investment made in them, this means for every dollar invested are capable of creating a net profit greater in terms of percentage of capital invested. If managers retained the earnings they will be reinvested in the business with a higher return rate, than if distributed has dividends by the shareholders. The ROE is obtained by three factors, the operational efficiency, the efficient use of assets, and the financial leverage, this can be seen in equation 27. The objective is choose companies with higher ROEs, where these value are obtained by the *profit margin* and *asset turnover*, and not obtained through excessive leverage (Hillier, Ross, Randolph, Jeffrey, & Jordan, 2010). This is the type of companies that have competitive advantage, normally monopolistic in their sector of activities, and are traded with high per share earnings.

$$ROE = \frac{Net\ Income}{Total\ Equity} = Profit\ margin \times Total\ asset\ turnover \times Equity\ multiplier \quad (27)$$

In table 8 are the values of ROE of some companies from different activities sector, can be noticed that companies that belong to the same industry as is the case of Apple and Microsoft have similar ROE.

| Stock | Average's 5 Year ROE |
|-------|----------------------|
| JNJ | 23,2 |
| KO | 31,1 |
| MSFT | 36,9 |
| AAPL | 36,3 |
| ADI | 18,2 |
| ADP | 23,1 |
| AMZN | 16,6 |

Table 8 - Average 5 years of ROE

3.4.3 Profit margin

This ratio measures the profitability of the company, calculating the percentage of the revenue retained after paying the operating, administrative, financial costs and taxes.

$$PM = \frac{Net\ Income}{Revenue} \quad (28)$$

When analyzing the evolution of the *Revenue* with the *Profit Margin* it is possible to make valid predictions about the future of the company. As observed in Figure 16, the growth trend of the PM along with the growth of Apple's revenue from 2009 to 2012 is reflected in the stock market with a valorization over 400% for the period in question, having been one of the biggest valorizations in Index S&P 500. This is a method to find the companies with the best businesses that have durable competitive advantage, allowing the increase of the profit margin at the same time with the increase of revenue (Schroeder, 2007).

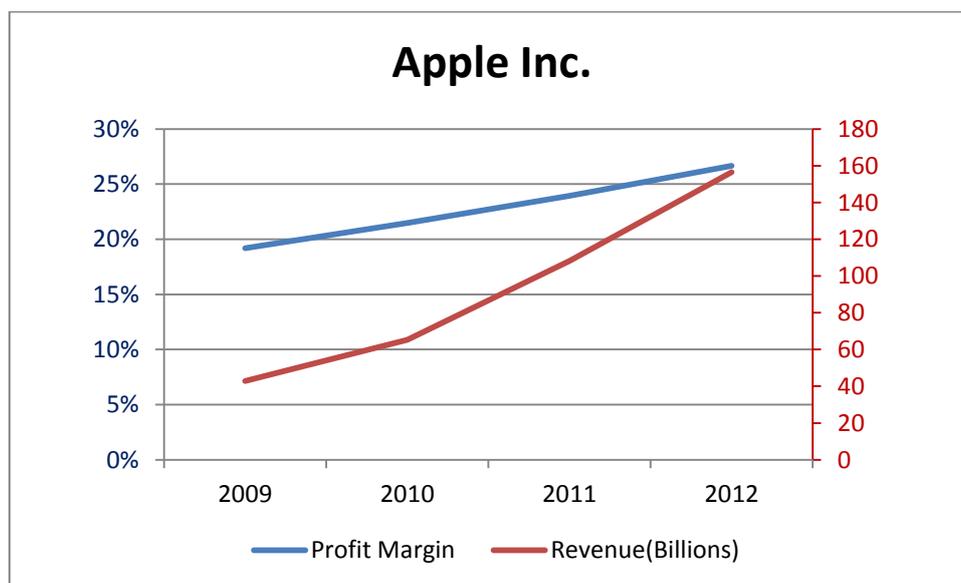


Figure 16 - Comparing Profit Margin and Revenue of Apple Inc

3.4.4 Price earnings ratio

Price earnings ratio established a relationship between the share price and the company's profits, is used to measure company levels of overvalued or undervalued. PER tends to be lower for companies such as slow growers and higher for fast growers because in the share price is incorporated investor expectations regarding to the future. The best use of the ratio is to find the cheapest company in a sector, but can be used to compare companies in different business sectors. It is possible to compare the historical record of PER in order to get a better perspective on the levels of overvaluation and undervaluation of each stock, avoid higher PER ratios is usually good investment policy, as it normally signal overvalued companies. For a company with a PER of 30 or greater it is necessary in the coming years achieve high rates of profit growth to be a profitable investment. This means that an investor are available to pay a higher price per earnings if expects higher earnings growth.

a) Price/Earnings to Growth (PEG-Ratio)

The PEG-ratio is a variation of PER that take in account the growth of earnings in the valuation of the company, meaning that a PEG below one signifies undervaluation. It is a good policy calculates the two ratios to have a better picture about the company, in terms of undervalue of the respective earnings, and in terms of growth.

$$PEG = \frac{PER}{Annual\ EPS\ Growth} \quad (29)$$

b) Relation between P/E ratio and ROE

Using Dividend Discount model of the equation 5 can be proved that the P/E ratio increases with the ROE. Considering the dividend is equal to earnings less a rate of reinvested earnings R_i in the company, the dividend is given by equation 30.

$$D_1 = EPS_1(1 - R_i) \quad (30)$$

The growth rate g is given by the rate of reinvestment of the earnings times the return of equity of the company (ROE).

$$g = R_i \times ROE \quad (31)$$

Substituting equation 30 and 31 in the equation 5 it is get the equation 32 that give the intrinsic price of the share for a company with constant growth rate. Equation 33 relates the PER to the ROE, by demonstrating that companies with goods ROE have higher PER.

Companies with higher reinvestment rates will have higher growth rates and this is reflected in the PER like said previously. It is common that high R_i , happened when the ROE of the company exceeds the discount rate, meaning that it is better for the shareholder that the company reinvest its earnings in the business, or in a new project, or buyback shares, because the money received as dividends when

reinvested at the better rate of return that the investor has is lower than return of capital given by the company.

$$P_0 = \frac{EPS_1(1 - R_i)}{R - ROE \times R_i} \quad (32)$$

$$PER = \frac{P_0}{EPS_1} = \frac{(1 - R_i)}{R - ROE \times R_i} \quad (33)$$

c) The P/E Market

The PER of the stock market is calculated by the aggregation of all PERs of the companies, it is indicator used to determining whether the market is overvalued or not. Interest rates influence the overall value of the market, and it's PER, low interest rates increase the real returns of future years, thereby contributing to an increase in investors' expectations, and causing a valuation of general market. In Figure 17 is represented the mensal values of EPS an prices of the index S&P 500 from 1988 to 2013, it can be seen that in the years of 2001 and 2007 when exist a major difference between the two trends, the market start a correction. Meaning that when the markets is sob valorized in relation to its earnings, or a bear market occurred, or a higher growth of the economy need to happen to maintain the higher prices.

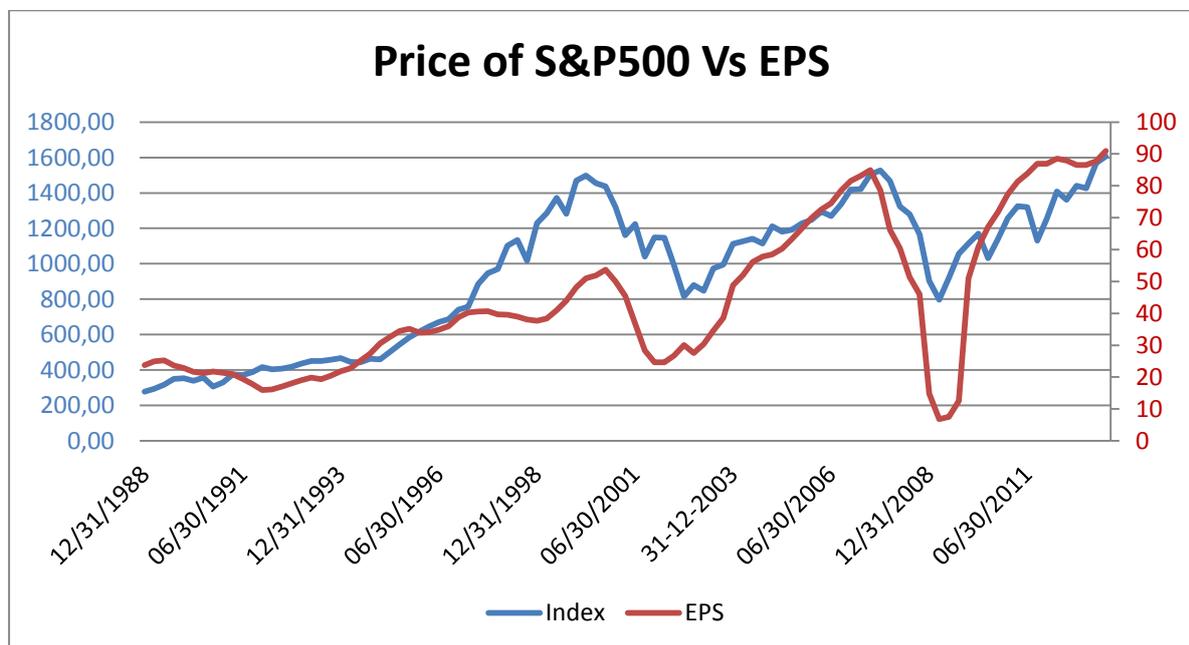


Figure 17 - EPS of the Market S&P 500 vs Index Price

3.4.5 Revenue growth rate (RG)

The percentage of revenue growth is an economic indicator that shows the evolution of the business, two factors are responsible for its increases, the company is a better competitor and is gaining market share to other competitors, or the company is inserted in a fast growing sector and are growing with it.

Although this indicator is important to analyze the company from the growth point of view, not give any indication of the profitability of the company. There may be a case where a company has a strong growth in revenues, but due to other factors the net earnings show a decrease, as is the case in recent years of Amazon.

$$RG = 100 * \frac{Revenue_{Actual} - Revenue_{Last Year}}{Revenue_{Last Year}} \tag{34}$$

In Figure 18 can be seen the case of amazon with revenue increases and share price, when other fundamental as the net income not improved.

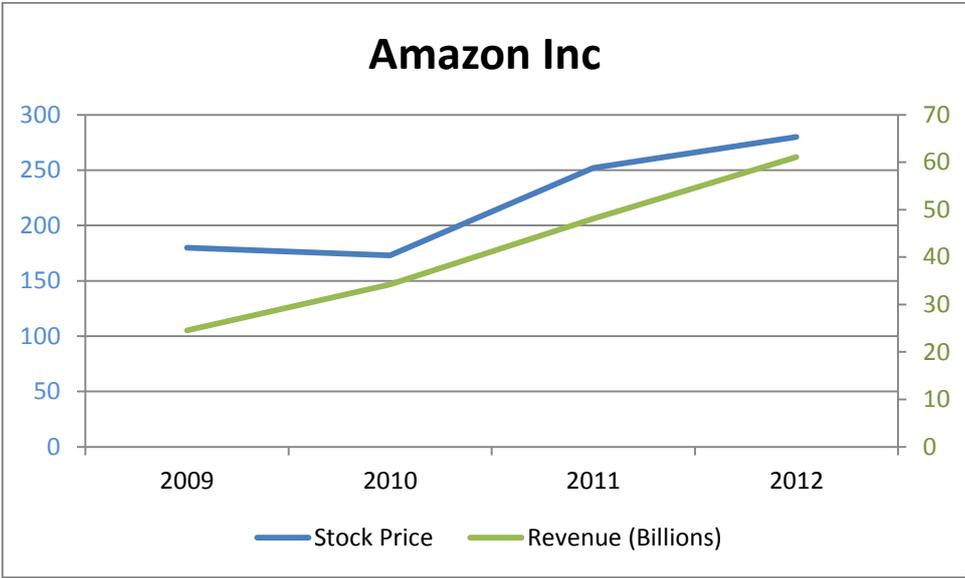


Figure 18 - Revenues vs stock price of Amazon

3.4.6 Rate of change in common stock outstanding

Common stock outstanding represents the fundamental ownership hold of the corporation by the shareholders, considering also shares held by the company managers. When the company issues shares this number is added to the previous value in the *Balance sheet*, representing an increase of total number of shares and a distributing of the company's value by a greater number of shares. The monetary value obtained with the selling of these new shares is registered as cash inflow in section of *Issuance (Retirement) of Stock, Net* in the *Cash Flow Statement*. In the case of company repurchases its shares the amount spent is recorded as cash outflow in the same section, this actions removes from circulation shares, and this reduction will be recorded in the *common stock outstanding*.

The reduction of the number of outstanding shares represents an increase in EPS and a decrease of PER ratio. The shares buy-back can be a sign that managers, who are better informed than the other

participants about the company's business and its future, are considering that the company is undervalued and are optimistic about the future of company.

The objective when using this indicator is to find companies that in recent years has repurchased its shares, thus reduced the number of shares outstanding without its stock price has suffered major change as demonstrated by Figure 19.

$$\Delta Common_{stocks} = \frac{\text{common stock outstanding}_{Actual} - \text{common stock outstanding}_{Last\ year}}{\text{common stock outstanding}_{Last\ year}} \quad (35)$$

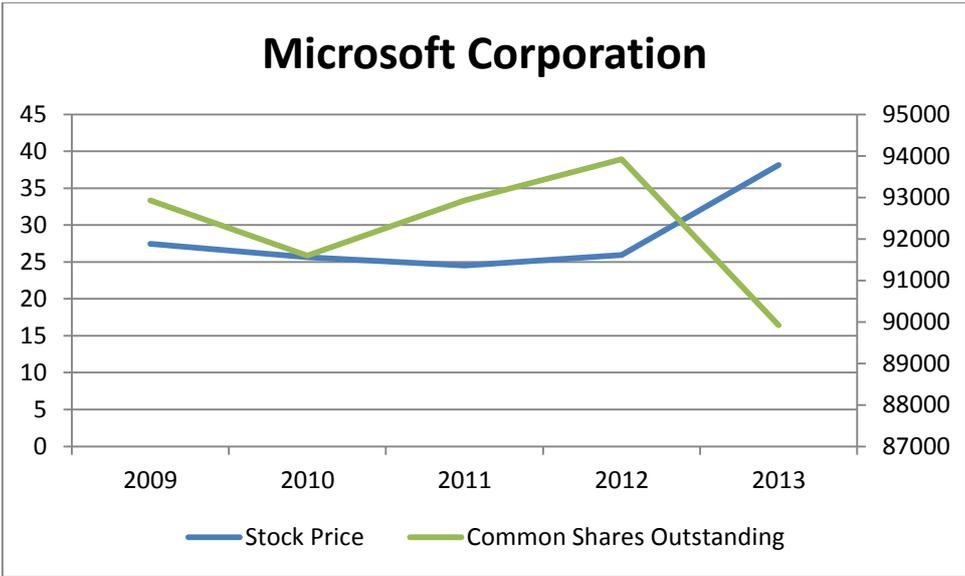


Figure 19 - Common shares Outstanding of Microsoft

3.4.7 Net Income Growth rate

A positive trend of net income shows that the company has consistency and durability in the profits, and a good result obtained in one year, are not just a result of favourable economic growth of the GDP or a financial engineering. When comparing the trends of net income and stock price, divergences can be found between both, and can be an indication of undervaluation of the company, if the net income trend shows a positive tendency and the stock prices do not show a similar behaviour. As observed in Figure 20 the Wal-Mart net income growth in the years from 2007 to 2012, but the stock price between 2008 and 2010 decline from their top.

$$\Delta Net\ Income = 100 * \frac{Net\ Income_{Actual} - Net\ Income_{Last\ Year}}{Net\ Income_{Last\ Year}} \quad (36)$$

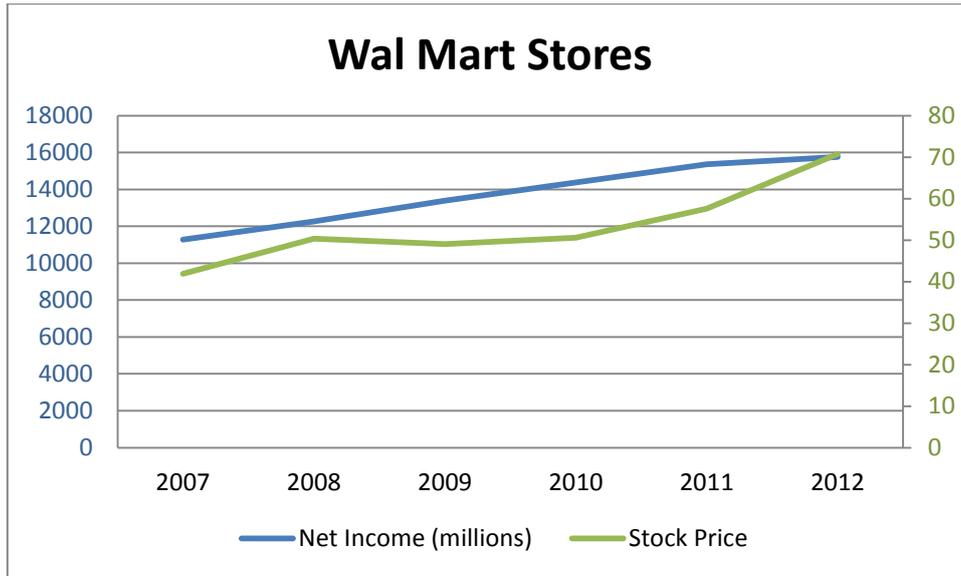


Figure 20 - Stock Price and net income of Wal-Mart Stores

3.5 Technical analysis

The function of technical analysis is predicting patterns in stocks prices and exploring it. The technical analysts believe that prices are gradually close of its intrinsic value and they use volume data as well as price history to predict the market. One of the principal applications of this analysis is the use of moving averages to find trends in the market prices.

3.5.1 Moving average

There exist a great number of different types of moving averages, depending of the calculated method, but it functions and interpretations are the same, in this work only the simple moving average (SMA) are discussed and utilized in the Evolutionary algorithms.

Moving average is a technical indicator composed by series of numbers that represent average value of a security's price, over a set period. The values are the means from several consecutive subsets from the set selected by the period size of the MA, as defined by equation 37. To calculate the mean values can be used the close, high, low, open and Adj.Close prices. The Adj.Close prices are used by the algorithms designed for calculating the SMA for the stocks of the Index.

$$SMA_t = \frac{P_t + P_{t-1} + \dots + P_{t-n}}{n} \quad (37)$$

The MA function is to eliminate noise in the stock price, thus allowing the analyst to better determine the trend. MA is also used in trading systems as trigger points to enter and exit the market. It is common to calculate MA of the volumes of transactions to filter the noise and thus have a better perception of activity in the assets under study.

The MA's more used by traders are the SMA with the time periods of 200, 150 and 21 days, calculated on the close daily prices (Matos, 2007).

The choice of the period is the most important factor in defining this indicator, since with a shorter period may incorporate an amount of noise still very high, and this will result in many more trades done by the trading system, while a longer period eliminates this inconvenience, but arise other problem, the entry point can be too late, and this mean lost a large part of the price movement. As such the best period choice depends more of the strategy used, the time frame of the trading system, and the financial instruments to trade.

There are two methods used in systems with MA's as trigger for the trade, the first method is the intersection between the MA and the price, in the case of open a long position, the market entry occurs when the price crosses the moving average from below to above. The second method is used when two MA crosses. The signal is given when the MA with shortest period cross from below to above the MA with longest period, and vice versa in the case of a short position.



Figure 21 - Trend of SP&500 with two SMA of 200 and 150 days

3.6 Conclusions

In this chapter were described, the architecture of the system, how to get the financial data used. It is done an explanation of each financial ratio used, its demonstrated why companies with good ROE show higher P/E Ratio, how to use PM ratio with increase of revenue to find the best business and to select a stock. It is explained the technical moving average indicator, and the use of it in a trading system.

4. Multi-Objective optimization

This chapter describes the Multi-objective approach used in this research. It will be presented in section 4.1 two chromosomes representing the investment models used, they have some differences in the trading systems, and thus it will be explaining the functionality of both. In section 4.2 is described the methods for making the evolution of populations namely crossover and mutation methods. Section 4.3 describes the operation of an external file to retain the best individuals found, in each iteration. The description of genes used in both chromosomes models that define the trading system will be presented in chapter 4.4. The restrictions used in the investment models are explained in section 4.5, the simulator is described its work in section 4.6, and the measures for evaluation of the portfolios in section 4.7.

4.1 Chromosome structure

The representation of the chromosomes is similar for the two investments models, where each individual of the population is composed of a sequence of values called genes.

The Chromosome is divided in two parts, the first group is the financial ratios weights and the second one is the trading parameters.

4.1.1 First Model of the Chromosome

This model only invests in companies that have some potential for a good return, using the first parameter of trading, the gene of *Min value for Portfolio* optimized, to filter the candidates to enter in the portfolio.

This gene defines the minimum valuation of the stock that a candidate needs to have to enter in the portfolio.

The first seven genes are financial ratios weights. The remaining genes define the behavior of the trading systems, in Figure 22 is represented the model of the chromosome. The seven weights are the DR, ROE, PM, PER, GR, Variation of Commons Stock outstanding, and Growth in Net income. It uses five parameters of a technical trading system the stop loss, take profit to determine the exit points, the number of days to use in a SMA for trigger the entry point, and a size position to define the percentage to allocate to each stock. In Figure 23 is represented the chromosome part that refers to the trading parameters. It is possible with this chromosome to have a lot of different strategies depending of the values of the genes corresponding to the parameters of trading. The system can have a more diversified portfolio, by using a lower value of position size, where its minimum value to allocate at each investment is 5%, this allow the portfolio to be composed with a maximum number of 20 stocks .

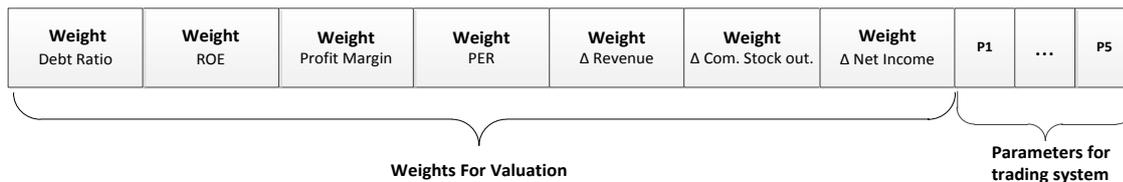


Figure 22 - Representation of the Chromosome

The combination of Stop Loss and Take Profit, define the profits and loss of each investment, the time horizon of each investment, and the accuracy of the system. For example a stop loss of 20% and a take profit of 50%, this is a system with a long time horizon, if the evaluation and criterion for enter in the market are good, it will can have an accuracy bigger of 50%, because the noise in the market not stop out many positions. But a system with the same take profit of 50% and a stop out of 5% will have lower accuracy, but when win a position can compensate for a great number of losses trades. Other type of strategies possible is a short-Term Trading, with a high number of transactions, due to having a short Stop Loss and Take Profit, for example can have both a value of 5%, this is translated in the market by a great number of trades, and if the accuracy is a 55% percentage or higher, is a system very profitable, with a low risk (Tharp, 2007).

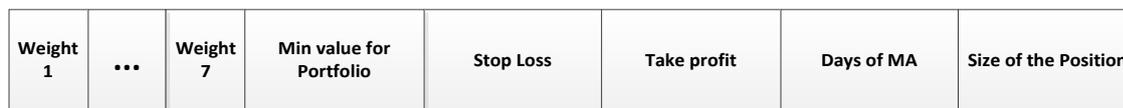


Figure 23 - Chromosome that represent the first model

4.1.2 Second Model of Chromosome

This model was developed to simulate a portfolio that always keeps a determined number of stocks in it. It has the first seven genes equal to the previous model, and the gene of the position size, and adds a new one, the limit number of stocks in the Portfolio.

The model operation is to maintain the stocks in the portfolio which have highest valuations, when one stock outside of the portfolio has a better valuation that the worst in the portfolio, this will be replaced by the first. In Figure 24 is represented the chromosome.



Figure 24 - Chromosome representing the second model

4.2 Reproduction Process or Evolution Process

For doing the process of evolution in the population are used crossover and mutation methods in each iteration, to the selected chromosomes.

4.2.1 Crossover

Crossover consists in recombine genetic parts of the chromosomes in the population selected in the mating selection, in order to create new and improved chromosomes. Next are presented the two methods used in the algorithms.

a) First Method of Crossover

This method consists in choosing two chromosomes randomly from the mating pool, and recombines them to generate two new chromosomes. It is divided each chromosome into two equal parts, the first six and last six genes of the chromosome, the child's are produced by joining the set of first part of the first chromosome with the second part of the second chromosome, to create the first child, and to the second child is the inverse process of selecting the parts. In Figure 25 is represented an illustration of the process.

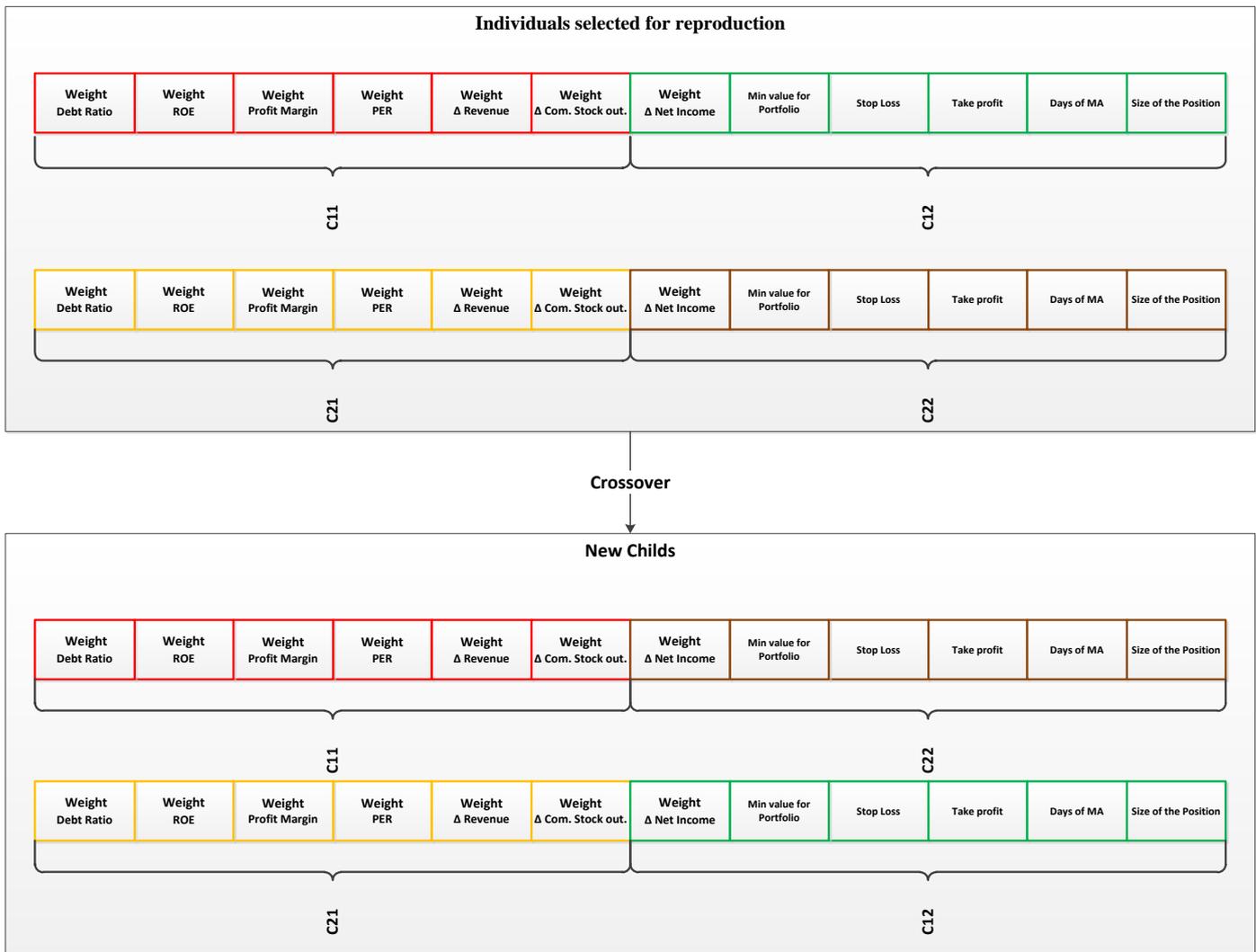


Figure 25 - Crossover Method using two cromossomes

This method was proved to be inefficient, since the population was very slow to evolve in terms of time resource, requiring a high number of iterations to generate meaningful population learning.

Considering that 20 iterations took about 24 hours of simulation, this method was only applied in the simulation of single objective in the first model to do a pre-training of the population.

b) **Second method of Crossover**

Due to problems identified in the previous method, it was necessary to improve the ratio of population learning by iteration for each simulation, so with a smaller number of iterations it is possible to obtain a better evolution. To achieve this it is selected from mating pool four chromosomes randomly, to provide their genes for a new chromosome, as illustrated in Figure 26. This method has the advantage to accelerate the evolution of the population and increasing its diversity, but it is necessary produce a sufficient number of new individuals, what means repeat the process in the same iteration more times that the method before.

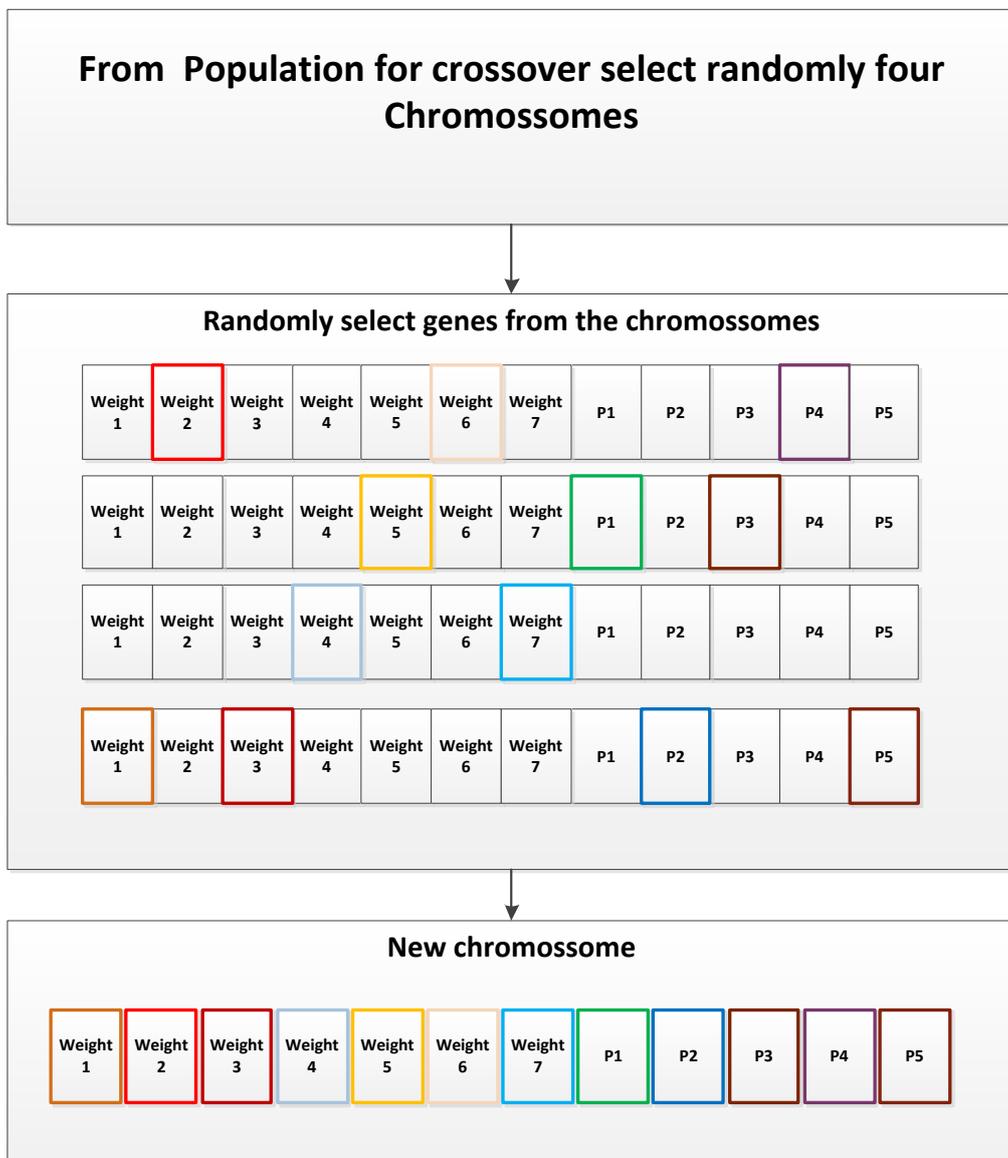


Figure 26 - Crossover Method with four parents a one Child

4.2.2 Mutation

Mutation is an operation performed on individuals randomly selected of the new population generated in the crossover and from the external file. It is an operation that creates a new individual by copy the selected chromosome to the mutation, and modifying some genes according with a mutation rate, that is randomly and limited inside one interval.

Next will be explaining the two methods applied in the algorithms to carry the mutation process.

a) First method of Mutation

This method choose randomly four genes from the chromosome, and each one is multiplied by a random mutation rate limited between 0.7 and 1.3, as illustrated in Figure 27.

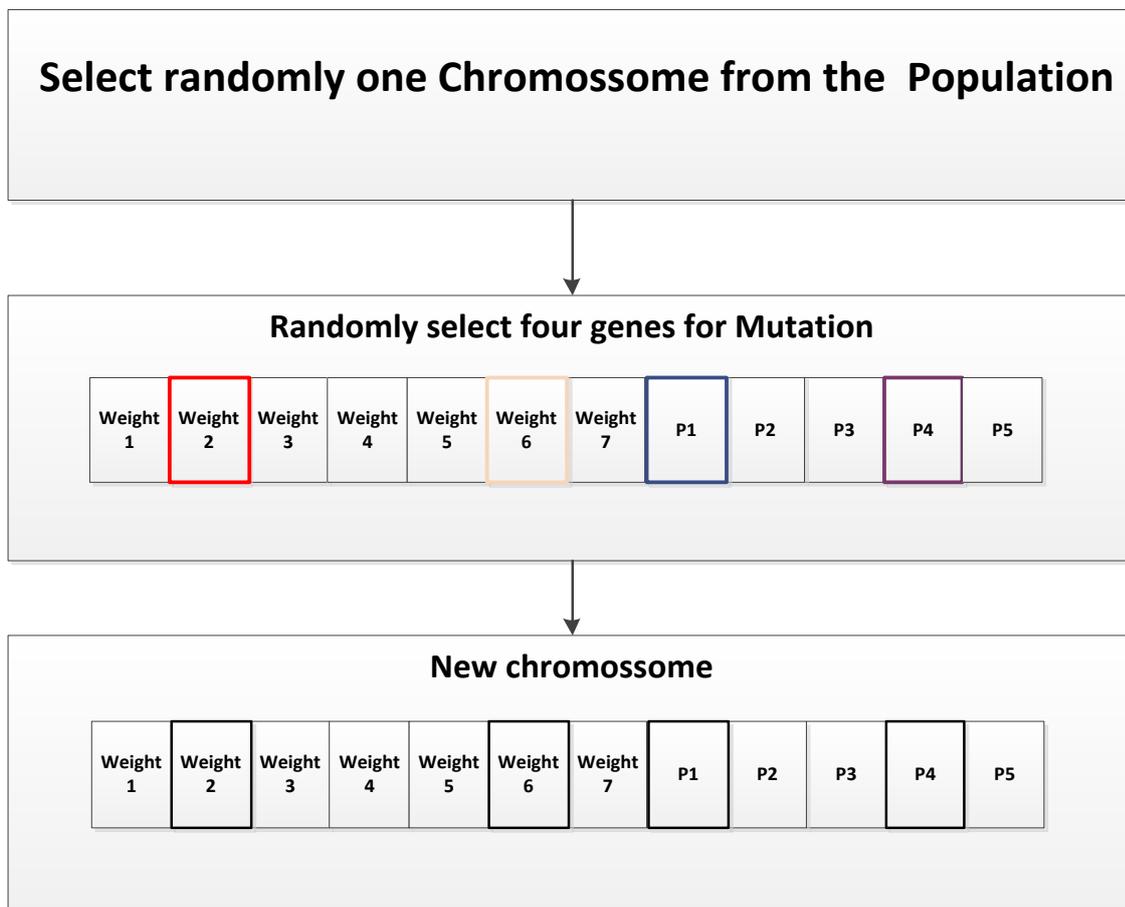


Figure 27 - Mutation Operation with four genes mutated

b) Second Method

After using the first method, it was observed that most of the mutated chromosomes, after evaluation not have conditions to belong to the external file due to poor evaluation. Because changing four genes normally improve one and deteriorate other, this not creating a better child and the population not evolves. Another method has been implemented which is similar to the previous one, but the number of genes to be mutated has become random, being limited between one to four genes, and the mutation rate is changed to the interval between 0.5 and 1.5.

4.3 External file

The external file holds the population composed by non-dominated solutions found along the search process, or in case of algorithm using single objective, keep in the file the chromosomes with higher valuation of the fitness function find in the training.

a) File Logic for Multi-objective Optimization

The function of the file is to record the individuals non-dominated and update them at each iteration of the training. For enter in the file each new individual is compared with all individuals in the file, to check if it dominates any member, the elements dominated are deleted from file and the individual is inserted, but if at least one element of the file dominates the new individual in consideration, it will not be added. In the case of an individual non-dominated by anyone in the file, it is added to the file.

In Figure 28 is represented the work of the file, in the *generation I* are recorded the last population produced from *generation I-1* and the solutions that compose the pareto set found. The process starts by compared the both sets, and selects the non-dominated chromosomes to go to the reproduction process (crossover and mutation) and update the file.

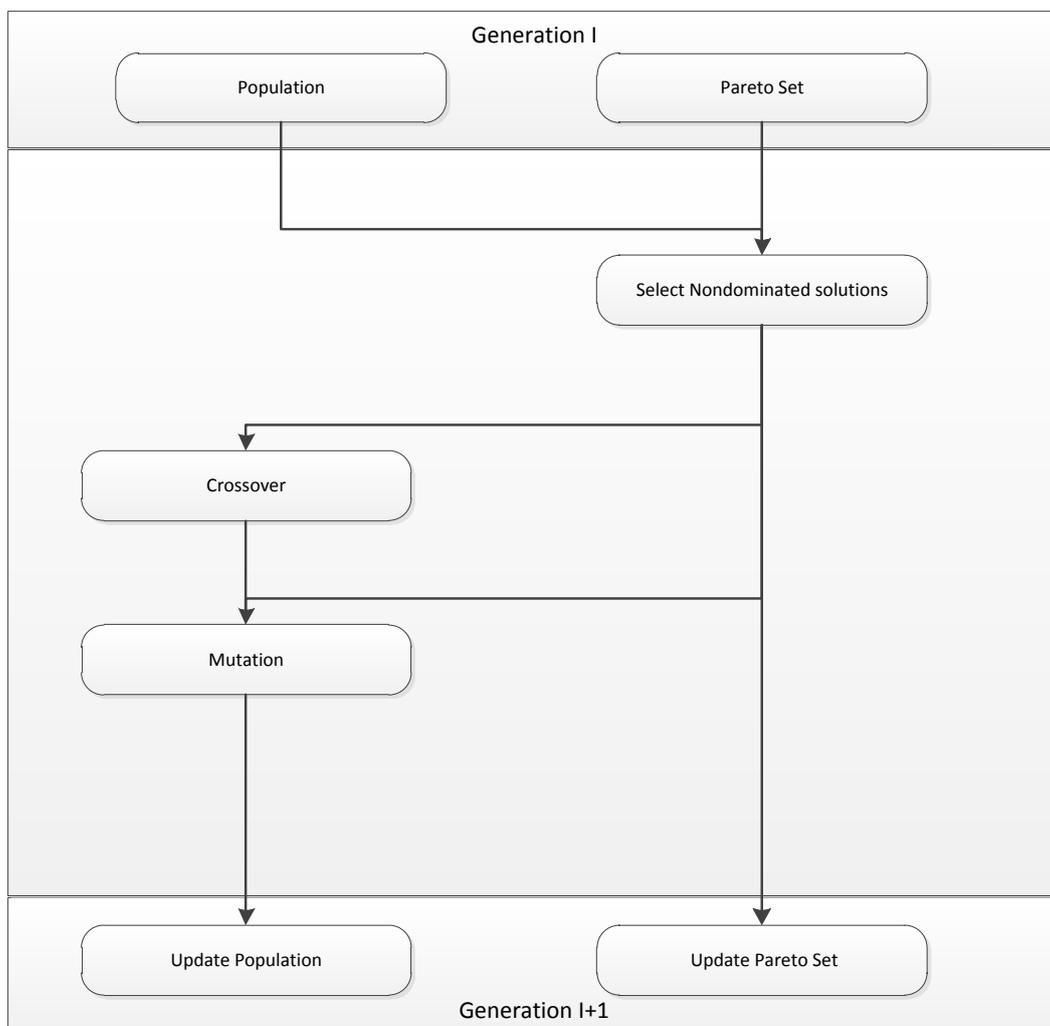


Figure 28 - Process diagram of the external file

b) File logic for single-objective Optimization

In case of one objective to optimize, or a function composed by two objectives like the Sharpe ratio the size of external file was limited in terms of space, for a determined number of solutions. Only enter in the file the best 80 individuals, and when is full if exists an individual with better fitness that the worst in the file, the worst will be deleted and the new one will be added to the file.

4.4 Parameters of the trading system

For defining the behavior of the trading system, like the entry and exit time, the money management rules, is used the part of the chromosomes called the parameters of the trading system, for the first model is used four genes as parameters, the Stop Loss, Take Profit, Position Size, number of days of SMA, and in the second model it's used the position size and the number of stocks.

4.4.1 Stop Loss

The gene corresponding to the stop loss percentage, defines the maximum loss for each investment, by defining the exit point, when price tendency not occurs as expected, the stop price is calculated using the equation 38.

$$P_{Stop} = P_{Entry} \times (1 - StopLoss) \quad (38)$$

4.4.2 Take Profit

The gene corresponding to take profit has the function of deciding the percentage profit gained for all investments made. The gene will determine the sale value of successful investments based on the purchase price of the shares. The sell price is calculated by the price of entry more the value of price change, determined by the gene, according to equation 39.

$$P_{profit} = P_{Entry} \times (1 + takeProfit) \quad (39)$$

4.4.3 Position Size

The position size gene determines the percentage of the current portfolio value to invest in each new stock, this define the level of concentration or diversification in the portfolio for each chromosome. For example an strategy that invest a low percentage in each investment, the portfolio can have a higher number of assets, if the value of position size is high, for example 20 % which is the maximum possible due to *Quantity constraint*, the Portfolio will be limited to 5 stocks. With a higher concentration, the portfolio has higher risk since it is exposed to a less number of unsystematic risks. The equations 40 and 41 are the implemented calculations in the algorithms to define the capital to invest.

$$\begin{aligned} \text{If } Capital_{available} > P_{zise} * (Balance_{portfolio} + Capital_{available}) \\ Invest_c = P_{zise} * (Balance_{portfolio} + Capital_{available}) \end{aligned} \quad (40)$$

$$\begin{aligned} \text{If } Capital_{available} < P_{zise} * (Balance_{portfolio} + Capital_{available}) \\ Invest_c = Capital_{available} \end{aligned} \quad (41)$$

4.4.4 Trigger Signal by SMA

This is a gene used to calculate the time period of the SMA to trigger the purchase action. Sometimes the share price of the company reached a point where is undervalue, but due to market volatility or general tendency of the market, the stock price can continue to fall or get in range during some time. To achieve a better timing of enter in a long position by the algorithm is used an SMA, with the time period optimized that avoids as many as possible of the two previous scenarios and enter in the market when a Bull tendency has been establishes. The operation is that when the stock price is above the moving average, this stock may enter in the portfolio.

The SMA was not used as trigger point in crossover as is more common in systems of trading, for two reasons, first may be cases where there is no capital available to make the purchase at the time, the second reason is the portfolio is full when the price crosses the moving average, but after can be sell one asset, and the stocks can enter in the portfolio.

4.4.5 Global Value input for portfolio

The Global Value of each stock is a valuation done daily using the financial ratios, to ranking the stocks for the algorithm make trading decisions, these values are calculated using the vector of weights of ratios of the chromosome multiplied by the respective ratios of the company, as demonstrated by equation 42.

$$Global_{value} = \sum_{i=1}^n weight_i * Ratio_i \quad (42)$$

a) Gene of the Min value for Portfolio

The *min value for portfolio* is the gene that defines a minimum score of global value for a stock to be accepted in portfolio, serves to ensure that only the better investments are taken when are at reasonable price.

The global value after being calculated for all firms, the companies with higher values have priority to enter in the portfolio, but only can enter if the value is higher that the value of the optimized gene. The objective of the optimization of this gene is to ensure that as the algorithm learns to choose the best companies by weighting the ratios, these companies will have better valuations.

In the Figure 29 is represented a flow chart of the process of valuation and order the stocks, this process is repeated for all stocks each day.

making a great diversification, because the investor have better knowledge about the company, its products, competitors, debt levels, and future long-term economic prospects of the business (Cunningham, 1998).

Quantity constraint was used to limit the gene position size, by putting a maximum and minimum value for the size of the position. The minimum limit is set to 5% of the portfolio value at the time of the transaction, and a maximum value is 20%. The lower limit is to avoid positions practically insignificant to the performance of the portfolio, and the upper limit is to avoid too much exposure or weight for any stock.

Long only constraint signifies that it is not allowed to perform short selling operations, meanings the weight invested in any stock is always positive.

It is a restriction used by a great number of value investors and institutional funds because the risks associated with the short selling it is considered higher than with long positions.

Transaction costs include brokerage fees, and bid-ask spreads, are costs occurred by doing the business. The transactions costs if are high, will affect the frequency of trading, by reduction the return of the portfolio (Hassan G. N., 2010).

In the first model is used a commission of 2% of the value of each transaction for incorporate real costs that happen in world of investments, but for the second model, the transaction costs are ignored, to allow the algorithm change the assets without restriction, and affecting to much the return.

4.6 Simulator

The simulator in each new day uses the adjusted daily closing prices of stocks, and the financial ratios, to rank in an increasing order the stocks with higher global value. The block of trading, make the transactions according with the investment model used and the values of the chromosome tested. Each transaction is recorded in the balance of the account, and it is calculated daily the monetary value of the Portfolio, to calculate the variance and return of the portfolio.

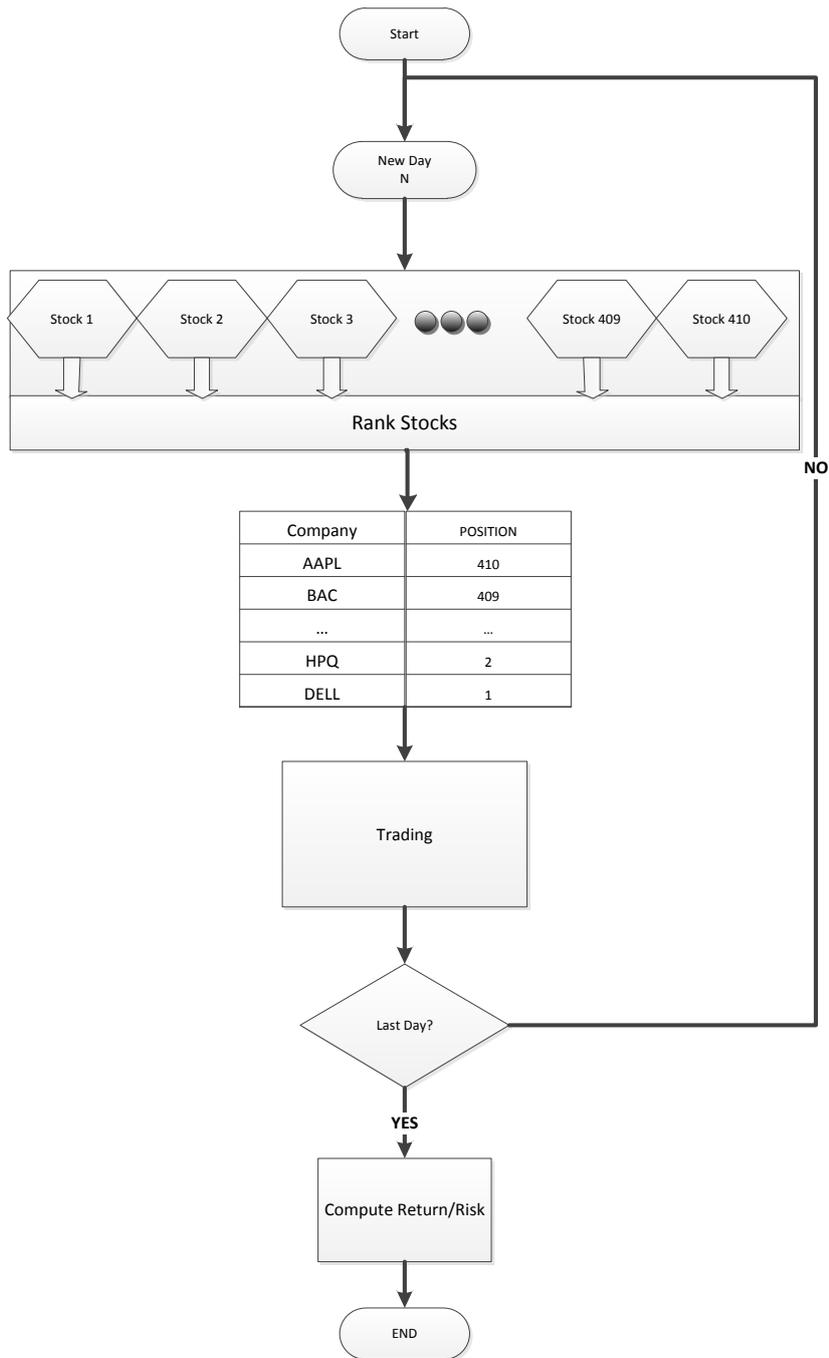


Figure 30 - Flow chart of the simulator of trading

4.7 Measures of Portfolio performance

The performances measures used to do the fitness of the solutions are the ROI, Mensal Variance of ROI, and the Sharpe-ratio. For the multi-objective simulations it is used the Pareto concept, with the ROI and Variance as objectives.

4.7.1 Calculation of Return of Investment

The return of investment in each simulation is calculated by the sum of the value of existing stocks (calculated by the current price of the stock P_i times the number of shares N_i), and the net cash at the end of the simulation, divided by the initial capital.

$$ROI = \frac{(\sum_{i=1}^n P_i * N_i) + Capital_{Last\ day\ of\ simulation}}{Inicial\ Capital} - 1 \quad (43)$$

4.7.2 Variance as measure of risk

The risk can be defined as the likelihood of deviation from the expected return of an asset in a time period. The dispersion or spread of a distribution measures how much a given return may deviate from their average, if dispersion is very high the returns that occur are very uncertain, and as such has a higher risk than a distribution with a lower dispersion, where returns are more accurate forecast.

The risk measure used is variance of the monthly returns of the portfolio that is calculated by using the equation 44 to calculate all monthly returns of the period, and uses them in equation 45 to obtain the variance.

$$Mensual\ Net\ return_i = (Return_i - Return_{(i-30)})/Return_{(i-30)} \quad (44)$$

$$\sigma^2 = \frac{1}{n} \times \sum_{i=1}^n (Mensual\ Net\ return_i - \overline{Mensual\ Net\ return})^2 \quad (45)$$

The risk definition described above can be seen in the Figure 31, the distribution function of monthly returns of Apple has a greater dispersion than the S&P 500, making the predictions of the returns more accurately for the Index.

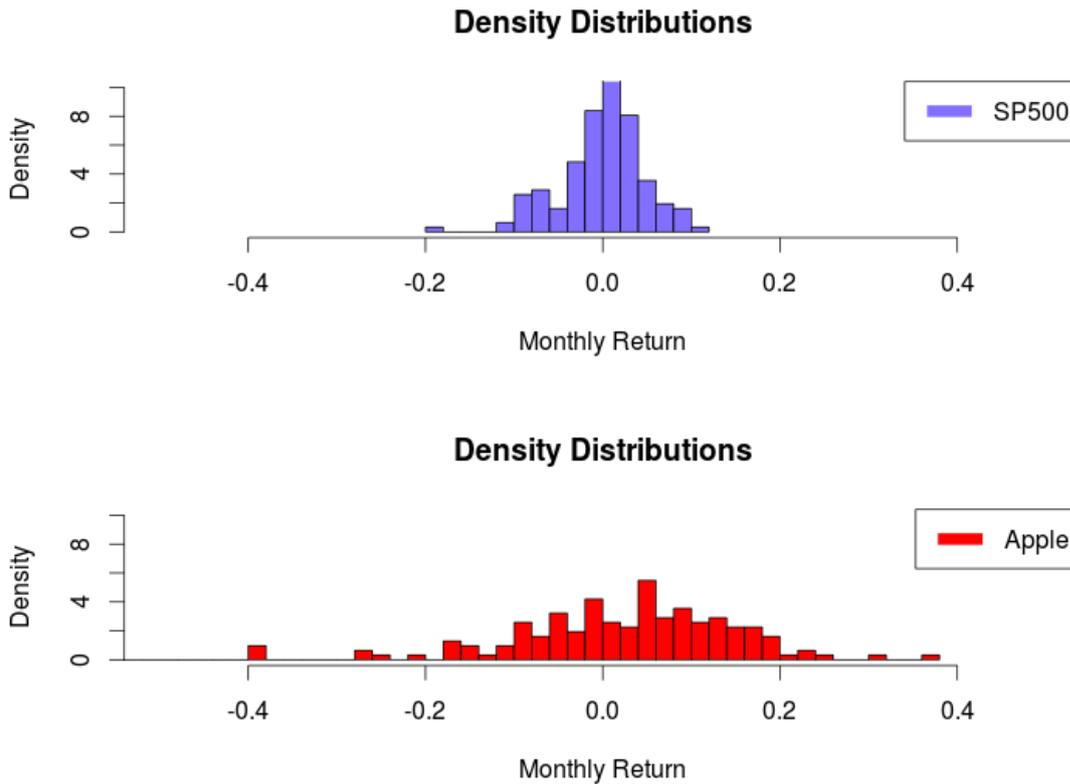


Figure 31 - Density Function of absolute returns of Apple and Index SP&500

4.7.3 Sharpe ratio

The Sharpe ratio was developed in 1966 by William F. Sharpe, is often used to evaluate the portfolios performance by pension funds and large investors. It is a simple measure of risk-adjusted performance, calculated by subtracting a risk-free rate to the return of the portfolio and dividing by the standard deviation of portfolio returns (Faith, 2007). It is a reward-to-variability ratio, or return per unit of volatility, it measures the return obtained by unit of risk used.

$$SR = \frac{R_p - R_f}{\sigma} \quad (46)$$

R_p – Mean Return of the Portfolio

R_f – Risk Free rate

σ – Portfolio Standard deviation

The Sharpe ratio analyze whether a portfolio's returns are due to smart investment decisions or a result of excess risk. This measurement is very useful because although one portfolio or fund can show higher returns, it is only a better investment if those higher returns do not come with too much additional risk. A negative Sharpe ratio indicates that a risk-free asset would perform better than the portfolio being analyzed.

a) **Risk Free-Rate**

The United States governmental Bond yield for 10 years observable on the graph of the Figure 32 shows for the period 2011 until 07.11.2013 a yield between 1.5 and 2.5%, these bonds are considered by the general investors as risk free asset. It was choose a constant rate of 2% in the risk free rate for calculating the Sharpe Ratio in the Simulations.



Figure 32 - Trend of 10 Year Bonds of USA

These bonds are considered risk-free because the main rating agencies give a score of triple A to this asset, so one of the safest assets in the world. They are issued in U.S. dollars and the risk of default by the government the USA is very small, since the North American Federal Reserve can create monetary mass for the government comply with its obligations. The real risk of holding treasury bills and bonds of USA to the investors is the depreciation of the dollar that can decrease the purchasing power of the assets in USA dollars.

4.7.4 Pareto Dominance Concept

The concept of Pareto dominance ($x_1 > x_2$), was used in the MOEA, to determine which individuals from the population that are non-dominated with respect to objectives ROI and variance, and thus can integrate the external file. The mathematical definition for the concept is given by the equations 47.

$$\begin{aligned}
 & x_1 > x_2 \text{ if,} \\
 & R_1 \geq R_2, \& Risk_1 < Risk_2 \\
 & \text{OR} \\
 & R_1 > R_2, \& Risk_1 \leq Risk_2
 \end{aligned}
 \tag{47}$$

R_i – Return of Portfolio i

Risk_i – Variance of Portfolio i

4.8 Conclusion

In this chapter was presented the models of investment used and the restrictions associated to them, it is demonstrated why a crossover method with random genes from four chromosomes is more reliability.

In the mutation process a bigger rate with random number of genes is better application that a one with a fixed number of genes. It is explained the trading system parameters, how implement a short and long term trading system, and the measures used to evaluate the performance of the portfolios.

5. Results

This chapter is divided in three parts, the first is section 5.1 where is presented the simulations performed by the first investment model, then in section 5.2 shows the experiences and results obtained for the second model. Section 5.3 presents the conclusions of the results obtained for the two models.

The training and real test simulations have been performed using 410 stocks of S&P 500 index, since these had complete data for the period under study. The Index is used as a benchmark for comparing the results obtained in the training and the real test. For all the simulations the algorithm starts with an initial capital of 100000 Dollars.

5.1 Simulations and Results of the first model

This section presents the simulations and the optimizations of single objective using the ROI, and the MOEA using the Pareto concept, for the model described in chapter 4 in section 4.1.1 with the constraints of section 4.5.

5.1.1 Single objective using ROI

In the simulation using ROI as objective to optimize, the population was trained during the time period from 2010-06-17 to 2012-01-03. The first training simulations performed show a tendency towards convergence of the population to a local maximum, in which the average return of the solutions found are very similar to the index return. To solve the problem, from iteration 30, it is changed the process of reproduction of the algorithm, from then on is used the second methods of crossover and mutation. In the end of the training the results obtained in terms of return and variance are shown in Figure 33, almost all solutions found has outperformed the index in terms of return.

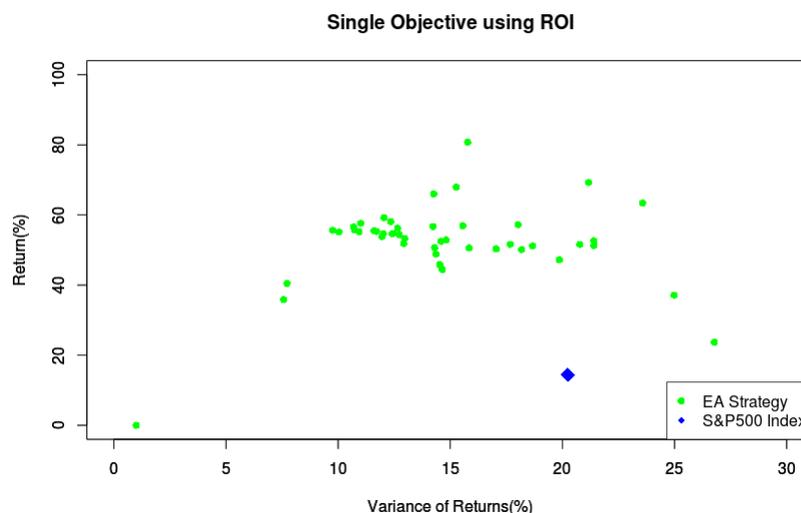


Figure 33 - Single Objective results obtained in training using ROI as Objective

a) Single-Objective Real Test

The population obtained from the training was used in a real test during the period between 04.01.2012 and 07.06.2013. The results show the returns of the chromosomes are in the same range of values of the training, thus demonstrating some adjustment between the training and the real test.

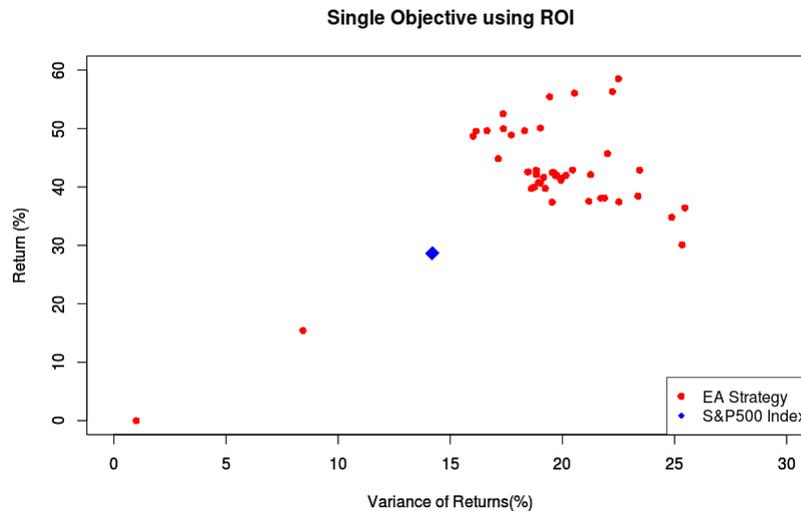


Figure 34 - Results of Real Test in Single Objective

It is concluded by analyzing the results that the objective ROI used in SOEA, allow to obtain a set of strategies with better returns using more risk than a passive management strategy that follow the S&P 500 Index.

5.1.2 Multi-objective using the Pareto Dominance Concept

In this experience the objectives to optimize are the return and the variance, where it is used the concept of pareto dominance in the environment selection for reproduction of the population, the training was conducted in the period from 2010-06-17 to 2012-01-03.

The curve in the Figure 35 obtained in the training has some discontinuities due to constraints, but by observing the results with the benchmark is concluded that set of solutions has better results for both objectives, and a great number of solutions dominates the Index.

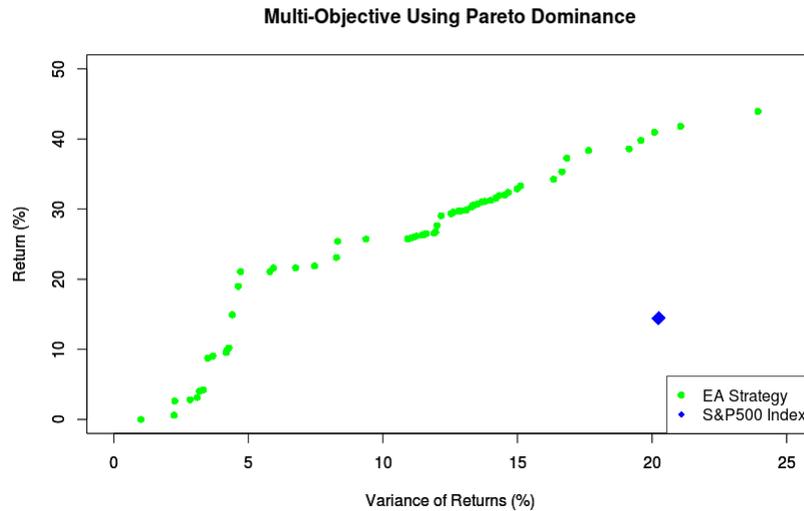


Figure 35 - MO results obtained in training using the Pareto dominance

a) **Multi-objective real test**

In figure 36 are the results of the population obtained from the training and used in a real test during the period between 04.01.2012 and 07.06.2013. These results have some differences from those obtained in the training in terms of shape of the curve, where now appear solutions that are dominated by others. The chromosomes demonstrate a capability to maintain a better performance than the benchmark, and the form of the Pareto curve is maintained.

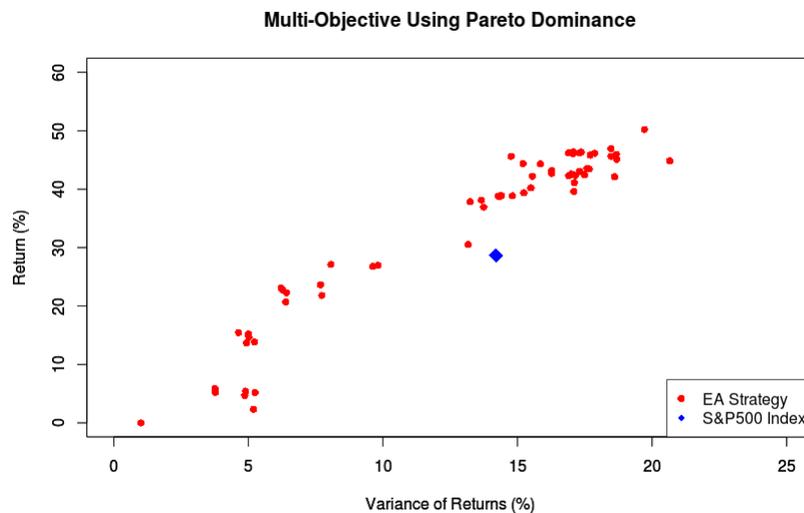


Figure 36 - Results obtained in real test

b) **Index compared with the results**

It was performed a comparison of the trend of S&P 500, and the trend obtained by a strategy that invest in all chromosomes of the population found in the training. By analyzing the figure 37 it can be seen the curve of the strategy have a better ROI, despite some effects as over fitting in training and the transaction costs occurred. It is possible to conclude that diversify the capital to invest in the

solutions found by the algorithm is a more viable strategy than a passive investment of buying a set of stock that are representative of the index.

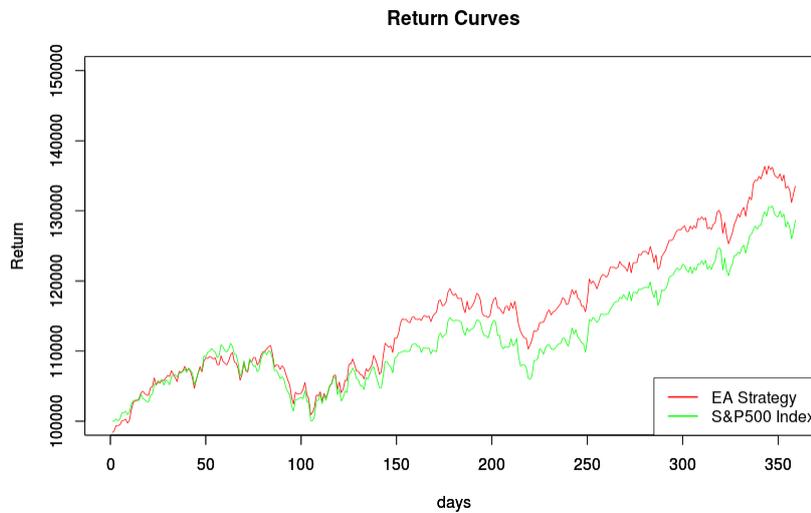


Figure 37 - Average return of the MOEA

c) Analyze of the Chromosomes

From this simulation were selected five results with different variances to analyze the chromosomes and get a better insight on the type of solutions found by the algorithm. Table 9 shows the results for the two objectives of the chromosomes selected.

| | Return (%) | Variance (%) |
|----------------------|-------------------|---------------------|
| S&P 500 | 28,685 | 14,20 |
| Chromosome ID | | |
| 1 | 50,24 | 19,71 |
| 27 | 45,83 | 17,70 |
| 36 | 38,81 | 14,28 |
| 51 | 13,66 | 4,94 |
| 54 | 22,73 | 6,25 |

Table 9 - Results of the real test

In Table 10 is presented the value of the genes for the same chromosomes. When analyzing the genes of the first chromosome it can be seen it has a preference for companies with high profit margins, earnings growth, and uses little diversification. The chromosome 54 is a strategy that allows greater diversification, it allocates 5% on each stock, the ratios with more importance are the ROE, and the net income growth rate. It is a more conservative strategy where it invests in established businesses with monopoly characteristics. The chromosome 51 is the strategy that has the higher frequency of trading, but obtained a lower return, due to the transactions costs. The analyses of the genes in the chromosome, demonstrates the higher returns are achieved by a better selection of companies with a higher concentration of the investment.

| Genes | Crom1 | Crom27 | Crom36 | Crom51 | Crom54 |
|---------------------------|-------|--------|--------|--------|--------|
| Debt ratio | 0,53 | 0,53 | 0,28 | 0,28 | 0,28 |
| ROE | 0,64 | 0,12 | 2,24 | 0,12 | 2,24 |
| Profit Margin | 3,01 | 0,61 | 0,49 | 0,96 | 1,01 |
| PER | 0,54 | 0,40 | 0,23 | 0,23 | 0,27 |
| Δ of Rev. | 0,34 | 1,13 | 0,62 | 0,65 | 0,57 |
| Δ common Stock out | 1,33 | 1,08 | 1,04 | 0,97 | 1,25 |
| Δ net income | 1,79 | 2,07 | 1,24 | 1,70 | 1,95 |
| Global Value | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| Stop Loss | 1,11 | 0,88 | 1,98 | 0,88 | 0,82 |
| Take Profit | 1,89 | 2,46 | 1,18 | 0,12 | 1,36 |
| Days of MMA | 1,00 | 5,00 | 4,00 | 1,00 | 1,00 |
| Position Size | 0,20 | 0,20 | 0,20 | 0,05 | 0,05 |

Table 10 - Genes of the Chromosomes

5.1.3 Multi-objective using the Pareto Dominance Concept with two years of training

In this experience was training an MOEA with the Pareto dominance concept between 17/06/2010 and 11/06/2012. The results obtained by the solutions found are in Figure 38, where exist some discontinuities in the curve, but the solutions found show some diversity along the curve.

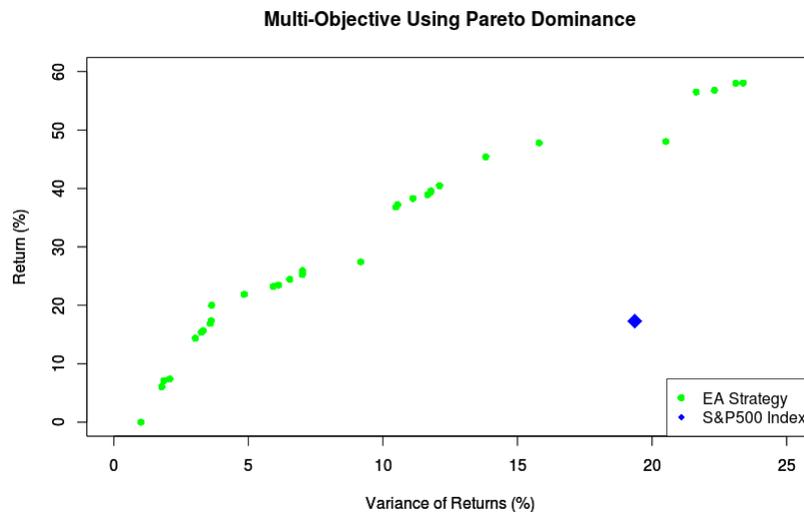


Figure 38 - Pareto curve obtained in training using the dominance of Pareto

a) Multi-objective real test

The real test performed using the population obtained before, from 12.06.2012 to 06.07.2013, are in figure 39. The results had a worse performance than the simulation after. This is due the over fitting effect, because the population was been trained in a period where most of the stocks were undervalued, and thereby the optimization of the gene of *min value enter in portfolio* put it with a higher value, than the *global value* for real test period of the stocks with better returns, and not allow them to entry in the portfolio. This happens because the stocks with higher valorization, in the training period, continue in the period of real test with higher valorization that the others.

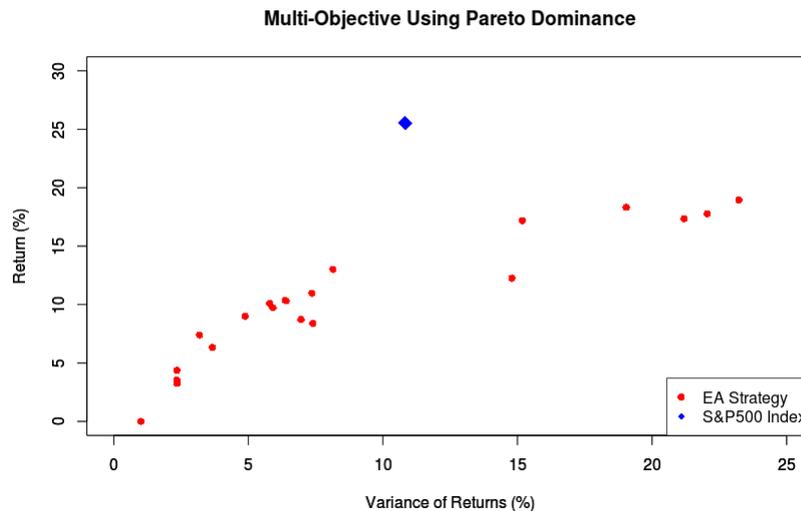


Figure 39 - Pareto curve obtained in real Test

5.2 Simulations and Results of the Second Model

In this section all the training simulations of the populations are done in the period 2010-06-17 to 2012-06-11 and the real tests are executed in 2012-06-12 to 2013-06-11.

It is used the second model, with Cardinality constraint, Quantity constraint, and Long only constraint, in the next simulations. The optimizations done are SOEA with the ROI, SOEA with the Sharpe ratio, and the MOEA with the dominance of pareto.

5.2.1 Single objective using ROI

SOEA was used for trained a population with the ROI objective, the results obtained in Figure 40 are in range between 60% and 80%, but with different degrees of risk. All solutions found outperform the index with a great difference in terms of return.

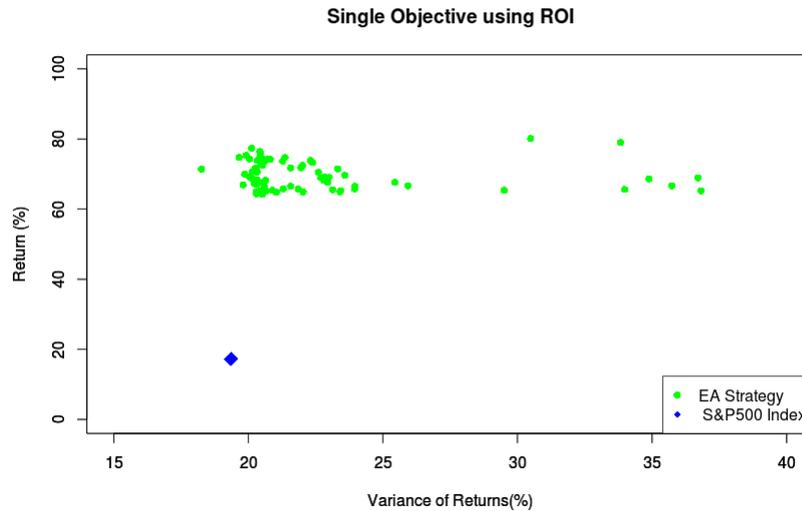


Figure 40 - Results obtained in training using ROI as Objective

a) Single objective Real test

By analyzing the Figure 41, it can be observed that the chromosomes have better results in terms of return and variance than the index, this is due the stocks selected to belong to the portfolio, and in general, they have a positive tendency with higher return than the index with less oscillation in the price trend. It is defined two accumulation zones of the results or clusters that are coincident with the results obtained in the training.

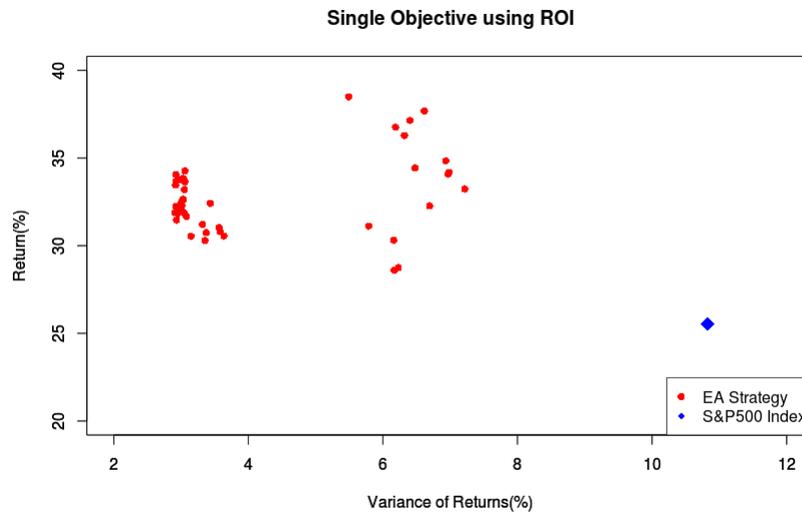


Figure 41 - Results of Real Test

b) Analyses of the chromosomes

As previously it was selected 5 chromosomes with similar returns but with higher difference in the variances, in order to draw conclusions about what can increase the risk of the portfolio.

| | Return (%) | Variance (%) |
|----------------------|--------------|--------------|
| SP500 | 25,55 | 10,82 |
| Chromosome ID | | |
| 29 | 34,07 | 2,92 |
| 47 | 28,75 | 6,23 |
| 72 | 32,59 | 3,02 |
| 79 | 37,69 | 6,62 |
| 80 | 36,76 | 6,18 |

Table 11 - Results of the real test

By comparing the genes and the results obtained is verified the differences comes from the companies selected to invest that are decided by the weights to calculate the *global value*. It is verified that chromosome 29 is able to invest in companies with lower risk than the others strategies.

From the analyzes is concluded to construct a portfolio where the objective is maximize the return, the ratios more important in selection of companies are the ROE, net income growth rate, and it is necessary to reduce the number of stocks in the portfolio to increase the concentration in the investments made.

| Genes | Crom29 | Crom47 | Crom72 | Crom79 | Crom80 |
|---------------------------|----------|---------|---------|----------|---------|
| Debt ratio | 0,00145 | 0,00189 | 0,00492 | 0,00040 | 0,00229 |
| ROE | 0,14888 | 0,08200 | 0,14888 | 0,14888 | 0,10070 |
| Profit Margin | 0,00002 | 0,00706 | 0,04570 | 0,00013 | 0,09422 |
| PER | 0,87995 | 0,87995 | 0,87995 | 0,76841 | 0,37111 |
| Δ of Rev. | 0,00351 | 0,07086 | 0,03352 | 0,03352 | 0,05643 |
| Δ common Stock out | 0,00186 | 0,00299 | 0,00911 | 0,00911 | 0,00911 |
| Δ net income | 1,49170 | 1,84461 | 1,69393 | 1,69390 | 0,80764 |
| Stocks Number | 20,00000 | 8,00000 | 8,00000 | 13,00000 | 8,00000 |
| Position Size | 0,20000 | 0,14353 | 0,20000 | 0,20000 | 0,20000 |

Table 12 - Genes of the Chromosomes

5.2.2 Single-objective using the Sharpe ratio as objective

The population in this simulation was trained using the Sharpe ratio as objective. The results obtained in Figures 42, show an increase of slope of the curve in relation to the increased variance, a behavior opposite to that obtained in relation to the Pareto curve.

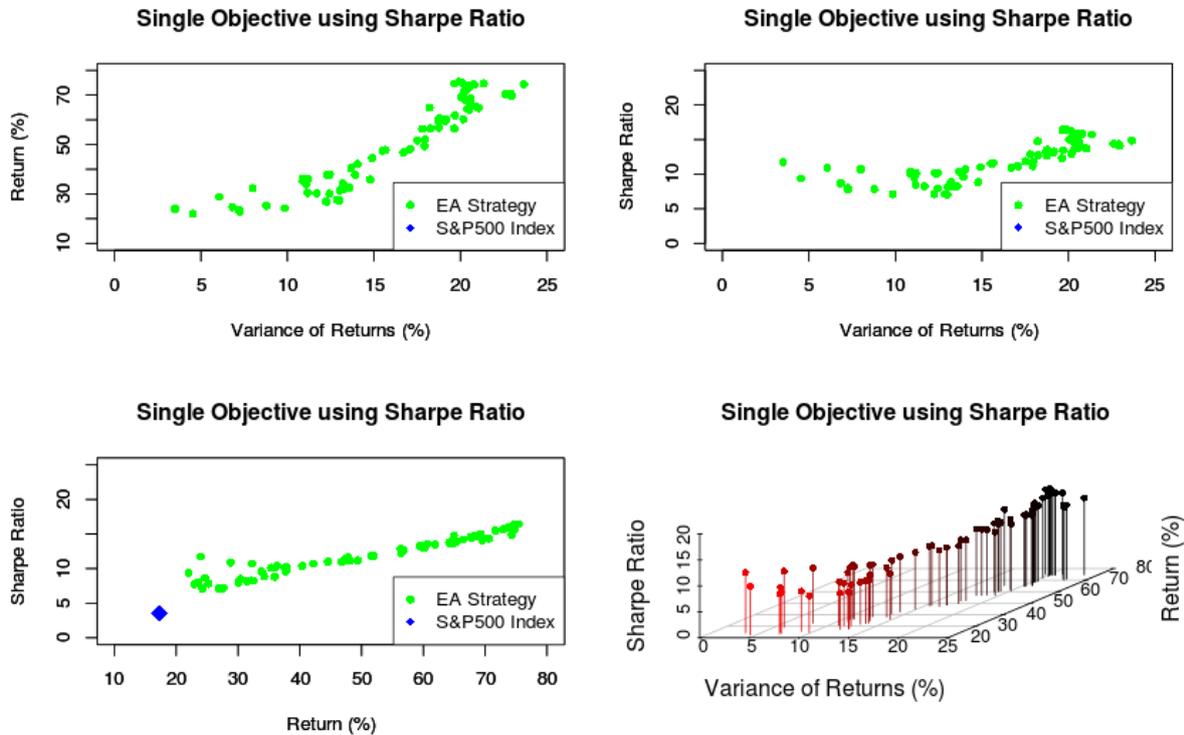


Figure 42 - Sharpe ratio training results

a) Real test using Sharpe ratio

The real test results in figure 43 and 44, show lower risk solutions in relation to the Index, thus its verifying the effect of efficiency in the use of risk by the solutions found by the algorithm when using the Sharpe ratio as objective. By comparing this results in relation to the experience with the MOEA using the concept of Pareto dominance, is concluded the results of the Sharpe ratio are in the same range in terms of returns, but with less variance.

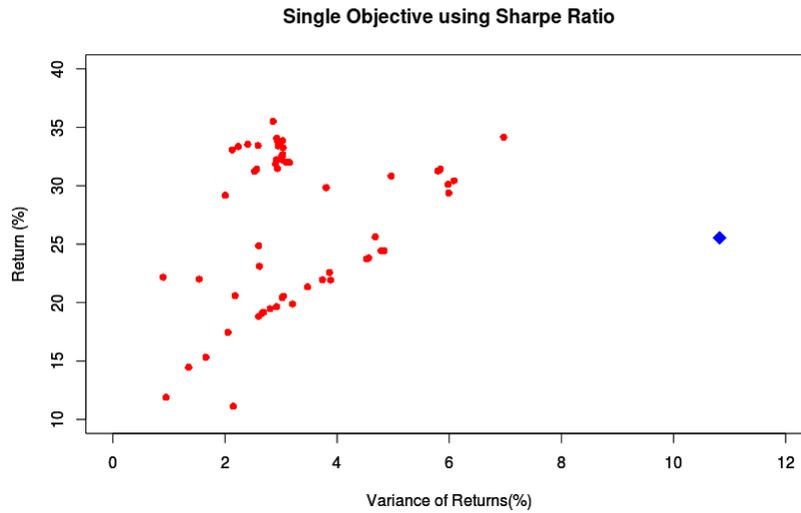


Figure 43 - Results obtained in the real test for the Sharpe ratio objective

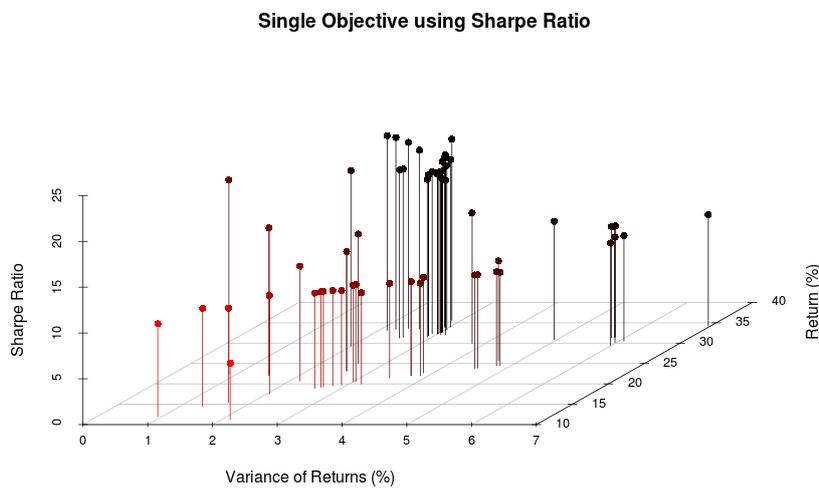


Figure 44 –Graphic 3D of real test for the Sharpe ratio objective

b) Analyzing strategies found

From the population of the training was selected 5 chromosomes based in the results that they obtained in the real test, that are in table 13.

| | Return (%) | Variance (%) |
|----------------------|------------|--------------|
| S&P 500 | 25,55 | 10,82 |
| Chromosome ID | | |
| 7 | 34,00 | 1,71 |
| 63 | 23,11 | 2,61 |
| 68 | 29,19 | 2,00 |
| 76 | 17,46 | 2,05 |
| 79 | 11,88 | 0,95 |

Table 13 - Results of the real test

By analyzing the genes in table 14 the difference in the results by the chromosome 7 in relation to the others are justified by more concentration of the investments, and the selection of companies with better ROE, and profit margin.

| Genes | Crom7 | Crom63 | Crom68 | Crom76 | Crom79 |
|--------------------|-----------|-----------|-----------|-----------|------------|
| Debt ratio | 0.0126327 | 0,005913 | 0,0072424 | 0,0072424 | 0,00039647 |
| ROE | 0,14888 | 0,002256 | 0,0245575 | 0,0150283 | 0,024558 |
| Profit Margin | 0,0457 | 0,00338 | 0,0069581 | 0,0075712 | 0,0059008 |
| PER | 0,87995 | 0,8799457 | 0,8799457 | 0,4179216 | 0,87995 |
| Δ of Rev. | 0,03352 | 1,0759832 | 0,0617327 | 0,03352 | 0,13537 |
| Δ common Stock out | 0,00911 | 0,0200047 | 0,0067419 | 0,009113 | 0,010056 |
| Δ net income | 1,49167 | 1,5465318 | 0,4998159 | 0,8128513 | 1,4917 |
| Stocks Number | 8 | 7 | 5 | 8 | 8 |
| Position Size | 0,2 | 0,0892165 | 0,1182039 | 0,0741898 | 0,05 |

Table 14 - Genes of the Chromosomes

5.2.3 Multi-objective using the Pareto Dominance Concept

In this simulation is used the concept of Pareto dominance in the MOEA, the curve of Figure 45 was obtained in the training of the population. The results show some discontinuities in the curve and little more density of solutions in the area with lower variance.

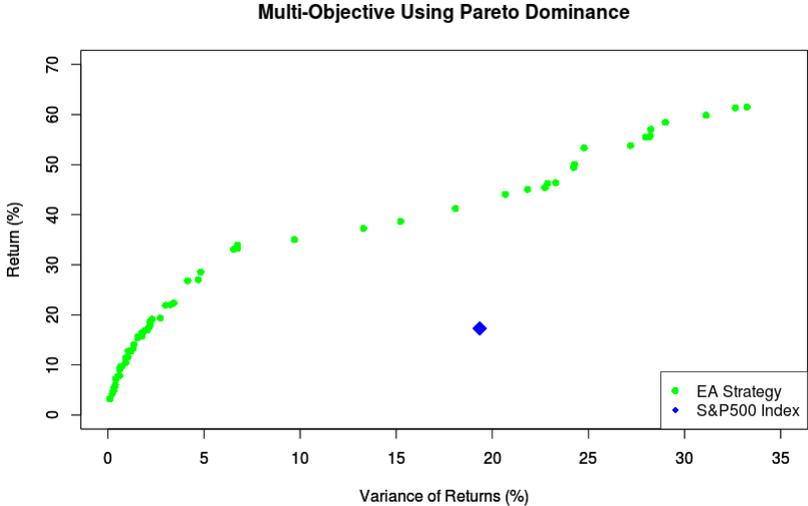


Figure 45 - Pareto curve obtained in training using the dominance of Pareto

a) Multi-objective real test

The population of the training obtained the following results in Figure 46 in the real test. The results demonstrate a better performance than the Index. Comparing the training of the MOEA and the real test curves, the results had a worse performance in real due to the over fitting effects. It can be seen the solutions as the ability to maintain better returns than the index for the same level of risk in both the training and real test.

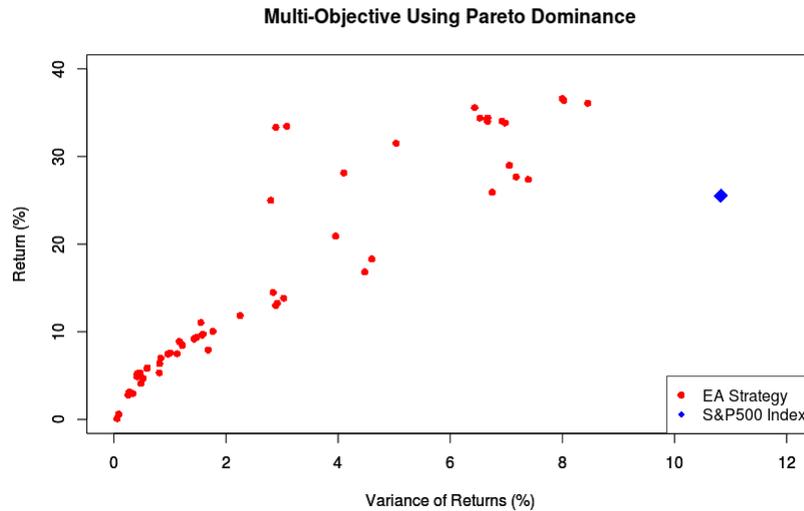


Figure 46 - Results obtained in the real test

b) Analyzing strategies found

From the population of the training was selected 5 chromosomes based in the results that they obtained in the real test, that are in table 15.

| | Return (%) | Variance (%) |
|----------------------|------------|--------------|
| SP&500 | 25,55 | 10,82 |
| Chromosome ID | | |
| 2 | 36.4 | 8.02 |
| 10 | 27,59 | 7,18 |
| 23 | 18.31 | 4.58 |
| 29 | 11,84 | 2,25 |
| 40 | 8,89 | 1.16 |

Table 15 - Results of the real test

By analyzing the genes in table 16 the difference in the results by the chromosome 2 in relation to the others are justified by more concentration of the investments, and the selection of companies with better ROE, and profit margin.

| Genes | Crom2 | Crom10 | Crom23 | Crom29 | Crom40 |
|--------------------|----------|----------|----------|--------|----------|
| Debt ratio | 0,000006 | 0,000002 | 0,017 | 0,0003 | 0,000002 |
| ROE | 0,00015 | 0,00015 | 0,31 | 0,0029 | 0,31 |
| Profit Margin | 0,3578 | 0,0059 | 0,02617 | 0,2156 | 0,00286 |
| PER | 1,11 | 1.11 | 2,099 | 0,8637 | 2.099 |
| Δ of Rev. | 0,0739 | 0,0739 | 0.8549 | 1.0786 | 0,01089 |
| Δ common Stock out | 1,3989 | 1.3983 | 1.9047 | 1.0967 | 1.5 |
| Δ net income | 2.9453 | 1,8516 | 0.331 | 0,043 | 0,638 |
| Stocks Number | 8 | 7 | 9 | 6 | 4 |
| Position Size | 0,18 | 0,18 | 0,063855 | 0,0638 | 0,0638 |

Table 16 - Genes of the Chromosomes

5.3 Conclusions

The fundamental approach to stock investment using evolutionary algorithms for optimization of single and multi-objectives demonstrates that the solutions found by the algorithms in general outperformed the Index S&P 500.

The Sharpe ratio simulation demonstrates that is a more efficient way to evaluate portfolio, and using it as fitness function in the process of reproduction lead to obtain chromosomes with better results in efficiency terms in the use of risk. The analysis of the solutions with higher returns with less variance in the portfolio prove this results are achieved by a higher concentration of the investments, and better selection of the companies based in the financial ratios more important, the ROE, profit margin, and net income growth rate.

The models of investment used allow a great number of trading strategies, tested in different time's frames, where the optimization finds strategies that outperform the index or have a good performance. The choices of the objectives to optimize are direct linked with the behavior of strategies obtained. If the objective is the ROI, is obtained a set of strategies with more concentration of investments and higher variances, in case of using the pareto concept is obtained a set of solutions describing an curve with different risks of the solutions, in case of Sharpe ratio the strategies are more efficient using the risk, and the slope of the curve increases with variance.

6. Conclusions and Future Work

The approach of using Intelligent Computing to forecast financial markets, represent a new field in financial markets investments. Develop this type of solution is not easy due to the different types of knowledge required, and the high competition in the markets.

The innovative approach implemented in this work uses the financial analysis in conjunction with technical analysis, trading rules, and money management to implement active portfolio management strategies to achieves superior results than a passive strategy.

The EA implemented uses data of the stocks that compose the Index S&P 500, and in the models of investment are incorporated real constraints used in portfolios. For evaluate the efficiency of algorithms the results obtained was compared with index.

6.1 Conclusion

The results obtained demonstrates this approach can be applied to the management of portfolios with great results due to the optimization of the strategies and the capacity of analyze hundreds of companies in a few seconds by the algorithms.

From the simulations of the different objectives, is realizing that the Sharpe ratio is better to train the populations with the goal to obtain portfolio management strategies with a more efficient use of the risk.

When using a model of investment that is always invested in the most undervalued stocks, the ratio more important is the PER for selecting the companies.

The ratios with more importance for select companies with a higher expected ROI and lower variance is the ROE and net income growth rate. It is concluded to obtain higher returns are necessary a higher concentration of the investments and to use a higher risk.

In this investigation is realized that this algorithms applied to real world investing using some human iteration to validate the investments by a more deep analyses of the companies and conditions of the markets, can improve the results obtained.

6.2 Future Work

Studies in this work took place during the period after the crash of the North American markets in 2008, and as such the periods tested represent a three years of Bull Market, leading to the over fitting of the solutions to this type of tendency.

In a future work, the challenge will be to test these algorithms strategies found in a bear market to check if still on average outperform the index, thereby ensuring adjustment of the chromosomes in the most adverse environment of the market. Future improvements to implement in this approach are described next:

- An implementation of other risk measures and performance of portfolios as objectives in the EA to study what objective functions give the best results, and allow to obtaining superior portfolio strategies.
- Allow short selling in conjunction with the longs positions, to the algorithms invest in the best companies in bull market, and short the worst in the bear market.
- Incorporate macroeconomic indicators in the algorithm to determine the general tendency of the market, for it selects the better strategies to use.
- Incorporate industrial indicators, and associate the companies that are affected by these indicators to try to extract conclusions about how to use them to find the best industries to invest and to short.
- Use adjustment measures, to avoid the effects of over fitting.

7. References

- Achelis, S. (2000). Technical analysis from A-to-Z. *Vision Book*.
- Anagnostopoulos, K., & Mamanis, G. (2011). The mean-variance cardinality constrained portfolio optimization problem: An experimental evaluation of five multiobjective evolutionary algorithms. *Expert Systems with Applications*, 38, 14208-14217.
- Blanchard, O. (2006). Macroeconomics. *PERSON Prentice Hall*.
- Bodie, Z., Kane, A., & Marcus, A. (2010). Investments. *McGraw-Hill*.
- Buffett, M., & Clark, D. (2008). Warren Buffett and the interpretation of financial statements. *SIMON & SCHUSTER*.
- Casanova, I. J. (2010). Trading-LCS: Dynamic stock portfolio decision-making assistance model with based machine learning. 1-8.
- Cesarone, F., Scozzari, A., & Tardella, F. (2009). Efficient Algorithms for Mean-Variance Portfolio Optimization With Hard-World Constraints.
- Chang, T., Meade, N., & Sharaiha, Y. (2000). Heuristics for cardinality constrained portfolio optimisation. *Computer & Operations Research*, 27, 1271-1301.
- Chen, M. R., Weng, J., & Li, X. (2009). Multiobjective Extremal Optimization for Portfolio Optimization Problem. *IEEE*, 552-556.
- Clack, C., & Patel, S. (2007). ALPS evaluation in Financial Portfolio Optimization. *IEEE*, 813-819.
- Cunningham, L. (1998). The essays of Warren Buffett: Lessons for corporate America.
- Faith, C. (2007). Way of the turtle. *McGraw-Hill*.
- Fonseca, C., & Fleming, P. (1993). Genetic Algorithms for Multiobjective Optimization: Formulation, discussion and generalization. *Proceedings of the Fifth International Conference on Genetic Algorithms*, 416-423.
- Ghang, T. J., Yang, S. C., & Chang, K. J. (2009). Portfolio optimization problems in different risk measures using genetic algorithm. *Expert Systems with Applications*, 36, 10529-10537.
- Goldberg, D., & Richardson, J. (n.d.). Genetic Algorithms with Sharing for Multimodal Function Optimization. *Proceedings of the 2nd International Conference on Genetic Algorithms*, 41-49.
- Golmakani, H. R., & Fazel, M. (2011). Constrained Portfolio Selection using Particle Swarm Optimization. *Expert Systems with Applications*, 38, 8327-8335.
- Gorgulho, A., Neves, R., & Horta, N. (2011). Applying a GA kernel on optimizing technical analysis rules for stock picking and portfolio composition. *Expert systems with Applications*, 38, 14072-14085.
- Graham, B., & Dodd, D. L. (1934). Security Analysis. *Mcgraw Hill*.
- Hagstrom, R. (1994). The Warren Buffett Way. *John Wiley & Sons, Inc*.
- Harlow, H. (1991). Asset allocation in a downside-risk. *Financial Analysts Journal*, 47, 28-40.

- Hassan, G. N. (2010). MultiObjective Genetic Programming for Financial Portfolio Management in Dynamic Enviroments. *Ph D Degree Thesis, University College London*.
- Hassan, G., & Christopher, C. (2009). Robustness of Multiple Objective GP Stock-Picking in Unstable Financial Markets. *Genetic and Evolutionary Computation (GECCO)*, 1513-1520.
- Hatzakis, I. (2004). Multi-Objective Evolutionary Methods for Time-Changing Portfolio Optimization Problems. *MS degree Thesis, Massachuttes Institute of Technology, Dep. of Mechanical Engineering*.
- Hillier, D., Ross, S., Randolph, W., Jeffrey, J., & Jordan, B. (2010). Corporate Finance. *Mcgraw-Hill, First European Edition ed*.
- Hirabayashi, A., Aranha, C., & Iba, H. (2009). Optimization of the Trading Rule in Foreign Exchange using Genetic Algorithm. *GECCO*, 1529-1536.
- Kaucic, M. (2012). Portfolio Management Using Artificial Trading Systems Based on technical Analysis. *Genetic Algorithms in Applications*.
- Kim, J. H., Han, J. H., Kim, Y. H., Choi, S. H., & Kim, E. S. (2012, Feb). Preference-Based Solution Selection Algorithm for Evolutionary Multiobjective Optimization. *Evolutionary Computation*, 16, 20-34.
- Liang, J., You, J., Han, G., & Li, L. (2012). Double Space Based Multiobjective evolutionary Algorithm. *IEEE*, 4, 1406-1411.
- Lin, D., & Wang, S. (2001). A multiobjective genetic algorithm for portfolio problem.
- Lozano, J., & Armañanzas, R. (2005). A Multiobjective Approach to the Portfolio Optimization Problem. (E. C. IEEE, Ed.) *IEEE*, 2, 1388-1395
- Lynch, P., & Rothchild, J. (1989). One Up On Wall Street. *SIMON & SCHUSTER*.
- Malkiel, B. (1999). A random walk down wall street.
- Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, 7, 77-91.
- Matos, F. B. (2007). Ganhar Em Bolsa. *Dom Quixote*.
- Metaxiots, K., & Liagkouras, K. (2012). Multiobjective Evolutionary Algrithms for Portfolio Management:A comprehensive literature review. *Expert Systems with Applications*, 39, 11685-14085.
- Noilublao, C., & Bureerat, S. (2009). Simultaneous topology, Shape and Sizing optimisation of Skeletal Structures Using Multiobjective Evolutionary Algorithms. *Evolutionary Computation*, 572.
- Pandari, A. R., Azar, A., & Shavazi, A. R. (2012). Genetic algorithms for portfolio selection problems with non-linear objectives. *African Journal of Business Management*, 6, 6209-6216.
- Pindyck, R., & Rubinfeld, D. (2006). Microeconomics. *PERSON Prentice Hall*.
- Robinson, T. R., Greuning, H. v., Henry, E., & Broihahn, M. A. (2009). International Financial Statement Analysis. *Wiley&Sons*.
- Samuelson, P. (1958). The fundamental approximation theorem of portfolio analysis in term of means variances and higher moments. *Review of Economic Studies*, 65-86.
- Schaerf, A. (2002). Local Search Techniques for Constrained Portfolio Selection Problems. *Computational Economics*, 20, 177-200.

- Schroeder, A. (2007). *The Snowball. Warren Buffett and Business of life. Actual Editora.*
- Schyns, M., & Crama, Y. (2003). Simulated annealing for complex portfolio selection problems. *European Journal of Operational Research*, 150, 546-571.
- Skolpsdungket, P., Dahal, K., & Harnpornchai, N. (2007). Portfolio Optimization using Multi-objective Genetic-Algorithms. *IEEE*, 516-523.
- Soleimani, H. (2007). Portfolio selection using genetic algorithm. *MS Degree Thesis, Armirkabir University of Technology, Industrial Engineering Department.*
- Tettamanzi, A., & Loraschi, A. (1995). An evolutionary algorithm for portfolio selection in a downside risk framework. *Working Papers in Financial Economics*, 6, 8-12.
- Tharp, V. (2007). *Trade to Financial Freedom. Mcgraw-Hill.*
- Zitzler, e., & Thiele, L. (1998). An Evolutionary Algorithm for Multiobjective optimization: The Strength Pareto Approach.
- Zitzler, E., Laumanns, M., & Bleuler, S. (2004). A Tutorial on Evolutionary Multiobjective. *Metaheuristics for Multiobjective Optimization*, 535, 3-37.

Appendix A-Results Table

It is presented in the next table the trades performed in the real test by the MOEA of the simulation done in section 5.1.2, for the chromosome 36.

Table 17 - Trades of chromosome 36

| CROM 36 | | | |
|------------|-------|----------|--------|
| DATE | STOCK | POSITION | AMOUNT |
| 4-01-2012 | DTV | BUY | 20414 |
| 4-01-2012 | ESRX | BUY | 16255 |
| 4-01-2012 | LLTC | BUY | 12922 |
| 4-01-2012 | CLX | BUY | 10283 |
| 4-01-2012 | YHOO | BUY | 8191,6 |
| 4-01-2012 | ACAS | BUY | 6512,5 |
| 4-01-2012 | EBAY | BUY | 5178,6 |
| 4-01-2012 | LSI | BUY | 4128,5 |
| 4-01-2012 | MRK | BUY | 3305,3 |
| 4-01-2012 | EOG | BUY | 2562,5 |
| 4-01-2012 | CBS | BUY | 2102,5 |
| 4-01-2012 | AES | BUY | 1663,3 |
| 4-01-2012 | NUE | BUY | 1304,7 |
| 4-01-2012 | WU | BUY | 1059,6 |
| 4-01-2012 | MS | BUY | 838,56 |
| 4-01-2012 | AKAM | BUY | 671,77 |
| 4-01-2012 | TEX | BUY | 529,87 |
| 31-01-2013 | TEX | SELL | 1130,7 |

Appendix B-Results Table

It is presented in the next table the trades performed in the real test by the MOEA of the simulation done in section 5.1.2, for the chromosome 51.

Table 18 - Trades of chromosome 51

| CROM 51 | | | |
|------------|----------|----------|--------|
| DATE | STOCK | POSITION | AMOUNT |
| 2012-01-04 | ESRX.csv | BUY | 5082,7 |
| 2012-01-04 | YHOO.csv | BUY | 4835,2 |
| 2012-01-04 | ACAS.csv | BUY | 4590,6 |
| 2012-01-04 | LSI.csv | BUY | 4361,3 |
| 2012-01-04 | EBAY.csv | BUY | 4123,7 |
| 2012-01-04 | MRK.csv | BUY | 3936,6 |
| 2012-01-04 | EOG.csv | BUY | 3690 |
| 2012-01-04 | AES.csv | BUY | 3534,4 |
| 2012-01-04 | DTV.csv | BUY | 3358,2 |
| 2012-01-04 | NUE.csv | BUY | 3202,4 |
| 2012-01-04 | CBS.csv | BUY | 3030,9 |
| 2012-01-04 | F.csv | BUY | 2867,2 |
| 2012-01-04 | HOG.csv | BUY | 2709,6 |
| 2012-01-04 | MOLX.csv | BUY | 2575,5 |
| 2012-01-04 | MET.csv | BUY | 2466,4 |
| 2012-01-12 | MET.csv | SELL | 2657,1 |
| 2012-01-13 | VLO.csv | BUY | 5133,1 |
| 2012-01-18 | LSI.csv | SELL | 4720 |
| 2012-01-19 | LSI.csv | BUY | 5190,5 |
| 2012-01-19 | F.csv | SELL | 3120,5 |
| 2012-01-20 | VLO.csv | SELL | 5516,1 |
| 2012-01-23 | ACAS.csv | SELL | 5006,1 |
| 2012-01-23 | MOLX.csv | SELL | 2766,7 |
| 2012-01-24 | VLO.csv | BUY | 5209,6 |
| 2012-01-25 | ACAS.csv | BUY | 5216,9 |
| 2012-01-25 | HOG.csv | SELL | 2965,1 |
| 2012-01-26 | LSI.csv | SELL | 5600,5 |
| 2012-01-30 | F.csv | BUY | 5236,9 |
| 2012-02-01 | LSI.csv | BUY | 5232,5 |
| 2012-02-03 | NUE.csv | SELL | 3447,5 |
| 2012-02-13 | EOG.csv | SELL | 3968,6 |
| 2012-02-14 | EOG.csv | BUY | 5294,9 |
| 2012-02-16 | ESRX.csv | SELL | 5463 |
| 2012-02-16 | AES.csv | SELL | 3863,1 |

| | | | |
|------------|----------|------|--------|
| 2012-02-17 | EBAY.csv | SELL | 4432,3 |
| 2012-02-21 | ESRX.csv | BUY | 5335,6 |
| 2012-02-21 | EBAY.csv | BUY | 5053,8 |
| 2012-02-21 | AES.csv | BUY | 4786 |
| 2012-03-07 | VLO.csv | SELL | 5618,9 |
| 2012-03-08 | VLO.csv | BUY | 5280,5 |
| 2012-03-08 | CBS.csv | SELL | 3272,1 |
| 2012-03-14 | LSI.csv | SELL | 5875,9 |
| 2012-03-20 | LSI.csv | BUY | 5373,4 |
| 2012-03-26 | DTV.csv | SELL | 3622,7 |
| 2012-04-11 | SE.csv | BUY | 5212,6 |
| 2012-04-17 | ESRX.csv | SELL | 5767,3 |
| 2012-04-19 | ESRX.csv | BUY | 5257,2 |
| 2012-04-19 | EBAY.csv | SELL | 5612,9 |
| 2012-04-25 | EBAY.csv | BUY | 5244,8 |
| 2012-04-26 | ACAS.csv | SELL | 5722,4 |
| 2012-04-27 | ACAS.csv | BUY | 5309,7 |
| 2012-07-09 | MRK.csv | SELL | 4252,8 |
| 2012-07-12 | MRK.csv | BUY | 5146,6 |
| 2012-07-20 | EBAY.csv | SELL | 5669,9 |
| 2012-07-26 | EBAY.csv | BUY | 5181,2 |
| 2012-08-08 | ACAS.csv | SELL | 5741,8 |
| 2012-08-09 | ACAS.csv | BUY | 5354 |
| 2012-08-27 | VLO.csv | SELL | 5673,5 |
| 2012-08-28 | VLO.csv | BUY | 5373,1 |
| 2012-09-07 | EBAY.csv | SELL | 5597,6 |
| 2012-09-13 | EBAY.csv | BUY | 5434,5 |
| 2012-10-18 | MRK.csv | SELL | 5579,2 |
| 2012-10-25 | MRK.csv | BUY | 5374,1 |
| 2012-11-19 | YHOO.csv | SELL | 5235,9 |
| 2012-11-21 | YHOO.csv | BUY | 5348,9 |
| 2012-12-13 | ACAS.csv | SELL | 5753,3 |
| 2012-12-17 | ACAS.csv | BUY | 5441,9 |
| 2013-01-02 | VLO.csv | SELL | 5919,6 |
| 2013-01-03 | CNP.csv | BUY | 5557,2 |
| 2013-01-03 | MIL.csv | BUY | 5275,4 |
| 2013-01-04 | SO.csv | BUY | 5553,3 |
| 2013-01-04 | ED.csv | BUY | 5286,5 |
| 2013-01-04 | F.csv | SELL | 5634,1 |
| 2013-01-08 | LUK.csv | BUY | 5546,9 |
| 2013-01-11 | MIL.csv | SELL | 5856,5 |
| 2013-01-15 | MIL.csv | BUY | 5564 |
| 2013-01-18 | EOG.csv | SELL | 5707,1 |
| 2013-01-24 | EBAY.csv | SELL | 5895,4 |
| 2013-02-11 | YHOO.csv | SELL | 5837,4 |
| 2013-02-12 | YHOO.csv | BUY | 5646,5 |
| 2013-02-12 | LUK.csv | SELL | 5985,3 |

| | | | |
|------------|----------|------|--------|
| 2013-02-13 | LUK.csv | BUY | 5656 |
| 2013-02-14 | ACAS.csv | SELL | 5859,3 |
| 2013-02-15 | ACAS.csv | BUY | 5659 |
| 2013-03-05 | CNP.csv | SELL | 5988,7 |
| 2013-03-08 | CNP.csv | BUY | 5736,3 |
| 2013-04-02 | YHOO.csv | SELL | 6082,4 |
| 2013-04-04 | YHOO.csv | BUY | 5733,7 |
| 2013-04-16 | LUK.csv | SELL | 6116,6 |
| 2013-04-19 | LUK.csv | BUY | 5748,4 |
| 2013-04-25 | ED.csv | SELL | 5683,5 |
| 2013-04-29 | ED.csv | BUY | 5846,7 |
| 2013-04-29 | CNP.csv | SELL | 6194,2 |
| 2013-04-30 | CNP.csv | BUY | 5840,5 |
| 2013-05-08 | YHOO.csv | SELL | 6185,8 |
| 2013-05-10 | YHOO.csv | BUY | 5856,5 |