Genetic algorithm evolved agent-based equity trading using Technical Analysis and the Capital Asset Pricing Model

Cyril Schoreels and Jonathan M. Garibaldi
Automated Scheduling, Optimisation and Planning Research Group
School of Computer Science and IT
University of Nottingham
Nottingham, UK
e-mail: czs@cs.nott.ac.uk

Abstract: *This paper investigates genetic algorithm evolved agents trading on real historical equity market data using technical analysis, the capital asset pricing model and a hybrid model of the two approaches. Three agent groups are generated, each using solely one of the two approaches or their hybrid to determine trading decisions. Each group consists of ten independently evolved populations over a thousand generations, whose elite's performances are consequently averaged and used to compare to the other approaches. Results indicated that the technical analysis based approach performed better than the capital asset pricing model based approach, while the hybrid approach in turn outperformed both. As part of ongoing research, the results would suggest significant benefits in performance oriented implementation of hybrid or multi-method based approaches in agent-based systems.*

Keywords: Agents, Genetic Algorithm, Technical Indicators, Capital Asset Pricing Model

1 Introduction

In previous research, agents were evolved using a genetic algorithm and employing Technical Analysis (TA) to make trading decisions. This proved successful compared to other benchmarks, such as in [16, 13] for example. In this investigation we extend this research by introducing the use of the Capital Asset Pricing Model (CAPM) [17], both individually as well as in a hybrid form using features from both the CAPM and TA based approaches.

Allen and Karjalainen published results using a genetic algorithm to evolve trading strategies [2]. Using daily closing price information and at varying transaction cost levels, they found that their evolved rules did not consistently earn excess returns over a simple buy-and-hold strategy. The buy-and-hold strategy being a common reference benchmark, where the entire starting capital is evenly invested in the available assets for the entire time period under consideration. They stated that there had been little formal analysis in financial economics of genetic algorithms and that the purpose of their paper was to demonstrate how genetic algorithms could be used to find technical trading rules. Further work was performed by Potvin et al, who used genetic programming to again evolve trees of technical trading rules [11]. They found that their trading rules performed well during stable or falling markets, while in appreciating markets buy-and-hold fared better. In general, genetic algorithms have been used in a wide variety of financial applications by many authors [3, 6, 9, 10].

The Technical Analysis used in this series of experiments is trend based, providing an indication as to the future movement of the time series under consideration. Research into the success of technical trading suggests that it can act as a good indicator for making trading decisions [7, 5, 1]. When compared to the CAPM, a main distinguishing feature is the introduction of risk, producing a more comprehensive analysis of the situation and a comparative ranking between alternatives. Forming part of ongoing research, the paper will first outline the system employed to evolve the agent populations, with emphasis on the TA, CAPM and hybrid methods employed. Following this, results are presented and analysed.

2 System Outline

Agent populations emulating equity market traders were evolved using a simple genetic algorithm, with three methods of analysing securities implemented. The first approach was based on previous work using TA, the second approach used the CAPM while a third hybrid approach of the previous two methods was used as well. In each, every agent represents an individual trader with a personal portfolio and capital holdings, collectively referred to as their total assets. With a fixed starting capital, at the end of every trading day each trader uses historical price data to make a decision on every security whether to buy, sell or hold. No limitations, apart from capital constraints, are set on the number of trades conducted. Additionally, every agent contains a genome which specifies the parameters used by that particular agent for its analysis and decision processes, effectively determining its trading behaviour. Depending on the method of analysis used, genome length varies with the number of parameters employed. Transaction costs were not included in the simulation and the environment is assumed discrete and deterministic in a liquid market.

2.1 TA based agents

Technical indicators are tools used in the technical analysis of financial markets, exploiting the existence of trends to determine potential buy, sell or hold conditions. In agent populations using TA, every agent used six technical indicators to generate eight trading signals for every security that is assessed. The indicators used were the Simple Moving Average, Relative Strength Index, Moving Average Convergence Divergence, Bollinger Bands, Price-Rate-of-Change and Stochastic Oscillator, with the latter two generating two signals each. More general information on technical indicators can be found in [1]. Depending on the signals returned, an overall decision of buy, sell or hold is made for a security. Allocation of funds between the securities selected for acquisition through TA is determined by one gene in the genome, allocating between 10%-100% of available capital equally between the buy options. Securities that are sold are converted into capital at their current closing price.

There are essentially three steps that every agent follows to determine whether or not to add a particular security to its acquisition or sale list. To analyse a security it first performs the indicator calculations, each using historical closing price information. Depending on the indicator, one or more values are returned which then need to be interpreted for the calculated values to gain meaning. Based on this interpretation, every agent then compiles a list of buy and sell signals for every security as the second step in this process. The final step is generating an overall buy, sell or hold decision for each security. The genome for each agent is of length 28 with varying cardinality depending on function. For further details see [12, 14].

2.2 CAPM based agents

The Capital Asset Pricing Model is based on Markowitz's portfolio theory [8] where a portfolio's overall exposure to risk is reduced through a balanced selection of assets. Essentially, the CAPM determines the attractiveness of a particular security in form of a numerical value, which then allows selection and comparison between possible investments. The higher the value, the more desirable the investment appears. The attractiveness of a security is its expected rate of return, an estimate as to what an investment in that security would yield within the specified time frame. In the agent populations that use the CAPM, both security selection and capital allocation between the securities selected for acquisition are determined by the model. A significantly reduced number of parameters are required in its calculation compared to the TA model introduced earlier, resulting in a smaller genome length. Specifically, three genes were used, with both G_1 and G_3 taking integer values of 1-10 inclusive and G_2 representing a boolean value. G_1 and G_3 are both further translated by multiplying with a constant value of 10, where G_1 is to represent the number of trading days taken under consideration in the calculation and G_3 the proportion of available capital to be invested, similar to the method used in the TA based approach. G_2 represents the choice between using ex-ante or a moving average for the prediction of the equity risk premium. The method employed in this investigation to calculate the expected rate of return of the available securities is shown in the following equation:

$$E(r) = R(f) + \beta(R(m) - R(f)) \tag{1}$$

where $E(r)$ is the expected return of the security under consideration, $R(f)$ is the risk free rate of return such as a fixed-yield bond, β represents the measure of risk and $R(m)$ is the return possible in the current market. In the CAPM, there are two methods of calculating $R(m)$, the estimated return of the market in percent over the time period determined by G_1. The first is to use historical data to calculate the average change over the period, while the second is to try to predict the future price by using the historical moving average. In this implementation the agent chooses between either method based on G_2. Once the expected return for all securities has been calculated, all securities with a negative value are dismissed for purchase and placed on the sales list. The available capital is then divided among those on the purchase list based on their relative expected rate of return. For example, if two securities A and B are to be purchased with expected returns of 5 and 15 respectively, 25% of the proportion of available capital to be invested based on G_3 would be invested in A and 75% would be invested in B.

2.3 TA and CAPM hybrid agents

A hybrid model of both TA and the CAPM was further implemented, where TA primarily determined the selection of securities, while the CAPM determined the allocation of capital among those selected. Additionally, securities selected through TA that were predicted a negative return by the CAPM were dismissed. To accommodate the increased complexity of the hybrid approach, the genome increased in size incorporating elements from both approaches, resulting in a length of 30. It is important to note however that as G_3 exists in both genomes and is identical in function, it is not duplicated. The technical indicators, agent decision types and all other elements described in section 2.1, [12, 14] and section 2.2 apply in the hybrid model.

2.4 Genetic algorithm

The genetic algorithm used was based on previous research [16, 12]. For selection, elitism is used, whereby a portion of the most successful agents carries forward unaltered every generation [4]. Immigration was also employed, where a portion of the worst performers are removed and replaced by a new randomly generated group of agents immigrating into the system, constantly introducing new genetic material. This facilitates greater coverage of the search space, while also avoiding premature convergence and non-exclusion of other possible solutions not present in the original base population's gene pool. Each agent in the mediocre population, those not removed or part of the elite, randomly selects another agent from the mediocre and elite population and uses two-point crossover to create an offspring. This offspring then replaces the parent from the mediocre population. The elite population and removed population are each respectively the top and bottom 25% of the entire agent population. Overall, the process aims at maintaining a high level of diversity and avoiding complete convergence of the elite population to avoid over-fitting and related drawbacks. Two randomly selected agents of the same decision type can mate, with a random part of the first agent's genome being replaced by the equivalent section from the second agent's genome, forming a new genome combination. Performance was measured as the area under an agent's total asset graph, as this avoids bias of results towards the cut-off date chosen.

3 Method and results

Historical financial data was taken from the German DAX-30. However the system does not include all securities from the DAX for its entire span, as due to changes in the constituents of the Index, daily closing price information was only available for 20 securities over the last few decades. Data used in the simulation was previously presented in [12, 13]. 1465 trading days worth of data was used representing six years from 01.01.1990 to 29.12.1995. Based on earlier research [12], this time period was assumed sufficient to evolve a competitive population. The out-of-sample period was taken over six years from 01.01.1997 to 31.12.2002.

For each approach, 10 populations of a 100 agents each were independently evolved from initially random populations over 1000 generations. The fitness function used the total area under the agent's total asset graph over the entire training duration. The elite from each population was exposed to the testing data with their performance averaged and a mean taken across the 10 populations for each approach. Beta correlation to the buy-and-hold strategy was also measured, being defined as how closely the performance followed that of the buy-and-hold strategy, with a value of 1 indicating synchronous movement, 0 indicating no correlation and -1 indicating asynchronous movement. The performance's volatility was also noted, with volatility being defined as the average daily relative change in value, giving an indication as to the fluctuations experienced on a day to day basis. Experiments were repeated for the two interest rates of 0% and 5%. A summary of these results is presented in Table 1.

Table 1: Results Summary - Means (Standard Deviation)

	Perf. (x10^6)		Beta		Volatility (%)	
	0%	5%	0%	5%	0%	5%
Buy-Hold	270.1		-		1.72	
CAPM	243.1 (1.9)	227.3 (1.5)	0.52 (0.07)	0.86 (0.02)	1.25 (0.03)	1.10 (0.02)
TA	241.6 (23.6)	270.2 (39.1)	0.90 (0.04)	0.83 (0.04)	1.16 (0.12)	1.21 (0.13)
Hybrid	288.6 (75.1)	279.7 (30.4)	0.84 (0.08)	0.80 (0.11)	1.17 (0.20)	1.08 (0.17)

Further, the CAPM based approach was compared to the TA based approach, the CAPM based approach to the hybrid approach and the TA based approach to the hybrid approach, to establish the significance of the difference in results. Again, each of these tests were repeated for both 0% and 5%. Results using Mann-Whitney U tests are shown in Table 2.

Table 2: Mann–Whitney U tests

Experiments compared		p-value
CAPM 0%	TA 0%	0.48125
CAPM 0%	Hybrid 0%	0.00008
TA 0%	Hybrid 0%	0.02881
CAPM 5%	TA 5%	0.01854
CAPM 5%	Hybrid 5%	0.00001
TA 5%	Hybrid 5%	0.73936
CAPM 0%	CAPM 5%	0.00001
TA 0%	TA 5%	0.06301
Hybrid 0%	Hybrid 5%	0.43587

4 Analysis and discussion

Considering the variations in interest rates first, a clear impact can be observed for the CAPM agents. At an interest rate of 0%, agents performed significantly better than those at a rate of 5%. As interest rates are not considered in the decision making process in TA, it is not suprising that no difference in results can be observed, despite a seemingly better performance at a 5% rate of interest. Similarly, no difference can be found for the hybrid model, with mean performance marginally lower at a 5% rate of interest than at 0%. Notably, disregarding an outlier in results for the hybrid method at 0%, having returned a performance of $494.4x10^6$, would result in a performance mean of $265.8x10^6$ with a standard deviation of $21.6x10^6$.

Comparing the individual methods to one another in each respective interest rate category, under and interest rate of 5% the populations using TA clearly outperformed those using the CAPM, while the hybrid approach seemingly outperformed both CAPM and TA based approaches. Using 0% however, little difference in performance exists between CAPM and TA based approaches, while the hybrid approach performed considerably better. In terms of volatility and beta factor, little difference exists between all approaches at both interest rates, with the exception of the CAPM at 0% displaying an extremely low beta value.

Looking at both Table 2 and Table 1, it appears that the CAPM based approach performs comparably or worse than the TA based approach, while the TA based approach can perform comparably or worse than the hybrid. The CAPM based approach always displayed worse performance than the hybrid. It is also important to note that the standard deviation observed for all three approaches varies significantly, with the greatest deviation experienced by the hybrid, while the TA based approach in turn had a significantly higher deviation compared to the CAPM based approach. In other words, consistency in performance dropped from the CAPM to TA to the hybrid approach.

Overall it appears that in this implementation agent populations that utilize the CAPM as their method for generating input into their decision models will generate consistent if comparatively below average returns. Particularly compared to the buy-and-hold strategy CAPM agents do not seem to perform very well. The agent populations that used TA to feed into their decision models perform comparatively to the buy-and-hold strategy as expected and repeatedly demonstrated in previous work [13, 16]. Lastly, agents using the hybrid method proved clearly superior to the others if significantly less consistent in their returns. With respect to the buy-and-hold strategy the hybrid method further appears marginally better, though this cannot be substantiated here.

In previous work an amalgamation process derived a single trading decision based on the decisions made by a population of agents using TA [15]. The investigation showed that the agents using the amalgamated decision consistently demonstrated lower performance than the optimized agent populations, suggesting that averaging optimized solutions does not offer any benefits in this situation. The work presented here extends this by introducing an alternative method of analysing securities, which in conjunction with the TA based agent population may in turn offer advantages to deriving a centralized decision. Future work will focus on expanding the pool of heterogeneous agents the amalgamation process may draw on and investigate the effects it may have on the derived decisions and decision process overall.

In conclusion, the hybrid method employed demonstrated better performance than its constituents, suggesting further hybridization of alternative methods. Future work implementing systems with increased complexity combining methodologically heterogeneous agent groups might also demonstrate similar benefits as demonstrated by the hybrid method in this simulation, as well as reduce the impact of inconsistency in performance.

References

[1] S. B. Achelis. *Technical analysis from A to Z.* McGraw Hill, second edition, 2000.

[2] F. Allen and R. Karjalainen. Using genetic algorithms to find technical trading rules. *Journal of Financial Economics*, 51(2):245–271, 1999.

[3] L. A. Becker and M. Seshadri. GP-evolved technical trading rules can outperform buy and hold. In *Procceedings of the Sixth International Conference on Computational Intelligence and Natural Computing*, pages 26–30, September 2003.

[4] D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Boston, MA, 1989.

[5] T. Hellström and K. Holmström. The relevance of trends for predictions of stock returns. *International Journal of Intelligent Systems in Accounting, Finance & Management*, 9(1):23–34, 2000.

[6] G. Kendall and Y. Su. Co-evolution of successful trading strategies in a simulated stock market. In *Proc. International Conference on Machine Learning and Applications*, pages 200–206, 2003.

[7] A. Lipton-Lifschitz. Predictability and unpredictability in financial markets. *Physica D*, 133:321–347, 1999.

[8] H. M. Markowitz. Portfolio selection. *Journal of Finance*, 7(1):77–91, March 1952.

[9] C. J. Neely. Using genetic algorithms to find technical trading rules: A comment on risk adjustment. Technical Report 99-015A, Federal Reserve Bank of St. Louis, October 1999.

[10] O. V. Pictet, M. M. Dacorogna, B. Chopard, M. Oudsaidene, R. Schirru, and M. Tomassini. Using genetic algorithm for robust optimization in financial applications. *Neural Network World*, 5:573–587, 1995.

[11] J.-Y. Potvin, P. Soriano, and M. Vallée. Generating trading rules on the stock markets with genetic programming. *Computers and Operations Research*, 31(7):1033–1047, 2004.

[12] C. Schoreels and J. M. Garibaldi. A preliminary investigation into multi-agent trading simulations using a genetic algorithm. In *Proceedings of the UK Computational Intelligence Workshop 2004 (UKCI 04)*, pages 62–69, September 2004.

[13] C. Schoreels and J. M. Garibaldi. A comparison of adaptive and static agents in equity market trading. In *Proceedings of the IEEE/WIC/ACM International Conference on Intelligent Agent Technology 2005 (IAT 05)*, pages 393–399. IEEE Computer Society, September 2005.

[14] C. Schoreels and J. M. Garibaldi. The effect of varying parameters on performance for adaptive agents in technical equity market trading. In *Proceedings of the 3rd IEEE International Conference on Computational Cybernetics (ICCC 05)*, April 2005.

[15] C. Schoreels and J. M. Garibaldi. Comparative study of central decision makers versus groups of evolved agents trading in equity markets. In *Proceedings of the 12th International Conference on Computing in Economics and Finance 2006 (CEF 06)*, June 2006.

[16] C. Schoreels, B. Logan, and J. M. Garibaldi. Agent based genetic algorithm employing financial technical analysis for making trading decisions using historical equity market data. In *Proceedings of the IEEE/WIC/ACM International Conference on Intelligent Agent Technology 2004 (IAT 04)*, pages 421–424. IEEE Computer Society, September 2004.

[17] W. F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442, September 1964.