

学校编码: 10384
学号: 27720131152819

分类号__密级__
UDC__

厦 门 大 学

硕 士 学 位 论 文

在技术分析系统中运用遗传算法进行选股
和管理投资组合

**Using Genetic Algorithms to Pick Stocks and Manage
Portfolio in Technical Analysis System**

王 振 华

指导教师姓名: 陈 灯 塔
专 业 名 称: 应用统计硕士
论文提交日期: 2016 年 3 月
论文答辩时间: 2016 年 5 月
学位授予日期: 2016 年 6 月

答辩委员会主席: _____
评 阅 人: _____

2016 年 5 月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

2016 年 5 月 15 日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

1.经厦门大学保密委员会审查核定的保密学位论文，于 年 月 日解密，解密后适用上述授权。

2.不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

2016 年 5 月 15 日

摘 要

在投资领域，如何高效地进行资产配置、获得高收益并控制风险是每个投资者进行投资决策时最为关注的问题，如何有效地选股和进行投资组合管理是一个广泛存在的难题。技术分析是主流的选股方法之一，是通过分析历史的市场数据如股价、成交量等，来预测股票未来的趋势的一种投资分析方法；遗传算法是模拟自然界物种进化的随机全局搜索最优解的算法，目前广泛应用于组合优化、机器学习、自适应控制、信号处理等领域；证券投资组合优化问题的实质就是有限的资产在具有不同风险收益特性的证券之间的优化配置问题。

本文基于遗传算法构建了在技术分析系统中选股和管理投资组合的量化投资模型。进行选股时，遗传算法被用来优化多个技术指标间的比重，综合各个技术指标进行模拟投资，从而选取表现最优的股票；进行投资组合管理时，遗传算法被用来优化投资组合的成分股之间的权重，以使投资组合的夏普比率最高。

本文选取了 A 股市场 1994 年至 2014 年 20 年间的沪深 300 成分股的个股数据和上证指数数据进行实证。上证指数数据用于比较遗传算法交易策略和购买持有（Buy and Hold）策略，沪深 300 个股数据用来选股，并将选出的股票构建投资组合，进而比较等权组合、马科维茨（Markowitz）权重组合和遗传算法（GA）权重组合的表现，包括收益、风险和夏普比率三个评价指标。实证结果表明：本文投资模型的交易策略要优于购买持有策略，且在投资组合管理上的表现要优于等权组合和马科维茨组合。

关键词：遗传算法； 选股； 投资组合

Abstract

In the field of investment how to efficiently allocate a capital, have better returns and control risks is the primary concern of every investor when making investment decisions, how to effectively choose stocks and manage portfolio is always a widespread problem. Technical analysis is one of the most popular methods for stock picking, by analyzing historical market data such as stock prices, volume, etc., we can predict the trend of a stock; Genetic Algorithms (GA) are stochastic global search method that mimics the metaphor of natural biological evolution, which is widely used in fields like portfolio optimization, machine learning, adaptive control and signal processing; Portfolio optimization problem actually is the allocation of securities that having different risk-return characteristics with limited capital.

In this paper a quantitative investment model is built in technical analysis system for stock picking and portfolio management based on Genetic Algorithms. In the stock picking part, GA is used to optimize the proportion between eight technical indicators to pick out the best performing stocks. In the portfolio management part, GA is used to optimize the weight of securities in the portfolio, to have a highest Sharpe ratio.

Data from A-share market between 1994 and 2014 is employed in this paper, which include the SSE Composite Index and Component stocks of CSI300. The SSE Composite Index is used to compare GA strategy with Buy and Hold strategy, while the Component stocks of CSI 300 is used for stock picking and portfolio management, then we compare the Sharpe ratio of three different portfolio strategies: equal weight, Markowitz weight and GA weight. The empirical results show that the proposed GA strategy is superior to Buy and Hold strategy, and has a better performance in portfolio management when compared to equal weight portfolio and Markowitz portfolio.

Key Words: Genetic Algorithms ; Stock Picking ; Portfolio Management

目 录

第一章 引言	1
1.1 背景和意义	1
1.2 主要方法和研究进展	2
1.3 本文的结构安排	3
第二章 文献综述	4
2.1 国外在该领域的研究发展与现状	4
2.2 国内在该领域的研究发展与现状	5
2.3 本文的研究与贡献	9
第三章 相关理论	10
3.1 遗传算法	11
3.1.1 遗传算法的优势	11
3.1.2 遗传算法基本流程	12
3.1.3 求解简单函数的例子	13
3.1.4 遗传算法在本文的应用	16
3.2 技术指标	16
3.2.1 技术分析介绍	17
3.2.2 技术分析的优势	17
3.2.3 技术指标	18
3.3 收益与风险计量	22
3.4 投资组合策略	24
3.4.1 等权组合	24
3.4.2 马科维茨权重组合	24
第四章 模型设计	27
4.1 流程设计	27
4.2 计分规则设计	28
4.3 交易规则设计	31
4.4 遗传算法模型设计	32
第五章 投资组合策略比较	34
5.1 等权组合策略	34
5.2 Markowitz 权重组合策略	34
5.2.1 Markowitz-等权组合收益点权重	35
5.2.2 Markowitz-最小方差点权重	35
5.2.3 Markowitz-最大夏普比率点权重	35

5.2.4	Markowitz-GA 收益点权重	36
5.3	遗传算法 (GA) 权重组合策略	36
第六章 A 股市场实证分析		38
6.1	数据安排	38
6.2	遗传算法策略与 Buy and Hold 策略在牛熊市的收益对比	39
6.3	选股	41
6.4	投资组合策略比较	43
第七章 结论		46
参考文献		48
附录.....		52
致 谢.....		58

Contents

Chapter 1 Introduction.....	1
1.1 Background and significance.....	1
1.2 Main method and the research progress	2
1.3 Structure of this paper	2
Chapter 2 Literature Review	4
2.1 Foreign literature review	4
2.2 Domestic literature review	5
2.3 Summary	9
Chapter 3 Related Theory	10
3.1 Genetic Algorithms.....	10
3.1.1 Advantages of GA.....	10
3.1.2 Basic procedure.....	11
3.1.3 A simple example solved by GA.....	12
3.1.4 GA used in this paper.....	16
3.2 Technical indicators.....	16
3.2.1 Introduction of technical analysis	16
3.2.2 Advantages of technical analysis	17
3.2.3 Technical indicators	17
3.3 Return and risk measure	22
3.4 Portfolio strategy	23
3.4.1 Equal weight portfolio	24
3.4.2 Markowitz weight portfolio	24
Chapter 4 Model Design.....	27
4.1 Data flow	27
4.2 Rules of score	28
4.3 Rules of trade	31
4.4 Genetic Algorithms design.....	32
Chapter 5 Portfolio strategies comparison.....	34
5.1 Equal weight strategy.....	34
5.2 Markowitz weight strategy	34
5.2.1 Markowitz-Equal weight	35
5.2.2 Markowitz-MVP weight	35
5.2.3 Markowitz-Sharpe weight.....	35

5.2.4 Markowitz-GA weight	36
5.3 GA weight strategy	36
Chapter 6 Empirical Analysis in A-share Market	38
6.1 Data.....	38
6.2 Comparison of GA strategy and Buy-and-Hold strategy	39
6.3 Stock picking.....	41
6.4 Portfolio strategies comparison.....	43
Chapter 7 Conclusion	46
References	48
Appendix.....	52
Acknowledgements	58

第一章 引言

随着我国证券市场的发展和金融市场的开放与深化，人们参与证券投资的意愿越来越强，股票投资是一个长期热门的领域。对投资者来说，如何选取股票和有效地管理投资组合是重中之重，如果投资标的选取不理想、投资组合分配不当，就不能产生理想的收益甚至亏损，这就需要选取投资标的、管理投资组合的有效方法。

1.1 背景和意义

随着现代金融理论的发展和计算机技术的广泛应用，用计算机编程来实现证券自动交易的量化投资方法得以不断发展。量化投资兴起于上个世纪 70 年代，凭借成本低、业绩稳定等特点广受关注，在近十年更是欧美资本市场发展的热点。相较于海外成熟市场的发展历程，A 股市场的发展为期较短，且市场上的投资者散户占多数，机构投资者的投资理念还不够成熟，未来在量化投资策略的广泛使用、投资策略多元化方面具有较大的空间。

A 股市场是一个弱有效市场，投资者除了要考虑股票本身的好坏，还要分析宏观政策、实体经济状况、投资者情绪、甚至市场操纵行为等因素。在这样的资本市场环境下，传统选股策略的表现往往差强人意，难以保证有效性。随着计算机技术的发展，投资者开始尝试对原先需要主观判断的现象、情形用数据和指标进行量化，采用计算机智能来代替人们在复杂的市场中逐一判断，建立了量化选股模型，自动寻找市场投资机会。

技术分析、基本面分析并称为两大主流的投资分析方法。基本面分析主要依靠对宏观形势、行业发展、公司财报信息的分析判断选择有潜力的行业、个股进行投资，也称“价值投资”。技术分析通过分析历史的市场数据如股价、成交量等来预测股票未来的趋势，通常表现在技术指标、K 线形态等上面。技术分析方法的核心是“择时”，即寻找买入、卖出的信号，在风险市场中通过仓位管理实现长期稳定的获利。技术分析系统一般包括交易规则、量价转换关系、技术指标或

技术形态等。

遗传算法是一种广泛应用与各大领域的智能优化算法。解决任何一个实际问题都可以看作是在潜在解空间搜索寻找一个解的过程，这实际上是一个优化过程，因为要不断比较好坏以寻求更优的解。对较小的空间，简单的枚举法就足够；而对大的空间，一般的方法效率太低，就需要使用人工智能的技术。遗传算法就是其中的一种：模拟适者生存的自然进化的随机搜索算法。由于其良好的性能和广泛的适用性，目前遗传算法已被广泛应用于投资组合优化、机器学习、自适应控制和信号处理等领域。在证券投资领域，同样存在大量的需要用到优化算法的难题，如实际投资中的选股、风险控制等，而在现代金融市场中，需要优化的问题更是复杂、庞大。

在计算机技术高速发展、金融市场日新月异的背景下，基于证券投资的需要和技术分析在投资领域的广泛应用，探讨研究选股和管理投资组合的有效方法有着十分重大的意义。本文研究的是在技术分析系统中应用遗传算法来构建选股和管理投资组合的量化投资模型。

1.2 主要方法和研究进展

本文使用的主要方法是遗传算法，应用技术指标进行选股和管理投资组合都是用到遗传算法进行优化。选股的部分是用遗传算法对八个技术指标间的权重进行优化，进而根据交易结果选取表现最优的股票，管理投资组合的部分是用遗传算法对投资组合里个股的权重进行优化，以使组合的收益、风险、夏普比率等指标综合达到最优。

本文分别对选股和管理投资组合的量化模型进行设计，并在实证部分对遗传算法策略和 Buy and Hold 策略进行了比较，结果显示遗传算法策略在投资中要优于 Buy and Hold 策略，遗传算法策略的投资组合的收益、夏普比率等指标较等权组合和马科维茨组合策略而言均有更好的表现。

1.3 本文的结构安排

本文的结构如下：

第一部分包括引言和文献综述，介绍国内外在该领域的相关研究的现状和本文研究的创新。

第二部分为模型设计，包括第三、四、五章。第三章为相关理论介绍，包括遗传算法的介绍、本文中用到的技术指标的介绍、风险的计量、Markowitz 投资组合理论；第四章为算法设计，介绍了基于遗传算法的量化选股模型的设计。第五章为投资组合策略模型设计，建立了本文实证部分要比较的三类投资组合策略模型。

第三部分为实证部分，即第六章 A 股市场的实证，展示了该模型在 A 股市场的应用结果，并对比分析了遗传算法和 Buy and Hold 两种交易策略的表现，在投资组合管理部分，给出了遗传算法组合与等权组合、马科维茨组合的比较分析。

第四部分为结论，总结本文所提出的基于遗传算法在技术分析系统中选股和管理投资组合的量化投资模型在 A 股市场的实证能够有较好的应用，并做出了进一步扩展的构思。

第二章 文献综述

本章回顾了国内外在遗传算法应用于选股和投资组合管理这个领域的学术研究现状，并总结本文的主要内容和所做的贡献。

2.1 国外在该领域的研究发展与现状

遗传算法是一种著名的寻求最优解的智能优化方法，最初由 Holland^[1,2]提出。遗传算法常被用来优化金融领域的各种问题，国外的研究已将其应用于证券投资领域的多个方面，本节主要对研究将遗传算法应用于投资领域的相关文献进行回顾介绍。

在投资领域的算法应用方面，Allen 和 Karjalainen^[3]提出了采用遗传算法来优化技术交易程序，结果表明由于与股票指数的收益存在低阶相关性，遗传算法得出的规则并不能比 Buy and Hold 有更稳定的超额收益。Jiang 和 Szeto^[4]提出了一种将遗传算法用于策略查找的方法来管理投资组合，该方法采用了市场分析中的收盘价的平均移动关系。Oh 等^[5]将遗传算法用于设计指数基金组合的优化管理。Acosta-González 等^[6]提出了协整平均这一新的指数跟踪 (Index Tracking) 方法用于选股和统计套利问题。Hassan 等^[7]研究了在不平稳的金融市场上用多目标遗传规划方法来选股的稳健性情况。Sinha 等^[8]设计了一个算法从股票池中选取股票来构建最优的投资组合。

在投资组合管理方面，Markowitz 的均值-方差模型^[9]是最为经典和广泛应用的理论，Markowitz 理论模型的主要内容是用历史收益数据的均值和方差来测算投资组合的期望收益和风险。后来的学者也提出了各种不同的优化方法，如目标规划^[10]和多目标规划^[11]来处理多目标投资组合问题。

在投资组合的优化问题上遗传算法吸引了一些学者进行研究^[12,13]。Arnone 等^[14]将遗传算法用于无约束的组合优化问题。Loraschi 等^[15]介绍了一种分布式的遗传算法，在并行虚拟机中处理投资组合选择问题。Shoaf 和 Foster^[16]采用遗传算法求解 Markowitz 投资组合选择问题，并发现了遗传算法的时间复杂度约为

$O(n \log n)$ ，要优于二次规划方法。然而大多数方法都忽略了最小交易量的存在，Lin 和 Liu^[17]在最小交易量的情况下比较了投资组合选择的三种模型，文中遗传算法被用来提升求解过程的效率，这篇研究表明了用遗传算法得出的投资组合和有效边界的组合非常接近。另一方面，Chang 等^[18]介绍了一种启发式的方法用来优化不同风险度量的投资组合问题，该研究的结果展示了在求解不同风险度量的投资组合的优化问题上，采用遗传算法具有很好的有效性。

除了遗传算法性能方面的优良表现，Chang 等^[18]论证了投资组合的个股数量规模越小收益越好。Reilly 和 Brown 的文章中^[19]有其他的很多因素会影响投资组合的表现，个股的风险和股票间的相关性都需要考虑。Gorgulho^[20]等应用遗传算法进行优化技术指标进行选股，作者在对技术指标进行计分之后进行选股，本文的选股模型部分是参考其思路并做了部分调整，本文在所有股票计分并交易之后，对收益率的表现进行选股。Chung^[21]等将遗传算法用于优化技术指标的参数设置和优化投资组合，表明了遗传算法的优化效果比其他投资组合管理方法要好，文章的侧重点在于对指标的参数进行优化和遗传算法进行改进。

2.2 国内在该领域的研究发展与现状

目前国内在选股模型、投资组合管理、遗传算法在投资领域的应用三个方向上均有若干相关研究文献，本节对三个方向的文献分别回顾介绍。

2.2.1 在量化选股方向上的研究

在量化选股模型方面，有基本面策略、技术面策略、动量反转等量化选股策略。李雪莲^[22]结合技术面和基本面构建了选股模型，分别从技术面和基本面对交易数据和财务数据进行了实证，并分析得出两者结合的量化选股模型更优于单个不结合的选股模型。黄伟男^[23]建立了动量反转策略的量化选股模型，分析了主板、中小板等不同层次的股票市场中动量反转策略的表现情况，构建了有卖空和无卖空机制条件下的量化选股模型并进行了实证分析。江方敏^[24]基于上市公司的主要财务指标建立多因子选股模型，在 A 股市场进行筛选出若干只股票构建投资组合，并对组合的收益和风险进行了简单评价。张利平^[25]构建了多因子量化选股模型，对上市公司的财务数据和交易数据进行分析检验，在其模型中去除冗余因

子得到了若干不冗余因子，并实证分析其有效性。李智^[26]尝试研究了不同层次的股票评价模型，从宏观经济、中观行业、微观公司经营三个层次分别分析股票型基金的量化投资策略，基本面分析指标和技术分析交易数据均有考虑。

随着计算机技术的不断发展应用，一些借助于计算机性能优势的选股策略逐渐有了研究和发展。苏治和傅晓媛^[27]探讨了支持向量回归机的选股策略，基于核主成分遗传算法构建了选股模型(KPCA-GA-SVR)，并对选股性能表现、稳健性和预测精确度做了实证分析，其构建选股模型的研究思路有一定新意。石煜^[28]从数据挖掘的角度探讨了基于现金流的量化选股模型，利用数据挖掘技术对数据库进行分类，将理论指标或属性抽象成一套可量化的规则，最终构建起分析选股平台，它可以对股票进行多维的分析和预测，具有一定的实践意义。毕斌^[29]采用文本挖掘的技术对门户网站的新闻、帖子等内容抓取文本，使用分布式计算机和大规模支持向量机对文本分类分析，不同于传统的基本面和技术面，该文章从股票关注度、情绪指标等角度分析股票的表现，并以此构建量化选股模型。张静^[30]研究了基于技术分析系统的选股模型，使用 Java、.NET 前台开发工具和 SQL 后台数据库对股票交易数据和技术指标进行分析，并模拟第二天的交易行情进行预测。

2.2.2 在投资组合管理方向上的研究

对于投资组合的风险度量问题，主要研究领域有投资组合风险的测度评价、风险管理等。高江^[31]在文章中引入了藤 Copula 函数，用 Montecarlo 模拟法来计算投资组合的 VaR，并通过返回检验方法测试该模型的 VaR 预测效果，实证结果显示该方法较传统的方差-协方差风险管理具有一定的优越性。张鹏^[32]比较了无卖空机制条件下的均值-VaR 和均值-方差两种投资组合模型，发现前者的投资组合有效边界是后者的有效边界的子集。周孝华等^[33]研究了基于 SV-GPD 的边缘分布模型对投资组合风险的测度能力，应用 Copula 函数处理资产的相关性，实证表明了该模型在风险管理上有一定的优越性能。苟红军等^[34]在商业银行风险管理的视角下，由 GARCH-EVT-COPULA 模型建立了 GARCH-EVT-Gaussian-COPULA 和 GARCH-EVT-t-COPULA 两种新的模型对投资组合进行风险度量，并与正态方法、历史模拟法、Montecarlo 法和极值方法分别度量的 VaR 进行了对比并作返回检验。陈辉等^[35]模拟统计保险投资组合的多元数据，对比分析 Copula

方法法和统近似方法(Add-VaR 法、N-VaR 法和 H-VaR 法)下投资组合的风险。吴振翔等^[36]结合 GARCH 模型和 Copula 函数建立投资组合模型。赵喜仓等^[37]则构建了半参数的 Copula-GARCH 模型,对风险资产的边缘分布进行拟合,并用基于 Copula 函数的 Montecarlo 模拟对投资组合进行 VaR 分析和检验。

对于在考虑某些条件下的投资组合优化策略问题,陈其安等^[38]的投资组合模型探讨了投资者情绪问题,研究了投资者情绪处于悲观或是乐观的情况下投资组合的表现情况。刘勇军等^[39]研究了资产收益率等因素为模糊变量的投资组合问题,建立了在有资产组合基数约束、投资比例约束、考虑流动性和分散化程度等条件下的投资组合优化模型,目标包括收益、偏度、风险及模糊性等。李腊生等^[40]对不同风险偏好的投资者的投资组合选择行为,在马科维茨均值-方差模型的理论最优解基础上,分别分析了风险规避、风险中性、风险追求三类投资者在证券市场下的运行表现。任大源等^[41]通过建立考虑交易成本、机会成本的投资组合模型,探讨了证券市场下的多期证券投资组合选择问题。李婷^[42]研究了考虑不同背景风险(健康状况、劳动收入、持有房产等因素)下投资组合的选择问题,研究了这些不能通过资产组合配置来分散风险的影响因素在金融市场中对人们投资行为的影响。金秀等^[43]从行为金融的层面在考虑投资者厌恶损失的心理条件下,建立了动态损失厌恶的投资组合模型,并与马科维茨组合、条件风险价值(CVaR)模型进行实证比较。高丽^[44]借助在线学习算法深入的研究了投资组合问题,基于启发式算法设计了反转在线投资策略,提出了基于均值回归非对称的分段损失函数,从而构建投资优化模型,并利用投资者预期和动量效应设计了不敏感损失函数,以此设计动量在线投资策略。张卫国等^[45]尝试寻找与最优定常再调整策略有相同的指数增长率的投资组合,基于线性学习函数研究了泛证券投资组合策略的应用。

投资组合的风险通常由方差来衡量,但方差的对称性使投资组合的收益受到限制,相比之下下偏距(LPM)不限制收益而只控制风险,凌爱凡等^[46]构建了鲁棒 LPM 投资组合模型,在与均值-方差模型进行实证比较的结果显示该模型具有更好的性能。在定量分析投资组合的风险是,风险资产间的非对称尾部相关性结构非常关键,任仙玲等^[47]用多元阿基米德 Copula 和核密度非参数估计对其进行

Degree papers are in the “[Xiamen University Electronic Theses and Dissertations Database](#)”.

Fulltexts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.