

- 南京大学工程管理学院文库 -

刘海飞 著

金融工程研究中的 交叉学科理论、应用与技术

社会网络视角与中国经验证据



南京大学出版社

金融工程研究中的 交叉学科理论、应用与技术

社会网络视角与中国经验证据

刘海飞 著



南京大学出版社

图书在版编目(CIP)数据

金融工程研究中的交叉学科理论、应用与技术：社会网络视角与中国经验证据 / 刘海飞著. — 南京：南京大学出版社，2017. 11

(南京大学工程管理学院文库)

ISBN 978-7-305-19608-9

I. ①金… II. ①刘… III. ①金融工程—研究 IV. ①F830.49

中国版本图书馆 CIP 数据核字(2017)第 286377 号

出版发行 南京大学出版社

社 址 南京市汉口路 22 号

邮 编 210093

出 版 人 金鑫荣

丛 书 名 南京大学工程管理学院文库

书 名 金融工程研究中的交叉学科理论、应用与技术
——社会网络视角与中国经验证据

著 者 刘海飞

责任编辑 唐甜甜

编辑热线 025-83594087

照 排 南京南琳图文制作有限公司

印 刷 江苏凤凰数码印务有限公司

开 本 710×1000 1/16 印张 12.5 字数 186 千

版 次 2017 年 11 月第 1 版 2017 年 11 月第 1 次印刷

ISBN 978-7-305-19608-9

定 价 69.80 元

网址: <http://www.njupco.com>

官方微博: <http://weibo.com/njupco>

官方微信号: njupress

销售咨询热线: (025) 83594756

* 版权所有, 侵权必究

* 凡购买南大版图书, 如有印装质量问题, 请与所购
图书销售部门联系调换

感谢国家自然科学基金面上项目“社会网络情境下资产选择与递单策略最优化研究——基于计算实验的方法”(71771116)、国家自然科学基金青年项目“流动性黑洞、订单提交策略与最优执行——基于高频数据与计算实验方法的理论与实证研究”(71101068)、国家自然科学基金面上项目“基于计算实验的指令驱动市场交易行为与演化机制研究”(71171109)、江苏省自然科学基金面上基金项目“基于社会网络分析的金融市场资产配置策略理论、方法与应用研究”(BK20161398)、江苏省金融工程重点实验室课题“金融大数据背景下投资组合理论及方法”(NSK2015-09)的资助!

南京大学工程管理学院文库

编委会

主 任 盛昭瀚

副主任 李心丹

委 员 李心丹 钱志新 周 晶 张旭革

周献中 沈厚才 肖条军 周跃进

朱华桂

前 言

世界金融随着时代发展,已进入了“再调整”的不稳定期,而中国金融领域也正在发生深层次变革,为应对新时期金融环境与趋势的挑战,金融市场的投资策略、产品创新、金融风险监管与控制等问题日益凸显,成为金融工程领域研究的学术重点及热点。

现实金融市场是复杂网络系统,不同类型参与主体、交易产品、交易机制等因素,以节点与节点的关联形成复杂的网络结构,即复杂社会网络。事实上,以微博、微信为代表的社会网络在国内的崛起,新媒体时代下金融讯息生成与扩散的完整传播链条,使得市场参与主体学习认知习惯、行为模式、风险控制,甚至金融系统演化涌现规律,发生了本质性变化。当今,金融热点不再是自上而下地通过门户网站、报纸、电视等传统媒体有选择性地层层传播,而是在覆盖在整个社会之上的扁平媒体网络中随机地自发性爆发。金融科技的进步、量化选股技术的发展,使得投资者在新时代新情境下选择投资组合时需要更为有效的方法。本书通过对科学问题的研究,在理论上揭示了中国金融市场中社会网络情境的重要性,及其对投资者决策、学习行为与市场质量的深层次影响,丰富了行为金融与市场微观结构的相关理论研究。在实践上更合理地为金融市场参与主体制定资产管理策略提供理论指导,对加强金融市场风险防范、促进制度创新有重要的科学意义。本书对有志于从事金融研究的硕士生、博士生,特别是基于交叉学科视角的金融研究方向学生具有参考价值。

本书共十章,主要分为两个部分。第1部分为第1、第2章,从现实背景与理论研究的双重角度出发,说明社会网络理论在金融领域应用的可行性与创新性。第1章为本书绪论,提出社会网络理论在金融市场中的应用背景,针对性地提出学术研究的热点问题,列示本书的研究思路、行文框架及重点内容,提纲挈领地说明本书研究方法与特色。第2章为社会网络在四大金融领域应用的理论综述,分别为社会网络理论在投资者决策、资产组

合、风险控制及资产定价领域的理论与实证应用,指明本书的深耕领域与研究方向。

第3至第10章为实证技术与应用。从社会网络的视角,利用金融工程学科的技术与方法,分层次递进式地研究如下科学问题:基于社会网络的互联网金融论文合作关系;基于社会网络的选股与指数复制应用;基于社会网络的互联网热点新闻对股价影响;基于社会网络的情感信息与资产组合建模及其分散化;基于社会网络的舆情信息与资产定价效率;基于社会网络的投资者关注与股价同步性;基于社会网络的线性组合优化及其分散化;基于社会网络的沪港通网络联动性及稳定性等。

本书在撰写过程中有幸得到国家自然科学基金青年项目“流动性黑洞、订单提交策略与最优执行——基于高频数据与计算实验方法的理论与实证研究”(项目编号:71101068)、国家自然科学基金面上项目“基于计算实验的指令驱动市场交易行为与演化机制研究”(项目编号:71171109)、江苏省自然科学基金面上项目“基于社会网络分析的金融市场资产配置策略理论、方法与应用研究”(项目编号:BK20161398)、江苏省金融工程重点实验室课题“金融大数据背景下投资组合理论及方法”(项目编号:NSK2015-09)的支持,在此表示衷心的感谢!此外,本书部分章节被《管理科学学报》《产业经济研究》等期刊发表,感谢期刊和审稿专家的帮助!

在本书即将付梓之际,诚挚地感谢我的博士生导师李心丹教授,在工作的道路上离不开恩师的指点;感谢南大工管院同事们的大力帮助。我还要感谢博士生柏巍以及研究生许金涛、孙宸、刘杰、郭剑雄、陈蕾、徐梓轩、查恒、钱泽宇、董令文、郑可在本书撰写整理过程中提供的帮助。

我要将特别的感谢致以我的母亲、岳母,以及我的爱人王利女士,爱子刘毅樊小朋友,感谢他们的理解与支持!

最后,我还要感谢南京大学出版社编辑唐甜甜等的帮助,他们的细致工作使本书得以及时顺利出版。由于学识所限,书中难免存在不足之处,敬请大家批评指正!

刘海飞

2017年4月于南京

目 录

前 言	i
第1章 绪 论	1
1.1 背景与问题提出	1
1.2 思路、框架与结构安排	3
1.3 研究方法与特色	4
第2章 国内外文献综述与评析	6
2.1 社会网络与投资者决策研究综述	6
2.2 社会网络与资产组合的研究综述	7
2.3 社会网络与风险控制的研究综述	7
2.4 社会网络与资产定价研究综述	8
2.5 现有文献的评析与展望	9
第3章 基于社会网络的互联网金融论文合作关系研究	11
3.1 理论回顾	12
3.2 数据采集与样本选择	13
3.2.1 数据采集	13
3.2.2 数据预处理	13
3.3 实证结果及分析	15
3.3.1 统计性描述分析	15
3.3.2 整体网络分析	18
3.3.3 子网络模式分析	22
3.4 本章小结	26
第4章 基于社会网络的选股与指数复制应用研究	28

4.1 理论回顾	28
4.2 模型构建	30
4.2.1 AAP 聚类选股模型	30
4.2.2 指数跟踪模型	31
4.3 数据采集与变量构建	32
4.4 实证结果及分析	33
4.4.1 沪深 300 关联网路分析	33
4.4.2 沪深 300 成分股连通网络聚类选股分析	34
4.4.3 稳健性检验	42
4.4 本章小结	45
第 5 章 基于社会网络的互联网热点新闻对股价影响研究	46
5.1 理论回顾	47
5.2 研究假设提出	49
5.3 研究设计	52
5.3.1 数据采集与处理	52
5.3.2 财经新闻异质性识别	52
5.3.3 异质性新闻有效性模型	54
5.4 实证结果及分析	56
5.4.1 政策扶持类新闻	57
5.4.2 兼收并购类新闻	59
5.4.3 再融资类新闻	61
5.4.4 盈利能力类新闻	63
5.4.5 违规处罚类新闻	64
5.5 本章小结	66
第 6 章 基于社会网络情感信息与资产组合建模及其分散化研究	68
6.1 理论回顾	69
6.2 研究设计	71
6.2.1 数据来源及样本选取	71

6.2.2 文本挖掘选股因子	72
6.2.3 投资组合构建	74
6.2.4 投资组合分散化水平分析	75
6.3 数据来源	77
6.4 实证结果及分析	78
6.4.1 文本挖掘选股因子	78
6.4.2 投资组合构建	80
6.4.3 投资组合有效前沿对比	82
6.4.4 投资组合分散化水平分析	82
6.5 本章小结	84
第7章 基于社会网络舆情信息与资产定价效率研究	85
7.1 理论回顾	86
7.2 研究设计	88
7.3 实证结果及分析	88
7.3.1 股票组合的日平均收益率检验结果及分析	88
7.3.2 股票组合的时间序列回归实证结果及分析	89
7.3.3 舆情因子的构建方法	94
7.3.4 舆情因子的具体计算方法	95
7.3.5 股票组合的形成	97
7.4 本章小结	97
第8章 基于社会网络投资者关注与股价同步性研究	99
8.1 理论回顾	100
8.2 研究假设提出	102
8.3 研究设计	103
8.3.1 数据来源	103
8.3.2 变量选取	103
8.3.3 计量模型设计	106
8.4 实证结果及分析	107

8.5 本章小结	110
第9章 基于社会网络线性组合优化及其分散化研究	112
9.1 理论回顾	113
9.2 社会网络选股和分散化理论	115
9.2.1 基于社会网络的标的选择	115
9.2.2 分散化理论	117
9.3 数据采集与处理	118
9.4 实证结果及分析	118
9.4.1 基于社会网络选股的实证研究	118
9.4.2 均值-方差分析	121
9.4.3 组合分散化程度分析	124
9.5 本章小结	130
第10章 基于社会网络的沪港通网络联动性及稳定性研究	131
10.1 理论回顾	132
10.2 模型构建	133
10.2.1 构建沪港通关联网络	133
10.2.2 沪港通关联网络的稳定性分析	135
10.3 数据采集与处理	136
10.4 实证结果及分析	138
10.4.1 沪港通市场关联网络	138
10.4.2 沪港通市场网络稳定性	143
10.5 本章小结	150
参考文献	152

表目录

表 3-1	论文篇数统计表	14
表 3-2	论文合著情况统计表	15
表 3-3	发表论文数的作者频数分布表	16
表 3-4	每篇论文作者数的论文频数分布表	17
表 3-5	合作者数的作者频数分布表	17
表 3-6	基本统计数据表	18
表 3-7	合作关系网络节点度的分布	20
表 3-8	合作关系网络节点强度的分布	20
表 3-9	子网络按其所包含的节点数汇总表	22
表 3-10	完备模式子网络分布统计表	24
表 3-11	前 10 个主要子网络的统计特征数据表	26
表 4-1	市值排序选股的股票信息统计表	38
表 4-2	权重排序选股的股票信息统计表	39
表 4-3	AAP 聚类选股的股票信息统计表	41
表 5-1	互联网异质性新闻的分类识别结果统计表	54
表 5-2	政策扶持类新闻事件的异常收益率检验统计表	57
表 5-3	兼收并购类新闻事件的异常收益率检验统计表	59
表 5-4	再融资类新闻事件的异常收益率检验统计表	61
表 5-5	盈利能力类新闻事件的异常收益率检验统计表	63
表 5-6	违规处罚类新闻事件的异常收益率检验统计表	65
表 6-1	股吧 VS 雪球网站信息对比表	71
表 6-2	利多、利空词库	73
表 6-3	利多新闻示例	74
表 6-4	贵州茅台 12 月份关注度指标与情感指标	78
表 6-5	神奇公式选股模型股票组合	80

表 6-6	文本挖掘选股模型股票组合	81
表 6-7	选股模型入选标的的边际风险贡献	83
表 6-8	投资组合分散化水平分析	84
表 7-1	不同舆论关注度的股票组合日均收益率	89
表 7-2	单指标舆情因子的实证检验	90
表 7-3	综合指标作为舆情因子的实证检验	92
表 7-4	股票板块模型实证检验	93
表 7-5	单项舆情指标的相关性分析	94
表 7-6	三因子的股票组合	95
表 8-1	社交网络平台微博信息质量评价指标构建	104
表 8-2	研究变量描述性统计表	107
表 8-3	研究变量相关性分析表	108
表 8-4	社交网络信息质量与股价同步性线性与非线性模型回归估计	108
表 9-1	数据处理后的三个市场数据统计	118
表 9-2	上证 A 股,2006 年 4 月 1 日至 2014 年 12 月 31 日社交网络最优聚类选股 结果	120
表 9-3	排序选股结果	122
表 9-4	不同限制条件上证 A 股长周期股票策略的 MV 分析	123
表 9-5	不同限制条件上证 A 股熊市股票策略的 MV 分析	124
表 9-6	不同限制条件上证 A 股长周期股票策略的 MD 分析	129
表 9-7	不同限制条件沪深 300 熊市股票策略的 MD 分析	129
表 10-1	沪股通股票行业分布统计表	137
表 10-2	沪股通、港股通、沪港通市场关联网中心度信息统计表	139

图目录

图 3-1 论文篇数走势图	14
图 3-2 单一模式子网络图	23
图 3-3 双核模式子网络图	23
图 3-4 完备模式子网络图	24
图 3-5 发展模式子图	25
图 4-1 沪深 300 关联网络图	34
图 4-2 沪深 300 成分股关联网络图	34
图 4-3 沪深 300 指数跟踪分析	36
图 4-4 不同选股方法下的日均跟踪误差曲线图	37
图 4-5 沪深 300 指数及各跟踪指数的收益率走势图	42
图 4-6 不同选股方法的日均跟踪误差曲面图	44
图 4-7 优势指标曲线图	45
图 5-1 事件研究法的研究窗划分示意图	55
图 5-2 政策扶持类新闻事件的异常收益率走势图	58
图 5-3 兼收并购类新闻事件的异常收益率走势图	60
图 5-4 再融资类新闻事件的异常收益率走势图	62
图 5-5 盈利能力类新闻事件的异常收益率走势图	64
图 5-6 违规处罚类新闻事件的异常收益率走势图	66
图 6-1 原始风险因子与最小扭转赌注、主成分赌注	77
图 6-2 贵州茅台文本挖掘买入卖出信号	79
图 6-3 神奇公式 VS 文本挖掘投资组合有效前沿对比	82
图 9-1 上证 A 股 2006 年 4 月 1 日至 2014 年 12 月 31 日社交网络聚类 结果	119
图 9-2 基于上证 A 股社会网络选股 21 只股票的均值-分散化分析	126
图 9-3 基于上证 A 股排序选股 21 只股票的均值-分散化分析	128

图 10-1 沪港通事件研究时间划分示意图	138
图 10-2 沪港通市场在四个不同时期的关联网络图	139
图 10-3 沪港通关联网络平均路径长度	143
图 10-4 沪港通市场关联网络的全局效率曲线图	144
图 10-5 沪港通市场关联网络的破碎指标走势图	145
图 10-6 沪股通市场关联网络的全局效率曲线图	146
图 10-7 沪股通市场关联网络的破碎指标走势图	147
图 10-8 港股通市场关联网络的全局效率曲线图	148
图 10-9 港股通市场关联网络的破碎指标走势图	148
图 10-10 沪港通、沪股通、港股通各时段较前一时段稳定性提高程度	149

第1章 绪论

1.1 背景与问题提出

互联网、内联网等社会网络改变了人们的沟通方式,打破了地域的限制,使得对传统媒体的依赖正在慢慢地转移。政府、社会机构、金融机构非常关注在网络媒体上表达的观点。以微博、微信为代表的社会网络在国内的崛起,新媒体时代下金融讯息生成与扩散的完整传播链条,使得市场参与主体的学习认知习惯、行为模式、风险控制,甚至金融系统演化及涌现规律,发生了本质性的巨大变化。如今金融热点不再是自上而下地通过门户网站、报纸、电视等传统媒体有选择性地层层传播,而是在覆盖在整个社会之上的扁平媒体网络中随机地自发性爆发^[1]。目前,中国散户投资者以20~50岁的白领和自由职业者为主,而这些人也正好是社会网络的主力适用人群。在新媒体时代,网民和散户这两个群体正在加速融合。社会网络不但极其敏感地反映出这部分市场主体的整体价值观与其所关注的兴趣点,更反过来凭借其强大的用户附着力^{①[2][3][4]},占据了这部分市场参与主体大量的注意力资源,进一步影响到他们投资决策理念、资产组合选择量化模式等,进而形成对金融市场的驱动力。

现如今随着金融科技迅猛发展,社会热点事件在发生后的短短几分钟内,即可通过智能手机等移动终端进入社会网络并爆炸式传播,而依附于这些信息的投资交易决策即是该资讯所关联的证券的买卖链接,点击下单即可完成交易。传统资产选择、组合构建完全凭投资主体经验,及其交易层面的手工操作模式,受到了资金规模扩大、市场风险加剧、波动频繁等问题的

① 用户附着力是指用户对产品,以及产品交易过程中工具的依赖。

挑战。而 IT 与通信技术的不断完善与成熟,使得快速交易递单策略模式成为现实,它为解决交易操作效率、风险管理等难题带来了曙光。谁能拥有对社会网络热点信息流行的实时监控和快速反应能力,谁就能在新媒体与金融科技时代情景下证券投资市场中占得先机。可见,互联网金融时代的异质性信息传递与资产间网络结构,正在发生根本性的革新,微观层面上金融资产交易模式也呈现模型化、自动化的趋势,以平衡市场冲击成本与时间等待风险为目标的递单交易算法的理论与实践被广泛关注。因此,在新媒体与金融科技的现今时代背景下,从金融资产间和投资主体间的拓扑网络视角,探究社会网络结构对资产选择、递单交易的行为决策模式,以及宏观市场质量的影响等,其重要性日益凸显。

20 多年以来,国内外学者对金融市场资产配置、组合分散化、互动学习、递单算法等投资者行为理论的相关金融问题的研究,已经跳出了传统的完全理性的研究框架,历经了理性到有限理性的研究路径。虽然这些研究增进了学者们对投资者行为理论与实践的了解,但是研究背景方面仍存在着一些局限性。现有研究仅仅从个体的视角,并未过多地注意到人人关联的社会网络这一重要因素,及其在金融市场造成的广泛影响,理论和应用领域等诸多方面存有不足与研究空白。实际上,社会学家舒尔茨(Myers Schultz)早在 20 世纪 50 年代,已将社会网络的理论和观点引入经济学的分析中。直到六度分割理论的提出,即单个主体和任何一个陌生人之间所间隔的人不会超过六个,社会网络的普适性与重要性被充分地展现出来^[5]。近几年,以社会网络的视角研究金融问题,也引起了国内外不少学者的重视。有关社会网络对金融影响的研究,诸如关于社会网络成员之间关联程度衡量研究^[6];基于网络舆情信息搜索的研究^[7];关于信息传染与主体趋同行为研究^[8];关于主体(个人或企业)之间风险分担的研究^[9];关注金融网络结构及其稳定性研究^[10];金融网络与主体投资行为的影响研究^{[11][12][13]}等方面,也充分地体现出社会网络理论与方法应用到金融研究的适应性与可行性。

另外,诸如资产配置、组合分散化、互动学习、递单算法等相关研究在分析方法、手段以及工具方面也有一定的局限性。近几年,以现代计算机技