

多维数据下的大数据择时策略研究

——互联网大数据挖掘系列专题之（八）

报告摘要:

● 网络热度择时策略思想

我国 A 股投资者中散户众多，散户投资者的情绪、对市场的关注程度通常对市场涨跌起到推波助澜的作用。这些投资者的“关注度”往往可以反映在热门财经网站的关注人数和搜索引擎的搜索量上，如百度、东方财富网、新浪财经、雪球网、和讯财经等，本篇专题报告称之为“网络热度”。

在大数据发展趋势下，互联网数据为量化投资提供了新的数据来源，搜索量、股吧情绪、新闻热度、个股关注度等新型的数据相比传统数据更加能够反映整个市场的情绪，本篇专题报告通过对这些热度数据进行深入研究，探讨这些数据在择时策略上的应用效果。

● 网络热度择时策略构建

利用网络的热度数据，构建了布林通道，根据均线设置上轨和下轨，将上轨、下轨分别作为预测沪深 300 指数次日涨跌的两个阈值。具体地，择时策略原理如下：

设置布林通道：以热度数据的 M 日均值构造均线，以 M 日热度数据的标准差的 N 倍构造上轨、下轨，即：

布林通道上轨=热度数据的 M 日均值+N*M 日热度的标准差；

布林通道下轨=热度数据的 M 日均值-N*M 日热度的标准差；

择时策略：当日热度突破通道上轨，则发出看多信号，次日开盘买入；当日热度突破通道下轨，则发出看空信号，次日开盘卖出；相连两日发出不同信号，先平仓后根据信号进行买卖；相连两日的第二日不发出信号，次日开盘直接平仓，考虑可以做空的收益；

● 实证结果

我们采用了三个网络源的热度数据进行实证分析。热度 1 发出的信号准确，测试区间累计净值达到 10.31，年化收益率 58.21%，胜率 57.61%，赔率 1.14，累计最大回撤仅有 18.11%，收益回撤比达到 3.23。热度 2 和热度 3 经过综合信号策略改进后，区间内实现的收益率相比单一热度择时更高，超额年化收益率达到 54.47%，累计最大回撤下降，收益回撤比显著提高。

● 风险提示

本报告提出的择时策略基于网络热度数据，热度数据仅仅是投资者投资情绪的一个方面表现，市场行情受到其他因素影响。

图 1 网络热度择时策略表现

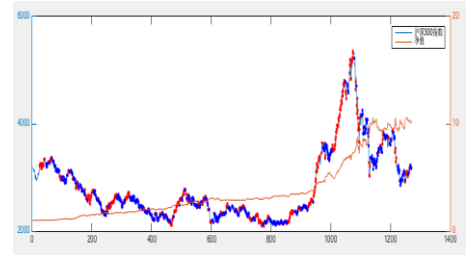


表 1 网络热度择时策略指标表现

累计净值	10.31
年化收益率	58.21%
超额年化收益率	58.55%
累计最大回撤	18.11%
胜率	57.61%
赔率	1.14

分析师：史庆盛 S0260513070004



020-87577060



sqs@gf.com.cn

相关研究：

- 《基于网络新闻热度的择时策略——互联网大数据挖掘系列专题(一)》 2014-06-25
- 《基于互联网挖掘的热点选股策略——互联网大数据挖掘系列专题之（五）》 2015-04-09
- 《基于大数据挖掘的 Smart Beta 策略——互联网大数据挖掘系列专题之（七）》 2016-03-23

联系人：陈原文 020-87576976
chenyuanwen@gf.com.cn

目录索引

一、前言	4
1.1 互联网大数据与量化投资	4
1.2 互联网大数据研究体系	6
1.3 互联网热度择时策略介绍	7
二、网络热度择时策略构建	8
2.1 样本数据	8
2.2 策略原理	11
三、实证结果	12
3.1 热度 1 实证结果	12
3.2 热度 2、热度 3 实证结果	13
3.3 参数敏感性测试	15
四、总结	17
4.1 网络热度是有效的量化择时数据	17
4.2 研究不足与未来方向	17

图表索引

图 1 网络热度择时策略表现	1
图 2 CAYMAN ATLANTIC 旗下大数据基金产品业绩	5
图 3 广发金工数据抓取体系	6
图 4 广发金工大数据挖掘策略框架	6
图 5 A 股成交额与沪深 300 指数走势	7
图 6 热度 1 与沪深 300 指数走势	8
图 7 网络热度择时策略表现	8
图 8 中证 500 网络情绪走势	9
图 9 行业网络情绪变化一览	9
图 10 网络热度一览	9
图 11 热度 1 与沪深 300 指数走势	10
图 12 热度 2 与沪深 300 指数走势	10
图 13 热度 3 与沪深 300 指数走势	10
图 14 热度策略择时表现	12
图 15 热度 1 热度 1 综合信号择时策略表现	13
图 16 热度 2 择时策略表现	14
图 17 热度 3 择时策略表现	14
图 18 热度 2、热度 3 综合策略表现	15
图 19 热度 1 敏感性测试结果	16
图 20 热度 2 敏感性测试结果	16
图 21 热度 3 敏感性测试结果	16
表 1 网络热度择时策略指标表现	1
表 2 大数据量化产品一览	5
表 3 互联网挖掘小工具一览	7
表 4 上交所交易额按投资者分类	7
表 5 热度与次日沪深 300 指数的相关系数	10
表 6 热度 1 择时策略指标表现一览	12
表 7 热度 1 择时策略分年度指标表现一览	13
表 8 热度 1 综合信号择时策略指标表现一览	13
表 9 热度 2 择时策略指标表现一览	14
表 10 热度 3 择时策略表现指标一览	14
表 11 热度 2、3 综合策略指标表现一览	15

一、前言

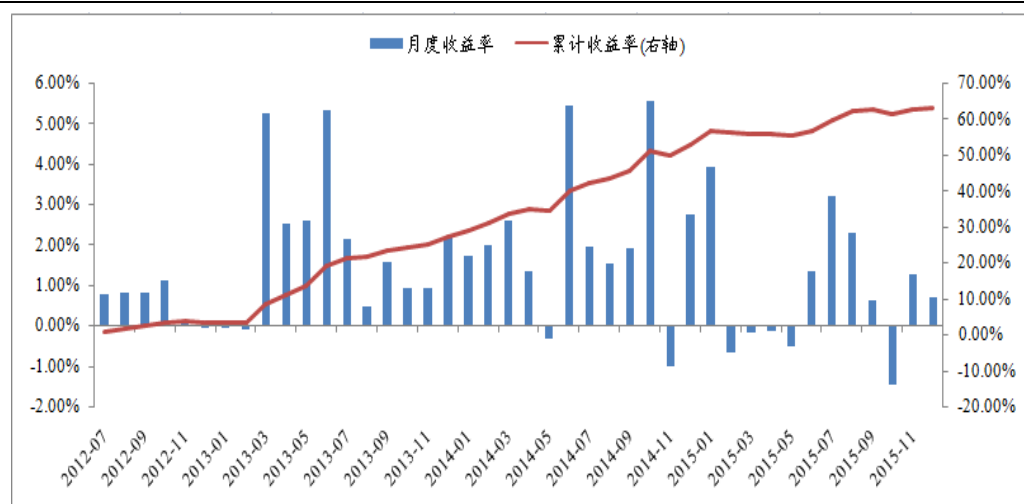
1.1 互联网大数据与量化投资

近十几年来，互联网在计算机技术的不断发展下得到了迅速的普及和发展，逐渐替代传统的纸质传播媒介。互联网的传播已经逐渐改变了人们的生活方式和习惯，尤其是在这个信息爆炸的时代，人们已经习惯于利用互联网搜索、了解和关注自己感兴趣的新闻、热点事件等。伴随着人们对互联网使用程度的增加，互联网上沉淀的数据规模已呈指数型速度增长，数据的存储结构形式也发生了多样化的概念，如何存储、管理以及挖掘这些数据中所隐藏的价值，逐渐成为一个新的领域，大数据的概念便伴随之产生了。

另一方面，传统的量化投资研究面临瓶颈。传统的量化投资一般从财务指标、行情数据、行业数据中挖掘有价值的投资信息，而随着量化投资领域的发展，这些传统数据中隐含的大部分投资信息已经被专业投资者所挖掘，专业投资者希望另辟蹊径，寻找新的数据维度来挖掘有用信息。互联网大数据由于具有数据量大 (Volume)、类型繁多 (Variety)、价值密度大 (Value)、时效性高 (Velocity) 的特征，为量化投资提供了新的数据来源，搜索量、股吧情绪、新闻热度、个股关注度等新型的数据相比传统数据更加能够及时甚至提前反映整个市场的情绪，从而能够为投资决策起到指导的作用。

纵观海内外，对大数据相关的量化投资策略的研究近几年刚刚兴起，但对互联网大数据相关的量化策略的研究已然成为量化投资新的研究方向和热门研究领域。2011年5月，英国对冲基金Derwent Capital Markets建立了规模为4000万美金的对冲基金，该基金是首家基于社交网络的对冲基金，通过分析Twitter的数据内容来感知市场情绪，从而指导进行投资。在2012年，CAYMAN ATLANTIC公司新成立了一只基于网络社交媒体Twitter、搜索引擎Google及其他新闻媒体数据的进行交易的对冲基金。根据其官网披露的基金历史收益数据，截止2015年12月，该基金累计收益率高达63.21%，年化收益率为15.02%，在基金运行的42月中有32个月获得了正收益，亏损的10个月当中，平均月亏损为-0.45%，而获得正收益的月份中，平均月收益率为2.12%。国内各大公募基金相继与互联网公司合作，推出了一些大数据基金产品，如广发基金与百度的中证百度百发策略100指数型基金、南方基金与新浪合作的大数据100指数型基金等等。目前国内大数据量化产品包括了指数型、股票型和混合型基金以及集合资产管理计划等类型。

图 2 CAYMAN ATLANTIC 旗下大数据基金产品业绩



数据来源：CAYMAN ATLANTIC 官网、广发证券发展研究中心

表 2 大数据量化产品一览

基金公司	合作方	大数据基金	产品类型
广发基金	百度	中证百度百发策略 100 指数型	指数型
博时基金	蚂蚁金服	中证淘金大数据 100 指数型	指数型
南方基金	新浪	大数据 100 指数型	指数型
东方资管	京东	东方红京东大数据灵活配置混合型	混合型
南方基金	新浪	大数据 300 指数型	指数型
博时基金	银联	中证银联智惠大数据 100 指数型	指数型
广发基金	百度	百发大数据策略精选灵活配置混合型	混合型
广发基金	百度	百发大数据策略成长灵活配置混合型	混合型
嘉实基金	腾讯	腾讯自选股大数据策略股票型	股票型
大成基金	360	中证 360 互联网+大数据 100 指数型	指数型
大成基金	雪球	中证雪球社交投资精选大数据指数型	指数型
鹏华基金	银联	中证银联智策消费大数据指数型	指数型
中欧基金	银联	中证银联智策大数据 100 指数型	指数型
博时基金	房天下	中证房天下大数据指数型	指数型
博时基金	雪球	中证雪球智选大数据 100 指数型	指数型
海富通基金	东方财富	东财大数据灵活配置混合型	混合型
富国基金	蚂蚁金服	中证娱乐大数据指数型	指数型
泰达宏利基金	同花顺	同顺大数据量化优选灵活配置混合型	混合型
天弘基金	蚂蚁金服	中证电商大数据指数型	指数型
天弘基金	雪球	中证雪球领先组合 100 指数型发起式	指数型
富国基金	蚂蚁金服	中证医药大数据指数型	指数型
广发资管	新浪网	广发资管玺智大数据精选	集合资产管理计划

数据来源：广发证券发展研究中心

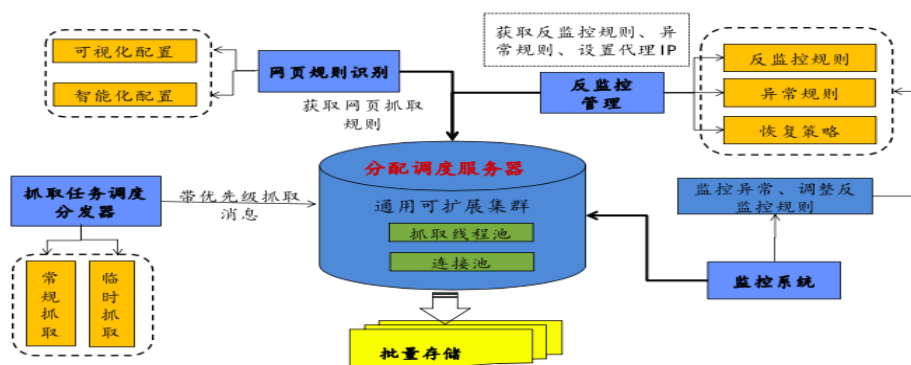
随着机构投资者对大数据这一块的重视性的逐渐加强，对互联网大数据的量化

投资研究将更加地多样化，包括对市场的择时、行业配置、选股、热点追踪等领域，甚至是实时的大数据量化监控系统。

1.2 互联网大数据研究体系

在互联网数据挖掘与研究方面，广发金工建立了完善的数据抓取平台以及对互联网大数据各个数据维度的研究视角。

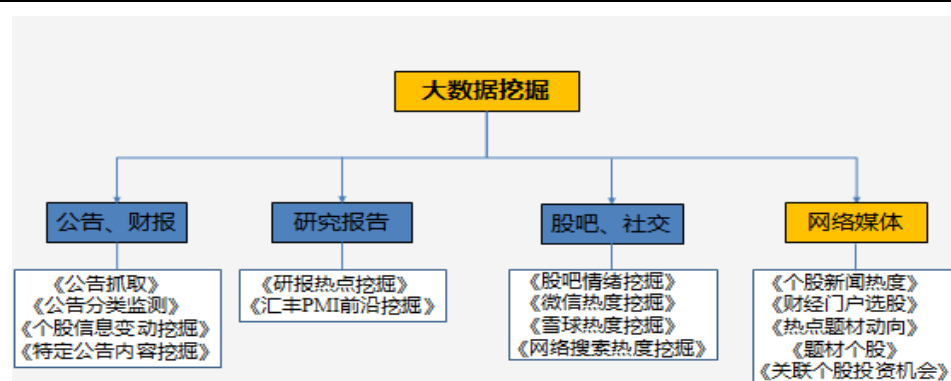
图 3 广发金工数据抓取体系



数据来源：广发证券发展研究中心

广发金工研究了多个维度的互联网数据，比如从公告、财报角度研究的报告《公告披露背后隐藏的投资机会——互联网大数据挖掘系列专题之（二）》等；从股吧、社交角度研究的报告《倾听股吧之声，洞察大盘趋势——互联网大数据挖掘系列专题之（三）》等；以及从网络媒体角度分析的报告《基于互联网挖掘的热点选股策略——互联网大数据挖掘系列专题之（五）》、《基于大数据挖掘的关联个股投资机会——互联网大数据挖掘系列专题之（六）》、《基于大数据挖掘的Smart Beta策略——互联网大数据挖掘系列专题之（七）》等。

图 4 广发金工大数据挖掘策略框架



数据来源：广发证券发展研究中心

基于对大数据的研究积累，广发金工也开发了一系列的互联网挖掘小工具，包括上市公司公告抓取、研究报告抓取、搜索量抓取、关注度抓取等小工具。

表 3 互联网挖掘小工具一览

A 股新闻热度搜索工具	A 股上市工具公告抓取工具
上市公司信息变更抓取	文本信息批量识别及处理
汇丰 PMI 实时监测工具	个股研报热点监测工具
特定公告实时监测工具	财经小编选股工具

数据来源：广发证券发展研究中心

1.3 互联网热度择时策略介绍

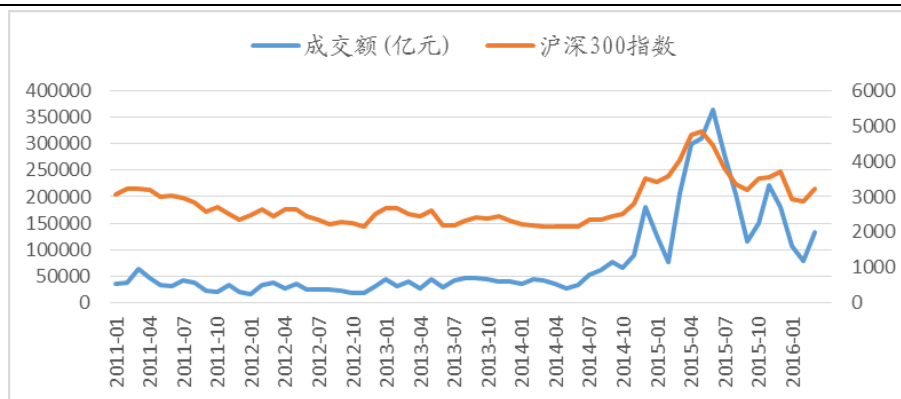
我国A股市场的散户投资者众多，散户投资者贡献了A股市场的80%以上的成交额。尽管散户投资者在整个A股市场总市值上占比不大，但是散户投资者的巨大成交额与股市涨跌有极强的相关性。2011年以来的统计表明，A股成交额与沪深300指数走势相近，二者之间的相关系数达到0.88。

表 4 上交所交易额按投资者分类

	2011 年	2012 年	2013 年
自然人	83.50%	80.80%	82.20%
一般法人	2.10%	2.10%	2.50%
专业机构	14.40%	17.10%	15.30%

数据来源：公开资料、广发证券发展研究中心

图 5 A股成交额与沪深300指数走势



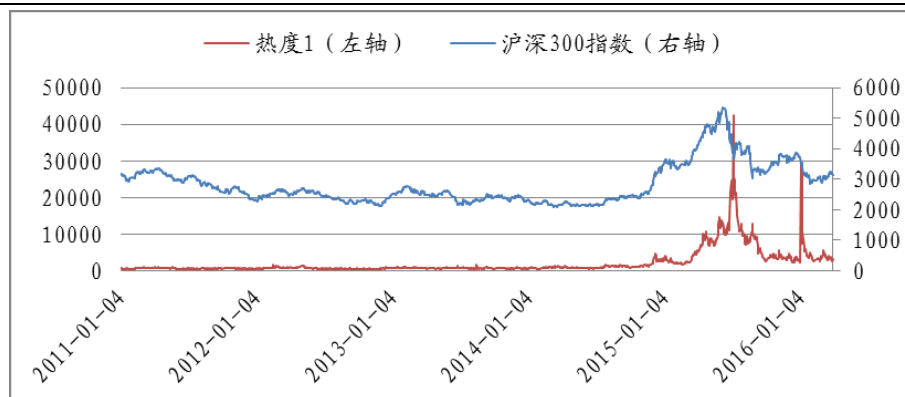
数据来源：WIND、广发证券发展研究中心

可见，股市的涨跌通常伴随着成交额的增减变动，而散户投资者作为A股成交额的主力军，他们的情绪对市场涨跌有着推波助澜的作用，是一种领先于市场行情的指标。当投资者对市场情绪较乐观时，会先搜索、关注相关标的的财务数据、行业信息等相关信息，然后进行投资交易，而行情数据是投资者交易的结果，已经反映了这些情绪的变动，是相对于投资者情绪的滞后数据。

我们基于一些热门专业的财经网站和搜索引擎，比如百度搜索引擎、东方财富网、新浪财经、雪球网、和讯财经等互联网数据源，利用网站上股票的关注人数、搜索量数据，得出散户投资者的“关注度”指标，本篇专题报告称之为“网络热度”。本篇专题报告通过对这些网络热度的数据进行分析，探讨网络热度数据在大盘指数上的择时应用。

以某网站的网络热度为例。2011年至今，该热度数据与沪深300指数走势相近，我们计算二者之间的相关性，热度数据与次日沪深300指数的相关系数为0.68，说明该热度数据的表现与次日沪深300指数的表现有很强的相关关系。

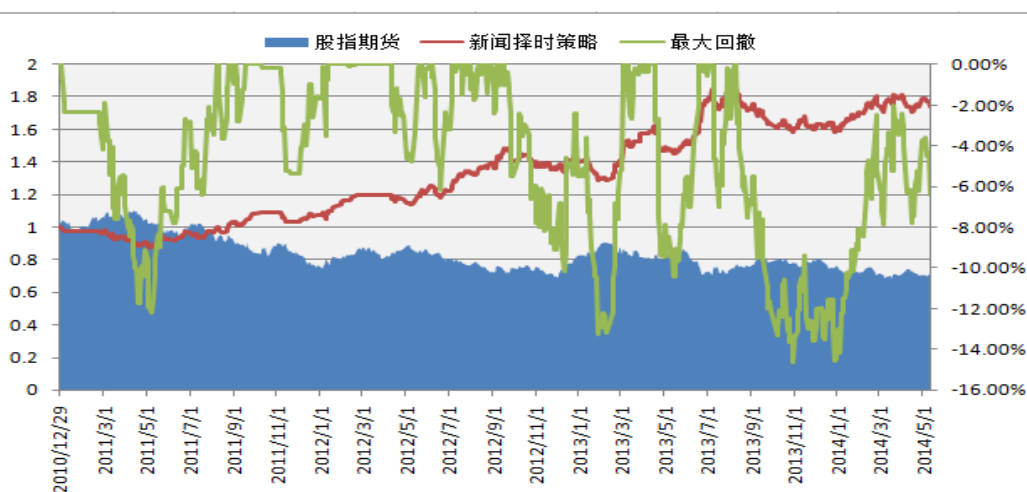
图 6 热度1与沪深300指数走势



数据来源：互联网、广发证券发展研究中心

广发金工曾经在报告《基于网络新闻热度的择时策略—互联网大数据挖掘系列专题(一)》中介绍过利用网络热度进行择时的策略方法，本报告是在这篇报告的基础上进一步研究网络搜索量、关注度等数据对大盘的影响。在本报告的后续部分，我们将使用不同来源的网络热度数据（热度1，热度2，热度3），构建我们的热度择时策略。

图 7 网络热度择时策略表现



数据来源：广发证券发展研究中心

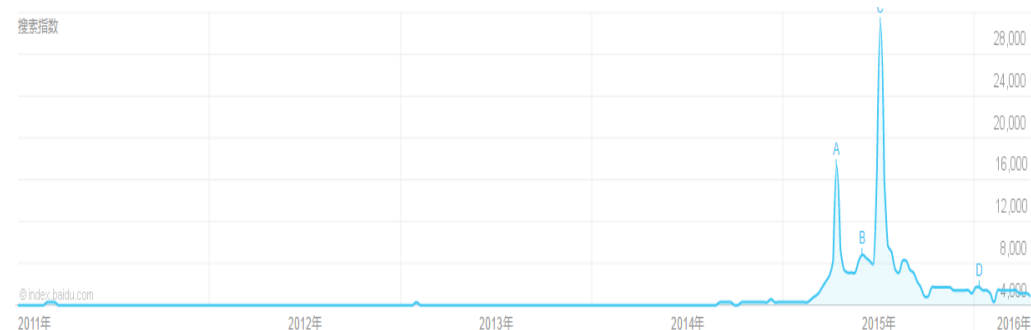
二、网络热度择时策略构建

2.1 样本数据

我们采用了三个典型的网络数据源的热度数据，分别记为：热度1，热度2，热度3。热度数据是基于A股上市公司的关注人数每日增量数据的加总或搜索量的每日

数据，反映的是整个市场的当日的热度。由于数据的可获取性，三个热度数据的时间区间不同。统计结果表明，三个热度与沪深300指数均有相近的走势，并且热度与次日指数相关系数均超过了0.5。（关于具体的热度数据来源，欢迎感兴趣的机构投资者通过邮件或电话详细咨询广发金工团队）。

图 8 中证 500 网络情绪走势



数据来源：互联网、广发证券发展研究中心

图 9 行业网络情绪变化一览

行业简称	关注度	关注度变化值	关注度变化幅度	排名趋势	行业平均关注度
工业工程	50872	11276 ↑	28.5% ↑	-	154
电子与电气设备	48945	10709 ↑	28.0% ↑	-	166
化工品（III）	38499	5725 ↑	17.5% ↑	-	141
软件与计算机服务	35123	6042 ↑	20.6% ↑	-	249
不动产	33338	7702 ↑	30.0% ↑	-	224
硬件与设备	28875	6658 ↑	30.0% ↑	1 ↑	298
有色金属	28034	11700 ↑	71.6% ↑	4 ↑	438
医药与生物科技	28010	9334 ↑	50.0% ↑	2 ↑	172
建筑与材料（III）	27733	4376 ↑	16.7% ↑	3 ↓	154
汽车与零配件（III）	23386	2347 ↑	11.2% ↑	2 ↓	246
食品生产（III）	22008	2631 ↑	13.6% ↑	2 ↓	193
采矿业	20450	7929 ↑	63.3% ↑	5 ↑	310
个人用品	19967	5087 ↑	34.2% ↑	1 ↑	215
银行业（III）	17468	1501 ↑	9.4% ↑	2 ↓	1092
普通金融服务	16251	2609 ↑	19.1% ↑	-	580
一般零售业	13747	1401 ↓	9.3% ↓	3 ↓	178
钢铁	13668	5750 ↑	72.6% ↑	6 ↑	325
辅助服务	12643	2802 ↑	28.5% ↑	3 ↑	253
传媒（III）	12450	532 ↑	4.5% ↑	1 ↓	271
电力（III）	12360	2407 ↑	24.2% ↑	-	203

数据来源：互联网、广发证券发展研究中心

图 10 网络热度一览



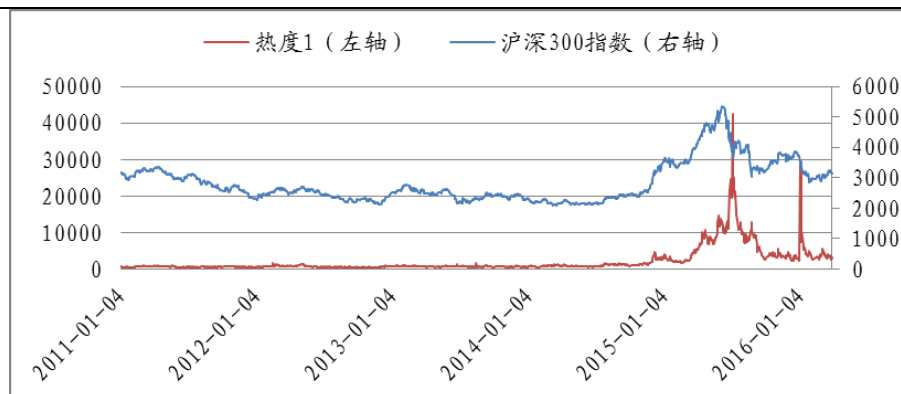
数据来源：互联网、广发证券发展研究中心

表 5 热度与次日沪深 300 指数的相关系数

	热度 1	热度 2	热度 3
相关系数	0.68	0.50	0.63

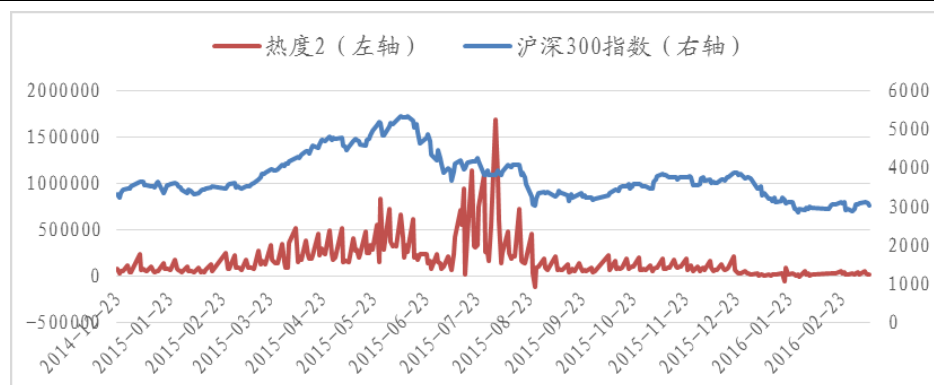
数据来源：互联网、广发证券发展研究中心

图 11 热度 1 与沪深 300 指数走势



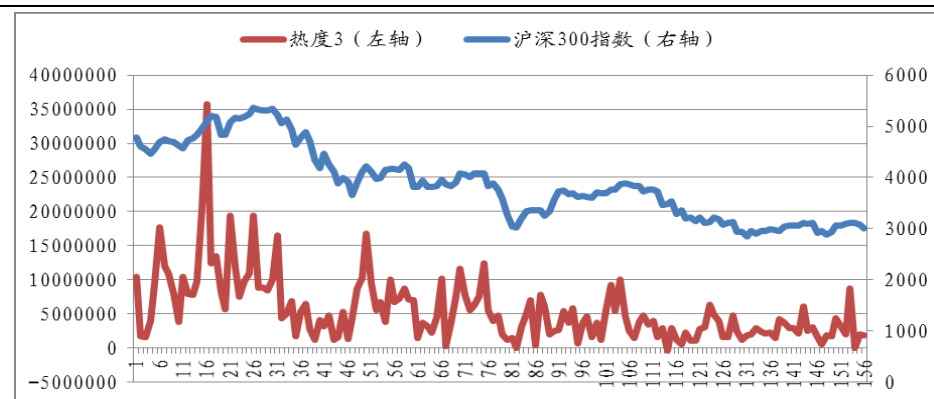
数据来源：互联网、广发证券发展研究中心

图 12 热度 2 与沪深 300 指数走势



数据来源：互联网、广发证券发展研究中心

图 13 热度 3 与沪深 300 指数走势



数据来源：互联网、广发证券发展研究中心

此外，本篇专题报告策略所需数据还包括样本区间的沪深300指数收盘价。本篇报告中将策略的收益表现与两个基准进行比较，第一个基准是同期沪深300指数表现，第二个基准是采用相同的参数下，以沪深300指数作为择时信号数据的收益表现。

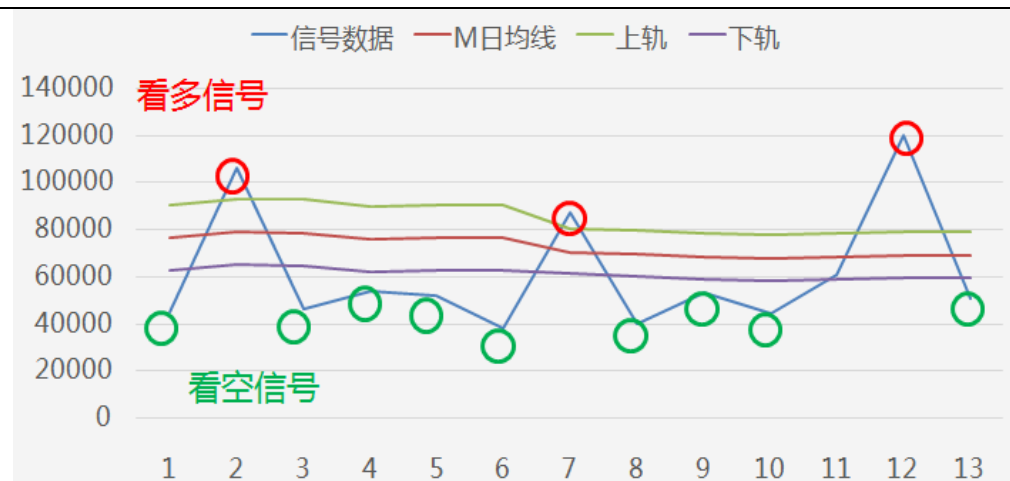
2.2 策略原理

基于上述对网络热度的分析，可以看出当当日热度较大时，次日股市上涨的可能性较大，当当日热度较小时，次日股市下跌的可能性较大。只有当网络热度短期内积累到一定程度时，次日股市上涨才会成为一个可预测的事件，同理，当网络热度短期内下跌到一定程度时，次日股市的下跌的可能性才更大。为了更好地描述这一过程，对网络热度设置一个阈值，高于或者低于某个阈值，才发出看多或者看空的信号。为了更够动态地描述网络热度的变化，因此，需要寻找到不同时点上的阈值，才能准确地运用热度数据预测次日市场的涨跌走势。

在本篇专题报告策略中，构建的择时策略为采用了布林通道，根据均线设置上轨和下轨，将上轨、下轨分别作为预测沪深300指数次日涨跌的两个阈值。具体地，专题策略的策略原理如下：

- **初始资金：**1
- **设置布林通道：**以热度数据的M日均值构造均线，以M日热度数据的标准差的N倍构造上轨、下轨，即：
布林通道上轨=热度数据的M日均值+N*M日热度的标准差；
布林通道下轨=热度数据的M日均值-N*M日热度的标准差；
- **设置参数：**M=25，N=0.3
- **择时策略：**当日热度突破通道上轨，则发出看多信号，次日开盘买入；当日热度突破通道下轨，则发出看空信号，次日开盘卖出；相连两日发出不同信号，先平仓后根据信号进行买卖；相连两日的第二日不发出信号，次日开盘直接平仓
- **假设可以做空，策略收益包含做空收益、亏损**

图 14 策略构建原理



数据来源：互联网、广发证券发展研究中心

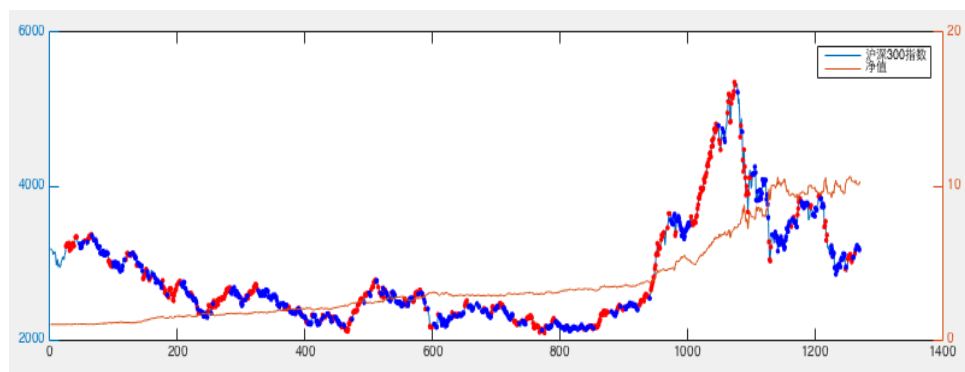
三、实证结果

3.1 热度 1 实证结果

以热度1数据进行实证，时间区间为2011年1月4日至2016年3月29日，实证结果如下图15所示。沪深300指数线上的红色符号刻画了看多信号，蓝色符号刻画了看空信号，另一条红色曲线表示净值走势。整体上看，热度1发出的信号准确，历史区间累计净值达到10.31，年化收益率58.21%，超额年化收益率58.55%，胜率57.61%，赔率1.14，累计最大回撤仅有18.11%，收益回撤比达到3.23。

我们同时将25日均线、0.3倍标准差的相同策略用在沪深300指数择时上。与热度1表现相比，择时策略的累计净值仅有1.49，绝对收益率、超额收益率均远远低于热度1择时表现，整体的收益回撤比仅有0.30，从实证结果上可以看到，互联网热度数据确实对指数的择时有较强的预测作用。

图 15 热度策略择时表现



数据来源：广发证券发展研究中心

表 6 热度 1 择时策略指标表现一览

	累计净值	年化收益率	超额年化收益率	胜率	赔率	累计最大回撤	收益回撤比	看多次数	看多胜率	看空次数	看空胜率
热度 1	10.31	58.21%	58.55%	57.61%	1.14	18.11%	3.23	457	59.30%	548	56.20%
沪深 300	1.49	8.21%	8.55%	51.80%	1.02	28.17%	0.30	574	53.66%	540	49.81%

数据来源：广发证券发展研究中心

热度1择时策略分年度测算，实证结果如下表7所示。在不同年度中，假设年初投资资金为1，采用与表6相同的策略和参数。从热度1策略表现优异，无论是绝对收益率还是相对收益率均为正数，并且累计最大回撤很低。

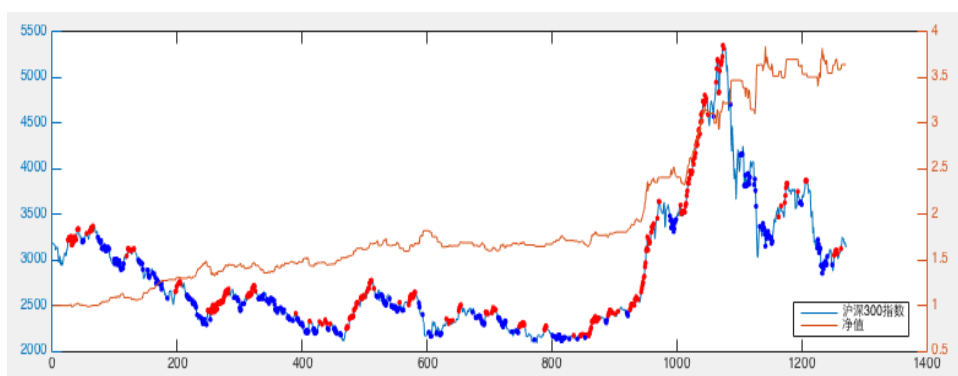
表 7 热度 1 择时策略分年度指标表现一览

年份	年化收益率	超额年化收益率	胜率	赔率	累计最大回撤	收益回撤比	看多次数	看多胜率	看空次数	看空胜率
2011	73.04%	100.14%	62.35%	1.22	6.41%	15.63	78	56.41%	92	67.39%
2012	49.45%	41.67%	56.46%	1.17	10.31%	4.04	97	57.73%	112	55.36%
2013	18.06%	26.08%	54.89%	1.00	10.73%	2.43	69	57.97%	115	53.04%
2014	57.83%	4.87%	58.85%	1.28	4.72%	1.03	108	60.19%	84	57.14%
2015	113.63%	107.91%	55.61%	1.25	18.11%	5.96	91	61.54%	114	50.88%
2016 至今	31.61%	84.98%	60.00%	0.81	12.47%	6.81	14	71.43%	31	54.84%

数据来源：广发证券发展研究中心

从上面实证结果中，可以看出热度1作为择时信号源数据的优质性。进一步，可以考虑综合信号策略，即将热度1和沪深300指数作为两个信号源，当且仅当两个信号源同时发出相同信号时，整体才发出看多或者看空的信号；参数设置与上面相同。综合信号策略的一个目的在于以热度1这一相对准确的信号源来优化指数的择时策略，通过双重信号确认，从信号的角度看，可以剔除指数择时上发出的一些错误信号。从图16以及表8的实证结果中可以看出，相比简单的指数择时的方法，综合信号的结果更优。

图 16 热度 1 综合信号择时策略表现



数据来源：广发证券发展研究中心

表 8 热度 1 综合信号择时策略指标表现一览

	累计净值	年化收益率	超额年化收益率	胜率	赔率	累计最大回撤	收益回撤比	看多次数	看多胜率	看空次数	看空胜率
综合信号	3.64	28.96%	29.30%	57.99%	1.12	12.14%	2.41	300	61.67%	307	54.40%
沪深 300	1.49	8.21%	8.55%	51.80%	1.02	28.17%	0.30	574	53.66%	540	49.81%

数据来源：广发证券发展研究中心

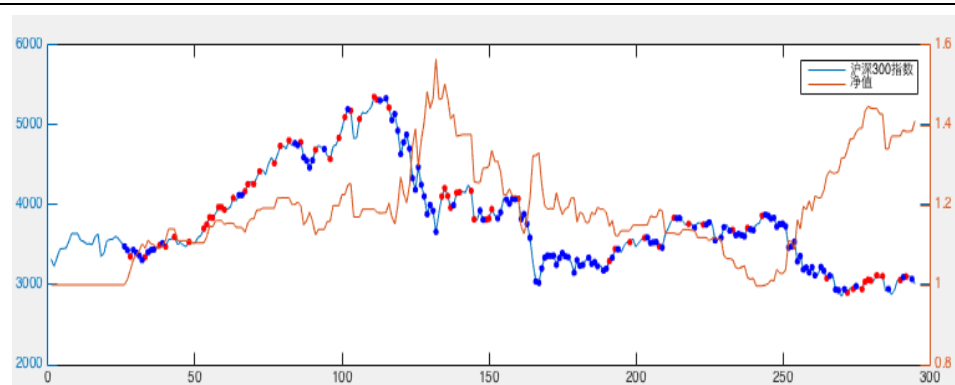
由上图和表格结果可以看到，经过热度1的信号确认，综合信号择时效果优于单指数择时效果，累计净值由1.49增加到3.64，收益回撤比由0.30增加到了2.41。综合以上，热度1择时策略在历史区间内具有优异的表现。

3.2 热度 2、热度 3 实证结果

策略同时分别对网络热度2和网络热度3进行了实证分析。热度2的实证时间区间是2014年12月22日至2016年3月10日，实证结果如下。热度2策略的累计净值为1.41，

实现年化收益率33.90%、超额年化收益率41.90%，但受到震荡行情影响，2015年四季度出现较大回撤，收益回撤比表现一般。

图 17 热度 2 择时策略表现



数据来源：广发证券发展研究中心

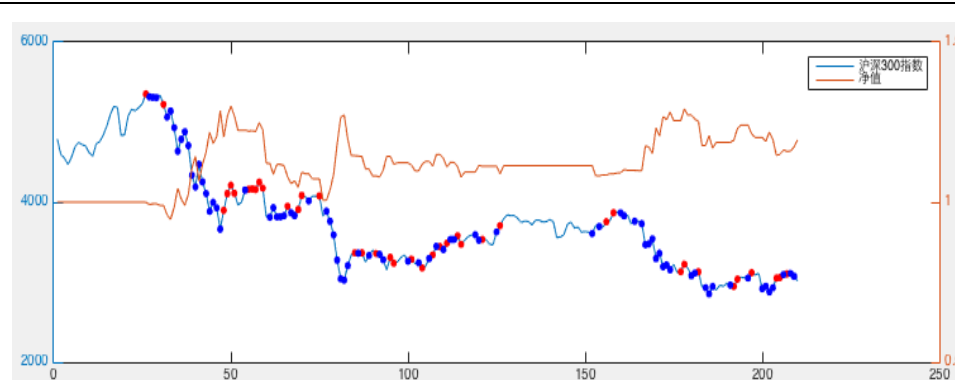
表 9 热度 2 择时策略指标表现一览

	累计净值	年化收益率	超额年化收益率	胜率	赔率	累计最大回撤	收益回撤比	看多次数	看多胜率	看空次数	看空胜率
热度 2	1.41	33.90%	41.90%	53.72%	1.09	36.34%	1.15	65	58.46%	123	51.22%
沪深 300	1.15	12.60%	20.61%	51.64%	1.03	28.17%	0.73	130	57.69%	114	44.74%

数据来源：广发证券发展研究中心

同样地，对热度3策略也采用相同的策略进行实证分析。热度3的实证区间是2015年5月4日至2016年3月10日，累计净值达到1.20，超额年化收益率为66.06%，但胜率较低，仅有47.86%，主要原因也是受到2015年四季度震荡行情影响。

图 18 热度 3 择时策略表现



数据来源：广发证券发展研究中心

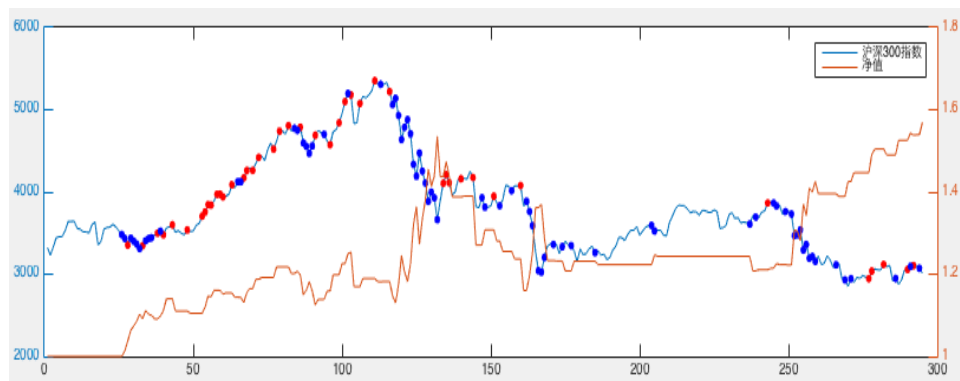
表 10 热度 3 择时策略表现指标一览

	累计净值	年化收益率	超额年化收益率	胜率	赔率	累计最大回撤	收益回撤比	看多次数	看多胜率	看空次数	看空胜率
热度 3	1.20	23.68%	66.06%	47.86%	1.29	22.61%	2.92	40	50.00%	77	46.75%
沪深 300	0.87	-15.10%	27.28%	46.43%	1.10	28.17%	0.97	66	48.48%	102	45.10%

数据来源：广发证券发展研究中心

从热度2、热度3的实证结果可以看出，策略的整体表现并不是太优异，为了能够综合运用各个网络热度的信息，对此采取了综合信号策略进行改进，目的是通过信号的双重确认，能够筛除一些不确定的投资者情绪，增加信号的准确度，减少震荡行情中的错误信号的干扰。

图 19 热度 2、热度 3 综合策略表现



数据来源：广发证券发展研究中心

表 11 热度 2、3 综合策略指标表现一览

	累计 净值	年化收益 率	超额年化收 益率	胜率	赔率	累计最大 回撤	收益回 撤比	看多 次数	看多胜率	看空 次数	看空胜率
综合信号	1.57	46.46%	54.47%	56.90%	1.12	24.42%	2.23	43	62.79%	73	53.42%
热度 2	1.41	33.90%	41.90%	53.72%	1.09	36.34%	1.15	65	58.46%	123	51.22%
沪深 300	1.15	12.60%	20.61%	51.64%	1.03	28.17%	0.73	130	57.69%	114	44.74%

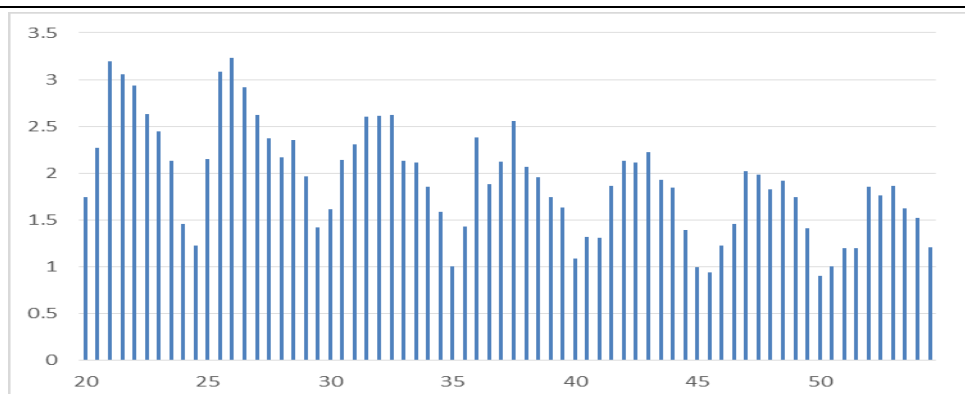
数据来源：广发证券发展研究中心

由图19、表11的结果中可以看出，综合信号策略下，震荡行情中发出的信号明显减少，区间内实现的收益率相比单一热度择时更高，超额年化收益率达到54.47%，累计最大回撤下降，收益回撤比为显著增加。综上，对于震荡行情区间，我们可以通过综合考虑两个信号源的策略来优化择时。

3.3 参数敏感性测试

在上述的实证分析中，策略的测算统一使用M=25，N=0.3的参数设置。为了寻找最优的参数设置，并且观察择时策略的表现对于参数的敏感性，进一步进行了参数的敏感性相关的测试。接下来，分别对三个网络热度源数据在不同的参数M、N下的收益回撤比进行了统计，其中，M的范围为20至50，以5为公差等差递增，N的范围为0.1至1，以0.1为公差等差递增。

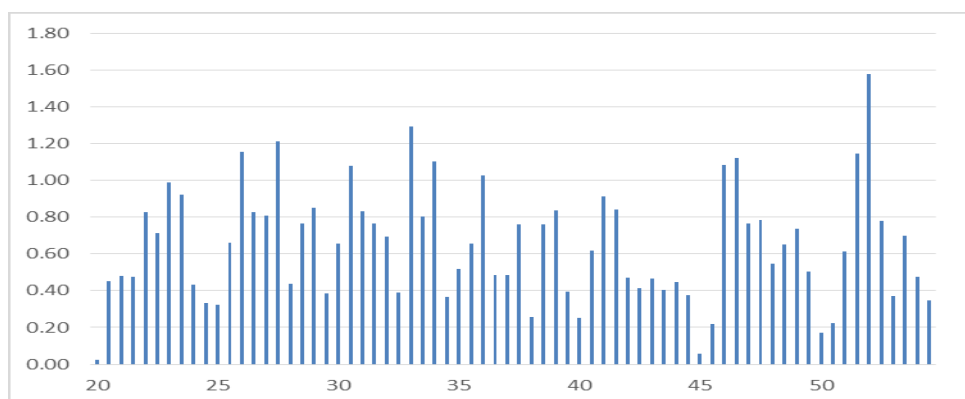
图 20 热度 1 敏感性测试结果



数据来源：广发证券发展研究中心

上图中，横轴表示不同M、N参数设置，例如20至25之间的10个分割点分别表示N等于0.1到1的情况，纵轴表示该参数设置下的收益回撤比。可以看到，热度1测试的收益回撤比主要分布在1~2.5之间，远高于沪深300指数的收益回撤比0.30，并且随着N的增加，收益回撤比呈现先增大后减小的规律。从图16的实证结果中可以观察到，热度1数据在M=20、25，N=0.3、0.4的设置下实证结果表现最优。

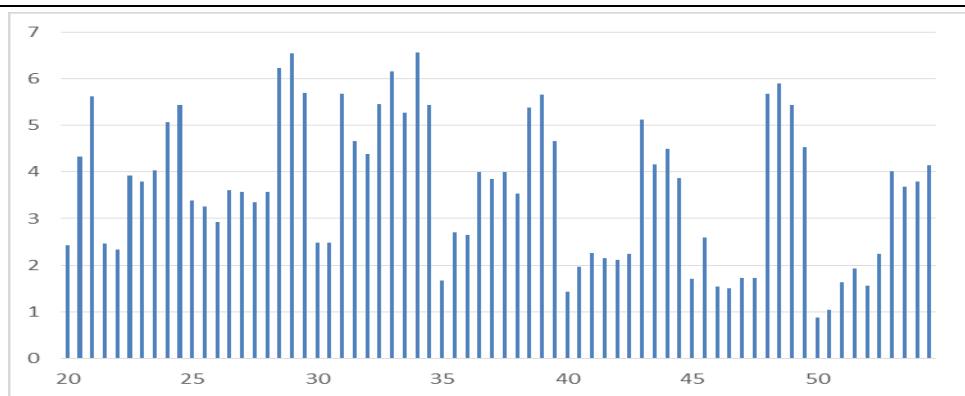
图 21 热度 2 敏感性测试结果



数据来源：广发证券发展研究中心

从图21的实证结果中可以看出，热度2的回测结果的收益回撤比主要分布在0.3~1之间，但相比热度1规律性较弱。且从实证的结果中可以观察到，热度2在M=25、30、45、50的设置下表现较好。

图 22 热度 3 敏感性测试结果



数据来源：广发证券发展研究中心

最后，对热度3的择时参数敏感性进行了测算。从图21的结果中可以看出，热度3对参数的敏感性较大，收益回撤比范围在0.87~6.56之间，波动较大，但大多数都高于300指数的择时策略收益回撤比0.97，且从图18中可以观察到，热度3在M=25、30的设置下表现相对较好。

综上，我们对三个热度数据进行了敏感性测试得出结论：

- 热度1、热度2的敏感性相对较小，回测结果比较稳定；
- 由于三个热度数据体现出不同的性质，最佳的参数设置也不相同；
- 整体而言，不同参数下三个热度数据测试的收益回撤比普遍都高于简单的300指数的择时策略，并且全部为正数。

四、总结

4.1 网络热度是有效的量化择时数据

综上所述，本篇专题报告以三个不同的网络数据源的热度数据构建了量化择时策略，实证结果表明网络热度作为择时信号源在历史区间内具有优异的表现，网络热度相对于大盘具有领先性质，基于网络数据源的数据，可以根据网络情绪的变动比较准确地预测市场的涨跌。总结一下，本篇专题报告中网络热度择时策略的要点包括：

- 首先，在热度择时策略的构建中，热度数据的质量是首要关键，本篇专题策略中选取了比较能够代表投资者情绪的典型财经网站、搜索引擎的热度数据；
- 其次，在专题策略中通过设置布林通道，寻找每日发出看多信号、看空信号的阈值；不同的热度数据体现出不同的性质，可以通过参数敏感性测试来选取最佳的参数设置；
- 最后，在震荡行情下，热度择时策略受到了影响，专题策略中提出了综合信号的改进方法，增加信号的准确度，减少仅依赖单一信号的偏差；

4.2 研究不足与未来方向

本篇专题报告对网络数据源在择时方面的应用进行了探讨，初步得到了一些实用性的研究结论，但也存在以下不足之处，同时也提出几个未来可以深入研究的方向：

- 首先，由于不同热度来源网站的用户群体不同，并且单一热度也难以代表全部的市场投资者，我们可以进一步尝试综合利用三个热度数据来开发更加复杂的模型；但是目前，受限于数据的时间长度，尤其是热度2、热度3的区间较短，样本量较小，策略的稳定性尚需继续研究；
- 其次，我们使用了统一的参数进行综合信号测试，但是由于不同热度在均值和波动性上具有不同的性质，我们可以尝试在综合信号测试时对不同信号设置不同的参数，充分挖掘热度数据的价值；

- 本报告是对市场择时的策略报告，未来还可以将这个策略推广到行业配置方面，通过分析不同行业的热度判断出恰当的配置时点。

风险提示

本专题的择时策略基于网络热度数据，热度数据仅仅是投资者投资情绪的一个方面表现，市场行情受到其他因素影响。

广发证券—行业投资评级说明

- 买入： 预期未来 12 个月内，股价表现强于大盘 10%以上。
- 持有： 预期未来 12 个月内，股价相对大盘的变动幅度介于-10%~+10%。
- 卖出： 预期未来 12 个月内，股价表现弱于大盘 10%以上。

广发证券—公司投资评级说明

- 买入： 预期未来 12 个月内，股价表现强于大盘 15%以上。
- 谨慎增持： 预期未来 12 个月内，股价表现强于大盘 5%-15%。
- 持有： 预期未来 12 个月内，股价相对大盘的变动幅度介于-5%~+5%。
- 卖出： 预期未来 12 个月内，股价表现弱于大盘 5%以上。

联系我们

	广州市	深圳市	北京市	上海市
地址	广州市天河区林和西路9号耀中广场A座1401	深圳市福田区福华一路6号免税商务大厦17楼	北京市西城区月坛北街2号月坛大厦18层	上海市浦东新区富城路99号震旦大厦18楼
邮政编码	510620	518000	100045	200120
客服邮箱	gfyf@gf.com.cn			
服务热线				

免责声明

广发证券股份有限公司具备证券投资咨询业务资格。本报告只发送给广发证券重点客户，不对外公开发布。

本报告所载资料的来源及观点的出处皆被广发证券股份有限公司认为可靠，但广发证券不对其准确性或完整性做出任何保证。报告内容仅供参考，报告中的信息或所表达观点不构成所涉证券买卖的出价或询价。广发证券不对因使用本报告的内容而引致的损失承担任何责任，除非法律法规有明确规定。客户不应以本报告取代其独立判断或仅根据本报告做出决策。

广发证券可发出其它与本报告所载信息不一致及有不同结论的报告。本报告反映研究人员的不同观点、见解及分析方法，并不代表广发证券或其附属机构的立场。报告所载资料、意见及推测仅反映研究人员于发出本报告当日的判断，可随时更改且不予通告。

本报告旨在发送给广发证券的特定客户及其它专业人士。未经广发证券事先书面许可，任何机构或个人不得以任何形式翻版、复制、刊登、转载和引用，否则由此造成的一切不良后果及法律责任由私自翻版、复制、刊登、转载和引用者承担。