

# 基于大数据挖掘的 Smart Beta 策略

## —互联网大数据挖掘之七

### 报告摘要:

#### ● Smart Beta 策略在国外兴起与互联网大数据发展

最近十几年,在海外市场,Smart beta 投资策略逐渐兴起。Smart beta 透明、低成本化的特点也使其成为一大类 ETF 产品的投资策略。据晨星(Morningstar)公司的统计数据,使用 Smart Beta 策略的基金规模已从 2006 年的 1030 亿美元飙升至 2015 年底的 6160 亿美元。

伴随着互联网的发展以及不断普及,人类每天产生的数据量正在呈现指数级的增长。同时伴随着互联网的快速发展,投资者也越来越倾向于从互联网上获取相关的个股信息。互联网大数据时代下,网络数据源的应用成为当前量化策略研究的一个新方向。搜索量、股吧情绪类、新闻热度类等新型因子也渐渐进入投资者研究的视野。

#### ● 另类数据切入点——关注度

传统的 Smart Beta 策略更多的从股票的基本面因子、波动率、动量等因子出发,构建策略,而这些因子所包含的价值信息已基本上被专业投资者所挖掘透彻。互联网的快速发展,给专业投资者研究提供了另一个数据研究的维度-大数据。本文试图从互联网上个股的用户关注度角度出发研究基于关注度选股的 Smart Beta 策略。

#### ● 基于关注度数据的策略构建

在历史回测期内,按照申万一级行业成分股将个股分为不同的行业股票池,根据每个时期个股在过去一段时间的关注度,对不同行业的股票按照行业内个股关注度进行排序,选择行业股票池内后 5% 的股票作为超配组合,同时选择前 5% 的股票作为低配组合,在下一个交易日买入或者卖出对应组合的个股。

#### ● 实证分析

实证结果表明,在历史回测期内基于热门网站个股关注度选股策略表现优异。在雪球-东财关注度全市场选股策略月频调仓下,对冲后的累计净值为 1.92,月度胜率为 100%,最大回撤为-6.03%。在中证 800 月频调仓下,对冲后的累计净值为 1.61,月度胜率为 92.86%,最大回撤为-2.9%。

#### ● 核心假设风险:

本篇报告所提供的个股关注度数据一定程度上能够反映投资者对个股的关注,但股票价格受诸多因素影响,历史的数据并不能精确地预测未来金融市场变化,注意市场其他风险的变化。

图 1 月频调仓下雪球-东财关注度全市场选股策略回测表现

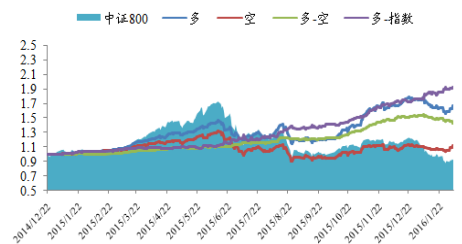
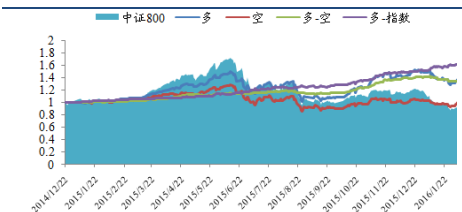


图 2 月频调仓下雪球-东财关注度中证 800 选股策略回测表现



分析师: 史庆盛 S0260513070004



020-87577060



sqs@gf.com.cn

#### 相关研究:

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(六)	2015-09-09
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
那些年一起追过的财经小编选股策略——互联网大数据挖掘系列专题之(四)	2014-08-21

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

## 目录索引

SMART BETA 策略兴起 .....	5
SMART BETA 定义 .....	5
SMART BETA 策略在国外兴起 .....	5
互联网大数据发展 .....	7
大数据时代的到来 .....	7
大数据应用量化产品逐渐丰富 .....	8
广发金工互联网大数据研究 .....	10
另类因子-关注度 .....	11
关注度简介及处理方法 .....	11
策略构建原理及框架 .....	13
策略构建原理 .....	14
实证分析 .....	15
数据来源 .....	15
雪球关注度选股实证结果 .....	15
东财关注度选股实证结果 .....	20
综合策略实证结果 .....	22
总结及展望 .....	26
风险提示 .....	26

# 图表索引

图 1 月频调仓下雪球-东财关注度全市场选股策略回测表现 .....	1
图 2 月频调仓下雪球-东财关注度中证 800 选股策略回测表现 .....	1
图 3 Smart Beta 与传统 Beta 以及 Alpha 的区别 .....	5
图 4 美国 ETF 市场规模 .....	6
图 5 2015 年美国市场新发行 Smart Beta ETF 规模占比 .....	6
图 6 美国市场 Smart Beta ETF 规模占比 .....	6
图 7 Smart Beta 产品特性 .....	6
图 8 大数据 4V 特征 .....	8
图 9 大数据投资优势-提前预判 .....	8
图 10 大数据在量化领域应用 .....	9
图 11 完善数据抓取体系 .....	10
图 12 广发金工大数据研究体系 .....	10
图 13 关注度数据选取常见指标 .....	11
图 14 个股关注度差异 1 .....	12
图 15 个股关注度差异 2 .....	12
图 16 关注度数据处理 .....	13
图 17 深深宝 A 收盘价与关注度走势 .....	13
图 18 通程控股收盘价与关注度走势 .....	13
图 19 博威合金收盘价与关注度走势 .....	14
图 20 振东制药收盘价与关注度走势 .....	14
图 21 关注度选股框架 .....	14
图 22 周频调仓下,雪球关注度选股策略表现 .....	15
图 23 月频调仓下,雪球关注度选股表现 .....	16
图 24 周频调仓下雪球关注度策略回测表现 .....	17
图 25 月频调仓下雪球关注度中证 800 选股策略回测表现 .....	18
图 26 周频调仓下雪球关注度沪深 300 选股策略回测表现 .....	19
图 27 月频调仓下雪球关注度沪深 300 选股策略回测表现 .....	19
图 28 周频调仓下东财关注度全市场选股策略回测表现 .....	20
图 29 月频调仓下东财关注度全市场选股策略回测表现 .....	21
图 30 周频调仓下东财关注度中证 800 选股策略回测表现 .....	21
图 31 月频调仓下东财关注度中证 800 选股策略回测表现 .....	22
图 32 周频调仓下雪球-东财关注度全市场选股策略回测表现 .....	23
图 33 月频调仓下雪球-东财关注度全市场选股策略回测表现 .....	24
图 34 周频调仓下雪球-东财关注度中证 800 选股策略回测表现 .....	24
图 35 月频调仓下雪球-东财关注度中证 800 选股策略回测表现 .....	25

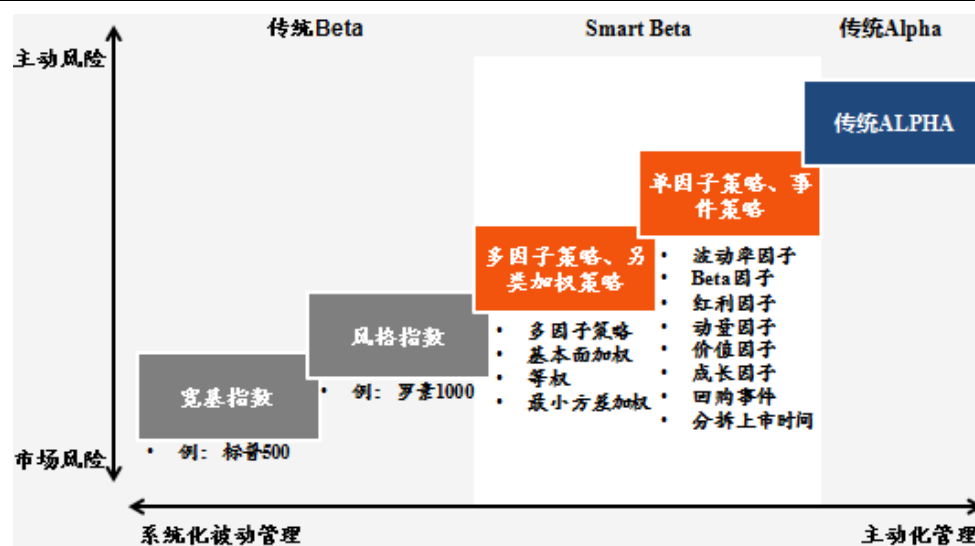
表 1 美国市场 Smart Beta ETF——规模前 10 的 Smart Beta ETF(截至 2015/12/31).....	7
表 2 公募基金大数据量化研究产品一览.....	9
表 3 传统 Smart Beta 研究体系.....	11
表 4 周频调仓下策略指标表现.....	15
表 5 月频调仓下策略指标表现.....	16
表 6 周频调仓下策略指标表现.....	17
表 7 月频调仓下中证 800 选股策略指标表现.....	18
表 8 周频调仓下沪深 300 选股策略指标表现.....	19
表 9 月频调仓下沪深 300 选股策略指标表现.....	20
表 10 周频调仓下东财关注度全市场选股策略指标表现.....	20
表 11 月频调仓下东财关注度全市场选股策略指标表现.....	21
表 12 周频调仓下东财关注度中证 800 选股策略指标表现.....	22
表 13 月频调仓下东财关注度中证 800 选股策略指标表现.....	22
表 14 周频调仓下雪球-东财关注度全市场选股策略指标表现.....	23
表 15 月频调仓下雪球-东财关注度全市场选股策略指标表现.....	24
表 16 周频调仓下雪球-东财关注度中证 800 选股策略指标表现.....	25
表 17 月频调仓下雪球-东财关注度中证 800 选股策略指标表现.....	25

## Smart Beta 策略兴起

### Smart Beta 定义

从广义上讲，Smart beta策略是指以获取各类风险因子的风险溢价为目的而设计出的投资策略，这种策略在增加对风险因子的暴露的情况下希望获得更高的风险收益。

图 3 Smart Beta与传统Beta以及Alpha的区别

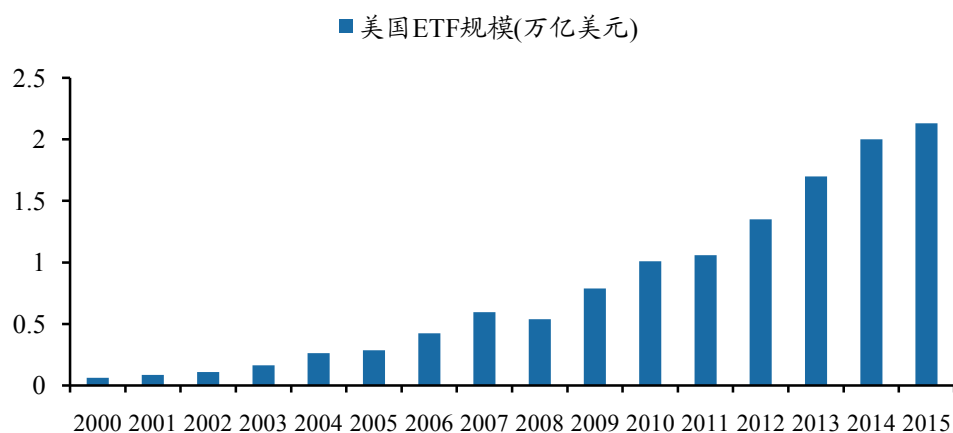


数据来源：广发证券发展研究中心

### Smart Beta 策略在国外兴起

最近十几年，在海外市场中，Smart beta投资策略逐渐兴起并成为炙手可热的投资策略，这类投资策略属于被动化投资策略，避免常规的指数市值加权的限制，试图增加一些复杂的因子来战胜市场，这为投资者提供了更加灵活、多样化的投资组合策略。Smart beta透明、低成本化的特点也让其成为一大类ETF产品的投资策略。据晨星(Morningstar)公司的统计数据，使用Smart Beta策略的基金规模已从2006年的1030亿美元飙升至2015年底的6160亿美元。

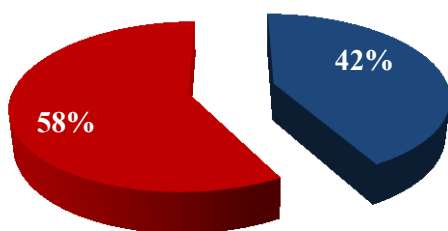
图 4 美国ETF市场规模



数据来源: Bloomberg,morningstar, 广发证券发展研究中心

图 5 2015年美国市场新发行Smart Beta ETF规模占比

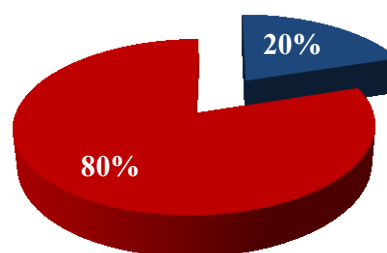
■ Smart Beta ETF ■ 传统Beta ETF



数据来源: Bloomberg,morningstar, 广发证券发展研究中心

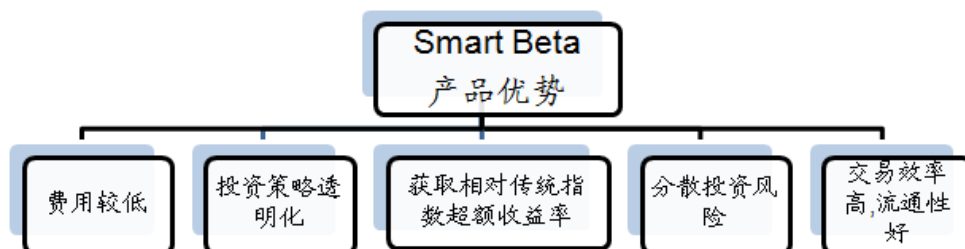
图 6 美国市场 Smart Beta ETF 规模占比

■ Smart Beta ETF ■ 传统Beta ETF



数据来源: Bloomberg,morningstar, 广发证券发展研究中心

图 7 Smart Beta产品特性



数据来源: Bloomberg, 广发证券发展研究中心

表 1 美国市场Smart Beta ETF——规模前10的Smart Beta ETF(截至2015/12/31)

ETF 代码	ETF 简称	上市日期	策略类型	规模 (亿美元)	管理费用率%
IWF US Equity	iShares 安硕罗素 1000 成长型 ETF	2000/5/26	成长型	315.90	0.2
IWD US Equity	iShares 安硕罗素 1000 价值型 ETF	2000/5/26	价值型	273.61	0.2
VIG US Equity	领航股利递增 ETF	2006/4/27	红利型	192.25	0.1
VTV US Equity	领航价值型 ETF	2004/1/30	价值型	186.48	0.09
HEDJUS Equity	WisdomTree 欧洲对冲股票基金	2009/12/31	红利型	169.41	0.58
DXJ US Equity	WisdomTree 日本对冲股票基金	2006/6/16	红利型	149.05	0.48
DVY US Equity	iShares 安硕精选红利 ETF	2003/11/7	红利型	132.65	0.39
SDY US Equity	SPDR 标普股利 ETF	2005/11/15	红利型	128.28	0.35
VYM US Equity	领航高股利收益 ETF	2006/11/16	红利型	114.90	0.1
RSP US Equity	Guggenheim 标普 500 等权重 ETF	2003/4/30	等权重	96.10	0.4

数据来源: Bloomberg, 广发证券发展研究中心

## 互联网大数据发展

### 大数据时代的到来

在最近十几年中, 互联网在计算机技术不断发展与普及下得到了快速的发展。传统的媒体(报纸、广播等)传播已越来越互联网传播所替代, 传统媒体所占的市场份额已越来越低。在互联网不断普及与发展的趋势下, 人们获取信息的渠道也越来越倾向于从互联网上获取。而伴着互联网的传播, 数据规模的沉淀速度已成指数型速度增长, 数据的存储结构形式也发生了多样化的概念, 以往主要以数值型的结构形式来存储数据, 而当前, 文本型、视频、音频等非结构化或者半结构化的数据存储形式已占据当前信息存储的主要形式, 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。

在传统的量化投资研究中, 专业的投资者在选股上的研究, 更多的从个股的财务数据、行情数据, 例如EPS、PE、收盘价、开盘价、成交量等, 以及行业数据, 例如行业指数涨跌幅、成交量等数据角度出发。而这些传统的数据研究维度中所隐含的大部分有价值的投资信息已被专业的投资者所挖掘。专业投资者在量化投资策略研究中也希望能另辟途径, 从其他新的数据维度中挖掘出对投资决策有用的信息。而随着大数据的发展, 专业的金融机构等也紧跟时代的步伐, 试图从互联网大数据这一新的数据维度中挖掘出对投资决策等有用的信息, 搜索量、网络舆情、股吧情绪、新闻热度、个股关注度等新型的数据进入专业投资者研究的视野。而当前, 对互联网大数据相关的量化策略的研究已然成为当前量化投资新的研究方向和热门研究领域。

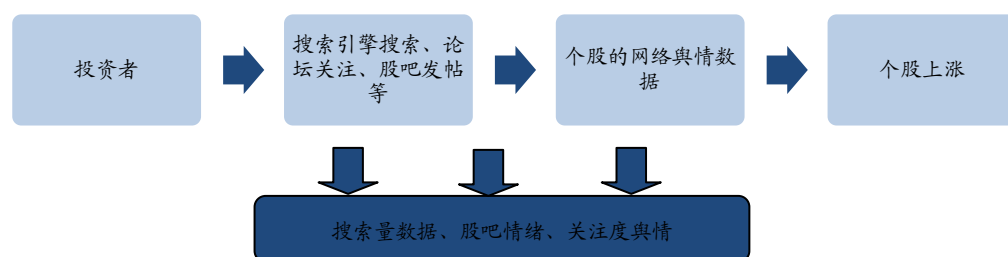


图 8 大数据4V特征



数据来源：Bloomberg，广发证券发展研究中心

图 9 大数据投资优势-提前预判



数据来源：Bloomberg，广发证券发展研究中心

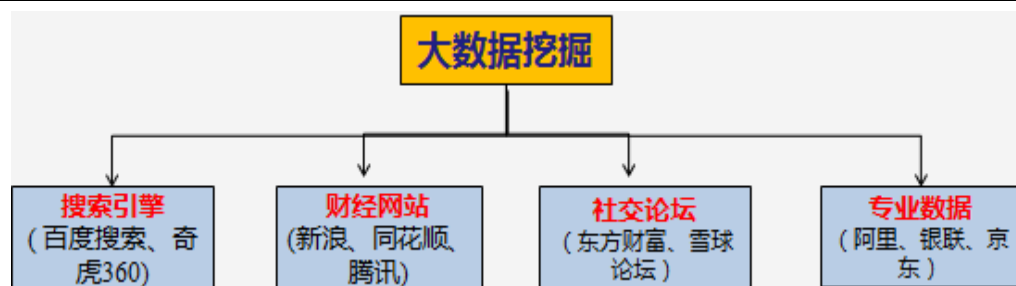
### 大数据应用量化产品逐渐丰富

当前，在国内量化投资领域中，对大数据应用中数据源的创新应用是其中的一大特色。在数据源上，目前主要的数据源包括搜索引擎、热门财经网站、社交网站论坛以及特定的专业数据网站，例如阿里巴巴的淘宝数据、银联的用户刷卡数据等。在这些大数据的应用中，一个明显的特点是数据的提供商占据着主导地位，数据提供商的数据多样化性、稳定性等特点决定着整个量化策略产品的有效性。

而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。



图 10 大数据在量化领域应用



数据来源：Bloomberg，广发证券发展研究中心

表 2 公募基金大数据量化研究产品一览

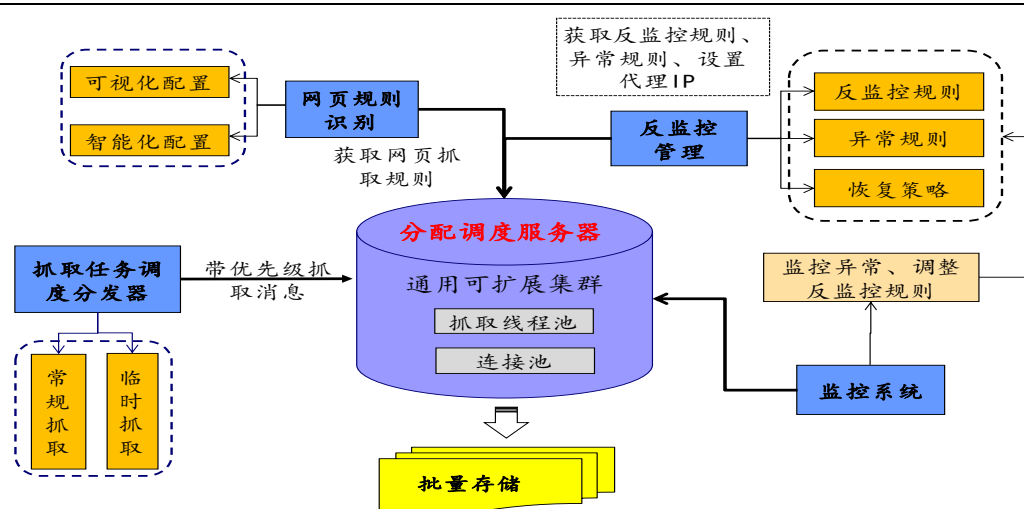
基金公司	合作方	大数据基金	产品类型
广发基金	百度	中证百度百发策略 100 指数型	指数型
博时基金	蚂蚁金服	中证淘金大数据 100 指数型	指数型
南方基金	新浪	大数据 100 指数型	指数型
东方资管	京东	东方红京东大数据灵活配置混合型	混合型
南方基金	新浪	大数据 300 指数型	指数型
博时基金	银联	中证银联智惠大数据 100 指数型	指数型
广发基金	百度	百发大数据策略精选灵活配置混合型	混合型
广发基金	百度	百发大数据策略成长灵活配置混合型	混合型
嘉实基金	腾讯	腾讯自选股大数据策略股票型	股票型
大成基金	360	中证 360 互联网+大数据 100 指数型	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

数据来源：广发证券发展研究中心

## 广发金工互联网大数据研究

对互联网大数据的挖掘与研究，广发金工在该领域也进行了比较深入的研究，搭建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。

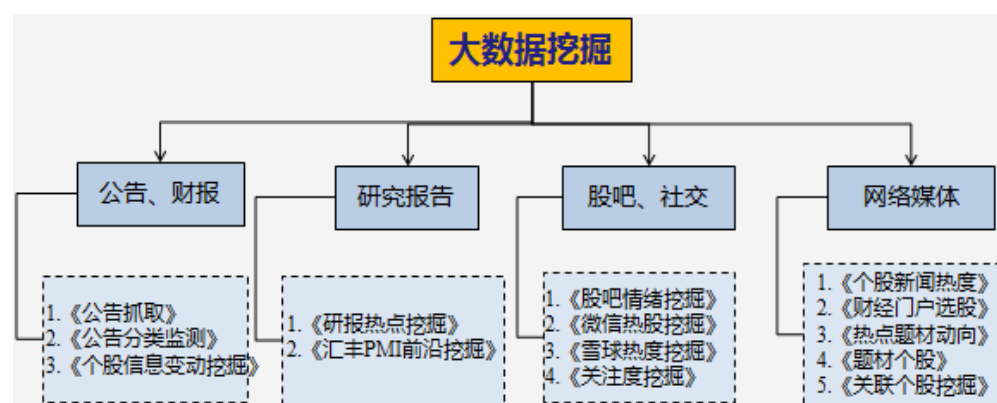
图 11 完善数据抓取体系



数据来源：广发证券发展研究中心

在对各个数据维度研究上，广发金工研究了多个维度的数据，包括对A股整个市场上历年上市公司公告公告的挖掘；卖方研究机构对个股、行业以及大盘的研究报告；个股股吧、社交网络的情绪、热度挖掘以及对网络媒体，例如热门财经网站的荐股、热点题材、个股关注度等的研究。

图 12 广发金工大数据研究体系



数据来源：广发证券发展研究中心

在专题策略上也取得了一系列的研究成果。例如之前采用文本挖掘的方法对上市公司公告披露背后的投资机会进行了统计分析以及实证，得到了较好的就结果，具体可见《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）》专题报告；从股吧、个股的新闻热度、上市公司信息变更、财经频道的荐股信息等角度对文本信息进行挖掘，具体可见《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）》、《倾听股吧之声，洞察大盘趋势—互联网大数据挖掘系列专题（三）》、《那些年一起追过的财经小编选股策略—互联网财经频道文本挖掘策略》、《上市公司披露信息变更隐含的投资机会—事件驱动策略之（十四）》；基于互联网海量的个股新闻信息对热点概念的选股可见《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）》、基于个股的共同关注度额的选股研究可见《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(六)》等相关的专题策略报告。

## 另类因子-关注度

在传统的Smart Beta框架下，在选股的因子上主要从波动率、红利、beta、动量、质量等因子角度出来考虑选股问题，在个股权重上往往采取另类的加权方式，例如等权、波动率倒数加权、最小方差加权以及基本面加权等方式。在传统的Smart Beta框架下，对传统的因子的挖掘已基本上都在内了，很难再从这些传统的因子中挖掘出相对较高的风险溢价。在当下大数据时代下，从互联网数据这一新的数据维度中挖掘出特定的因子获取风险溢价或许是Smart Beta策略的另一个方向。

在传统的研究框架中，对于关注度的阐述，往往使用个股的换手率、成交量等替代指标，而在大数据时代下，网络数据作为一种新的数据研究维度，因而可以用网络上个股的相关数据作为一个新的替代指标，例如论坛上个股的发帖量、转发量、评论量或者是个股的关注人数等。本篇报专题告从互联网热门论坛上，个股股吧的关注人数这一新的数据维度出发，试图从个股的关注度的变化中寻找新的选股策略。

表 3 传统Smart Beta研究体系

Smart Beta 策略	具体策略举例
单因子策略	波动率、Beta、红利、动量、质量、成长因子等
多因子策略	多类因子结合
另类加权策略	等权、最小方差加权、基本面加权等
事件策略	回购、分拆上市等

数据来源：广发证券发展研究中心

图 13 关注度数据选取常见指标



数据来源：广发证券发展研究中心

## 关注度简介及处理方法

对于关注度高低的见解，不同的投资大师有不同的见解。《彼得·林奇的成功投资》一书中对高关注度的个股有这样的一说“如果说有一种股票我避而不买的话，那它一定是最热门行业里面最热门的股票。这种股票受到大家最广泛的关注，每个投资者上下班途中在汽车上或在火车上都会听到人们谈论这种股票，一般人往往由于禁不住这种强大的社会压力就买入了这种股票”。本杰明·格雷厄姆在其经典的著

作《证券分析》中对低关注度的个股有这么一说“标准证券或主要证券对其报告利润几乎总是迅速做出反应以至于夸大年收益变化在股票市场上的影响。知名度较低的证券的表现主要取决于市场专业人士的态度。如果对其缺乏兴趣，价格将远远滞后于其统计表现。如果该证券受到关注（这种关注或者是人为操纵，或是自然出现的），则结果可能完全相反，公司业绩的变化将最大限度地反映在其股票价格上”。从对关注度的相对高低变化中寻找相关的规律从而做出选股的决策是本篇专题研究的一个核心。

个股之间因为上市时间、所在板块、所在行业等的不同，造成在财经网站或者在股吧上关注人数的不同。为了能够使得不同的个股在关注人数上既能与其自身历史的关注水平进行对比，同时也能够与其他个股进行比较，需要对关注数据进行标准化处理，从而保证个股在关注度对比变化较大时，均有可能被选入到组合中，从而防止因个别个股因关注人数较多导致其他个股在选股时难以被选到的情况发生，错过一些投资机会。

图 14 个股关注度差异1



数据来源：广发证券发展研究中心

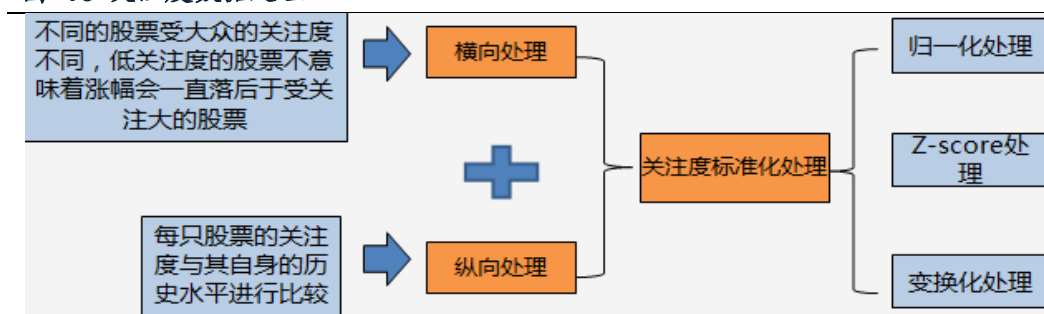
图 15 个股关注度差异2



数据来源：广发证券发展研究中心

在对关注度数据进行处理时，可以对关注度数据进行归一化处理、正态化处理或者是其他方式的变换等处理。

图 16 关注度数据处理



数据来源：广发证券发展研究中心

对数据的归一化处理方式,常见的有min\_attention-max\_attention对原始数据进行线性变换,设min\_attention和max\_attention分别为原数据的最小值和最大值,将原数据的一个原始值attention通过 标准化映射成在区间[0,1]中的值,其公式为:新数据=(原数据-最小值)/(最大值-最小值)。对数据的正态化处理方式,常见的有Z-score处理,经过处理后数据符合标准正态分布,即均值为0,标准差为1,其公式

为:新数据=(原数据-均值)/标准差,  $standard\_attention = \frac{attention - u}{\sigma}$ , 本

篇专题对关注度的实证采用对关注度进行Z-score处理,用股票过去一段时间,周频或者月频下的关注度的均值作为个股的均值,同时使用过去一段时间,周频或者月频下关注度的标准差作为个股的标准差。在个股历史关注度时间序列足够多的情况下,可以考虑用个股关注度全部数据作为样本计算该股票关注度的均值和方差。

## 策略构建原理及框架

通过对各个热门财经网站上个股的关注度进行标准化处理后,得到关注度的标准化的数据。从图16、17、18、19中可以看出,经过标准化处理后的个股关注度的走势与个股的股价走势之间存在一定的关系。因此,从这些标准化的个股关注度数据中挖掘选股机会,是本篇专题报告研究的重点。

图 17 深深宝A收盘价与关注度走势



图 18 通程控股收盘价与关注度走势





数据来源：互联网、广发证券发展研究中心

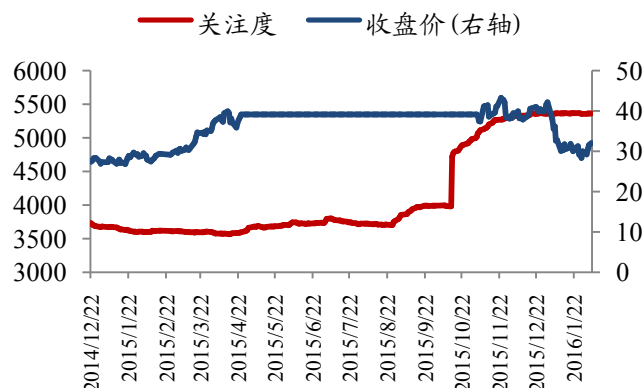
数据来源：互联网、广发证券发展研究中心

图 19 博威合金收盘价与关注度走势



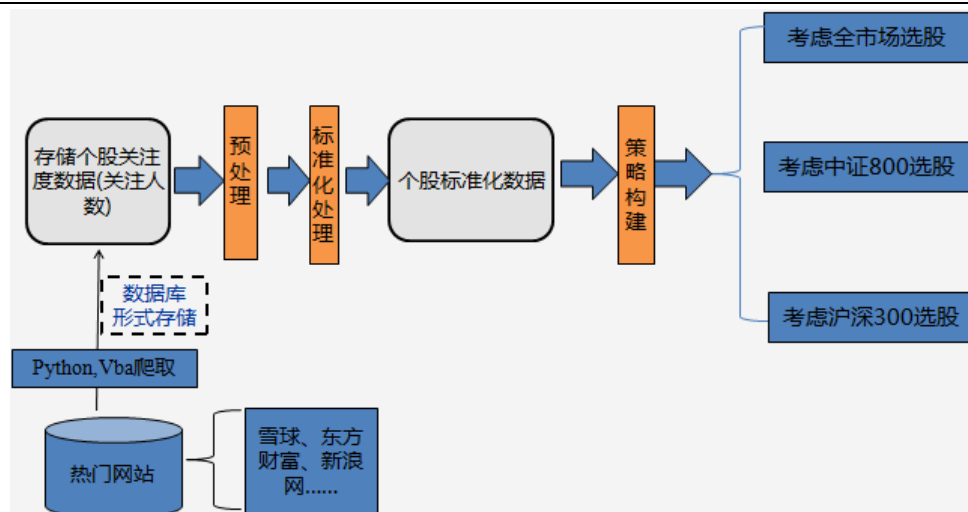
数据来源：互联网、广发证券发展研究中心

图 20 振东制药收盘价与关注度走势



数据来源：互联网、广发证券发展研究中心

图 21 关注度选股框架



数据来源：互联网、广发证券发展研究中心

## 策略构建原理

经过上述对关注度数据的讨论，建立本专题的关注度选股策略的策略原理：在历史回测期内，按照对应时期申万一级行业成分股将个股分为不同的行业股票池，根据每个时期个股在过去一段时间的关注度数据，对不同行业的股票按照行业内股票的关注度进行排序，选择排序完后的行业股票池内后5%的股票作为超配组合，同时选择行业股票池内前5%的股票作为低配组合；在下一个交易日买入或者卖出对应组合的个股，考虑个股的涨跌停影响；行业权重按照中证800或300成分股行业权重，行业内个股等权；考虑周频或月频调仓；交易费用千分之二；初始净值为1；



## 实证分析

### 数据来源

- ◆ 关注度数据：雪球、东财等；
- ◆ 个股数据：全市场A股；
- ◆ 指数数据：申万一级行业成分股以及沪深300、800指数；
- ◆ 数据区间：2014年至2016年至今；

依据上述的策略原理以及原始的关注度等数据，对东财、以及雪球等热门网站上的关注度数据进行实证分析。

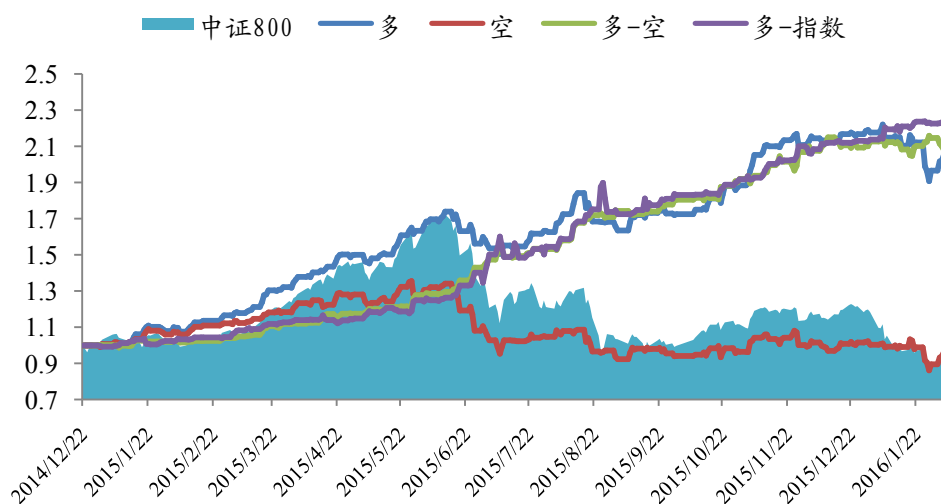
### 雪球关注度选股实证结果

#### 全市场选股实证

在回测过程中，考虑关注度在周频或者月频调仓下的策略表现结果。

从图22以及表4中的结果可以看出，在全市场选股的情况下，周频调仓下，策略的表现结果优异，超配的组合在回测期内累计净值为2.04，年化收益率为85.25%，周频下的胜率为67.24%，最大回撤为-14.15%，多空对冲下，在回测期内累计净值为2.1，年化收益率为89.97%，周胜率为75.86%，最大回撤为-4.95%；在中证800指数对冲的情况下，回测期内的累计净值为2.24，年化收益率为101.19%，胜率为77.59%，最大回撤为-9.22%。

图 22 周频调仓下,雪球关注度选股策略表现



数据来源：广发证券发展研究中心

表 4 周频调仓下策略指标表现

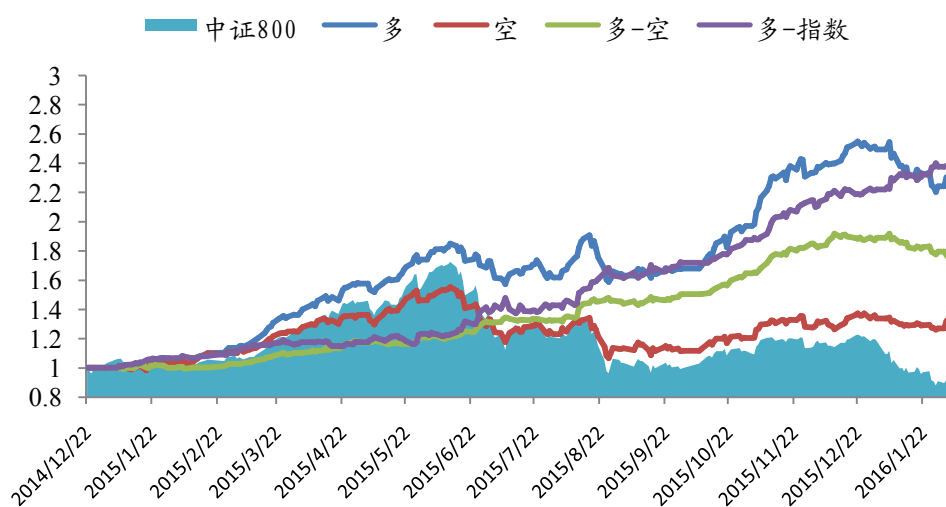
指标	多头	多-空	多-指数
累计净值:	2.04	2.10	2.24
年化收益率:	85.25%	89.97%	101.19%

年化波动率:	21.09%	13.97%	19.51%
胜率:	67.24%	75.86%	77.59%
最大回撤:	-14.15%	-4.95%	-9.22%

数据来源：广发证券发展研究中心

从图23以及表5的结果中可以看出，基于雪球关注度在全市场的选股，在月频调仓下的表现优异。在回测期内，超配组合的累计净值为2.33，年化收益率为107.9%，月胜率为78.57%，最大回撤为-16.79%；在假设低配股票可以融券卖空的情况下，多空对冲的累计净值为1.74，年化收益率为61.77%，月胜率为85.71%，最大回撤为-9.61%。在中证800指数对冲的情况下，策略对冲后的累计净值为2.4，年化收益率为113.6%，月胜率为100%，最大回撤为-7.19%。

图 23 月频调仓下，雪球关注度选股表现



数据来源：广发证券发展研究中心

表 5 月频调仓下策略指标表现

指标	多头	多-空	多-指数
累计净值:	2.33	1.74	2.40
年化收益率:	107.90%	61.77%	113.16%
年化波动率:	25.34%	12.27%	17.46%
胜率:	78.57%	85.71%	100.00%
最大回撤:	-16.79%	-9.61%	-7.19%

数据来源：广发证券发展研究中心

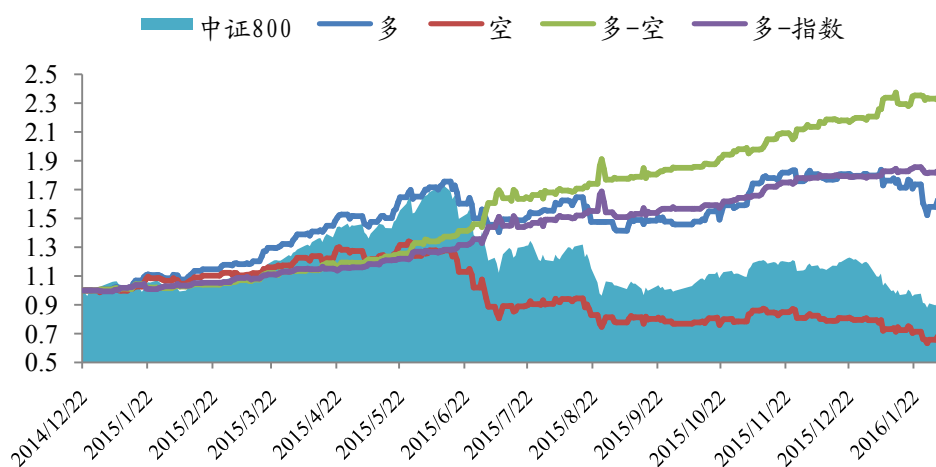
综合周频以及月频调仓策略下的表现可以看出，在月频调仓下的策略表现更加优异。

### 中证800成分股选股实证

其次在回测过程中，考虑在中证800成分股中，关注度在周频或者月频调仓下的策略表现结果。

从图24以及表6的结果中可以看出，在周频调仓下，策略在回测期内的表现优异。在回测期内，基于雪球关注度的选股策略超配组合的累计净值为1.64，年化收益率为53.43%，周胜率为65.52%，最大回撤为-20.05%。在假设低配个股可以融券卖空的情况下，多空对冲后的累计净值为2.32，年化收益率为107.08%，胜率为82.76%，最大回撤为-7.45%；在中证800指数对冲的情况下，策略对冲后的累计净值为1.83，年化收益率为69.06%，周胜率为75.86%，最大回撤为-10.46%。

图 24 周频调仓下雪球关注度策略回测表现



数据来源：广发证券发展研究中心

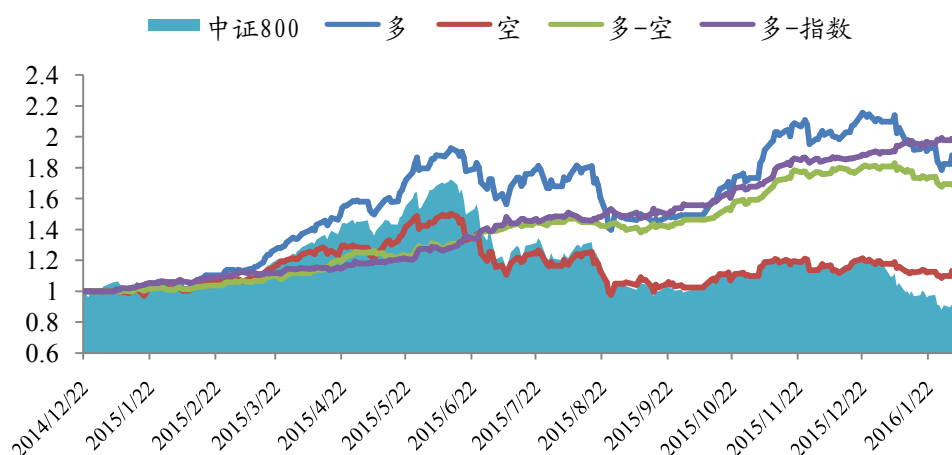
表 6 周频调仓下策略指标表现

指标	多头	多-空	多-指数
累计净值:	1.64	2.32	1.83
年化收益率:	53.43%	107.08%	69.06%
年化波动率:	26.88%	18.23%	16.31%
胜率:	65.52%	82.76%	75.86%
最大回撤:	-20.05%	-7.45%	-10.46%

数据来源：广发证券发展研究中心

从图25以及表7的结果中可以看出，基于雪球关注度在中证800成分股中的选股策略，在月频调仓下的回测表现优异。在历史回测期内，超配组合的累计净值为1.9，年化收益率为74%，月胜率为71.43%，最大回撤为-27.36%；在假设低配股票可以融券卖空的情况下，多空对冲的累计净值为1.7，年化收益率为58%，月胜率为85.71%，最大回撤为-8.33%。在中证800指数对冲的情况下，策略对冲后的累计净值为2，年化收益率为82%，月胜率为100%，最大回撤为-3.86%。

图 25 月频调仓下雪球关注度中证800选股策略回测表现



数据来源：广发证券发展研究中心

表 7 月频调仓下雪球关注度中证800选股策略指标表现

指标	多头	多-空	多-指数
累计净值:	1.90	1.70	2.00
年化收益率:	74%	58%	82%
年化波动率:	32.16%	12.75%	12.57%
胜率:	71.43%	85.71%	100.00%
最大回撤:	-27.36%	-8.33%	-3.86%

数据来源：广发证券发展研究中心

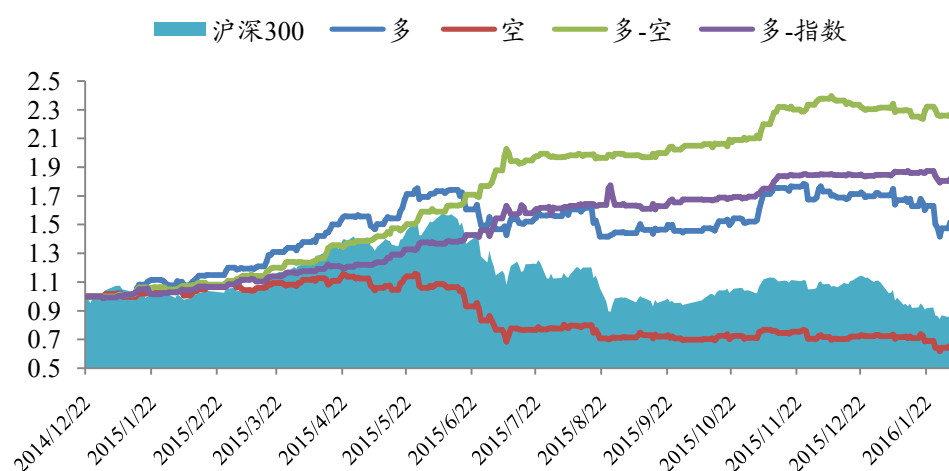
在中证800成分股中的选股策略中可以看出，月度调仓整体表现优异，多空策略周频调仓更好。

#### 沪深300成分股选股实证

基于雪球关注度数据，最后考虑在沪深300成分股中周频或者月频调仓下的策略表现。

从图26以及表8的结果中可以看出，在周频调仓下，策略在回测期内的表现优异。在回测期内，基于雪球关注度的选股策略超配组合的累计净值为1.54，年化收益率为47%，周胜率为63.79%，最大回撤为-20.65%。在假设低配个股可以融券卖空的情况下，多空对冲后的累计净值为2.31，年化收益率为110%，胜率为70.69%，最大回撤为-6.6%；在沪深300指数对冲的情况下，策略对冲后的累计净值为1.85，年化收益率为72%，周胜率为72.41%，最大回撤为-9.45%。

图 26 周频调仓下雪球关注度沪深300选股策略回测表现



数据来源：广发证券发展研究中心

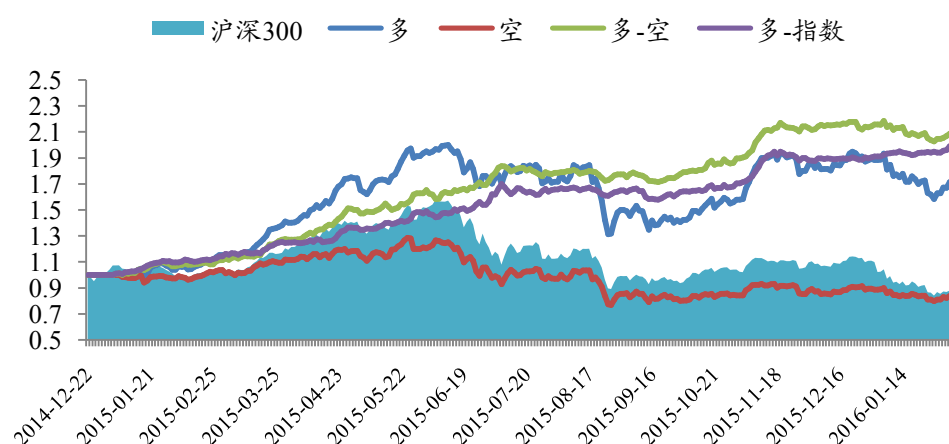
表 8 周频调仓下雪球关注度沪深300选股策略指标表现

指标	多头	多-空	多-指数
累计净值:	1.54	2.31	1.85
年化收益率:	47%	110%	72%
年化波动率:	28%	15%	16%
胜率:	63.79%	70.69%	72.41%
最大回撤:	-20.65%	-6.60%	-9.45%

数据来源：广发证券发展研究中心

从图27以及表9的结果中可以看出,在月频调仓下,策略在回测期内的表现优异。在回测期内,基于雪球关注度的选股策略超配组合的累计净值为1.7,年化收益率为61%,月度胜率为64.29%,最大回撤为-34.41%。在假设低配个股可以融券卖空的情况下,多空对冲后的累计净值为2.1,年化收益率为94%,月度胜率为85.71%,最大回撤为-7.22%;在沪深300指数对冲的情况下,策略对冲后的累计净值为1.99,年化收益率为85%,月度胜率为85.71%,最大回撤为-7.17%。

图 27 月频调仓下雪球关注度沪深300选股策略回测表现



数据来源：广发证券发展研究中心

表 9 月频调仓下沪深300选股策略指标表现

指标	多头	多-空	多-指数
累计净值:	1.70	2.10	1.99
年化收益率:	61%	94%	85%
年化波动率:	36%	15%	14%
胜率:	64.29%	85.71%	85.71%
最大回撤:	-34.41%	-7.22%	-7.17%

数据来源：广发证券发展研究中心

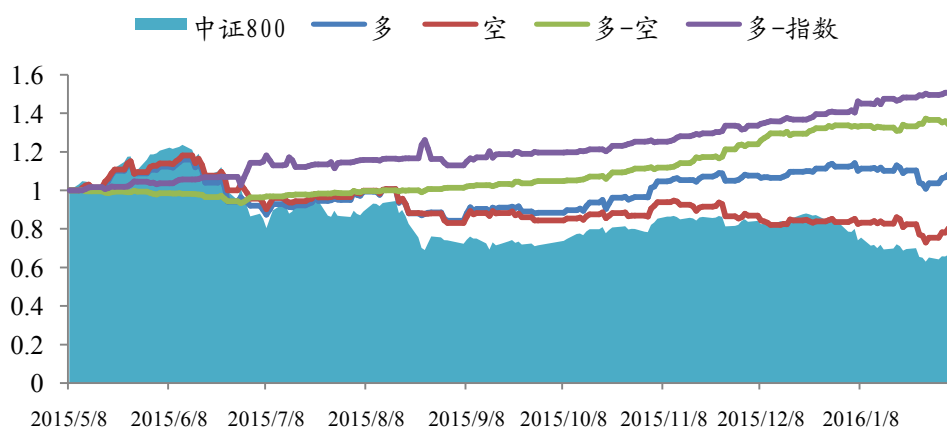
## 东财关注度选股实证结果

### 全市场选股实证

在回测过程中，考虑关注度在周频或者月频调仓下的策略表现结果。

从图28以及表10的结果中可以看出，基于东财关注度选股策略在周频下的回测表现结果较差。

图 28 周频调仓下东财关注度全市场选股策略回测表现



数据来源：广发证券发展研究中心

表 10 周频调仓下东财关注度全市场选股策略指标表现

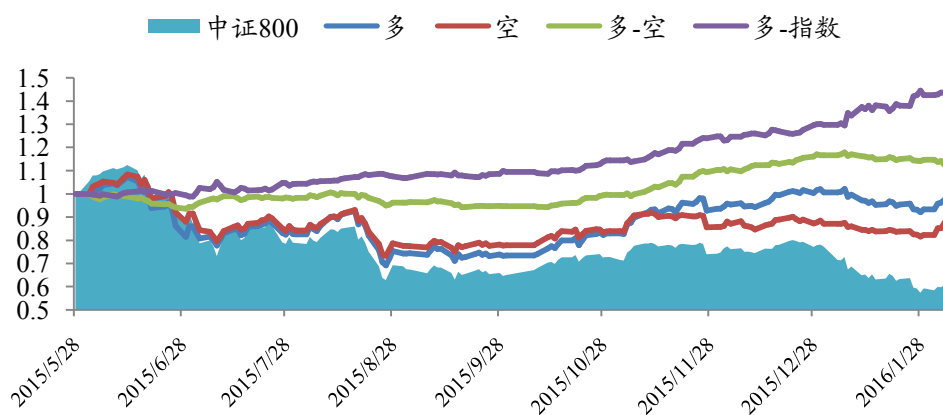
指标	多头	多-空	多-指数
累计净值:	1.07	1.34	1.51
年化收益率:	9.70%	46.41%	69.59%
年化波动率:	29.99%	10.38%	18.54%
胜率:	64.10%	74.36%	84.62%
最大回撤:	-27.42%	-6.80%	-10.36%

数据来源：广发证券发展研究中心

从图29以及表11的结果中可以看出，基于东财关注度选股策略在月频下的回测表现结果也较差。



图 29 月频调仓下东财关注度全市场选股策略回测表现



数据来源：广发证券发展研究中心

表 11 月频调仓下东财关注度全市场选股策略指标表现

指标	多头	多空	多-指数
累计净值:	0.97	1.12	1.43
年化收益率:	-4.58%	17.28%	65.46%
年化波动率:	35.87%	10.46%	13.15%
胜率:	44.44%	44.44%	88.89%
最大回撤:	-35.52%	-6.36%	-4.41%

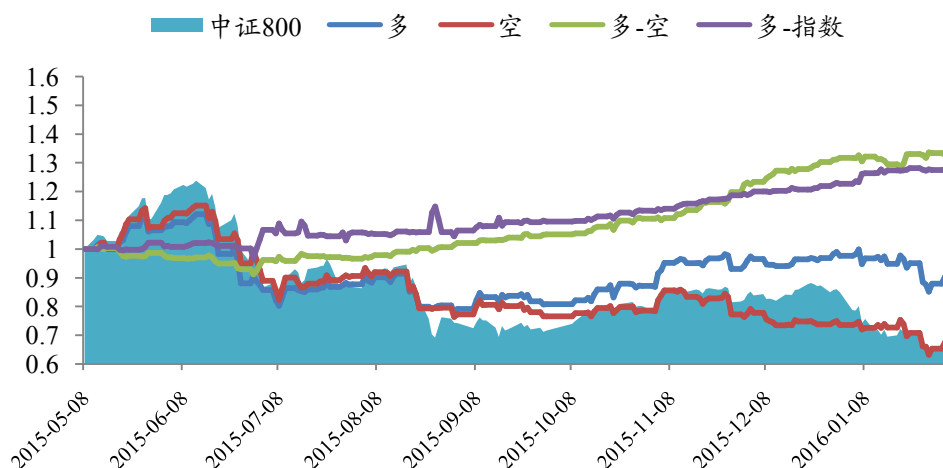
数据来源：广发证券发展研究中心

#### 中证800成分股选股实证

其次在回测过程中，考虑在中证800成分股中，关注度在周频或者月频调仓下的策略表现结果。

从图30以及表12的结果中可以看出，基于东财关注度选股策略在周频下在中证800选股的回测表现结果也较差，在回测期内的累计净值为0.91，最大回撤为-30.68%。

图 30 周频调仓下东财关注度中证800选股策略回测表现



数据来源：广发证券发展研究中心

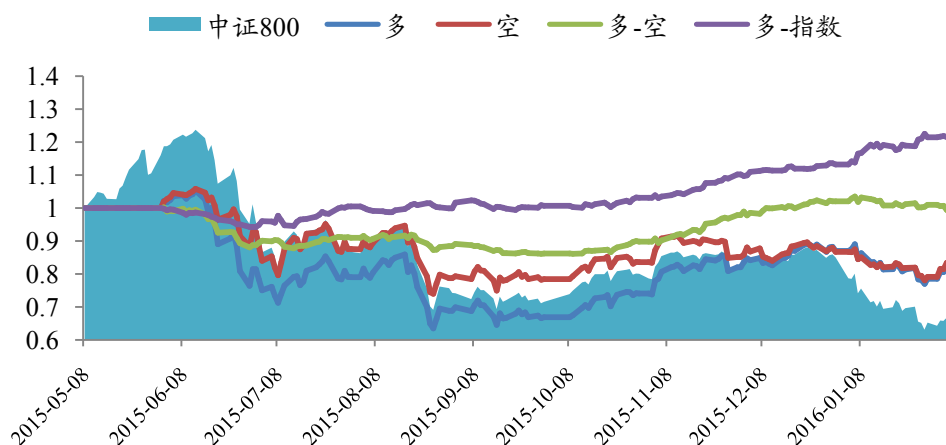
表 12 周频调仓下东财关注度中证800选股策略指标表现

指标	多头	多-空	多-指数
累计净值:	0.91	1.31	1.28
年化收益率:	-11.93%	40.79%	36.61%
年化波动率:	32.21%	11.37%	17.13%
胜率:	58.97%	76.92%	71.79%
最大回撤:	-30.68%	-9.34%	-9.05%

数据来源：广发证券发展研究中心

从图31以及表13的结果中可以看出，基于东财关注度选股策略在月频下在中证800选股的回测表现结果也较差，在回测期内的累计净值为0.81，最大回撤为-39.7%。

图 31 月频调仓下东财关注度中证800选股策略回测表现



数据来源：广发证券发展研究中心

表 13 月频调仓下东财关注度中证800选股策略指标表现

指标	多头	多-空	多-指数
累计净值:	0.81	0.99	1.21
年化收益率:	-25.77%	-1.72%	31.42%
年化波动率:	40.33%	11.70%	9.97%
胜率:	44.44%	44.44%	77.78%
最大回撤:	-39.70%	-14.03%	-5.69%

数据来源：广发证券发展研究中心

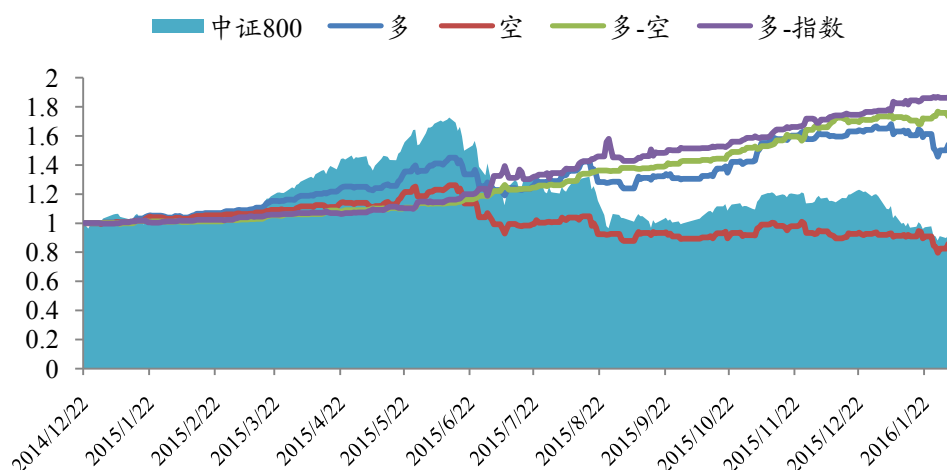
### 综合策略实证结果

从上述基于雪球以及东财的关注度选股策略回测结果可以看出，在历史回测期内在市场以及800的表现相差较大，为了降低整个策略的波动以及提高策略的胜率，综合运用雪球以及东财的数据进行选股。综合策略构建原理与上述选股的策略原理一致，在期初将资金二等分分别回测在雪球以及在东财上的选股策略。

### 全市场选股实证

从图32以及表14的结果中可以看出，在周频调仓下，综合策略在回测期内的表现优异，回撤、胜率都等到了改善。在回测期内，综合选股策略超配组合的累计净值为1.56，年化收益率为46.66%，月度胜率为68.96%，最大回撤为-17.91%。在假设低配个股可以融券卖空的情况下，多空对冲后的累计净值为1.72，年化收益率为60.02%，月度胜率为79.31%，最大回撤为-3.74%；在中证800指数对冲的情况下，策略对冲后的累计净值为1.87，年化收益率为72.27%，月度胜率为81.03%，最大回撤为-9.65%。

图 32 周频调仓下雪球-东财关注度全市场选股策略回测表现



数据来源：广发证券发展研究中心

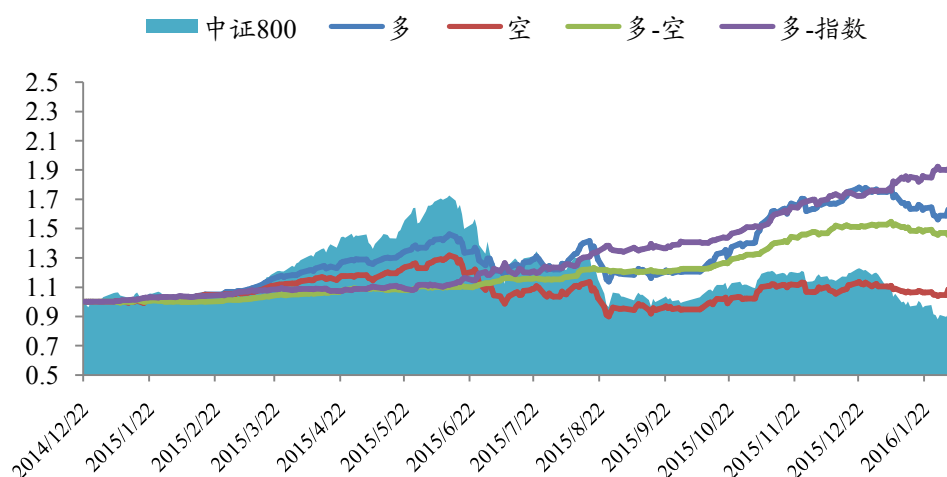
表 14 周频调仓下雪球-东财关注度全市场选股策略指标表现

指标	多头	多-空	多-指数
累计净值:	1.56	1.72	1.87
年化收益率:	46.66%	60.02%	72.27%
年化波动率:	21.13%	9.65%	16.67%
胜率:	68.97%	79.31%	81.03%
最大回撤:	-17.91%	-3.74%	-9.65%

数据来源：广发证券发展研究中心

从图33以及表16的结果中可以看出，在月频调仓下，综合策略在回测期内的表现优异，回撤、胜率都等到了改善，并且月频下的策略回测结果表现优于周频下的策略表现。在回测期内，综合选股策略超配组合的累计净值为1.65，年化收益率为54.09%，月度胜率为71.43%，最大回撤为-22.04%。在假设低配个股可以融券卖空的情况下，多空对冲后的累计净值为1.43，年化收益率为36.46%，月度胜率为85.71%，最大回撤为-7.6%；在中证800指数对冲的情况下，策略对冲后的累计净值为1.92，年化收益率为75.58%，月度胜率为100%，最大回撤为-6.03%。

图 33 月频调仓下雪球-东财关注度全市场选股策略回测表现



数据来源：广发证券发展研究中心

表 15 月频调仓下雪球-东财关注度全市场选股策略指标表现

指标	多头	多-空	多-指数
累计净值:	1.65	1.43	1.92
年化收益率:	54.09%	36.46%	75.58%
年化波动率:	24.23%	7.84%	13.03%
胜率:	71.43%	85.71%	100.00%
最大回撤:	-22.04%	-7.60%	-6.03%

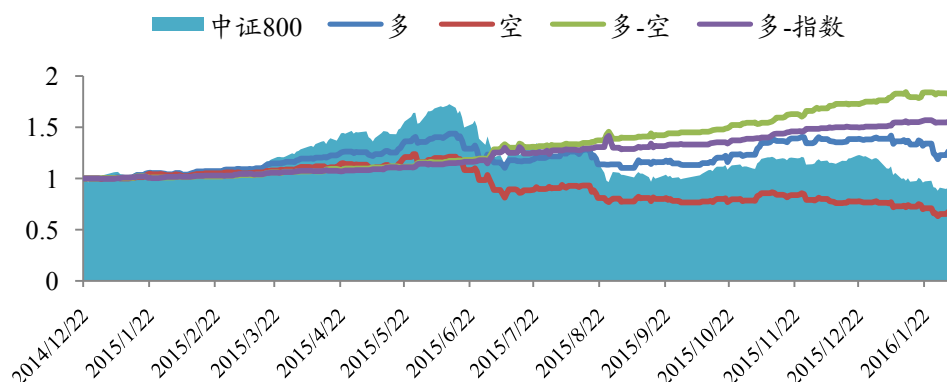
数据来源：广发证券发展研究中心

#### 中证800成分股选股实证

其次在回测过程中，考虑在中证800成分股中，雪球-东财关注度在周频或者月频调仓下的策略表现结果。

从图34以及表16的结果中可以看出，在周频调仓下，综合策略在中证800选股回测期内的表现优异，回撤、胜率都等到了改善。在回测期内，综合选股策略在中证800指数对冲的情况下，策略对冲后的累计净值为1.55，年化收益率为46.53%，月度胜率为82.76%，最大回撤为-9.41%。

图 34 周频调仓下雪球-东财关注度中证800选股策略回测表现



数据来源：广发证券发展研究中心

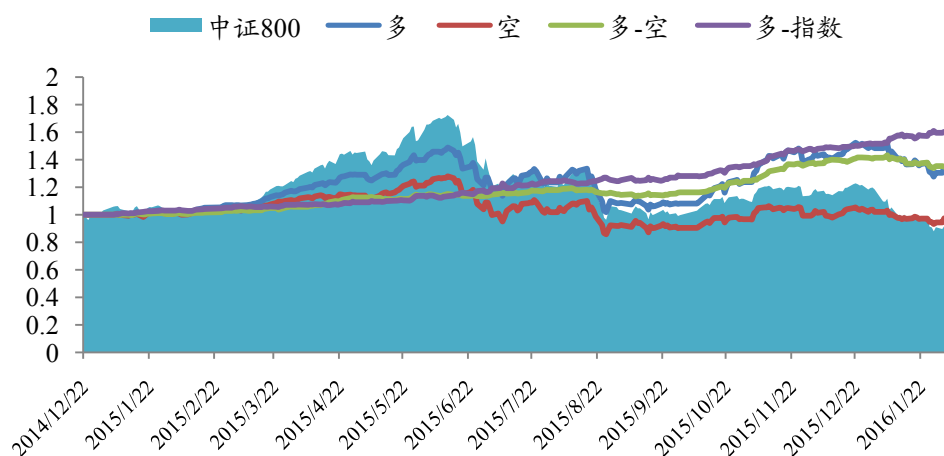
表 16 周频调仓下雪球-东财关注度中证800选股策略指标表现

指标	多头	多-空	多-指数
累计净值:	1.27	1.81	1.55
年化收益率:	23.21%	67.30%	46.53%
年化波动率:	25.53%	12.35%	14.47%
胜率:	65.52%	82.76%	82.76%
最大回撤:	-23.34%	-4.62%	-9.41%

数据来源：广发证券发展研究中心

从图35以及表17的结果中可以看出，在月频调仓下，综合策略在中证800选股回测期内的表现不错，回撤、胜率都等到了改善。在回测期内，综合选股策略在中证800指数对冲的情况下，策略对冲后的累计净值为1.61，年化收益率为50.69%，月度胜率为92.86%，最大回撤为-2.9%。

图 35 月频调仓下雪球-东财关注度中证800选股策略回测表现



数据来源：广发证券发展研究中心

表 17 月频调仓下雪球-东财关注度中证800选股策略指标表现

指标	多头	多-空	多-指数
累计净值:	1.35	1.34	1.61
年化收益率:	29.86%	29.10%	50.69%
年化波动率:	29.88%	8.44%	9.00%
胜率:	64.29%	71.43%	92.86%
最大回撤:	-31.72%	-6.52%	-2.90%

数据来源：广发证券发展研究中心

## 总结及展望

本专题基于互联网上个股的关注度数据进行了实证研究。实证结果表明，策略在回测期内表现优异，基于互联网关注度指标能够作为一个有效的选股因子；不同的热门网站上因关注的人的层次等不同，导致关注度的选股策略效果表现各异。

在大数据研究的方向上，未来可能更多地基于关注度，加入更多的因素考虑个股之间的动态变化构建策略以及在大盘、行业整体上考虑关注度变化研究行业配置、择时策略，同时针对机构投资者的个性化定制需求；

## 风险提示

本篇报告所提供的个股关注度数据一定程度上能够反映投资者对个股的关注，但其中的关注度数据并不能代表所有投资者的观点，股票价格受诸多因素影响，历史的数据并不能精确地预测未来金融市场变化，注意市场其他风险的变化。



## 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。