

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）

报告摘要:

● 互联网金融大数据强势来袭

近年来，随着互联网大数据及行为金融的发展，人们越来越关注非结构化的数据给投资者的投资决策带来的影响。在非结构化数据中，上市公司公告是其中一个组成部分。上市公司所披露的某些公告常常隐藏着许多关键信息，从而带给投资者巨大的想象空间，导致股价大幅波动。然而由于公告的发布具有非定期性以及内容多样性等特征，导致我们很少能够对其背后隐藏的投资机会进行系统的识别分类并对未来利好信息的精准把握。本文的目的是通过对公告进行重新分类，并结合公告的披露频率以及公告日个股股价涨跌情况来对相应个股进行涨跌预测！

● 公告披露事件选股策略构建

本专题中，针对上市公司最新披露的公告信息进行事件选股策略构建时，我们主要考虑了以下几方面的信息：

（1）**考虑公告类型**：公告主要划分为上述八大类，此外对于财务报表我们还单独划分了业绩预增与快报大幅超预期两类公告；

（2）**考虑长期未出公告**：在前期报告《此时无声胜有声——长期不出公告股票的投资机会》中，我们发现上市公司若长期未出任何公告，则在发布公告之后其股价上涨概率较高；沿用该思路，我们继续对个股的细分公告类型进行发布期限跟踪，若上市公司长期未出某一类型的公告，则跟踪其发布该类公告后的个股表现情况

（3）**考虑公告披露当日个股表现**：同一类型的公告，若不细分公告的具体内容实际上还难以界定该类公告属于利好还是利空，同样也难以与其对股价带来的影响；而考虑公告日个股股价表现实际上是参考了公告日投资者对该公告的普遍预期偏差情况，因此我们将不同的公告类型在公告披露当日的个股表现情况列为重点区分标准之一。

● 重大事项及股东大会公告等策略选股能力突出

历史回测结果表明，财务报表、股东大会、增发公告以及重大积极事项等公告策略的选股效果相当突出，均能获得较高的超额收益。

其中，建议买入披露财务报告当日股价窄幅震荡的个股，持有一周左右，平均能获得年化超额收益 37%，信息比 1.02；

建议买入发布股东大会公告的个股，若个股当日个股大涨则效果更佳，持有期为一个月左右，平均年化超额收益 20%；

增发公告当日，若个股大涨则坚决买入，持有期限为一个月左右，需要注意的是若个股当日没有大涨，则规避该类个股；

重大积极事项披露同样带来个股显著超额收益，可持有 10 个交易日左右。

识别风险，发现价值

图 1：财务公告事件策略实证结果

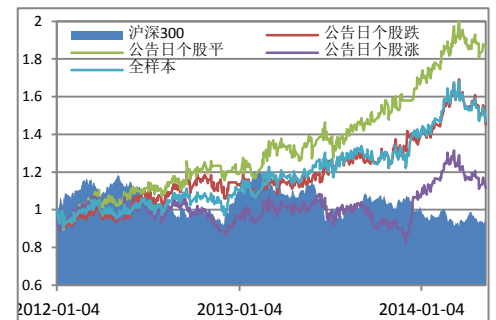


图 2：股东大会公告事件策略实证结果

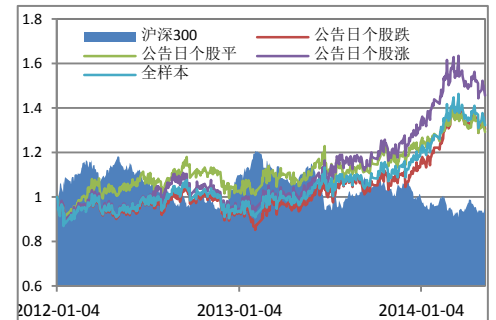


表 1：财务公告事件策略实证结果

	跌	平	涨	全部
年化收益	19.2%	36.7%	4.9%	19.9%
信息比	0.66	1.02	0.31	0.68
回撤	14.1%	12.9%	23.9%	12.8%
胜率	55.0%	54.6%	51.2%	54.4%

数据来源：广发证券发展研究中心

分析师：史庆盛 S0260513070004



02087555888618



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之（一）	2014-06-25
广发证券金融工程：事件驱动量化选股周报	2014-06-23
此时无声胜有声——长期不出公告股票的投资机会	2012-03-21

请务必阅读末页的免责声明

目录索引

一、互联网大数据挖掘体系介绍.....	4
1.1 什么是互联网大数据?.....	4
1.2 互联网金融数据获取.....	5
1.3 互联网大数据抓取体系	6
二、公告数据来源及抓取.....	8
2.1 公告类型	8
2.2 数据来源	9
2.3 公告信息披露平台结构	10
三、公告披露事件驱动选股策略构建	11
四、实证分析	12
4.1 数据说明	12
4.2 实证结果	12
4.3 工具推介	16
五、总结	17
风险提示	18

图表索引

图 1: 财务公告事件策略实证结果	1
图 2: 股东大会公告事件策略实证结果	1
图 3: 互联网数据来源	5
图 4: 文本信息挖掘流程	6
图 5: 互联网数据抓取体系	7
图 6: 上市公司公告类型划分	8
图 7: 公告数据抓取首页	11
图 8: 公告数据抓取内容	11
图 9: 财务公告事件策略实证结果	13
图 10: 股东大会公告事件策略实证结果	14
图 11: 增发公告事件策略实证结果	15
图 12: 重大积极公告事件策略实证结果	16
图 13: A 股公告抓取工具展示	17
表 1: 财务公告事件策略实证结果	1
表 2: 著名 IT 公司文本挖掘项目	5
表 3: 上市公司公告类型划分明细	9
表 4: 财务公告事件策略实证结果	13
表 5: 股东大会公告事件策略实证结果	14
表 6: 增发公告事件策略实证结果	15
表 7: 重大积极公告事件策略实证结果	16

一、互联网大数据挖掘体系介绍

1.1 什么是互联网大数据？

随着云时代的来临，大数据（Big data）也吸引了越来越多的关注，人们用它来描述和定义信息爆炸时代产生的海量数据，并命名与之相关的技术发展与创新。正如《纽约时报》的一篇专栏中所称，“大数据”时代已经降临，在商业、经济及其他领域中，决策将日益基于数据和分析而作出，而并非基于经验和直觉。

实际上，大数据从很早以前就一直存在着，然而只是光只有数据大是没有用处的。世界上最大的数据估计和互联网一点关系都没有，今天我们所普遍关注的数字不仅仅是大，最重要的是这些大数据是以在线的形式存在了，这个恰恰是互联网的特点。所有东西在线这个事情，远远比“大”更反映本质。

像快的打车要用一个交通的数据，如果这些东西不在线，是没有用的。

又好比今天的淘宝数据和大众点评数据，因为他在线了，所以就值钱！写在磁带、写在纸上的数据，根本没有用，也没法用。

反过来讲，互联网也让数据搜集变得非常容易。过去美国谁要做总统，需要做盖勒普调查，去街上拦2000个人，在纸上打个勾，预测就很准了。现在不用做这个事情，只要在twitter上分析每个人发的东西，就可以知道总统会是谁了。

可见，互联网行业的“大数据”才称得上是有使用价值和可操作性的“大数据”！这些数据的规模是如此庞大，以至于不能用G或T来衡量，因此也常常称为“巨量数据”或“海量数据”，具有数量巨大、结构复杂、类型众多等特点。那么互联网大数据到底有多大？一组名为“互联网上一天”的数据告诉我们，一天之中，互联网产生的全部内容可以刻满1.68亿张DVD；发出的邮件有2940亿封之多（相当于美国两年的纸质信件数量）；发出的社区帖子达200万个（相当于《时代》杂志770年的文字量）；卖出的手机为37.8万台，高于全球每天出生的婴儿数量37.1万……

目前，互联网上的数据量已经从以往的TB（1024GB=1TB）级别跃升到PB（1024TB=1PB）、EB（1024PB=1EB）乃至ZB（1024EB=1ZB）级别。国际数据公司（IDC）的研究结果表明，2008年全球产生的数据量为0.49ZB，2009年的数据量为0.8ZB，2010年增长为1.2ZB，2011年的数量更是高达1.82ZB，相当于全球每人产生200GB以上的数据。而到2012年为止，人类生产的所有印刷材料的数据量是200PB，全人类历史上说过的所有话的数据量大约是5EB。IBM的研究称，整个人类文明所获得的全部数据中，有90%是过去两年内产生的。而到了2020年，全世界所产生的数据规模将达到今天的44倍。

互联网发展已有二十年，这二十年就是各行各业被互联网冲击的二十年，最先受到冲击的是媒体，然后是零售业、旅游，刚刚开始的是金融业！

受冲击行业的每一次涅槃，也是一次重生！例如：媒体行业诞生了以雅虎为代表的门户，谷歌、百度为代表的搜索引擎，Twitter为代表的社交媒体，至今这个行业的变化还在继续；在零售业，产生了阿里巴巴、亚马逊、易贝这样的电商公司；在旅游行业则诞生了携程、去哪儿网这样的公司；而**金融行业的故事则刚刚开始……**

1.2 互联网金融数据获取

伴随着互联网的快速发展,信息的提供者与使用者之间的界限已经越来越模糊。在互联网金融时代的背景下,金融信息的来源渠道也越来越丰富多样。投资者有更多的渠道来获取相关的金融信息,过去上百年的金融研究,往往局限于对以数字形式存在的数据的研究,而忽略了对非结构化的金融“数据”的研究,在当前互联网时代下,分析师的研究报告、股吧论坛帖子的信息、新闻媒体的新闻以及微博和微信等非结构化文本信息往往能够反应当前市场上投资者对股市的投资情绪,而这些信息往往又对投资者的投资决策起到潜移默化的作用。

采用互联网文本挖掘的方法对这些非结构化文本形式存在的金融信息进行挖掘,并从中提取出对投资决策有利的信息是该专题报告研究的重点。

图3: 互联网数据来源



数据来源: 广发证券发展研究中心

基于上述三类数据来源,我们将采用互联网文本挖掘技术来获取相关的信息。

“文本信息挖掘”的概念最早由Ronen Feldman博士提出,并倡导将非结构化的内容转变为有价值的商业智能行业中,即文本驱动商务智能概念。目前许多IT巨头已经纷纷在不同领域针对大数据开展了文本挖掘的项目。

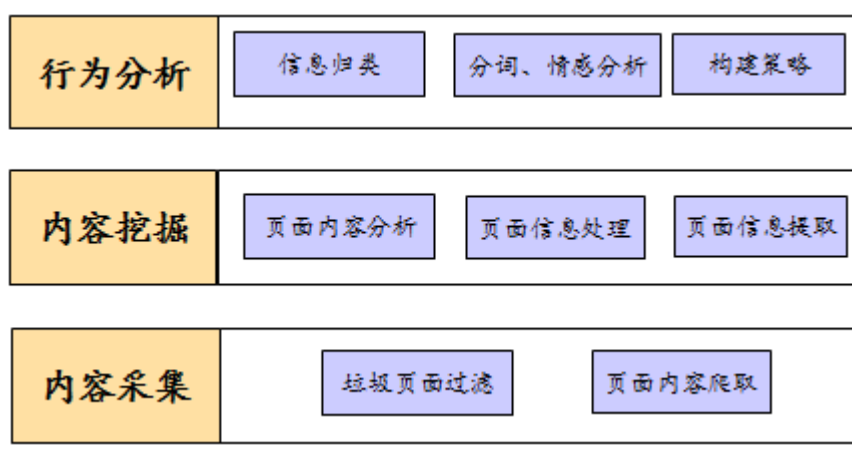
表2. 著名IT公司文本挖掘项目

公司	文本挖掘项目
谷歌	X Lab, 谷歌大脑项目
微软	TextFlow项目
脸书	深度学习研究小组预测用户行为
百度	创建深度学习研究院
腾讯	成立中文处理研究室, 从事文本挖掘研究

文本信息挖掘是通过分析用户数据,从大量数据中寻找其规律的技术,主要有数据准备、规律寻找和规律表示等步骤。移动互联网数据具有数据量大、数据结构复杂、数据内容分散等特点,呈现出爆炸性增长的趋势。因此,为了从浩如烟海的数据中提取出有效信息,必须选择合适的数据挖掘策略。

信息挖掘是一个复杂的过程，需要进行大量的数据采集和运算等。按照基本功能，可以将整个信息挖掘流程划分成内容采集、内容挖掘和行为分析3个环节。

图4：文本信息挖掘流程



数据来源：广发证券发展研究中心

（一）内容采集

进行互联网文本数据挖掘的基础是数据的真实性和有效性，内容采集主要包括以下两个方面。

a) 页面内容爬取。这是将网页的内容通过爬虫获取的部分，分析页面代码格式，进行网页代码的编码转换等，尽可能获取自己需要的信息。

b) 页面垃圾过滤。页面中不可避免地会存在大量的垃圾信息，这些信息严重干扰到对后期信息挖掘的准确性，页面垃圾过滤机制会找出包括广告在内的段落以及其他对内容挖掘无效的部分，并将其清除，不进入内容挖掘部分。

（二）内容挖掘

主要是对需要的特定信息进行提取，该阶段处理后的文本数据是后期进行分词、情感分析的基础；

（三）行为分析

整个文本挖掘过程的重点和难点是分词和情感分析，对于海量文本信息而言，程序的处理速度也是至关重要的一点。

1.3 互联网大数据抓取体系

对于大规模的互联网信息抓取而言，单线程的数据获取是一项非常局限的事情，因此多线程、分布式的信息抓取平台是必须搭建的。因此数据的抓取平台的搭建是一项基础性的工作。

整体上，我们的大数据抓取平台可分为三部分：

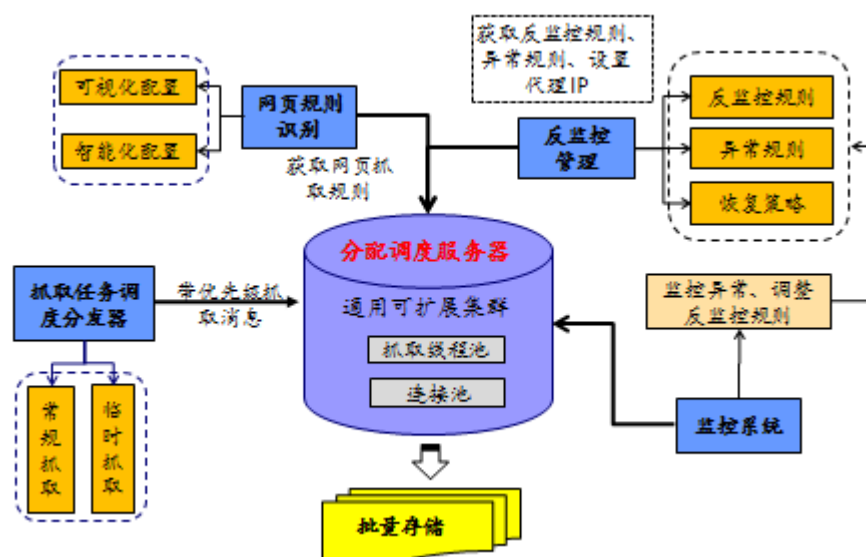
首先是搜索热门网站，对需要提取信息的网站的网页编码格式、网页制作规则进行研究，提取出需要提取的信息的网页编码格式；

第二是对需要提取信息的网站进行分布式配置，多线程爬取特定的信息，并将信息存储到指定的数据库中；

第三，监控数据爬取过程，防止网站数据异常带来的数据抓取缓慢等问题，主

要是防止对指定网站的频繁访问导致IP被限制访问的问题。

图5：互联网数据抓取体系



数据来源：广发证券发展研究中心

上图刻画了在大规模的互联网数据爬取过程中的整体框架，对于单个网站的的数据抓取框架可以简化成以下几个步骤：

- 1) 首先对需要批量爬取信息的网页结构进行研究，主要是研究需要提取的特定信息的网页格式进行研究，这步骤可以借助在IE浏览器下打开需要提取信息的网页，然后调用出开发人员工具，找到需要批量提取的特定信息的网页代码规则，或者利用Firefox浏览器下的XPATH工具以及View Source Chart工具；
- 2) 找到不同个股的网页url之间的关系，利用url之间的关系，抓取所有个股的网页url；
- 3) 由于网页抓取的数据量往往非常庞大，出于效率的考虑，往往采用分布式、多线程的方式进行，若条件允许需要同时几十部甚至上百部电脑多个线程地抓取需要的内容。此时，就需要一台主程序电脑控制其他电脑的程序运转；在数据抓取过程中由于网络不稳定、电脑突然死机等问题，需要对抓取的过程进行监控，记录下每个进程运行的过程，防止数据抓取的缺漏，同时处于对网站安全的考虑，需要实时监控程序的运行，防止网络访问限制问题，设置适当的断线重连机制，IP切换机制等；
- 4) 对于抓取到的信息统一批量存储到指定的数据库中，构成后续分析的底层的数据库。

二、公告数据来源及抓取

上市公司所披露的某些公告常常隐藏着许多关键信息，从而带给投资者巨大的想象空间，导致股价大幅波动。然而由于公告的发布具有非定期性以及内容多样性等特征，导致我们很少能够对其背后隐藏的投资机会进行系统的识别分类并对未来利好信息的精准把握。

考虑到A股上市公司每日公告分别在上交所及深交所的信息披露平台上实时公布的，我们自主开发编写了一套网络公告文本抓取系统并维护相应的公告数据库，实现同步维护交易所披露的公告，从而能够第一时间把握上市公司披露的报告，并通过对公告类型及公告内容等重要信息进行分析，从而挖掘公告披露背后所隐含的投资机会。

信息渠道：上海证券交易所信息披露平台& 深圳证券交易所信息披露平台

技术要求：VBA+Matlab+多线程

数据特点及信息价值：非结构化金融数据、半结构化数据，对公告信息量化挖掘，获取有用价值

2.1 公告类型

什么类型的公告披露之后更加容易带来股价的上涨？由于公告类型较多，我们针对常见的公告分别建立不同的关键词库，用于识别公告类型，并将其分为8大类，其中，前三类公告分别为：财务报表、股东大会以及重大利好，这三大类公告一直以来都受到了许多投资者的重视，许多量化投资者及传统主动投资者都尤其关注这些公告的发布，我们不对具体的公告内容进行区分，而同一默认将其划分为“利好”公告；

第4-6类是大家所常见的“利空”公告，包括个股增发、交易异常信息提示、配股公告等，这些公告披露的同时，是否一定利空个股？

最后两大类公告则没有明确的多空界定，包括股权股本相关业务公告以及公司重大事项等。

在下文中我们将对上述公告类型进行一一测算，在区分公告类型的同时，也将结合公告披露当日个股股价的表现以及上市公司多长时间内未披露过该类公告一起来进行区分和定义，分析公告披露背后所隐藏的投资机会。

图6：上市公司公告类型划分

1. 财务报表	年报、中报、季报、业绩预告、业绩快报
2. 股东大会	股东大会
3. 重大利好	资产重组、重大合同、股权激励、政策刺激等
4. 增发	增发预案、增发获准、增发上市等
5. 交易提示	交易异动、澄清公告、风险提示
6. 配股	配股预案、配股获准、配股上市等
7. 股权股本	权益变动、质押冻结、预定回购等
8. 重大事项	利润分配、资金投向、委托理财、人事变动等

数据来源：广发证券发展研究中心

表3. 上市公司公告类型划分明细

财务报告	股东大会	积极事项	增发	交易提示	配股	股权股本	重大事项
业绩预告（细分） 业绩快报（细分） 年报 季报 半年报	股东大会	资产重组 收购兼并 重大合同 股权激励 政策影响	增发预案 增发说明书 增发获准 增发发行 增发上市	交易异动 澄清公告 风险提示	配股预案 配股说明书 配股获准 配股发行 配股上市	权益变动 股本变动 质押冻结 质押式回购 回购股权 约定购回 股权分置改革 权证公告	利润分配 增减持 资金投向 资产重组 收购兼并 重大合同 股权激励 关联交易 借贷担保 委托理财 违纪违规 政策影响 人事变动

数据来源：广发证券发展研究中心

2.2 数据来源

上市公司披露重要的报告，隐含巨大的想象空间，因而常常导致股价大幅波动。因此，及时、准确地获取到上市公司的公告就显得非常地重要。获取上市公司重要公告一般有两种途径：一种为通过上市公司的官方网站，一种为通过上市公司公告披露的指定平台，例如上海证券交易所、深圳证券交易所或者巨潮资讯平台等。通过上市公司官方网站获取公告，需要批量地获取到所有上市公司的官方网站地址，工作量大，效率低，不利于后期公告抓取的维护，而且需要对所有上市公司官方网站信息披露平台的网站需要提取信息的网页编码进行研究，无疑也是效率非常低的

事情；因此通过上海证券交易所或者是深圳证券交易所的公告披露平台获取到上市公司的公告就显得较为简单了，因为在交易所网站上往往能够查询所有上市公司历史时间段内的各类重要的公告信息，例如业绩披露公告、上市公司重组公告、交易异常公告等等。基于对公告获取平台效率的权衡，本专题获取的上市公司历史公告信息分别来源于深圳证券交易所公告披露平台以及巨潮资讯网公告披露平台。

2.3 公告信息披露平台结构

基于对效率的权衡，获取上市公司的历史公告，对于在深圳证券交易所上市的公司，我们采用深圳证券交易所的信息获取平台获取个股的重要公告信息，而对于在上海证券交易所上市的个股公告信息，我们采用巨潮资讯网的公告披露平台获取个股的重要公告信息。

以下以深圳证券交易所上市公司公告披露平台为例，解析公告获取过程：

- 1) 首先获取到上市公司公告披露平台网址：
<http://disclosure.szse.cn/m/search0425.jsp>;
- 2) 从获取到的上市公司公告披露平台上可以看到在网站上可以查询到不同上市公司在某一时间段内或者是某日的特定公告类型；
- 3) 通过2)中的解析后有两种方法可以获取到上市公司公告，一种为通过每次查询后的结果获取到结果的url，然后比较不同url之间的固定参数以及可变参数的关系，找到url之间的规律，另一种为通过模拟鼠标操作以及输入的方法获取到需要查询的信息；

以第一种方法为例，解析查询后获取的网址的规律。

通过对网页编码的解析，我们可以获取到每次查询后的网址的实际网址信息，例如在股票代码输入框中输入000001，时间段查询选择：2014-06-01至2014-06-22后点击确定后可以看到查询后的网址为：

<http://disclosure.szse.cn/m/search0425.jsp>与之前的网址没有发生变化，但是实际的查询网址为：

<http://disclosure.szse.cn/m/search0425.jsp?leftid=1&lmid=drgg&pageNo=1&stockCode=000001¬iceType=&startTime=2014-06-01&endTime=2014-06-22>，对不同的股票代码以及查询时间进行比较后，可以找到不同url之间的关系，具体解析如下：

固定参数部分：<http://disclosure.szse.cn/m/search0425.jsp?leftid=1&lmid=drgg>

可变参数部分：pageNo、stockCode、noticeType、startTime、endTime等

其中pageNo表示如果查询结果返回的是多页情况，可以通过pageNo的变化获取到不同页面的结果，stockCode表示需要查询信息的股票代码，noticeType表示需要查询的公告类型，startTime表示查询的开始时间段，endTime表示查询的结束时间段。

- 4) 通过3)中对查询网址的规律的查询，我们可以根据url的规律批量获取到个股在历史时间段内的公告，并将公告信息提取出来。

基于上述对公告获取规则的查询，我们可以得到不同个股查询结果后的公告信息。对需要提取的信息的网页编码格式进行解析后，可以获取到不同的个股公告信

息。

图7：公告数据抓取首页

数据来源：广发证券发展研究中心

上图是上交所信息披露平台上的公告搜索界面。

图8：公告数据抓取内容

上市公司公告全文

证券代码: 600000 检索范围: ☒ 全部 ☐ 临时公告 ☐ 定期报告 GO 减少条件

披露日期: 2014-03-22 至 2014-06-22 公告类型: 全部

公告名称	证券代码	公告类型	年份	披露日期	XBRL文档
浦发银行2013年度利润分配实施公告	600000	其它	2014	2014-06-17	
浦发银行第五届董事会第二十二次会议决议公告	600000	其它	2014	2014-06-11	
浦发银行第五届监事会第二十三次会议决议公告	600000	其它	2014	2014-06-11	
浦发银行关于与上海国际集团有限公司关联交易的公告	600000	其它	2014	2014-05-30	
浦发银行第五届董事会第二十次会议决议公告	600000	其它	2014	2014-05-30	
浦发银行第五届监事会第二十二次会议决议公告	600000	其它	2014	2014-05-30	
浦发银行关于关联交易的待审议意见	600000	其它	2014	2014-05-30	

数据来源：广发证券发展研究中心

上图是给定代码和日期区间后搜索得到的相应公告信息。

三、公告披露事件驱动选股策略构建

基于上述获取到的所有上市公司的历史公告信息，将个股的公告信息依据一定的标准进行分类，例如分类为财务报表、股东大会、重大利好、增发、配股、交易

提示等不同的类型，然后基于不同的公告类型、发布时间，发布时间间隔等建立不同的事件驱动策略。

本专题中，针对上市公司最新披露的公告信息进行事件选股策略构建时，我们主要考虑了以下几方面的信息：

(1) 考虑公告类型

公告主要划分为上述八大类，此外对于财务报表我们还单独划分了业绩预增与快报大幅超预期两类公告；

(2) 考虑长期未出公告

在前期报告《此时无声胜有声——长期不出公告股票的投资机会》中，我们研究发现上市公司若长期未出任何公告，则在发布公告之后其股价上涨概率较高；沿该思路，我们继续对个股的细分公告类型进行发布期限跟踪，若上市公司长期未出某一类型的公告，则跟踪其发布该类公告后的个股表现情况

(3) 考虑公告披露当日个股表现

同一类型的公告，若不细分公告的具体内容实际上还难以界定该类公告属于利好还是利空，同样也难以与其对股价带来的影响；而考虑公告日个股股价表现实际上是参考了公告日投资者对该公告的普遍预期偏差情况，因此我们将不同的公告类型在公告披露当日的个股表现情况列为重点区分标准之一。

四、实证分析

4.1 数据说明

样本区间：2012/1/4-2014/5/12。

股票范围：全部A股。

对冲标的：沪深300股指期货。

策略建仓：符合上述事件策略3个条件的个股次日开盘价全部买入，等权分配资金持有一段时间于到期日以收盘价平仓。同时买空股指期货进行完全对冲，获取个股超额收益。

交易费用：按照股指期货交易费用，双边万分之2。

策略参数：每一类事件不同，具体见实证结果。

4.2 实证结果

经实证，我们挑选了一下几类效果较好的事件选股策略进行推荐。

(一) 财务预告发布日个股窄幅震荡

实际上，若对财务预告的内容进行细分，我们会发现大幅预增预增类个股具有较高的超额收益，详情请参考前期研究报告《广发证券金融工程：事件驱动量化选股周报(20140623)》。

然而本策略我们不希望对公告的具体内容和类型进行细分，而是通过公告的大类划分来寻找投资者是否容易受到单纯的公告发布事件较大冲击，而对于公告的具体内容是利好还是利空，我们通过对公告日股价的表现来进行区分。

该类策略具体规则如下：

公告类型：财报（包括预告与快报）披露；

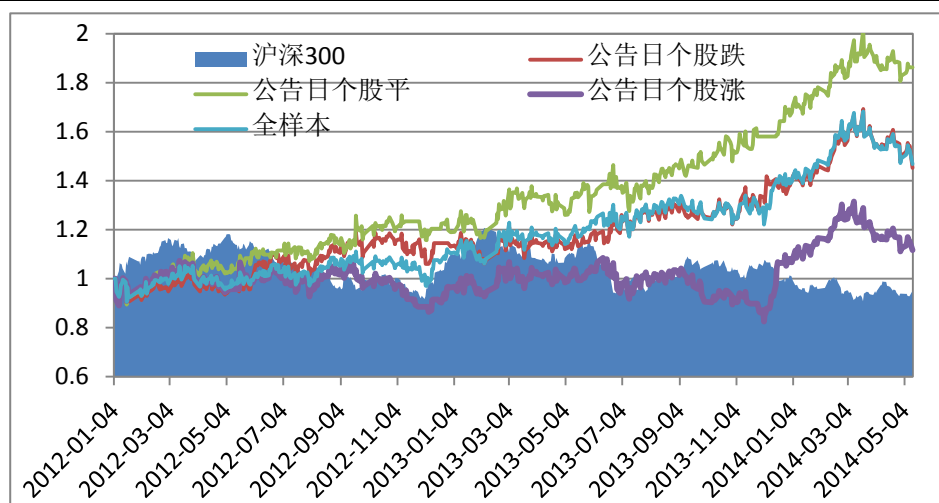
上次公告披露时间间隔要求：1个交易日以上；

公告披露日股价表现要求：窄幅震荡；

股票持有时间：4个交易日。

由于财务报表以及业绩预告、业绩快报等披露时间相对比较固定，且间隔较长，因此该类型公告对披露时间间隔没有任何要求；公告日最好是股价窄幅震荡，尽量规避当日出现大幅上涨的个股；相应个股买入持有4天为最佳，平均来看，持有期在一周以内的表现通常不错。

图9：财务公告事件策略实证结果



数据来源：广发证券发展研究中心

表 4：财务公告事件策略实证结果

	跌	平	涨	全部
年化收益率	19.2%	36.7%	4.9%	19.9%
年化信息比	0.66	1.02	0.31	0.68
回撤	14.1%	12.9%	23.9%	12.8%
胜率	55.0%	54.6%	51.2%	54.4%

数据来源：广发证券发展研究中心

（二）股东大会

股东大会形式上分为年会和临时会议两种。年会每年按时召开一次，上市公司的股东大会年会应于上一会计年度完结之后的6个月内举行。《公司法》规定了召开

临时会议的情形：1、董事人数不足《公司法》规定的人数或者公司章程所定的数的2/3时；2、公司未弥补亏损达到股本总额的1/3时；3、单独或合并持有公司有表决权股份总数10%以上的股东书面请求时；4、董事会认为必要时；5、监事会提议召开时。出现上述情形之一的，应当在2个月内召开临时股东大会。

客观地说，股东大会本身并必然意味着利好或利空，要看会议具体讨论的内容和做出的决策，但根据经验，A股投资者往往对即将召开的股东大会充满期待，从股价的表现上来说也常常表现为利好。

该类策略具体规则如下：

公告类型：股东大会召开公告披露；

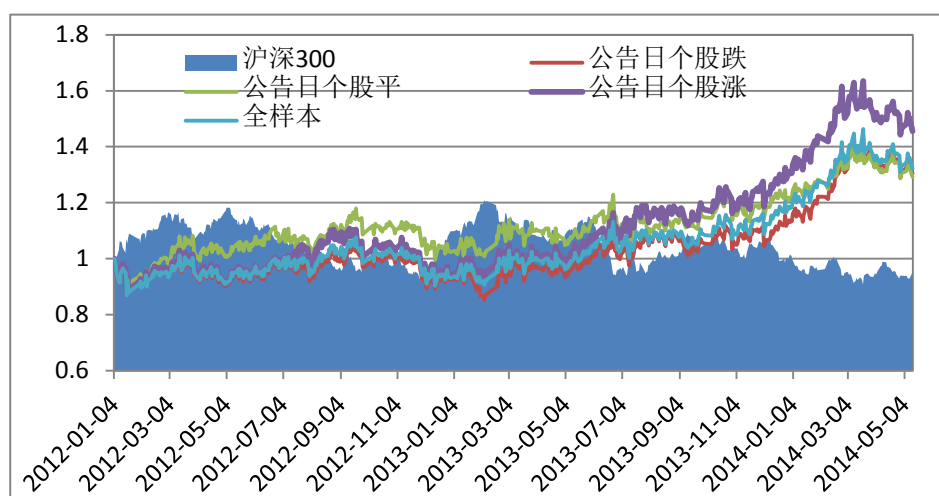
上次公告披露时间间隔要求：60个交易日以上；

公告披露日股价表现要求：无严格要求；

股票持有时间：20个交易日。

由于股东大会分为年会和临时会议，年会时间间隔较长，而临时会议可能会出现间隔较短的情况，因此该类型公告对披露时间间隔要求60个交易日以上，实际上是限制了规避频繁召开的临时股东大会，而只关注年会或者第一次临时大会的驱动效应；公告日最好是股价上涨；个股买入持有时间较长，20个交易日最佳。

图10：股东大会公告事件策略实证结果



数据来源：广发证券发展研究中心

表 5：股东大会公告事件策略实证结果

	跌	平	涨	全部
年化收益率	12.9%	12.4%	19.4%	13.7%
年化信息比	0.52	0.52	0.66	0.54
回撤	19.6%	15.6%	17.5%	16.1%
胜率	53.7%	54.6%	54.0%	53.8%

数据来源：广发证券发展研究中心

（三）增发公告

增发公告一般分为五大类，增发预案、增发说明书、增发获准、增发发行及增发上市，分别对应增发方案的不同阶段。实际上，定增方案的推出往往被市场解读为较大利空，尤其是定增方案确定之后，常常伴随着股价的下调，前期研究报告《广发证券金融工程：事件驱动量化选股周报(20140623)》中，我们提出了一类定增破发事件驱动策略。然而经实证发现，在满足一定条件下，定增事件也可能带来超额收益。

该类策略具体规则如下：

公告类型：增发方案相关公告披露；

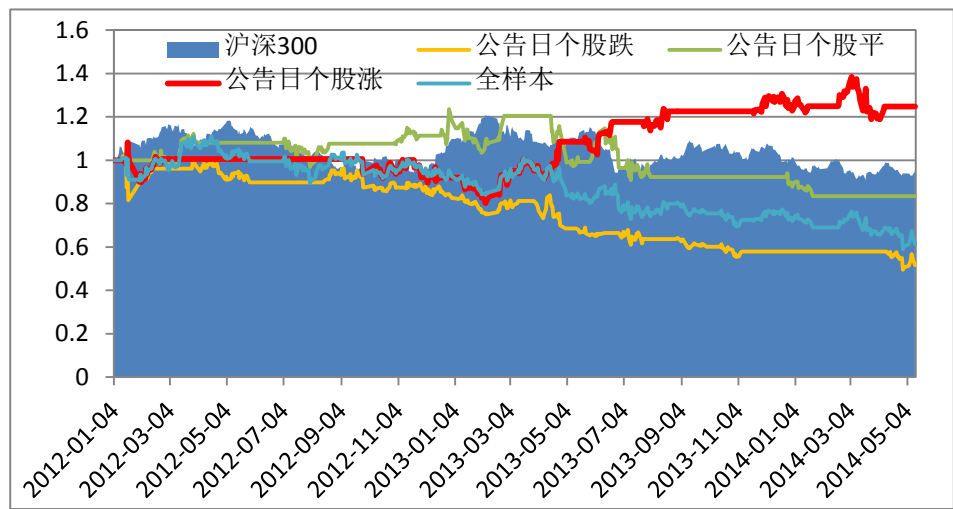
上次公告披露时间间隔要求：无要求；

公告披露日股价表现要求：严格要求个股大涨；

股票持有时间：20个交易日左右。

该类型公告对披露时间间隔无要求；但公告日要求股价必须大涨；个股买入持有时间较长，20个交易日左右最佳。

图11：增发公告事件策略实证结果



数据来源：广发证券发展研究中心

表 6：增发公告事件策略实证结果

	跌	平	涨	全部
年化收益率	-20.6%	-7.1%	10.5%	-16.4%
年化信息比	-0.89	-0.19	0.49	-0.47
回撤	51.9%	32.6%	26.0%	46.9%
胜率	51.8%	53.7%	53.6%	54.9%

数据来源：广发证券发展研究中心

(四) 重大积极事件公告

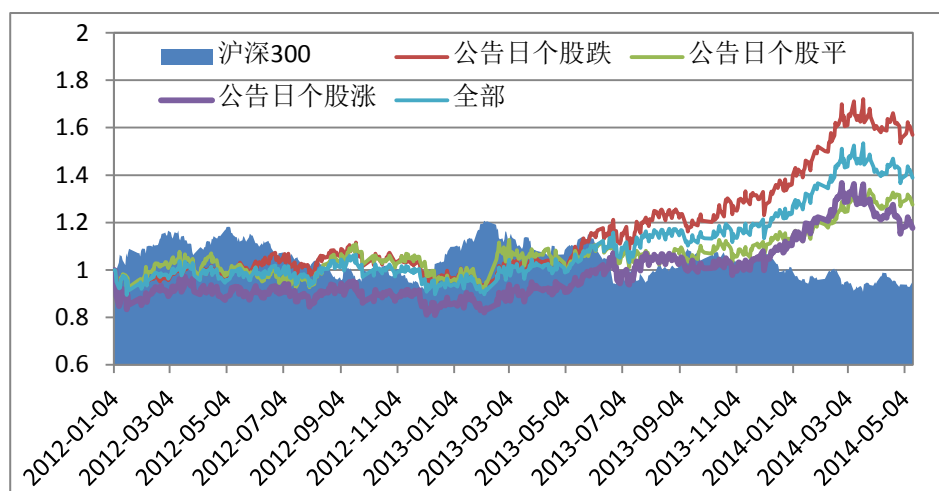
该类公告主要包含五大类，资产重组、收购兼并、重大合同、股权激励及政策影响等。其中“股权激励”公告事件我们在前期研究报告《广发证券金融工程：事件驱动量化选股周报(20140623)》中进行过单独测算，选股效果非常突出。

该类策略具体规则如下：

公告类型：五类重大事项公告披露；
上次公告披露时间间隔要求：无要求；
公告披露日股价表现要求：个股下跌；
股票持有时间：10个交易日左右。

该类型公告对披露时间间隔无要求；公告日若股价下跌则公告的后续利好效应更加凸显；个股买入持有时间10个交易日左右最佳。

图12：重大积极公告事件策略实证结果



数据来源：广发证券发展研究中心

表 7：重大积极公告事件策略实证结果

	跌	平	涨	全部
年化收益率	24.2%	11.7%	7.5%	16.5%
年化信息比	0.78	0.49	0.37	0.61
回撤	17.0%	17.7%	19.1%	15.7%
胜率	54.6%	53.6%	54.3%	54.7%

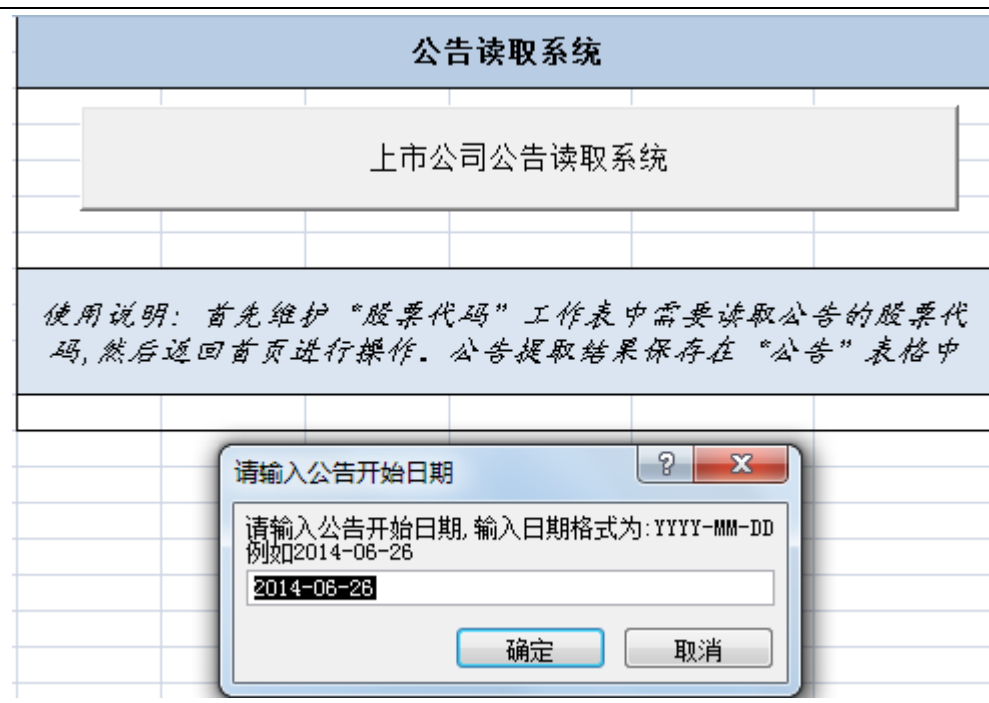
数据来源：广发证券发展研究中心

4.3 工具推介

基于对巨潮资讯公告披露平台以及深圳证券交易所信息披露平台的网页格式

的研究，我们开发出了上市公司公告读取系统，该系统能够批量读取上市公司任意时间段内的公告信息，并将与公告相关的信息，譬如公告标题、公告日期、公告下载地址等信息提取出来。目前该开发工具处于继续完善阶段，改善的方向有根据用户需求，选择是否将公告正文下载下来、根据提取出的公告标题直接将公告归类等

图13: A股公告抓取工具展示



数据来源: 广发证券发展研究中心

若您对该抓取工具感兴趣，欢迎来电与我们交流，或者发邮件向我们索取。

五、总结

在互联网大数据以及互联网金融时代背景下，信息量的快速增加为投资者进行投资决策提供了越来越多的信息。投资者关注的信息来源越来越丰富多样，信息的提供者与使用者的界限也越来越模糊。而影响投资决策的因素也越来越复杂，近年来，随着行为金融的发展，人们越来越关注非结构化的数据给投资者的投资决策带来的影响。在非结构化数据中，上市公司公告是其中一个组成部分。上市公司所披露的某些公告常常隐藏着许多关键信息，从而带给投资者巨大的想象空间，导致股价大幅波动。

那么什么类型的公告披露之后更加容易带来股价的上涨？由于公告类型较多，将其分为八大类进行一一测算，在区分公告类型的同时，也将结合公告披露当日个股股价的表现以及上市公司多长时间内未披露过该类公告一起来对事件进行区分和定义，分析公告披露背后所隐藏的投资机会。历史回测结果表明，财务报表、股东大会、增发公告以及重大积极事项等公告策略的选股效果相当突出，均能获得较高的超额收益。

风险提示

本模型从公告类型、公告披露时间间隔以及披露当天股价表现等角度来构建事件选股策略，属于纯量化的研究方法，所推荐的个股未必具有实质性的利好，其股价表现还受到更多因素影响，请结合基本面及自身判断进行恰当使用。

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	广州市天河北路 183 号 大都会广场 5 楼	深圳市福田区金田路 4018 号安联大厦 15 楼 A 座 03-04	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	510075	518026	100045	200120
客服邮箱	gfyf@gf.com.cn			
服务热线	020-87555888-8612			

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。