

## 基于深层次文本挖掘的策略研究——概念板块事件驱动套利策略一

金融工程

### ◆概念板块套利：基于文本挖掘的事件驱动策略

众所周知，财经新闻、股票论坛、股票研究报告等媒介信息会对投资者的投资情绪以及投资策略起到潜移默化的影响，这些文本信息中蕴含了大量未被利用的信息，如果能够合理地利用这些信息，将为我们理解市场提供一个全新的角度。基于这些文本信息，我们开发了光大关注度因子、普通投资者情绪指标、光大机构多空指标、经济担忧指标等系列产品，取得了显著的成果。而如何更深层次地应用这些信息，以及如何进一步地去量化这些文本信息，生成有效的投资策略，将成为我们一个重要的量化研究方向。在此系列报告中，我们将依托于光大金工强大的中文处理平台和丰富准确的历史数据积累，针对这些文本信息进行一系列的尝试性研究，并希望从中获取一定的信息回报。

在此报告中，我们开发基于概念热度的事件驱动套利策略。具体我们以周为单位，回测时间为 2010-09-01 至 2012-08-17，在此期间内，所有概念板块的平均收益率为-30%，而我们针对于各概念板块的择时策略按不同数据源考虑做空后平均收益为 38%-40%。在此基础上构造的概念板块事件驱动套利策略在同期大盘收益率仅为-20%的情况下取得了近 80%的绝对收益率，最大回撤则为约 15%。

### ◆概念板块轮动的基本统计规律：关注冲激与股价位置的双重信号

在概念轮动中，概念经历产生，活跃，静寂再活跃再寂静如此反复，而一个概念无论是明显上升还是下跌均需要足够的关注量。通过光大中文文本挖掘系统，我们为每个概念板块建立关键词列表，通过这些关键词列表则可以时刻关注着这些概念板块在财经新闻、股票论坛、股票研究报告等中出现的热度，并以此构造了光大概念热度指数。

统计表明，当概念热度指数出现尖峰（表明此时该概念板块受到大量的关注）后，概念板块会有中短期趋势改变，方向以概念板块指数所在高低位进行判断。具体的，如果以周为单位，当概念热度指数出现短期高峰，如果该指数价格处于短期低位，则产生做多信号，处于高位则产生做空信号，无论做多还是做空，产生后的前一周收益率均为负，此后转为正并逐渐上升，做多信号至第七周达到最大，做空信号则至第九周达到峰值，综合来看，在第八周达到峰值。

### ◆套利组合：两年累积收益 80%：

我们通过如下方法构造套利策略：每周对各概念指数进行过滤，触发做多信号的概念在一周后进入多头持仓名单，触发做空信号则进入空头持仓名单，每个概念指数在没有产生反向信号之前持仓六周后自动平仓，如六周内产生反向信号则同样于一周后反向。多数情况下，投资组合同时持有多头与空头仓位，对冲了大部分的市场风险，回溯测试结果也表明这是一个收益较好，风险中低的策略：近两年累积收益 80%，最大回撤 15%。

### 分析师

倪蕴韬 (执业证书编号：S0930512070002)  
021-22169338  
[niyt@ebsecn.com](mailto:niyt@ebsecn.com)

张斯会 (执业证书编号：S0930512020001)  
021-22169104  
[zhangsihui@ebsecn.com](mailto:zhangsihui@ebsecn.com)

### 联系人

吴晶  
021-22167203  
[wujing@ebsecn.com](mailto:wujing@ebsecn.com)

### 相关研报

互联网数据挖掘系统，行为金融新领域 2010-09  
量化择时之三——情绪拐点择时 2012-02

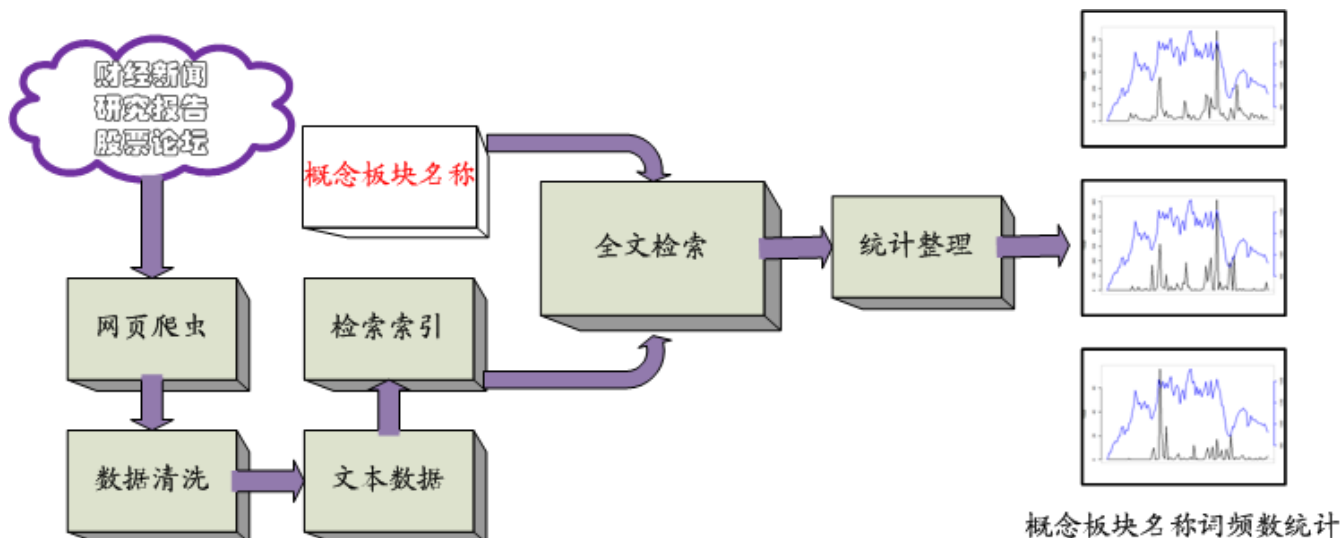
## 图表目录

图表 1: 概念热度指数应用系统架构 .....	3
图表 2: 基于研究报告文本的“石墨烯”概念热度指数与石墨烯板块指数.....	3
图表 3: 基于财经新闻的“石墨烯”概念热度指数与石墨烯板块指数.....	3
图表 4: 基于股票论坛文本的“石墨烯”概念热度指数与石墨烯板块指数.....	4
图表 5: 事件信号表现（基于研究报告摘要文本） .....	5
图表 6: 事件表现（基于财经新闻文本） .....	6
图表 7: 事件表现（基于股票论坛帖子文本） .....	6
图表 8: 随着持有期的持续增长，事件信号胜率变化情况。 .....	7
图表 9 随着持有期的持续增长，各信号平均收益率的变化情况。 .....	7
图表 10: 持有期为 6 周时，各概念板块择时策略表现（基于研究报告摘要文本） .....	8
图表 11: 持有期为 6 周时，各概念板块择策略表现（基于财经新闻文本） .....	10
图表 12: 持有期为 6 周时，各概念板块择时策略表现（基于股票论坛帖子文本） .....	11
图表 13: 基于研究报告文本的事件驱动套利策略表现 .....	13
图表 14: 基于财经新闻文本的的事件驱动套利策略表现 .....	14
图表 15: 基于股票论坛文本的的事件驱动套利策略表现 .....	14

## 1、光大概念热度指数

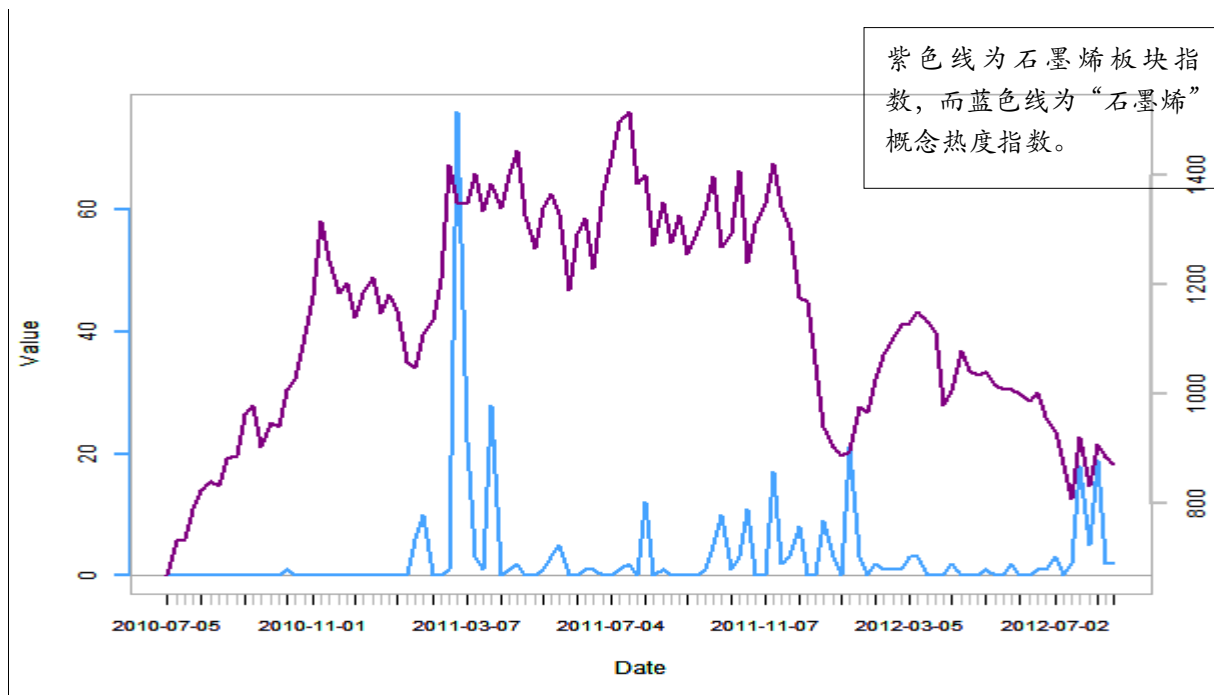
我们通过网络抓取技术采集了2010年6月至2012年7月的约36万篇研究报告摘要、300万篇财经新闻、2500万条论坛帖子分别构造我们的文本资料库。而概念板块数据主要获取于Wind标准的概念板块分类，具体为石墨烯、锂电池、智能交通、稀土永磁、舟山新区等74个概念板块。针对于每一个概念板块，我们制定了与该板块相关联的一系列关键词列表，然后分别在研究报告、财经新闻、论坛帖子等不同的文本资料库中搜索该关键词组合，最后按时间统计搜索结果，构造基于各文本资料的概念热度指数

图表 1：概念热度指数应用系统架构



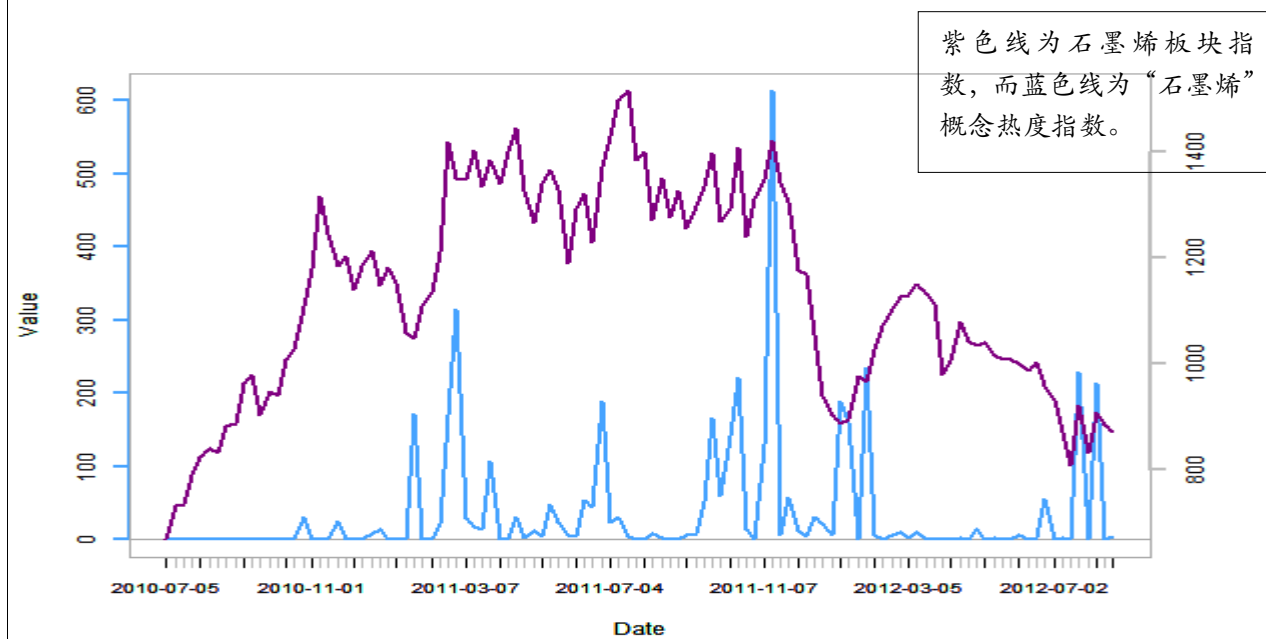
数据来源：光大金工中文云 [www.chinesecloud.net](http://www.chinesecloud.net)

图表 2：基于研究报告文本的“石墨烯”概念热度指数与石墨烯板块指数



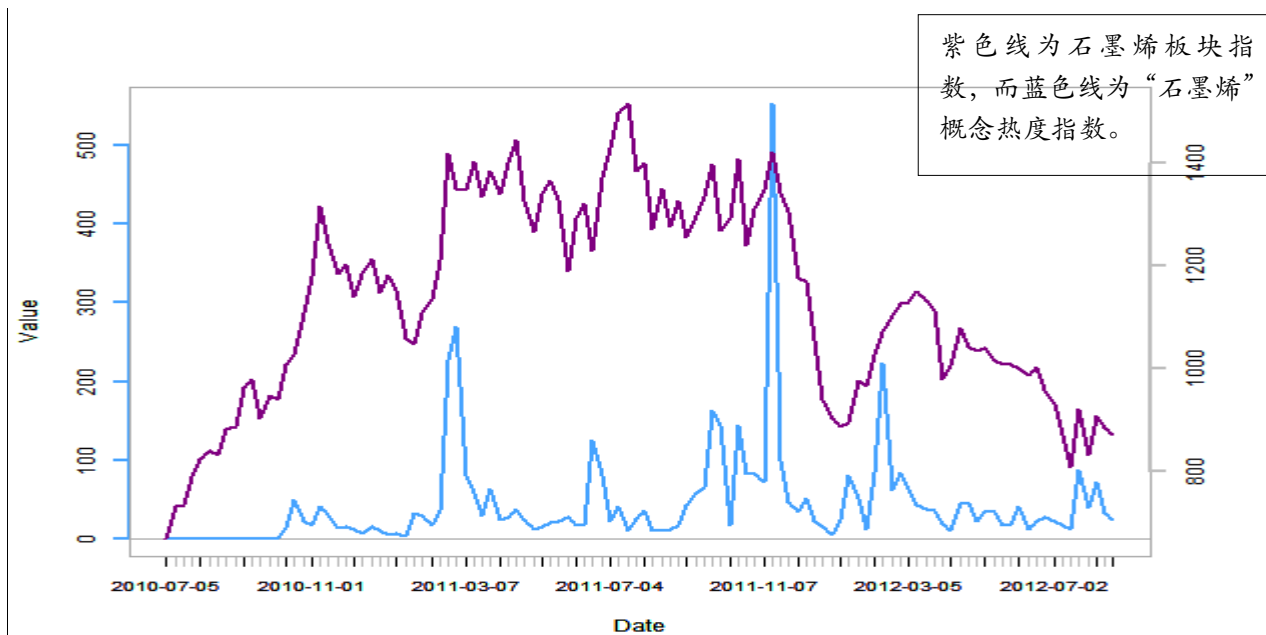
数据来源：WIND 资讯，光大金工中文云 [www.chinesecloud.net](http://www.chinesecloud.net)

图表 3：基于财经新闻的“石墨烯”概念热度指数与石墨烯板块指数



数据来源：WIND 资讯，光大金工中文云 [www.chinesecloud.net](http://www.chinesecloud.net)

图表 4：基于股票论坛文本的“石墨烯”概念热度指数与石墨烯板块指数



数据来源：WIND 资讯，光大金工中文云 [www.chinesecloud.net](http://www.chinesecloud.net)

### 3、概念板块事件驱动

在此模型中，我们基于的假设是：某个概念板块名称在各类媒介信息中出现的次数急剧上升，是由于该概念板块有突发性事件发生，该突发性事件导致了该概念板块名称频繁的出现，该板块名称频繁的出现传递给越来越多的投资者，这将会影响到一部分投资者的投资策略，一些投资者会考虑有一定的机会而持续跟进，另一些投资者则会感觉到一定的风险而选择退出。那么当这种事件发生时，作为投资者到底应该如何处理呢？

通过历史数据测试，我们发现当概念指数处于区域底部时，若概念板块名称大量地出现在媒介信息中，即概念热度指数达到一个高位时，会导致该概念

板块在未来的一段时间里保持一定的强势，并且随着持有时间的增加，胜率会逐渐增大。而当概念指数处于区域顶部时，会导致该概念板块在未来的一段时间内保持一定的弱势，并且随着持有时间的增加，胜率也在逐渐增大。而且在两种情况下，均是在持有时间为 8 周时收益达到最优。

为了具体量化概念热度指数达到一个高位这一现象，我们以其过去 10 周的均值加 0.5 倍标准差构造其上轨线，若在某一周里概念热度指数超过其上轨线，则我们认为满足热度指数达到一个高位这一现象。同样，以各概念指数过去 10 周的均值加减 1 倍标准差构造上下轨线，若概念指数位于其上轨线之上，则我们认为该概念板块指数处于高位；若概念指数位于其下轨线之下，则我们认为该概念板块指数处于低位。当概念热度指数达到一个高位，并且该概念指数处于低位时，那么将在下一个交易周中触发做多事件信号。当概念热度指数达到一个高位，并且该概念指数处于高位时，那么将在下一个交易周中触发做空事件信号。

#### 4、事件信号结论分析

根据上述产生的事件信号，分别对各概念板块进行测试，时间窗口为 2010 年 9 月至 2012 年 8 月。在此时间段内，基于研究报告摘要文本的概念板块事件信号共产生 417 次做多信号，408 次做空信号；基于财经新闻文本的概念板块事件信号共产生 440 次做多信号，348 次做空信号；基于论坛帖子文本的概念板块事件信号共产生 340 次做多信号，453 次做空信号。这里，我们主要考虑的是当事件信号产生后立即进行交易，则不同的持有周期下的事件信号的收益情况，具体分析如下：

图表 5：事件信号表现（基于研究报告摘要文本）

持有期	信号	次数	获胜	胜率	平均回报	最大回报	最差回报
1 周	做多	416	177	42.55%	-0.42%	10.78%	-12.10%
2 周	做多	416	193	46.39%	0.06%	26.02%	-20.45%
3 周	做多	416	217	52.16%	0.94%	23.19%	-25.11%
4 周	做多	416	239	57.45%	1.62%	28.56%	-24.42%
5 周	做多	416	247	59.38%	2.95%	30.47%	-19.30%
6 周	做多	416	264	63.46%	3.55%	33.49%	-22.01%
7 周	做多	416	258	62.02%	3.86%	32.54%	-25.30%
8 周	做多	416	232	55.77%	3.05%	30.30%	-31.63%
1 周	做空	408	202	49.51%	0.00%	14.92%	-37.73%
2 周	做空	408	240	58.82%	1.14%	13.40%	-28.43%
3 周	做空	408	276	67.65%	2.37%	20.55%	-31.71%
4 周	做空	408	287	70.34%	2.93%	25.67%	-24.93%
5 周	做空	408	292	71.57%	3.64%	40.20%	-29.47%
6 周	做空	408	306	75.00%	4.67%	48.42%	-34.64%
7 周	做空	408	308	75.49%	6.03%	47.04%	-45.79%
8 周	做空	408	336	82.35%	7.88%	38.42%	-47.90%
1 周	做多\做空	824	379	46.00%	-0.21%	14.92%	-37.73%
2 周	做多\做空	824	433	52.55%	0.59%	26.02%	-28.43%
3 周	做多\做空	824	493	59.83%	1.65%	23.19%	-31.71%
4 周	做多\做空	824	526	63.83%	2.27%	28.56%	-24.93%
5 周	做多\做空	824	539	65.41%	3.29%	40.20%	-29.47%
6 周	做多\做空	824	570	69.17%	4.10%	48.42%	-34.64%

7 周	做多\做空	824	566	68.69%	4.93%	47.04%	-45.79%
8 周	做多\做空	824	568	68.93%	5.44%	38.42%	-47.90%

数据来源：WIND 资讯，光大金工中文云 [www.chinesecloud.net](http://www.chinesecloud.net)

图表 6：事件表现（基于财经新闻文本）

持有期	信号	次数	获胜	胜率	平均回报	最大回报	最差回报
1 周	做多	438	193	44.06%	-0.63%	10.68%	-12.77%
2 周	做多	438	184	42.01%	-0.44%	26.02%	-18.94%
3 周	做多	438	213	48.63%	0.12%	23.06%	-19.82%
4 周	做多	438	229	52.28%	1.01%	24.81%	-22.98%
5 周	做多	438	245	55.94%	2.04%	25.61%	-19.69%
6 周	做多	438	260	59.36%	2.39%	31.64%	-22.01%
7 周	做多	438	256	58.45%	2.55%	29.94%	-19.07%
8 周	做多	438	232	52.97%	1.74%	32.46%	-38.87%
1 周	做空	348	162	46.55%	-0.30%	14.92%	-37.73%
2 周	做空	348	205	58.91%	1.04%	13.40%	-40.16%
3 周	做空	348	233	66.95%	2.00%	20.55%	-31.71%
4 周	做空	348	223	64.08%	2.51%	25.67%	-34.38%
5 周	做空	348	242	69.54%	3.66%	40.20%	-29.47%
6 周	做空	348	259	74.43%	4.89%	48.42%	-42.52%
7 周	做空	348	271	77.87%	6.51%	47.04%	-45.79%
8 周	做空	348	293	84.20%	8.37%	38.42%	-40.95%
1 周	做多\做空	786	355	45.17%	-0.48%	14.92%	-37.73%
2 周	做多\做空	786	389	49.49%	0.21%	26.02%	-40.16%
3 周	做多\做空	786	446	56.74%	0.95%	23.06%	-31.71%
4 周	做多\做空	786	452	57.51%	1.67%	25.67%	-34.38%
5 周	做多\做空	786	487	61.96%	2.76%	40.20%	-29.47%
6 周	做多\做空	786	519	66.03%	3.50%	48.42%	-42.52%
7 周	做多\做空	786	527	67.05%	4.30%	47.04%	-45.79%
8 周	做多\做空	786	525	66.79%	4.68%	38.42%	-40.95%

数据来源：WIND 资讯，光大金工中文云 [www.chinesecloud.net](http://www.chinesecloud.net)

图表 7：事件表现（基于股票论坛帖子文本）

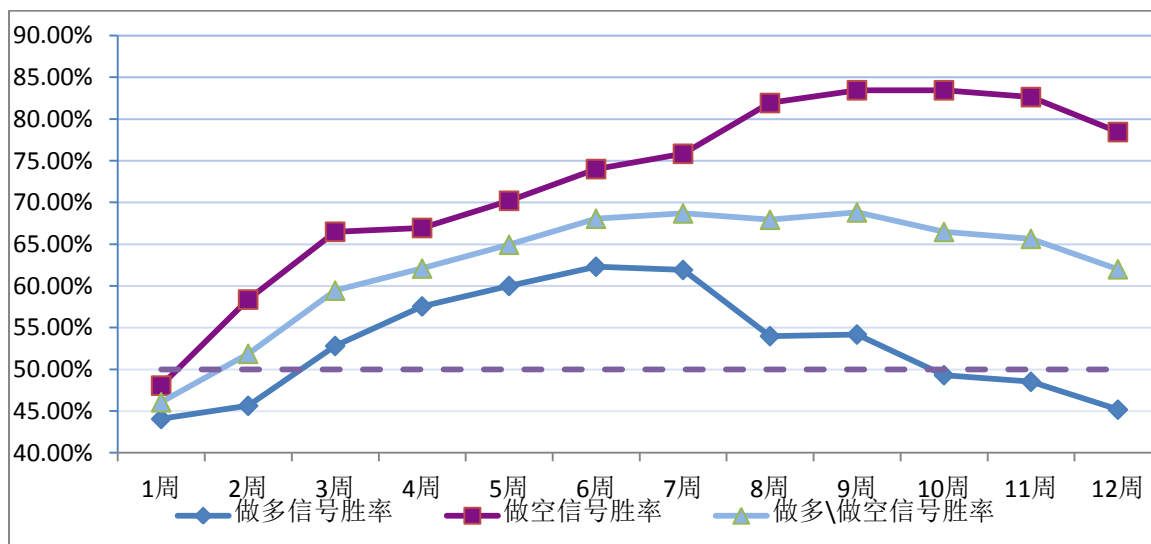
持有期	信号	次数	获胜	胜率	平均回报	最大回报	最差回报
1 周	做多	340	155	45.59%	-0.33%	16.67%	-11.34%
2 周	做多	340	165	48.53%	0.11%	26.02%	-14.22%
3 周	做多	340	196	57.65%	1.49%	23.82%	-15.91%
4 周	做多	340	214	62.94%	2.53%	27.50%	-19.04%
5 周	做多	340	220	64.71%	3.71%	28.92%	-17.08%
6 周	做多	340	218	64.12%	3.67%	30.21%	-22.01%
7 周	做多	340	222	65.29%	3.46%	27.47%	-21.59%
8 周	做多	340	181	53.24%	2.28%	28.81%	-38.87%
1 周	做空	453	218	48.12%	-0.02%	14.92%	-37.73%



2 周	做空	453	260	57.40%	0.94%	13.40%	-40.16%
3 周	做空	453	294	64.90%	1.69%	20.55%	-31.71%
4 周	做空	453	301	66.45%	2.45%	25.67%	-34.38%
5 周	做空	453	315	69.54%	3.09%	40.20%	-46.59%
6 周	做空	453	329	72.63%	4.21%	48.42%	-38.62%
7 周	做空	453	336	74.17%	5.82%	47.04%	-41.43%
8 周	做空	453	359	79.25%	7.34%	38.42%	-47.90%
1 周	做多\做空	793	373	47.04%	-0.15%	16.67%	-37.73%
2 周	做多\做空	793	425	53.59%	0.59%	26.02%	-40.16%
3 周	做多\做空	793	490	61.79%	1.60%	23.82%	-31.71%
4 周	做多\做空	793	515	64.94%	2.48%	27.50%	-34.38%
5 周	做多\做空	793	535	67.47%	3.36%	40.20%	-46.59%
6 周	做多\做空	793	547	68.98%	3.98%	48.42%	-38.62%
7 周	做多\做空	793	558	70.37%	4.81%	47.04%	-41.43%
8 周	做多\做空	793	540	68.10%	5.17%	38.42%	-47.90%

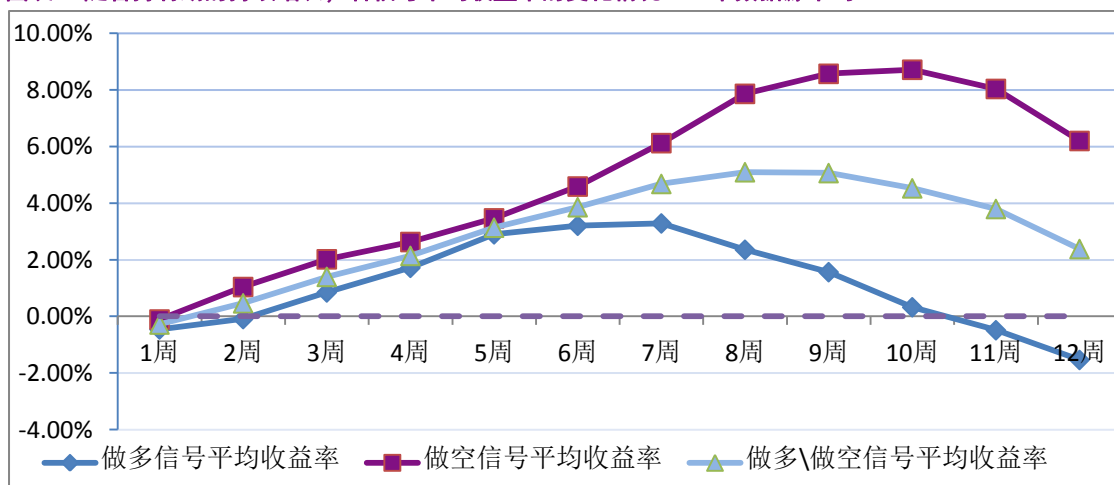
数据来源：WIND 资讯，光大金工中文云 [www.chinesecloud.net](http://www.chinesecloud.net)

图表 8：随着持有期的持续增长，事件信号胜率变化情况（三个数据源平均）



数据来源：WIND 资讯，光大金工中文云 [www.chinesecloud.net](http://www.chinesecloud.net)

图表 9 随着持有期的持续增长，各信号平均收益率的变化情况（三个数据源平均）



数据来源：WIND 资讯，光大金工中文云 [www.chinesecloud.net](http://www.chinesecloud.net)

从上图中，我们可以看出，当上述事件信号触发后，在随后的两周内，无论是做多还是做空，我们的信号收益率均是小于0的。而且做多做空信号的收益率在持有期为8周的时候达到极大值。那么我们将事件信号滞后两周得到交易信号，并设置当信号产生后，我们持有各概念指数的期限为6周。

## 5、交易信号结论分析

在按照上述方式产生交易信号后，对于单个概念板块的择时策略我们简单的设置为当在持有期若产生了一相反操作的信号，则我们进行相应的换仓操作。

在该策略下，基于三个不同的文本资料源的各概念板块的择时策略分别能够获取38%、39.4%和37.8%的平均收益率，而同期，所有概念板块的平均收益率仅为-30%。

图表 10：持有期为6周时，各概念板块择时策略表现（基于研究报告摘要文本）

概念板块	期间内该概念板块指数	择时策略收益 (做多\做空)	持有周数	获取正收益周数	信号产生次数	信号获取正收益次数
石墨烯指数	-25.36%	39.63%	39	21	11	8
锂电池指数	-44.12%	24.11%	44	26	13	9
智能交通指数	-25.67%	-5.89%	44	24	10	4
稀土永磁指数	-12.46%	82.75%	52	33	18	14
舟山新区指数	-23.83%	14.83%	45	25	13	7
网络游戏指数	-36.00%	39.29%	24	17	7	4
新疆区域振兴指数	-45.25%	58.32%	36	28	9	8
高铁指数	-42.16%	-5.90%	41	19	9	4
安防监控指数	-27.97%	12.73%	36	19	11	7
核高基指数	-31.64%	29.25%	34	19	8	7
水利水电建设指数	-20.79%	30.57%	54	28	18	10
苹果指数	3.57%	53.88%	52	27	17	12
文化传媒指数	-12.14%	65.54%	48	33	15	12
IPv6 指数	-31.25%	17.91%	42	25	11	7
概念板块龙头指数	-39.24%	0.00%	0	0	0	0
云计算指数	-35.93%	34.13%	46	27	10	8
港珠澳大桥指数	-16.53%	38.65%	47	30	12	7
分拆上市指数	-33.59%	40.80%	37	23	9	8
三网融合指数	-34.33%	85.27%	43	28	9	7
股权激励指数	-23.37%	40.38%	44	25	14	10
新材料指数	-33.69%	15.58%	29	16	8	7
宽带提速指数	-31.95%	59.70%	40	24	12	9
中日韩自贸区指数	-29.07%	32.40%	20	15	5	4
煤层气页岩气指数	4.80%	23.28%	43	28	12	7
绿色节能照明指数	-43.73%	4.89%	40	22	12	9
合同能源管理指数	-49.71%	33.71%	41	23	7	5
上海本地重组指数	-27.86%	42.64%	52	30	14	10
3G 指数	-33.07%	56.47%	57	34	15	10
移动互联网指数	-38.97%	32.86%	42	24	11	6
环保指数	-15.36%	132.71%	56	37	12	9
建筑节能指数	-44.08%	58.61%	55	35	13	10
温州金融改革指数	-14.20%	12.07%	57	33	18	14



海峡西岸指数	-22.81%	32.74%	38	22	10	7
移动支付指数	-46.43%	21.20%	43	25	11	8
核能核电指数	-34.20%	53.81%	31	21	10	7
抗癌指数	-25.08%	53.35%	22	17	7	6
物联网指数	-36.09%	44.91%	38	25	11	8
皖江城市带指数	-24.35%	41.44%	27	19	7	6
太阳能发电指数	-44.60%	55.14%	56	38	14	8
新能源汽车指数	-42.32%	49.57%	32	22	9	6
节能元器件 IGBT 指数	-40.65%	42.89%	40	26	13	11
卫星导航指数	-42.93%	-10.04%	38	17	8	2
新三板指数	-21.24%	83.96%	48	29	15	14
装饰园林指数	-9.56%	65.95%	43	27	11	7
外资并购国际板指数	-20.95%	20.24%	34	21	8	4
氢能指数	-25.58%	44.69%	46	30	14	9
高端装备制造指数	-36.90%	-23.38%	51	22	10	4
钢结构指数	-3.60%	-36.30%	55	25	16	6
垃圾发电指数	-29.39%	27.25%	38	24	7	6
通用航空指数	-51.31%	20.37%	23	14	6	4
涉矿指数	-46.51%	25.90%	49	30	13	11
新能源指数	-43.09%	39.84%	36	25	11	9
央企重组指数	-32.32%	55.65%	50	33	11	8
浦东新区指数	-26.57%	23.79%	49	30	12	8
参股金融指数	-27.41%	26.42%	32	18	8	7
成渝特区指数	-33.06%	0.66%	21	12	6	4
航母指数	-39.14%	29.27%	42	26	10	8
充电桩指数	-38.62%	94.85%	32	23	10	8
深圳前海新区指数	-26.30%	7.00%	38	20	8	3
生物育种指数	-33.57%	47.32%	46	27	11	8
保障性住房指数	-25.76%	35.35%	36	25	13	9
数字电视指数	-35.13%	75.59%	41	29	11	8
触摸屏指数	-30.36%	33.72%	40	23	11	7
清洁煤炭发电指数	-30.19%	45.20%	49	29	15	13
风力发电指数	-37.79%	19.20%	48	28	13	9
智能电网指数	-44.65%	51.96%	26	17	7	6
新型煤化工指数	-25.93%	59.57%	54	34	18	16
迪士尼指数	-35.66%	51.21%	47	32	12	10
黄金珠宝指数	-19.77%	77.46%	57	37	16	13
滨海新区指数	-33.96%	11.38%	51	30	14	10
乙醇汽油指数	-16.87%	100.80%	54	33	16	11
海南旅游岛指数	-30.46%	30.83%	24	16	6	4
北部湾经济区指数	-40.75%	54.25%	48	31	14	10
三沙指数	-26.11%	32.30%	11	7	3	3

数据来源：WIND 资讯，光大金工中文云 [www.chinesecloud.net](http://www.chinesecloud.net)

图表 11：持有期为 6 周时，各概念板块择策略表现（基于财经新闻文本）

概念板块	期间内该概念板块指数	择时策略收益 (做多\做空)	持有周数	获取正收益周数	信号产生次数	信号获取正收益次数
石墨烯指数	-25.36%	82.86%	38	20	8	7
锂电池指数	-44.12%	63.26%	39	25	14	10
智能交通指数	-25.67%	91.00%	33	21	8	7
稀土永磁指数	-12.46%	48.73%	52	31	18	12
舟山新区指数	-23.83%	6.38%	43	24	10	5
网络游戏指数	-36.00%	31.22%	43	22	8	4
新疆区域振兴指数	-45.25%	50.97%	41	28	10	7
高铁指数	-42.16%	36.70%	38	23	13	6
安防监控指数	-27.97%	18.52%	47	27	14	8
核高基指数	-31.64%	14.02%	46	25	10	5
水利水电建设指数	-20.79%	31.11%	54	29	17	11
苹果指数	3.57%	172.69%	62	39	21	16
文化传媒指数	-12.14%	9.57%	40	25	14	11
IPV6 指数	-31.25%	56.45%	34	20	6	4
概念板块龙头指数	-39.24%	0.00%	0	0	0	0
云计算指数	-35.93%	78.12%	45	28	12	8
港珠澳大桥指数	-16.53%	44.44%	46	28	11	9
分拆上市指数	-33.59%	47.41%	31	19	9	8
三网融合指数	-34.33%	55.37%	49	28	17	12
股权激励指数	-23.37%	20.35%	40	22	12	8
新材料指数	-33.69%	49.85%	44	26	15	11
宽带提速指数	-31.95%	9.36%	25	13	5	4
中日韩自贸区指数	-29.07%	38.82%	24	18	7	6
煤层气页岩气指数	4.80%	47.22%	40	27	10	9
绿色节能照明指数	-43.73%	42.23%	23	16	4	3
合同能源管理指数	-49.71%	23.47%	50	29	16	13
上海本地重组指数	-27.86%	34.12%	44	24	9	7
3G 指数	-33.07%	74.87%	55	31	21	15
移动互联网指数	-38.97%	60.14%	45	24	11	7
环保指数	-15.36%	73.08%	47	28	18	12
建筑节能指数	-44.08%	13.41%	40	22	7	3
温州金融改革指数	-14.20%	36.11%	45	24	11	9
海峡西岸指数	-22.81%	19.39%	41	22	10	7
移动支付指数	-46.43%	39.13%	53	32	14	9
核能核电指数	-34.20%	73.99%	46	30	11	8
抗癌指数	-25.08%	57.08%	27	20	7	7
物联网指数	-36.09%	61.23%	44	28	12	10
皖江城市带指数	-24.35%	33.91%	25	16	7	4
太阳能发电指数	-44.60%	27.95%	50	29	16	9
新能源汽车指数	-42.32%	40.18%	29	19	6	5
节能元器件 IGBT 指数	-40.65%	46.11%	24	16	6	6
卫星导航指数	-42.93%	10.13%	26	11	5	2
新三板指数	-21.24%	84.83%	43	25	13	12
装饰园林指数	-9.56%	48.33%	62	35	21	15

外资并购国际板指数	-20.95%	57.22%	34	22	8	6
氢能指数	-25.58%	23.64%	22	14	6	5
高端装备制造指数	-36.90%	-19.61%	41	20	10	3
钢结构指数	-3.60%	-30.00%	51	22	17	7
垃圾发电指数	-29.39%	48.68%	37	24	8	6
通用航空指数	-51.31%	34.12%	51	29	10	5
涉矿指数	-46.51%	43.27%	39	26	9	5
新能源指数	-43.09%	14.62%	36	18	13	7
央企重组指数	-32.32%	42.30%	31	21	6	5
浦东新区指数	-26.57%	27.72%	37	22	13	9
参股金融指数	-27.41%	32.36%	33	18	8	7
成渝特区指数	-33.06%	13.82%	18	11	3	2
航母指数	-39.14%	49.48%	35	22	10	7
充电桩指数	-38.62%	81.74%	38	25	8	4
深圳前海新区指数	-26.30%	6.56%	14	9	3	2
生物育种指数	-33.57%	19.67%	43	25	8	5
保障性住房指数	-25.76%	6.08%	46	26	17	11
数字电视指数	-35.13%	26.48%	60	33	16	10
触摸屏指数	-30.36%	25.04%	44	24	15	10
清洁煤发电指数	-30.19%	31.14%	20	13	4	4
风力发电指数	-37.79%	-0.97%	32	16	8	5
智能电网指数	-44.65%	39.44%	35	21	14	9
新型煤化工指数	-25.93%	36.32%	45	27	12	12
迪士尼指数	-35.66%	35.35%	37	24	8	7
黄金珠宝指数	-19.77%	64.80%	36	23	12	11
滨海新区指数	-33.96%	4.37%	39	25	8	5
乙醇汽油指数	-16.87%	63.23%	38	25	9	8
海南旅游岛指数	-30.46%	49.66%	29	20	6	5
北部湾经济区指数	-40.75%	43.15%	59	37	19	11
三沙指数	-26.11%	47.13%	17	12	4	4

数据来源：WIND 资讯，光大金工中文云 [www.chinesecloud.net](http://www.chinesecloud.net)

图表 12：持有期为 6 周时，各概念板块择时策略表现（基于股票论坛帖子文本）

概念板块	期间内该概念板块指数	择时策略收益 (做多\做空)	持有周数	获取正收益周数	信号产生次数	信号获取正收益次数
石墨烯指数	-25.36%	28.47%	40	21	12	7
锂电池指数	-44.12%	39.84%	40	25	12	9
智能交通指数	-25.67%	12.21%	27	14	6	5
稀土永磁指数	-12.46%	31.17%	47	28	17	12
舟山新区指数	-23.83%	36.72%	31	21	10	6
网络游戏指数	-36.00%	37.57%	35	21	8	8
新疆区域振兴指数	-45.25%	54.53%	46	32	13	11
高铁指数	-42.16%	-4.78%	32	16	8	4
安防监控指数	-27.97%	18.74%	48	27	13	10
核高基指数	-31.64%	53.02%	52	30	12	8
水利水电建设指数	-20.79%	23.20%	54	28	19	11

苹果指数	3.57%	73.08%	54	30	21	15
文化传媒指数	-12.14%	62.89%	32	20	10	8
IPV6 指数	-31.25%	81.47%	42	28	9	7
概念板块龙头指数	-39.24%	17.46%	18	12	3	2
云计算指数	-35.93%	48.14%	44	28	13	12
港珠澳大桥指数	-16.53%	9.61%	52	28	15	10
分拆上市指数	-33.59%	26.46%	37	23	9	7
三网融合指数	-34.33%	26.89%	37	23	11	7
股权激励指数	-23.37%	24.10%	51	29	15	11
新材料指数	-33.69%	23.41%	47	26	13	11
宽带提速指数	-31.95%	25.49%	19	12	4	3
中日韩自贸区指数	-29.07%	43.47%	19	15	4	3
煤层气页岩气指数	4.80%	63.27%	50	33	15	10
绿色节能照明指数	-43.73%	0.70%	25	15	6	5
合同能源管理指数	-49.71%	99.79%	41	27	8	7
上海本地重组指数	-27.86%	24.53%	45	26	14	9
3G 指数	-33.07%	40.31%	38	24	10	7
移动互联网指数	-38.97%	75.70%	57	34	13	9
环保指数	-15.36%	140.78%	48	30	15	13
建筑节能指数	-44.08%	62.60%	42	29	10	8
温州金融改革指数	-14.20%	-6.00%	40	22	12	7
海峡西岸指数	-22.81%	89.00%	51	36	12	10
移动支付指数	-46.43%	40.31%	50	30	12	8
核能核电指数	-34.20%	41.17%	43	26	13	9
抗癌指数	-25.08%	40.42%	31	18	11	9
物联网指数	-36.09%	76.36%	31	23	11	10
皖江城市带指数	-24.35%	30.61%	29	20	7	6
太阳能发电指数	-44.60%	8.37%	48	28	15	10
新能源汽车指数	-42.32%	124.38%	49	37	12	10
节能元器件 IGBT 指数	-40.65%	37.18%	46	30	15	13
卫星导航指数	-42.93%	45.77%	39	26	9	7
新三板指数	-21.24%	33.43%	53	27	16	11
装饰园林指数	-9.56%	27.36%	30	20	10	7
外资并购国际板指数	-20.95%	24.13%	27	17	7	5
氢能指数	-25.58%	-3.44%	26	16	5	3
高端装备制造指数	-36.90%	32.84%	41	22	10	7
钢结构指数	-3.60%	-37.68%	29	9	10	3
垃圾发电指数	-29.39%	65.73%	39	27	7	5
通用航空指数	-51.31%	32.51%	33	21	7	4
涉矿指数	-46.51%	-13.00%	44	21	13	6
新能源指数	-43.09%	27.46%	41	26	13	11
央企重组指数	-32.32%	40.93%	26	16	8	6
浦东新区指数	-26.57%	18.13%	49	28	15	9
参股金融指数	-27.41%	40.35%	49	29	11	9
成渝特区指数	-33.06%	28.68%	42	27	10	6
航母指数	-39.14%	25.33%	49	28	11	9
充电桩指数	-38.62%	79.19%	32	24	8	7
深圳前海新区指数	-26.30%	15.29%	28	16	6	4

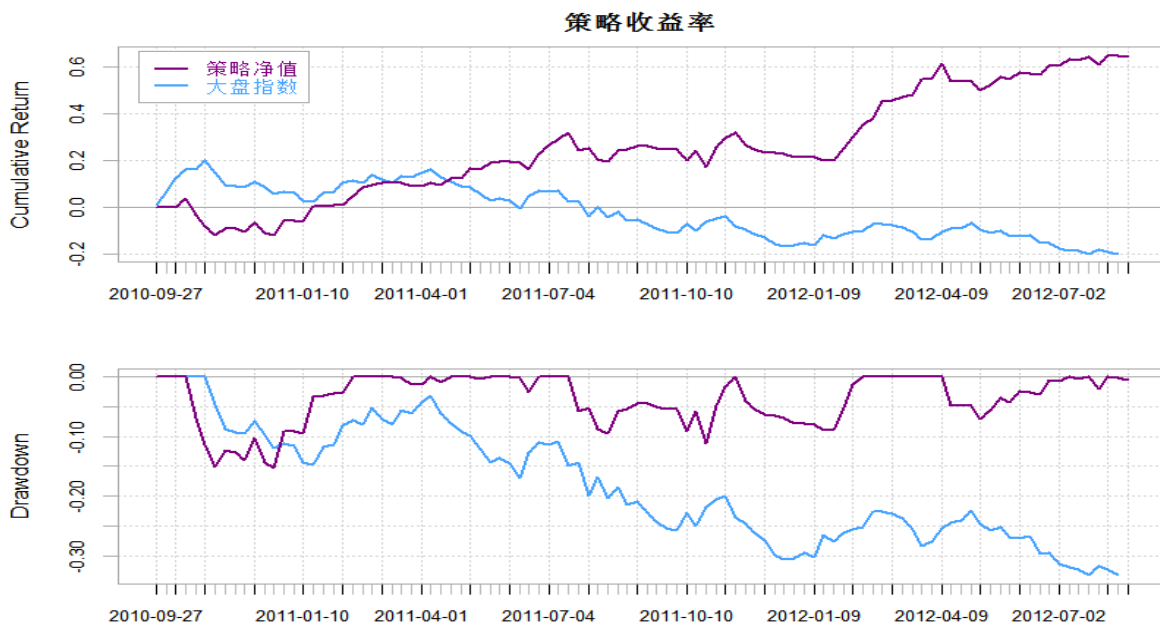
生物育种指数	-33.57%	-30.40%	52	25	11	3
保障性住房指数	-25.76%	30.56%	43	28	12	7
数字电视指数	-35.13%	16.94%	48	29	13	9
触摸屏指数	-30.36%	48.32%	33	23	10	7
清洁发电指数	-30.19%	0.00%	0	0	0	0
风力发电指数	-37.79%	27.57%	44	26	11	6
智能电网指数	-44.65%	147.66%	50	38	12	11
新型煤化工指数	-25.93%	50.10%	47	28	14	12
迪士尼指数	-35.66%	28.50%	36	24	10	8
黄金珠宝指数	-19.77%	75.68%	38	26	8	7
滨海新区指数	-33.96%	11.80%	36	21	9	7
乙醇汽油指数	-16.87%	56.73%	32	21	10	9
海南旅游岛指数	-30.46%	39.81%	36	24	10	8
北部湾经济区指数	-40.75%	23.92%	39	25	11	8
三沙指数	-26.11%	3.83%	23	12	6	3

数据来源：WIND 资讯，光大金工中文云 [www.chinesecloud.net](http://www.chinesecloud.net)

## 6、构造概念板块事件驱动套利策略，基于不同的文本数据源的绝对收益率约为 60%-80%，而同期大盘收益率为-20%。

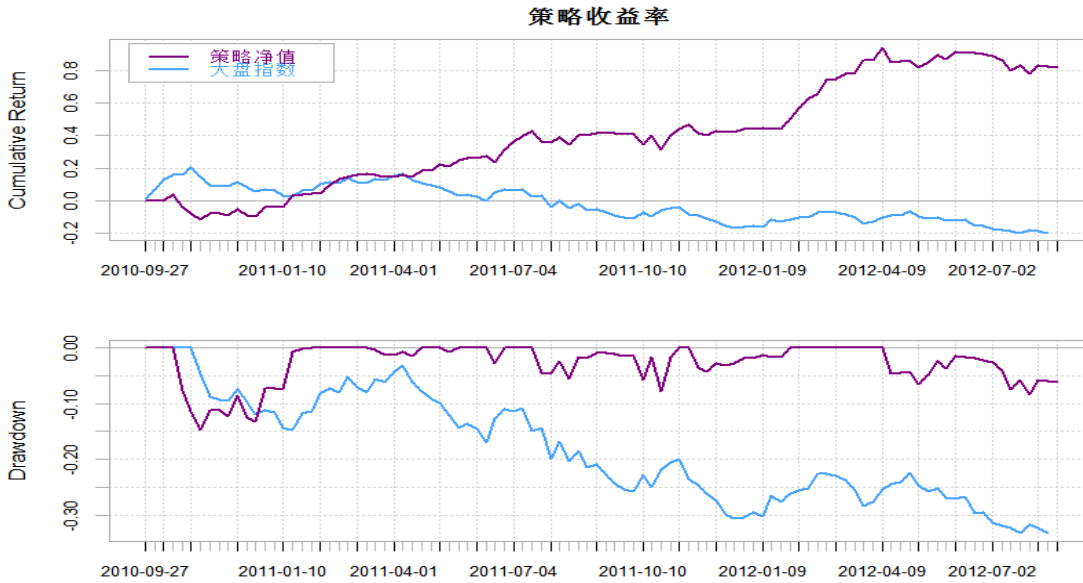
我们按照如下方式构造投资组合：在每一周中，按照上述方式产生各个概念板块的择时策略信号，仅考虑有做多、做空信号产生或者为持有期的概念板块，等权地做多、做空或者持有相对应的概念板块指数，该组合收益如下：

图表 13：基于研究报告文本的事件驱动套利策略表现



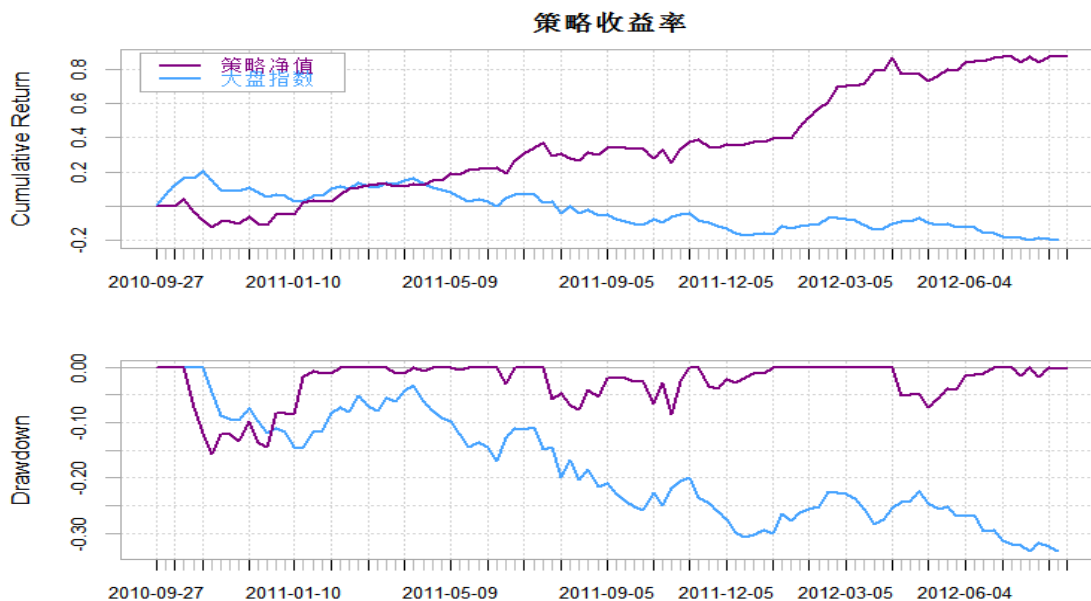
数据来源：WIND 资讯，光大金工中文云 [www.chinesecloud.net](http://www.chinesecloud.net)

图表 14：基于财经新闻文本的事件驱动套利策略表现



数据来源：WIND 资讯，光大金工中文云 [www.chinesecloud.net](http://www.chinesecloud.net)

图表 15：基于股票论坛文本的事件驱动套利策略表现



数据来源：WIND 资讯，光大金工中文云 [www.chinesecloud.net](http://www.chinesecloud.net)

## 7、策略分析

在上述的策略中，基于各文本资料的策略均能够取的良好收益，而且结论相差并不大。三个策略所产生的最大回撤也均是发生在 2010 年 10 月 25 至 2010 年 11 月 8 日的三周内，在该段时间内，基本上每周亏损近 5%。我们检查这段时间的具体交易清单如下：

图表 16：最大回撤时的交易清单

日期	指数收益	信号	信号收益	概念板块
2010/10/25	7.31%	-1	-7.31%	成渝特区板块



2010/11/1	6.15%	-1	-6.15%	石墨烯板块
2010/11/1	8.73%	-1	-8.73%	苹果板块
2010/11/1	2.39%	-1	-2.39%	新三板板块
2010/11/1	3.89%	-1	-3.89%	钢结构板块
2010/11/1	2.34%	-1	-2.34%	成渝特区板块
2010/11/8	11.03%	-1	-11.03%	石墨烯板块
2010/11/8	-7.82%	-1	7.82%	稀土永磁板块
2010/11/8	3.34%	-1	-3.34%	苹果板块
2010/11/8	3.68%	-1	-3.68%	港珠澳大桥板块
2010/11/8	4.65%	-1	-4.65%	新三板板块
2010/11/8	5.68%	-1	-5.68%	钢结构板块
2010/11/8	5.15%	-1	-5.15%	参股金融指数
2010/11/8	4.80%	-1	-4.80%	成渝特区板块
2010/11/8	3.78%	-1	-3.78%	清洁煤发电板块
2010/11/8	5.58%	-1	-5.58%	新型煤化工板块
2010/11/8	9.54%	-1	-9.54%	迪士尼板块
2010/11/8	1.52%	-1	-1.52%	黄金珠宝板块

数据来源：WIND 资讯，光大金工中文云 [www.chinesecloud.net](http://www.chinesecloud.net)

分析其主要原因为：一，集中持仓；在前两周的交易中，持仓的概念指数仅分别有 2 只和 5 只，这在一定的程度上集中了风险，导致了当周的巨大亏损。二，非完全中性；在目前的策略中，虽然多数时间都会同时持有多头与空头头寸以实现对冲的目的，但是目前的机制并不完全保证多空配比的平衡，在少数极端持有单向头寸时就会存在较大的风险暴露，例如在 2010 年 11 月 8 日那周里，我们所持的 12 个概念指数均是空头，但是市场疯狂则在持续，使得策略产生了显著的回撤。

针对上述两个问题，我们会在之后的报告进行改进：首先，控制持仓数，每一期，对所有板块进行强制比较，优选做多与做空板块，如果数量不足，则考虑持有现金；其次，对做多与做空板块进行持仓调整，尽可能实现市场风险中性。

## 8、总结.

在此报告中，我们主要介绍了在量化投资领域的一项探索性研究，从上述的图表中，我们能够感觉到这项研究的价值与意义。尽管报告中所列举的策略还不够成熟，仅仅考虑了在概念板块受到投资者的强烈关注时，判断概念指数所处的位置来确定做多和做空指数的操作。但这也说明了一个浅显的道理，即当某概念板块指数处在相对高位，并且受到投资者的强烈关注时，往往意味着该概念板块将大概率地进入一个调整的阶段，而且是随着持有时间的持续，调整的概率也在逐渐地增大；而当某概念板块指数处在相对低位，并且受到投资者的强烈关注时，往往意味着该概念板块将大概率地进入一个上升阶段。未来，我们将不断在此基础上优化该模型，以期得到更好的风险收益表现。

## 分析师声明

负责准备本报告以及撰写本报告的所有研究分析师或工作人员在此保证，本研究报告中关于任何发行商或证券所发表的观点均如实反映分析人员的个人观点。负责准备本报告的分析师获取报酬的评判因素包括研究的质量和准确性、客户的反馈、竞争性因素以及光大证券股份有限公司的整体收益。所有研究分析师或工作人员保证他们报酬的任何一部分不曾与，不与，也将不会与本报告中的具体的推荐意见或观点有直接或间接的联系。

## 分析师介绍

倪蕴韬，瑞典哥德堡大学金融学硕士,复旦大学世界经济系学士，2010 年加入光大证券研究所任金融工程分析师。  
主要研究方向：投资时钟量化框架、算法交易、高频策略。

## 行业及公司评级体系

买入—未来 6-12 个月的投资收益率领先市场基准指数 15%以上；  
增持—未来 6-12 个月的投资收益率领先市场基准指数 5%至 15%；  
中性—未来 6-12 个月的投资收益率与市场基准指数的变动幅度相差-5%至 5%；  
减持—未来 6-12 个月的投资收益率落后市场基准指数 5%至 15%；  
卖出—未来 6-12 个月的投资收益率落后市场基准指数 15%以上。  
市场基准指数为沪深 300 指数。

## 分析、估值方法的局限性说明

本报告所包含的分析基于各种假设，不同假设可能导致分析结果出现重大不同。本报告采用的各种估值方法及模型均有其局限性，估值结果不保证所涉及证券能够在该价格交易。

## 特别声明

光大证券股份有限公司（以下简称“本公司”）创建于1996年，系由中国光大（集团）总公司投资控股的全国性综合类股份制证券公司，是中国证监会批准的首批三家创新试点公司之一。公司经营业务许可证编号：z22831000。

公司经营范围：证券经纪；证券投资咨询；与证券交易、证券投资活动有关的财务顾问；证券承销与保荐；证券自营；为期货公司提供中间介绍业务；证券投资基金代销；融资融券业务；中国证监会批准的其他业务。此外，公司还通过全资或控股子公司开展资产管理、直接投资、期货、基金管理以及香港证券业务。

本证券研究报告由光大证券股份有限公司研究所（以下简称“光大证券研究所”）编写，以合法获得的我们相信为可靠、准确、完整的信息为基础，但不保证我们所获得的原始信息以及报告所载信息之准确性和完整性。光大证券研究所可能将不时补充、修订或更新有关信息，但不保证及时发布该等更新。

本报告根据中华人民共和国法律在中华人民共和国境内分发，仅供本公司的客户使用。

本报告中的资料、意见、预测均反映报告初次发布时光大证券研究所的判断，可能需随时进行调整。报告中的信息或所表达的意見不构成任何投资、法律、会计或税务方面的最终操作建议，本公司不就任何人依据报告中的内容而最终操作建议作出任何形式的保证和承诺。

在法律允许的情况下，本公司及其附属机构可能持有报告中提及的公司所发行证券的头寸并进行交易，也可能为这些公司提供或正在争取提供投资银行、财务顾问或金融产品等相关服务。投资者应当充分考虑本公司及本公司附属机构就报告内容可能存在的利益冲突，不应视本报告为作出投资决策的唯一参考因素。

在任何情况下，本报告中的信息或所表达的建议并不构成对任何投资人的投资建议，本公司及其附属机构（包括光大证券研究所）不对投资者买卖有关公司股份而产生的盈亏承担责任。

本公司的销售人员、交易人员和其他专业人员可能会向客户提供与本报告中观点不同的口头或书面评论或交易策略。本公司的资产管理部和投资业务部可能会作出与本报告的推荐不相一致的投资决策。本公司提醒投资者注意并理解投资证券及投资产品存在的风险，在作出投资决策前，建议投资者务必向专业人士咨询并谨慎抉择。

本报告的版权仅归本公司所有，任何机构和个人未经书面许可不得以任何形式翻版、复制、刊登、发表、篡改或者引用。

## 光大证券股份有限公司研究所 销售交易部 机构业务部

上海市新闻路1508号静安国际广场3楼 邮编200040

总机：021-22169999 传真：021-22169114、22169134

销售交易团队	姓名	办公电话	手机	电子邮件
北京	王汗青(总经理)	010-56513035	-	wanghq@ebsecn.com
	郝辉	010-56513031	-	haohui@ebsecn.com
	黄怡	010-56513050	-	huangyi@ebsecn.com
	梁晨	010-56513153	-	liangchen@ebsecn.com
企业客户	孙威(执行董事)	010-56513038	-	sunwei@ebsecn.com
	吴江	010-56513030	-	wujiang@ebsecn.com
	杨月	010-56513033	-	yangyue1@ebsecn.com
	顾超	021-22169485	-	guchao@ebsecn.com
上海	李大志(销售交易部总经理助理)	021-22169128	-	lidz@ebsecn.com
	严非(执行董事)	021-22169086	-	yanfei@ebsecn.com
	周薇薇	021-22169087	-	zhouww1@ebsecn.com
	徐又丰	021-22169082	-	xuyf@ebsecn.com
	韩佳	021-22169491	-	hanjia@ebsecn.com
	冯诚	021-22169083	-	fengcheng@ebsecn.com
深圳	黎晓宇(副总经理)	0755-83024434	-	lix1@ebsecn.com
	黄鹏华(执行董事)	0755-83024396	-	huanglh@ebsecn.com
	张晓峰	0755-83024431	-	zhangxf@ebsecn.com
	江虹	0755-83024029	-	jianghong1@ebsecn.com
富尊财富中心	罗德锦	0755-83024064	-	luodj@ebsecn.com
	濮维娜(副总经理)	021-62152373	-	puwn@ebsecn.com
	陶奕	021-62152393	-	taoyi@ebsecn.com
	戚德文	021-22169152	-	qidw@ebsecn.com