

인터넷 검색트렌드와 기업의 주가 및 거래량과의 관계에 대한 연구

A Study on the Relationship between Internet Search Trends and Company's Stock Price and Trading Volume

구평희(Pyunghoi Koo)*, 김민수(Minsoo Kim)**

초 록

본 논문에서는 인터넷 검색 추세와 주식시장 사이에 어떤 관계가 있는지를 알아보고자 한다. 관심 기업의 정보를 얻기 위하여 투자자가 인터넷 검색엔진을 활용하고 이것이 실제 투자로 이어질 수 있다는 가정에서, 기업에 대한 검색량의 변화가 해당 기업의 주가 및 거래량 변동과 어떤 관계성이 있는지를 실제 데이터를 통해 분석하였다. 검색량의 변화를 기초로 한 검색트렌드 투자전략을 대기업 그룹과 중소기업 그룹에 적용하여, 두 그룹의 수익률 등락과 주식거래량에 대한 상관관계를 분석하였다. 7년(2007년~2013년)간의 데이터를 기초로 KOSPI와 KOSDAQ 모두에서 검색트렌드 투자전략이 시장의 평균 수익률 이상을 실현하고, 대기업 보다는 중소기업에서 더 투자효과가 높다는 결과를 얻었다. 검색량과 주식거래량의 관계 또한 대기업보다는 중소기업이 더 영향을 받는다는 것을 알 수 있었다.

ABSTRACT

In this paper, we investigate the relationship between Internet search trends and stock market. Under the assumption that investors may use Internet search engine to obtain information for companies of their interests before taking actual investment actions, the relationship between the changes on Internet search volume and the fluctuation of trading volume as well as stock price of a company is analyzed with actual market data. A search trend investment strategy that reflects the changes on Internet search volume is applied to large enterprises' group and to small and medium enterprises' (SMEs) group, and the correlation between profit rate and trading volume is analyzed for each company group. Our search trend investment strategy has outperformed average stock market returns in both KOSPI and KOSDAQ markets during the seven-year study period (2007~2013). It is also shown that search trend investment strategy is more effective to SMEs than to large enterprises. The relationship between changes on Internet search volume and stock trading volume is stronger at SMEs than at large enterprises.

키워드 : 검색트렌드, 투자전략, 주식시장, 거래량

Search Trends, Investment Strategy, Stock Market, Trading Volume

이 논문은 부경대학교 자율창의학술연구비(2014년)에 의하여 연구되었음.

* First Author, Division of Systems Management and Engineering, Pukyong National University
(phkoo@pknu.ac.kr)

** Corresponding Author, Division of Systems Management and Engineering, Pukyong National University
(minsky@pknu.ac.kr)

Received: 2015-01-12, Review completed: 2015-05-07, Accepted: 2015-05-13

1. 서론

정보화 시대에 디지털 데이터의 홍수 속에서 빅데이터 및 데이터마이닝 기술을 다양한 사회 분야에 적용하여 그로부터 발견되는 정보의 의미와 가치를 제시하는 연구들이 활발히 발표되고 있다. 여기에는 주로 Google과 같은 검색 엔진의 검색내역[3, 7, 9, 12, 17], Facebook과 같은 SNS(Social Network Service) 사이트를 통해 파악되는 연결 관계[12, 18, 20], 특정한 사회적 이슈에 대해 Twitter나 블로그 등을 통해 나타나는 대중의 반응 행태[2, 6, 10, 14], 기타 다양한 유비쿼터스 장비로부터 수집되는 위치나 트랜잭션 기록 등이 활용되고 있다. 물론 빅데이터 자체의 응용 분야별 특성에 대한 연구 또한 꾸준히 이루어지고 있다[1, 13]. 이런 기존 연구의 많은 부분이 검색 엔진, 블로그, SNS 등을 통해 입력되는 단어들의 추세적인 흐름을 활용하여 관련 대상들의 사회적 혹은 경제적 경향을 예측하는 것과 관련되어 있다. 그러나 검색결과에 대한 조회 서비스를 일찍부터 시작한 외국과는 달리, 국내에서는 비교적 최근에 검색 추세에 대한 조회 서비스가 제공되기 시작하였으며, 검색 가능 기간 또한 짧아서 빅 데이터와 관련된 국내 연구들은 비교적 최근에 본격적으로 진행되기 시작했다.

본 논문에서는 국내의 인터넷 검색 추세와 증권시장에서의 투자활동 간의 관련성을 다루고자 한다. 다수의 참여자가 존재하는 증권시장은 개별 참여자들의 의사결정이 종합되어 나타나는 곳으로, 참여자의 행태에 대한 전체적인 이해가 시장동향의 예측에 중요한 요소가 되기 때문에 과거부터 많은 연구가 이루어져 온 분야이기도 하다[3, 10, 11, 12, 16, 18].

본 연구는 인터넷 검색 엔진에서 조회되는 용어를 기존의 많은 연구와 달리, 시장 전반에 대한 관심을 파악하기 위한 것이 아니라 개별 기업에 대한 관심도를 파악하기 위해 각 기업의 이름으로 한정하였다. 특히 기업의 규모에 따른 차이를 살펴보았다는 점에서 기존 연구와 큰 차이가 있다. 우선 기업명의 검색량 증감 추세를 기초로 주식투자전략을 세우고, 이러한 검색트렌드 투자전략의 기업별 투자 효과를 수익률을 통해 분석하였다. 또한 검색트렌드와 주식거래량과의 관계를 분석하여 인터넷 상의 관심이 주식거래로 이어질 수 있다는 개연성에 대해 살펴보았다. 특히, 투자대상을 대기업과 중소기업으로 구분하고 동일한 전략을 활용했을 때 기업의 규모에 따라 수익률이 어떻게 변화하고, 주식거래량과는 어떤 차이를 보이는지 비교·분석하였다.

본 논문의 구성은 다음과 같다. 제 2장에서는 인터넷 검색트렌드를 활용한 대표적인 빅데이터 관련 연구와 본 논문에서 제시하는 검색트렌드 투자전략에 대하여 설명하였다. 제 3장에서는 검색트렌드 투자전략의 수행도를 분석하였고, 제 4장에서는 검색트렌드와 거래량과의 관계를 분석하였다. 마지막 제 5장에서는 제 3장과 제 4장의 실험결과를 정리하고 이와 관련된 추가 논의사항을 제시하였다.

2. 관련 연구 및 검색트렌드 전략

빅 데이터와 관련된 연구는 최근에 사회 각계에서 많은 관심을 받고 있는데, 이것은 디지털화된 정보의 축적이 장기화되면서 이로부터 파생되는 가치를 보다 적극적으로 이해하고

획득하기 위한 자연스러운 움직임이라고 판단된다. 빅 데이터 관련 연구에는 실제로 방대한 분량의 데이터를 다루는 것뿐만 아니라 데이터의 규모는 작지만 높은 빈도로 생성되는 다차원의 고해상도 정보를 활용하는 연구도 함께 포함된다[13, 15, 22, 23]. 특정 기업이나 단체의 내부에서 축적된 데이터뿐만 아니라 최근에는 인터넷 상의 온라인 데이터를 활용한 연구도 활발히 발표되고 있는데, 특히 인터넷 검색 추세를 활용한 연구가 많은 관심을 받고 있다.

2.1 검색트렌드 관련 빅 데이터 연구

인터넷 검색 추세를 활용한 대표적인 연구로는 독감 예보에 지역별 검색 데이터를 활용한 Google Influenza를 들 수 있겠다[9, 17]. 이외에도 과거와 미래의 연도별 검색량 비율로 파악해본 국민들의 검색 성향과 해당 국가의 GDP와의 상관관계를 살펴 본 연구[18, 19], 특정 단어 집합에 대한 검색율의 증감이나 Twitter 및 Wikipedia 사용패턴을 통해 금융시장 참가자들의 행태를 파악하고자 하는 연구[2~5, 8, 10, 12, 14, 18, 20], 후보자에 대한 검색수와 선거 결과의 연관성을 찾는 시도[23] 등 매우 다양한 아이디어들이 실험되어 발표되고 있다. 이러한 인터넷 검색 추세를 활용한 연구의 대부분은 주로 인과관계의 틀 속에서 해석하기 어려운 현상을 동시성이나 개연성을 통해 이해하려는 시도를 담고 있다고 하겠다.

인터넷 검색과 증권시장과의 연관성을 연구하는 논문도 다수 발표되고 있는데, Preis 등은 인터넷 검색 엔진을 이용한 기업의 주간 검색량이 해당 기업의 다음 주 주식거래량과 밀접

한 관계를 보이고 있음을 발표하였다[20]. 이후, Bordino 등은 하루 동안의 인터넷 검색량이 당일뿐만 아니라 다음날과 그 이후 며칠 동안의 주식거래량에 영향을 미친다는 결과를 발표하였다[3]. Preis 등은 구글트렌드(Google Trends)를 이용하여 용어 별 검색수의 추이가 주식시장의 향후 움직임을 예측하는데 활용될 수 있다는 가정 하에서, 98개의 단어를 선별하여 이 단어 집합의 검색량 증감을 기초로 한 투자전략을 세우고, 이 전략이 시장평균 이상의 수익률을 보인다는 결과를 제시하였다[18]. 김민수와 구평희[11]는 Preis[18]의 연구를 국내 KOSPI 시장에 동일하게 적용할 때의 투자 수익률을 비교·분석하는 연구를 발표하였다. 우선 영어로 된 검색 용어를 한글로 바꾸고, Naver 검색엔진에서의 검색 추이를 기반으로 하여 Preis가 제시한 투자전략을 적용하여 KOSPI 지수를 거래한다. 과거 수년간의 KOSPI 지수와 Naver 검색추이 데이터를 이용하여 실험한 결과, 국내에서 검색트렌드를 이용한 투자전략은 미국에서 검색트렌드를 이용한 투자전략과는 상이한 결과를 낼 수 확인하였다. 구체적으로, 구글트렌드를 이용하여 시장평균 이상의 수익률을 보이는 검색어가 전체의 65% 가량을 차지하였지만, 네이버트렌드를 이용한 경우에는 약 25% 정도에 불과하였다. 따라서 이러한 인터넷 데이터를 활용하기 위해서는 해외와 국내의 상황이 많은 차이를 보인다는 점을 미리 이해할 필요가 있다고 주장하였다.

2.2 검색트렌드 투자전략

본 연구는 기존 Preis 등의 연구[18]와, 김민수 및 구평희 연구[11]의 연장선상에서, 검색용

어를 일반용어가 아닌 특정 기업명으로 대체하여 해당 기업의 개별주가 및 거래량과의 연관성을 살펴보고자 한다. 주식시장에서 투자자가 어떤 기업의 주식을 사거나 팔 때 기업 정보를 추가적으로 얻기 위하여 투자 전에 인터넷의 검색엔진으로 해당 기업명을 검색할 것이라는 인식하에 검색용어로 기업명을 택하였다. 물론 이러한 연관관계에 대한 인과론적 검증은 보다 폭넓은 사회과학적 연구가 필요한 주제로, 본 연구의 범위를 벗어난다. 본 연구에서는 대상들 간에 그러한 인과관계가 존재한다면, 이를 데이터간의 추세적 연관성 속에서 발견하고자 하는 시도로, 이것이 대상들 간의 인과적 관계를 증명하는 것은 아니다. 이러한 개연성을 설명할 수 있는 가설들의 가능성만을 제기하고자 한다.

기업명에 대한 검색량 추세 데이터를 확보하기 위하여, 2007년 이후의 검색 데이터에 대해 용어별로 검색횟수에 대한 상대적인 추이 정보를 주간 단위로 보여주는 네이버트렌드 기능을 이용하였다. 인터넷 통계데이터(<http://trendlogger.co.kr>)에 의하면 2013년 1년 동안 국내 검색사이트의 검색점유율은 네이버가 78.6%를 차지하고 있다. 따라서 본 논문에서는 신뢰도 측면에서 가장 검색수가 많은 네이버트렌드의 추이데이터를 택하여 투자전략을 세웠다. 투자의 주기는 네이버트렌드에서 데이터를 얻을 수 있는 주기인 주간 단위로 하고, 매주 초에 지난주 회사명에 대한 검색량을 보고 검색수가 그 이전 일정기간보다 증가하면 투자자의 관심이 증대되었다고 판단하고 해당회사의 주식에 대해 사자 포지션을 취하고, 반대로 검색수가 감소하면 팔자 포지션을 취한다. 검색량의 변동이 없는 경우에는 투자하지 않고 쓴다.

일반적으로 투자자의 관심과 높은 주식가치는 서로 상승작용을 보이는 것으로 인식된다는 점[16]에서 이러한 투자전략을 근거로 실험하는 것이 타당하다고 할 수 있겠다. 본 실험에서는 매주 초에 반복적으로 투자 포지션에 대한 의사결정을 내린다. 네이버의 검색트렌드를 기반으로 한 구체적인 투자 전략은 아래와 같다. 이후부터는 이를 검색트렌드(Search Trends) 투자전략으로 나타내기로 하였다.

투자 의사결정을 해야 하는 현시점을 t 주 초라 하고, 지난주, 즉 $(t-1)$ 주에 검색된 특정한 단어에 대한 상대적인 검색수를 $n(t-1)$ 이라고 하자. 그 이전의 Δt 주, 즉 $(t-1-\Delta t)$ 주부터 $(t-2)$ 주까지의 평균 검색수를 $N(t-2, \Delta t)$ 로 표현하면,

$$N(t-2, \Delta t) = \frac{n(t-2)+n(t-3)+\dots+n(t-1-\Delta t)}{\Delta t}$$

에 의해서 계산할 수 있다. 여기서, $n(t-1)$ 과 $N(t-2, \Delta t)$ 를 비교하여 $n(t-1) > N(t-2, \Delta t)$, 즉 이전 주 $(t-1)$ 의 검색수가 그 이전 Δt 주 동안의 평균 검색수보다 증가했다면 t 주의 첫 번째 거래일 증가 $p(t)$ 에 해당 기업의 주식을 매수하고 다음 주, 즉 $(t+1)$ 주 첫 거래일의 증가 $p(t+1)$ 에 해당주식을 매도하여 거래를 소멸시키는 전략을 취한다. 반대로, 이전 Δt 주 동안의 평균 검색수와 비교하여 검색수가 감소했다면 t 주의 첫 번째 거래일 증가에 해당 기업의 주식을 매도하고 다음 주 첫 거래일의 증가에 매수하여 거래를 소멸시킨다.

여기서 제시된 검색트렌드 투자전략은 검색용어가 네이버트렌드를 통해 얻은 국내 기업의 명칭이고, 주가가 국내 주식시장에서의 가

격이라는 점을 제외하고는 Preis 등의 연구[18]에서 사용된 기존 투자전략과 동일한 형태이다. 본 연구에서는 기존 연구와는 달리 수익률을 종합주가지수를 통해 파악하지 않고 검색대상이 되는 개별 기업의 수익률로 파악하였으며, 대기업군과 중소기업군의 수익률은 대상군에 속하는 기업들의 수익률을 합산하여 계산하였다.

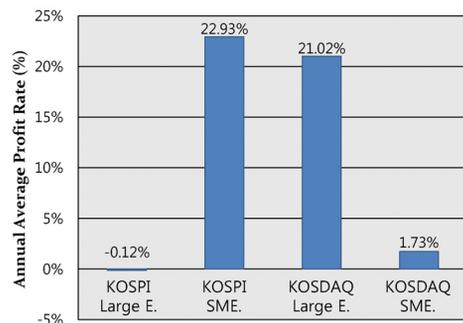
3. 실험 결과 분석

대기업과 중소기업에 대해 앞에서 제시한 검색트렌드 투자전략을 동일하게 적용하고, 그 결과를 비교하기 위하여 KOSPI 대기업 30곳과 중소기업 30곳, 그리고 KOSDAQ 대기업 30곳과 중소기업 30곳을 선정하였다. KOSPI와 KOSDAQ의 대기업은 주가총액을 기준으로 상위 30곳을 선정하였고, 중소기업으로는 시가총액 기준으로 하위 20% 내외의 기업이 선택되도록, KOSPI에서는 시가총액 300억 이상, KOSDAQ에서는 시가총액 200억 이상인 기업 중 시가총액이 작은 순서로 각각 30개의 기업을 선정하였다. 기업만을 대상으로 하기 위해 종목명 중에서 ETF 등은 제외하였다. 투자기간은 네이버트렌드의 데이터가 존재하는 2007년 초부터 2013년 말까지 7년 동안으로 정하였다.

3.1 기업규모별 검색트렌드 투자전략의 수익률

<Figure 1>은 KOSPI와 KOSDAQ의 대기업 및 중소기업 30개에 대하여 검색트렌드 투자전략을 적용할 경우의 연간 평균 수익률을

보여주고 있다. 여기에서 연간수익률은 주간수익률을 누적하여 구했다. 주간수익률의 계산은 t 주에 매도포지션을 취하는 경우 $p(t)$ 에 팔고 $p(t+1)$ 에 사므로 수익률이 $p(t)/p(t+1)-1$ 로 계산되지만, 본 논문에서는 기존의 연구[11, 18]에서와 마찬가지로 매도포지션이나 매수포지션에 따른 누적수익률의 계산을 간편하게 할 수 있도록, 자연로그함수를 사용하였다. 매도포지션에서의 수익률은 $\log(p(t))/\log(p(t+1))$ 로, 매수포지션에서의 수익률은 매도포지션과는 반대로 $\log(p(t+1))/\log(p(t))$ 을 사용한다. 주간수익률을 이렇게 계산하면 연간수익률은 주간수익률을 더하여 간단히 구할 수가 있다는 장점이 있다. <Figure 1>을 보면, 검색트렌드 투자전략을 적용하는 경우 KOSPI 대기업(Large E.)에서는 연간수익률이 0%에 가까운 반면에, KOSDAQ 대기업에서는 20% 이상의 높은 수익률을 기록하고 있음을 알 수 있다. 중소기업(SME.)의 경우, KOSPI에서 약 23%의 수익률을 보이지만 KOSDAQ에서는 약 2%의



<Figure 1> Annual Average Profit Rate of Search Trends Investment Strategy over KOSPI and KOSDAQ Market for Large Enterprises and SMEs

수익률만을 보이고 있다. 이 네 값을 모두 평균하면 검색트렌드 투자전략을 적용하여 약 11.4%의 연간수익률을 네 시장 환경에서 결과적으로 얻을 수 있다고 하겠다.

분석기간(2007~2013년) 동안에 주식이격 자체의 등락으로 인한 수익률의 변화가 발생할 수 있으므로, 이로 인한 영향을 파악하여 살펴볼 수 있도록 Preis 등의 연구[18]에서 벤치마크 투자전략으로 사용했던 buy-and-hold (BH) 투자전략을 그대로 적용한 경우와 검색트렌드(Search Trend: ST) 투자전략을 적용한 경우의 연간 평균수익률을 다음의 <Table 1>에 각각 비교하였다.

BH 투자전략에서는 주초 첫 거래일 증가로 항상 주식을 매수하고, 그 다음 주 첫 거래일 증가로 매도 청산하는 전략이다. 즉, BH 투자 전략은 주가가 오르면 이익을 보고 내리면 손실을 입는 전략이다. <Table 1>에서 확인할 수 있듯이 검색트렌드 투자전략과 BH 투자전략 사이에 상당한 수익률 차이가 있음을 알 수 있다. 즉, 검색트렌드 투자전략을 적용하는 경우 KOSPI 대기업에서는 BH 투자전략과 비교하여 수익률이 -8.55%로 저조한 반면에, KOSDAQ 대기업에서는 BH 투자전략과 비교하여 31.18% 더 높은 수익률을 실현하고 있음을 알 수 있다. 반면에, 중소기업에 대해서는 KOSPI와 KOSDAQ 모두에서 검색트렌드 투자전략을 적용하는 것이 BH 투자전략을 적용하는 것보다 수익률이 각각 31.18%와 11.77% 높다는 것을 알 수 있다. 본 논문에서는 주식거래 비용은 고려하고 있지 않으므로, 검색트렌드 전략에서는 매주 거래가 발생할 때마다 거래비용이 발생할 것이고, buy-and-hold 전략에서는 거래비용이 발생하지 않는다. 따라서 실제로 거래비용을 고

려해야 하는 상황에서는 트렌드 투자전략의 수익률에서 거래비용을 차감해야 할 것이다.

<Table 1> Annual Average Profit Rates of Two Investment Strategies

	ST	BH	ST-BH*
KOSPI Large E.	-0.12%	8.43%	-8.55%
KOSPI SME.	22.93%	-8.25%	31.18%
KOSDAQ Large E.	21.02%	14.28%	6.74%
KOSDAQ SME.	1.73%	-10.04%	11.77%

* ST-BT is the difference between the profit rate of search trend (ST) strategy and the profit rate of buy-and-hold (BH) strategy.

검색트렌드 투자전략을 적용하는 경우 중소기업이 대기업보다 더 높은 수익률을 얻은 이유에 대하여 명확한 설명이 불가능하나 아래와 같은 요인이 투자 결과에 영향을 주었을 가능성이 있다고 판단된다.

- (1) 대기업은 투자 목적이 아닌 다른 목적(예를 들면, 뉴스, 제품이나 취업 관련 정보 등)으로 인터넷을 검색하는 경우가 상대적으로 많아서 검색트렌드 투자전략에 noise의 역할을 했을 가능성이 많고,
- (2) 대기업은 시가총액의 크기가 크므로 민감도가 중소기업보다 상대적으로 덜 한 반면, 중소기업은 투자자의 관심에 더 민감할 수 있으며,
- (3) 대기업은 기관이나 외국인 등의 대형 투자자가 많은 반면, 중소기업은 개인투자자가 많이 참여하여 상대적으로 정보가 부족한 개인 투자자들이 투자 전에 해당 회사에 대한 정보를 얻기 위하여 인터넷을 검색할 가능성이 높다.

하여 매주 사자와 팔자 의사결정을 했을 경우에, 실제로 주식 등락을 맞춘 적중률을 나타낸다. 주가의 무작위성을 고려하여 본다면 주가등락의 적중률이 50%에 수렴하는 것이 정상적이라 할 수 있겠다. <Table 2>는 검색트렌드 투자전략이 양 주식시장에서 기업규모별로 주가 등락을 맞춘 비율과 통계적 유의성을 99% 신뢰구간으로 살펴 본 것이다. 예를 들어, KOSPI 중소기업의 경우 30개 기업을 대상으로 2007년 초부터 2013년 말까지 361주 동안 10,830번의 투자 기회 중 검색트렌드 투자전략에 의해 사자 또는 팔자 주문을 낸 9,572회의 투자 중 등락을 올바르게 예측한 횟수는 4,527회로 그 비율은 52.71%이다. 신뢰 구간은 등락 예측의 정확도가 50%라고 가정하고, 올바른 등락예측의 횟수는 이항분포를 따른다는 가정 하에 9,572회를 투자했을 때 99%의 확률로 등락을 올바르게 맞춘 비율의 범위를 나타낸다. 실험결과 등락을 맞춘 횟수의 비율이 이 범위를 벗어나면 유의수준 1%로 검색트렌드 투자전략의 적중률이 50%보다 크다고 할 수 있다. <Table 2>에서 모든 기업 규모에 있어 등락적중률이 50%를 넘어서고 있고, 중소기업은 KOSPI와 KOSDAQ 두 시장 모두에서, 대기업의 경우는

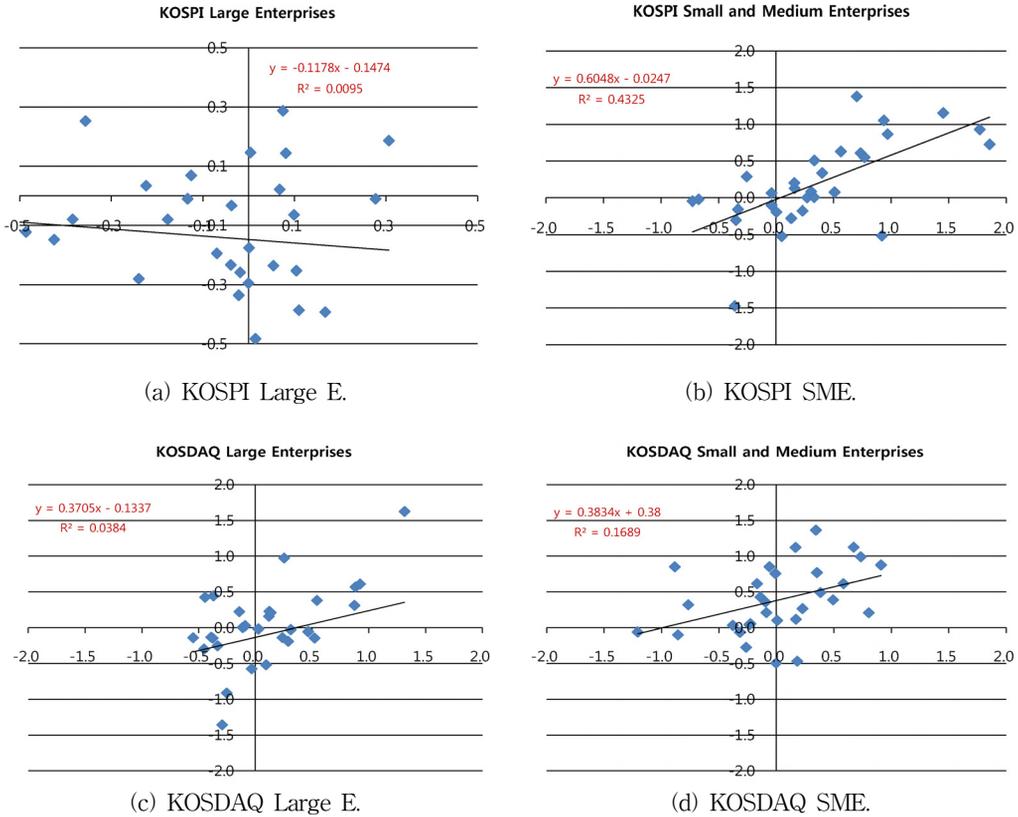
KOSDAQ 시장에서 99%의 신뢰구간을 벗어나 높은 확률로 주가등락을 맞춘다는 것을 알 수 있다.

검색트렌드 투자전략의 안정성 및 연속성을 보기 위하여 전체 분석기간 동안의 데이터를 학습기간(2007년에서 2011년까지)과 검증기간(2012년에서 2013년까지)으로 나누어, 두 기간 동안의 조정연수익률간의 관계를 분석하였다. <Figure 3>은 검색트렌드 투자전략을 적용하여 투자한 경우 KOSPI와 KOSDAQ의 대기업과 중소기업에 대해 두 기간 동안의 투자수익률 관계를 나타내는 산점도이다. 그림에서 x축은 학습기간 동안의 연조정수익률을 나타내며, y축은 검증기간 동안의 연조정수익률을 나타낸다.

KOSPI와 KOSDAQ 두 시장에서 모두 기업명에 대한 검색수와 대상기업의 연 수익률의 상관관계가 중소기업의 경우에 대기업 보다 상대적으로 높게 나오는 것을 알 수 있다. KOSPI 대기업을 제외한 나머지 세 그룹에 있어서도 모두 의미 있는 양의 상관관계가 있음을 볼 수 있는데, 이것은 검색트렌드 투자전략이 KOSPI 대기업을 제외한 나머지 기업군에 대해서 학습기간과 검증기간 전체에 걸쳐 비교적 안정적인 투자전략으로 활용될 수 있음을 의미한다. 두 시장에서 모두 중소기업의 경우 대기업보다 상관계수가 높게 나왔는데, 이는 중소기업의 경우 주식을 거래하기 위해 사전에 인터넷 검색 사이트에서 해당기업을 검색하는 경향이 강하다는 것으로 해석할 수 있겠다. 이것은 <Table 1>에서 검색트렌드 투자전략을 적용한 경우 중소기업이 상대적으로 높은 수익률을 보이는 결과와도 같은 현상으로 이해될 수 있겠다.

<Table 2> Hit-rate and Confidence Interval (99%) of Search Trend Investment Strategy

Company Category	Hit-rate	99% Confidence Interval
KOSPI Large E.	50.39	[48.49, 51.51]
KOSPI SME	52.71	[48.47, 51.53]
KOSDAQ Large E.	52.74	[48.43, 51.57]
KOSDAQ SME	51.62	[48.44, 51.56]

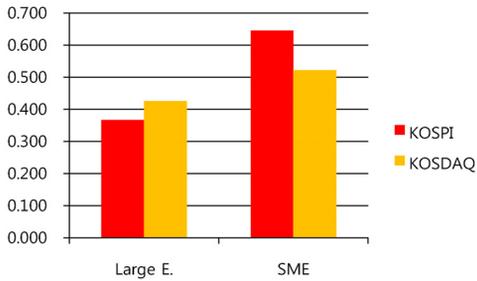


<Figure 3> Scatter plot between Adjusted Annual Profit Rates of Search Trend Investment Strategy during learning period (2017~2011) and verification period (2012~2013) for Top 30 and Middle 30 Companies in KOSPI and KOSDAQ Market (x-Axis is for learning period, y-Axis is for verification period. Direct line shows linear regression of each case and R^2 is a coefficient of determination)

3.2 검색추세와 주식 거래량의 관계

인터넷 검색 추세와 주가 변동과의 연결 관계를 좀 더 살펴보기 위해 해당 주식의 거래량 변동을 가능한 매개요인으로 살펴보았다. 즉, 투자 기업에 대한 관심은 인터넷 검색을 통해 나타나고, 이런 관심이 실제 해당 기업 주식의 거래를 수반하여 궁극적으로 주가를 변화시킬 수 있다는 개연성을 살펴보기 위해서이다.

검색엔진에서 기업명에 대한 검색량의 변화와 해당 주식의 거래량과의 관계를 알아보기 위해서, 앞서의 실험과 동일하게 KOSPI와 KOSDAQ의 대기업 및 중소기업 30개 회사를 선정하여 2007년부터 2013년까지의 검색 추세와 주식거래량 데이터를 분석하였다. <Figure 4>는 KOSPI와 KOSDAQ 시장에서 기업규모별로 기업명에 대한 인터넷 검색량과 주식거래량간의 상관관계를 나타낸 것이다.

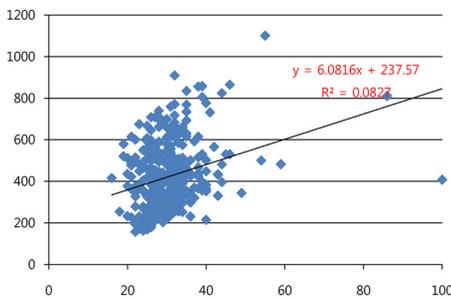


〈Figure 4〉 Correlation Coefficient between Search Volume and Trading Volume at the Company Groups in the KOSPI and KOSDAQ Markets

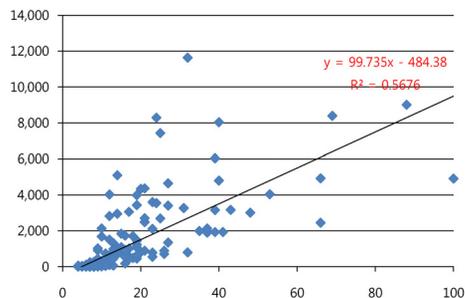
그림을 통해 크게 두 가지의 사실을 확인할 수 있는데, 첫째는 KOSPI와 KOSDAQ 두 시장 모두에서 검색빈도와 주식거래량 사이에 양의 상관관계(약 0.4~0.6의 상관계수)가 존재한다는 점이다. 즉, 인터넷 검색엔진에서 기업명에 대한 검색량이 증가하게 되면, 그 다음 주에 바로 주식거래가 늘어나게 되고, 이것이 주가에 영향을 미치게 되어서 결과적으로 앞서의 <Table 1>에서 살펴본 것과 같이 검색트렌드 투자전략의 높은 연간 평균 수익률을 보이게 된다고 설명할 수 있겠다. 두 번째는 이러한

관계성이 대기업 보다는 중소기업에서 보다 크게 나타난다는 점이다. 즉, 인터넷에서 검색량의 증감이 중소기업의 경우에는 바로 주식거래로 이어지는 정도가 대기업에 비해 높다는 의미이고, 이것은 <Table 1>에서 살펴본 것과 같이 검색트렌드 투자전략의 효과가 대기업보다 중소기업에서 더 높게 나타나는 현상에 대한 하나의 가능한 설명이 될 수 있겠다.

〈Figure 5〉는 KOSPI의 대기업인 삼성전자와 중소기업인 모나미를 예로해서, 인터넷 검색량과 주식거래량과의 관계를 보여주고 있다. 대기업인 삼성전자의 경우 상관계수는 0.29로 약한 양의 상관관계를 보여주는 반면에, 중소기업인 모나미의 상관계수는 0.753으로 인터넷 검색량과 주식거래량 사이에 높은 양의 상관관계가 있다는 것을 알 수 있다. 이러한 결과는 앞서 언급하였듯이 대기업의 경우 뉴스나 다양한 매체를 통해 해당 기업에 대한 정보를 쉽게 접할 수 있는 것에 반해, 중소기업의 경우에는 다른 매체로부터의 정보가 상대적으로 부족한 상황에서 인터넷 정보검색이 투자 결정에 보다 더 많은 직접적인 영향을 미치게 된 결과라고 해석할 수 있겠다.



(a) Samsung Electronics



(b) Monami

〈Figure 5〉 Correlation between Search Volume and Trading Volume of Company (x-Axis is for search volume of company name, y-Axis is for trading volume of that company)

5. 결론 및 논의 사항

본 논문에서는 인터넷 검색엔진에서 기업명 검색의 추이를 이용하여 투자전략을 세우는 방안을 제시하고, 이러한 투자전략이 수익률에 어떻게 영향을 주는지를 실제 데이터를 이용하여 분석하였다. 특히, 이러한 투자전략이 국내 KOSPI와 KOSDAQ 두 주식시장의 대기업과 중소기업에 적용되는 경우의 투자성과를 비교하였다. 2007년부터 2013년까지 7년간의 검색 데이터와 주식데이터를 가지고 분석한 결과, KOSPI, KOSDAQ 모든 경우에 수익률이 시장 평균 이상이었다. 특히 KOSPI와 KOSDAQ 두 시장 모두에서 검색트렌드 투자전략의 투자 성과가 대기업 보다는 중소기업에 대상으로 하는 경우에 더 크게 기대된다는 것으로 조사되었다. 또한, 인터넷 검색량과 주식거래량의 관계를 조사한 결과 두 시장 모두에서 기업규모와 무관하게 양의 상관관계가 있는 것으로 나타났다. 특히 투자수익률의 경우와 마찬가지로 주식거래량도 대기업 보다는 중소기업이 검색 추세에 더 많은 영향을 받는다는 것을 볼 수 있었다. 본 연구는 과거의 데이터를 기초로 거래비용을 배제하고 진행되었기에, 이러한 결과를 실제의 주식투자에 바로 적용하는 것은 위험이 따르고, 현실적으로 수익을 보장할 수도 없다. 그러나 본 연구의 결과를 다른 여러 투자 전략과 함께 의사결정시에 보완적인 데이터로 유용하게 사용할 수 있을 것으로 기대한다.

본 논문은 검색 추세를 이용한 투자전략이 주식투자 수익률에 영향을 주는지에 대한 연구를 다루고 있다. 본 연구는 아직까지는 초보적인 단계로, 실제의 시장 투자전략까지 도출해 내기 위해서는 향후 아래와 같은 내용을 보

완 및 개선하는 연구가 추가되어야 할 것이다. 첫째, 본 논문의 투자전략은 검색 용어로서 해당기업의 명칭만을 이용하였다. 기업명을 검색할 때 연관된 다른 용어들과 함께 검색하는 경우 이러한 용어들의 성질(예를 들면 긍정적인 용어인지, 부정적인 용어인지)에 따라 주가 가격의 등락이 영향을 받을 수 있을 것으로 판단된다. 이러한 감성분석을 활용하여 투자전략의 개선이 필요할 것이다. 둘째, 인터넷 검색용어와 더불어, 인터넷에서 실시간으로 얻을 수 있는 신문 기사를 투자전략에 포함시킨다면 이 또한 제시한 투자전략을 개선하는 효과가 있을 것으로 기대된다. 셋째, 본 연구의 분석 결과에서 볼 수 있듯이 검색 추세가 대기업과 중소기업에 미치는 영향에 차이가 있다. 따라서 기업규모별로 차별화 된 투자전략을 수립하는 것이 필요할 것이다. 넷째, 데이터의 수집과 분석 기간인 2007년~2013년 중에서 2008년과 2009년은 세계금융위기 및 회복으로 주가지수의 변동성이 심한 시기였다. 이러한 변동성이 큰 상황이 투자전략의 성과에 어떤 영향을 주는지에 대한 연구도 의미가 있다 하겠다. 마지막으로, 인터넷 검색량의 추이 통계가 주간 단위로 발표되기 때문에 본 논문에서는 주간 단위의 투자결정을 실시하였다. 일간 단위의 통계나 실시간 검색량 추이를 활용할 수 있다면 투자주기를 짧게 하여 좀 더 최근 정보를 활용한 투자 전략을 수립할 수 있을 것이다.

References

- [1] Biswas, S., Yoo, J. H., and Jung, C. Y.,

- “A Study on Priorities of the Components of Big Data Information Security Service by AHP,” *Journal of Society for e-Business Studies*, Vol. 18, No. 4, 2013.
- [2] Bollen, J., Mao, H., and Zeng, X. J., “Twitter mood predicts the stock market,” *Journal of Computational Science*, Vol. 2, No. 1 pp. 1-8, 2011.
- [3] Bordino, I., Battiston, S., Caldarelli, G., Cristelli, M., Ukkonen, A., and Weber, I., “Web search queries can predict stock market volumes,” *PLOS One*, Vol. 7, No. 7, pp. 1-17, 2012.
- [4] Choi, H. and Varian, H., “Predicting Initial Claims for Unemployment Insurance Using Google Trends,” Technical Report, Google, 2009.
- [5] Choi, H. and Varian, H., “Predicting the present with Google Trends,” *The Economic Record*, Vol. 88, pp. 2-9, 2012.
- [6] Choi, S. and Kwon, O., “The Study of Developing Korean SentiWordNet for Big Data Analytics: Focusing on Anger Emotion,” *Journal of Society for e-Business Studies*, Vol. 19, No. 4, pp. 1-19, 2014.
- [7] Cooper, C., Mallon, K., Leadbetter, S., Pollack, L., and Peipins, L., “Cancer Internet Search Activity on a Major Search Engine, United States 2001-2003,” *Journal of Medical Internet Research*, Vol. 7, No. 3, 2005.
- [8] Ettredge, M., Gerdes, J., and Karuga, G., “Using Web-based search data to predict macroeconomic statistics,” *Communications of the ACM*, Vol. 48, No. 11, pp. 87-92, 2005.
- [9] Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., and Brilliant, L., “Detecting influenza epidemics using search engine query data,” *Nature*, Vol. 457, pp. 1012-1014, 2009.
- [10] Goel, S., Hofman, J. M., Lahaie, S., Pennock, D. M., and Watts, D. J., “Predicting consumer behavior with Web search,” *Proceedings of the National Academy of Sciences*, Vol. 7, No. 41, pp. 17486-17490, 2010.
- [11] Kim, M. and Koo, P., “A Study on Big Data Based Investment Strategy Using Internet Search Trends,” *Journal of the Korean Operations Research and Management Science Society*, Vol. 38, No. 4, pp. 53-63, 2013.
- [12] McLaren, L. and Shanbhogue, R., “Using internet search data as economic indicator,” *Quarterly Bulletin*, Q2, pp. 134-140, 2011.
- [13] Min, G. Y. and Jeong, D. H., “Research on Assessment of Impact of Big Data Attributes to Disaster Response Decision-Making Process,” *Journal of Society for e-Business Studies*, Vol. 18, No. 3, 2013.
- [14] Moat, H. S., Curme, C., Avakian, A., Kenett, D. Y., Stanley, E., and Preis, T., “Quantifying Wikipedia usage patterns before stock market moves,” *Scientific Report*, Vol. 3, pp. 01801 : 1-5, 2013.

- [15] National Information Society Agency (Korea), New Value-Creation Engine, Big Data's New Possibility and Response Strategy, IT & Future Strategy, Vol. 18, 2011.
- [16] Parikh, P., VALUE INVESTING and BEHAVIORAL FINANCE: Insights into Indian Stock Market Realities, Tata McGraw-Hill Education, 2009.
- [17] Polgreen, P. M., Chen, Y., Pennock, D. M., and Nelson, F. D., "Using Internet Searches for Influenza Surveillance," Healthcare Epidemiology, Vol. 47, pp. 1443-1448, 2008.
- [18] Preis, T., Moat, H. S., and Stanley, H. E., "Quantifying trading behavior in financial markets using Google Trends," Scientific Report, Vol. 3, pp. 01684 : 1-5, 2013.
- [19] Preis, T., Moat, H. S., Stanley, H. E., and Bishop, S. R., "Quantifying the Advantage of Looking Forward," Scientific Report, Vol. 2, pp. 00350 : 1-2, 2012.
- [20] Preis, T., Reith, D., and Stanley, H. E., "Complex dynamics of our economic life on different scales: insights from search engine query data," Philosophical Transactions of the Royal Society, Vol. 368, pp. 5707-5719, 2010.
- [21] Song, M., Business Future Map That Big Data Builds, Hans Media, 2012.
- [22] Suzuki, R., The Age of Big Data Business, The Soup, 2012.
- [23] Yoon, H., Now Its Big Data Era, eBiz Books, 2012.

저 자 소 개



구평회
1992년
1996년
1999년~현재
관심분야

(E-mail: phkoo@pknu.ac.kr)
Purdue 대학교 산업공학과 (공학석사)
Purdue 대학교 산업공학과 (공학박사)
부경대학교 시스템경영공학부 교수
SCM, Lean System, Data Mining 등



김민수
1996년
2002년
2004년~현재
관심분야

(E-mail: minsky@pknu.ac.kr)
서울대학교 산업공학과 (공학석사)
서울대학교 산업공학과 (공학박사)
부경대학교 시스템경영공학부 부교수
BPM, 기술경영, Big Data 등