[UCI]I804:24011-200000263558

2013년 2월
석사학위논문

Statistical Analysis of Individual Companies listed on KOSPI100 index by using Google Trends

조선대학교 대학원

경 영 학 과

염　영　란

구글트렌드를 이용한 KOSPI100 인덱스 개별기업의 통계적 분석

Statistical Analysis of Individual Companies listed on KOSPI100 index by using Google Trends

2013년 2월 25일

조선대학교 대학원

경 영 학 과

염 영 란

Statistical Analysis of Individual Companies listed on KOSPI100 index by using Google Trends

지도교수 김 규 영

이 논문을 경영학 석사학위신청 논문으로 제출함

2012년 10월

조선대학교 대학원

경 영 학 과

염 영 란

염영란의 석사학위 논문을 인준함

위원장 조선대학교 교수 이한재 (인)

위 원 조선대학교 교수 이계원 (인)

위 원 조선대학교 교수 김규영 (인)

2012년 11월

조선대학교 대학원

Table of Contents

ABSTRACT

List of Figures

List of Tables

ABSTRACT

Statistical Analysis of Individual Companies listed on KOSPI100 index by using Google Trends

Youm Young Ran
Advisor : Prof. Kim Kyou-Young, Ph.D.
Department of Business,
Graduate School of Chosun University

본 논문은 소셜네트워크의 한 부분인 구글 트렌드를 이용하여 KOSPI100 인덱스의 개별기업의 통계적 분석을 연구한 논문이다. 자기상관(Autocorrelation)함수를 이용한 구글트렌드에서의 KOSPI100 개별기업의 상관관계를 살펴보았다. 또한, Detrended fluctuation analysis(DFA)의 Hurst 지수을 이용하여 KOSPI100 개별기업의 수익률과 변동성을 살펴보았으며, 같은 식을 이용하여 구글트렌드의 Hurst 지수값과 Hurst 지수의 절대값을 살펴보았다. 두 결과 모두 Hurst 지수값의 경우 0.5에 가까운 Anti-correlation을 발견할 수 있었으며, 절대값을 취한 경우 0.5 이상 0.8 이하의 상관관계가 있음을 살펴볼 수 있었다.

Chapter 1

Introduction

In the past, many people were think of psychology did not affect real economics. However, these days, there some facts that refute this. The first fact is social network. Many of us use tablet PC and smart phone now. Moreover, this product helps searching information and using Social Network Service(SNS) more easily and using anytime, anywhere. Also, Social Network Service(SNS) can be collect for researching data. Moreover, some studies indicate that it is effective. One of Social Network Service(SNS), which name is twitter. Using twitter data for predicting the spread to influenza-like dieseases [Polgreen et al. 2008 and Ginsberg et al. 2009]. Twitter can predicting the stock market [Johan et al. 2011]. With this, this paper work data from parts of Social Network Service(SNS), which is Google Trends[1]. Previous work was focus on search engine query data can show economic life on different scales [Tobias et al. 2010] and Google trends possible to predicting the present with Google Trends [Choi et al. 2012].

The second fact is behavioral economics. Behavioral economics are at the base of the theory. In our opinion, social network based on behavioral economics. Amos Tversky and Danieal Kahneman are consider as father of behavioral economics. behavioral economics has relation between psychology and economics, behavioral economics frame is humans are not rational[A Tversky and D Kaneman. 2000].

1) http://www.google.com/trends.

If someone want to know or buy something, they will be using internet search engine. After that, searching information and gather it. At those behavior helps makes information aggregation and decision making. When you click on the internet search engine (for example, Google, Yahoo, Naver etc.), in this case, Google make aggregation of data. It is Google Trends. We were focusing on Google Trends search volume index also we called number of hitting. Tobias's[2010] paper shows connection between Google Trends and S&P500 index. This paper was created based on the ideas of Tobias's work. We thought, maybe Google Trends and KOSPI100 index number of hitting have connection. Accordingly, we have been researching into possible relation for Google Trends and KOSPI100 index.

As the result, Hurst Exponent value is close to 0.5 or under the 0.5. This means that, Hurst Exponent value shows anti-correlation(short-term memory). In addition, Hurst Exponent absolute value is upper than 0.5, less than 0.8. This result means that in the absolute value have long range correlation.

1.1 Report Structure

This paper's construction of total 5 chapter and following this step. A detailed explanation of our background, methods, results, and conclusion.

Chapter 2, introduces the behavioral economics, background to our thesis including social mood and social network service, and Google Trends. This chapter also discusses other related topics in order to accomplish desired our goals.

Chapter 3, we discuss the methods used to linear autocorrelation and linear cross correlations.

Chapter 4, we analyze the results from our work.

Chapter 5, summarized the results. Finally, we finish our thesis by suggesting future work.

Chapter 2

Background

In this chapter we introduce the main idea and theories relating to our work in order to explain our motivation. First, we explain the behavioral economics. theory of social mood and social network service. Next, we continue explain about theory of social mood and social network service. Lastly, let we start by describe one of the concept Google Trends came up with.

2.1 Behavioral Economics

Behavioral Economics is the combination of psychology and economics. Over the years, microeconomics was close to psychology. In addition, behavioral economics is related the real market, behavioral finance. This part is studying the effects of social and emotional factors on the economic life's decision making of individuals and highly important for prices and return. For example, when people do buy or sell's in the market, the affect not only by neighbours but also by news usually realized by an external field. If some people hear bad news, people may be tempted to sell. Thus, the state of any subunit is a function of the states of all the other subunits and of a field parameter [Preis and Stanley. 2010].

Previous behavioral economics work started from cognitive psychology. Cognitive psychology focused on the brain as an information processing device. Afterward, two psychologists whose Amos Tversky and Danieal Kahneman were growing the behavioral economics. They studied uncertainty to economic of rational behavior and their frame is humans are not rational. To be short, behavioral economics close to psychology with the economic (also, finances) to determine when we make rational decision.

2.2 Social Mood and Social Network Service(SNS)

From psychological research point of view, a mood[2)] is an emotional state. Generally, moods have a positive(good) mood or negative(bad) mood valence. Positive(good) mood is people have a positive things when they succeed in something, have had good time, and feel no stress in their life. On the other hand, negative(bad) mood is contrary to the positive(good) mood. For example, people seem to be in the blues, stressful, and anxious about something. Basically, social mood is based on state of mind.

Nofsinger[2005] and Olson[2006] argue that changes in social mood cause people to make different decisions, because people have a different opinion and mood. In addition, psychological state and behavioral aspects of masses representing the social mood. Briefly, the main idea of social mood is collectively shared state of mind. And then, this feeling is usually expressing by using social network service(SNS).

Boyd and Ellison [2008] paper attempted to social network site(SNS) define is as follows:

> "We define social network sites as web-based services that allow individuals to (1) construct a public or semi-public profile within

2) http://en.wikipedia.org/wiki/Mood_(psychology)#Social_mood

a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system. The nature and nomenclature of these connections may vary from site to site."

Social network service(SNS) such as myspace[3], YouTube[4], facebook[5], and twitter[6] have many user and popular services. They have a common thread. For example, above social network service have a individual user web-pages and upload to their own opinion, image, and video.

MySpace is social entertainment and connetcting people to the movies, games, TV, and celebrities. YouTube is video contents, also many of people watch and share originally or created videos. All users can upload up to 15 minutes. Facebook and twitter are most popular micro blogging community in the world and they have a billion active users. On the other hand, there are some different concept between facebook and twitter. Facebook users can create profiles with photos, personal interests, and the other information. Moreover, facebook have a privacy settings and who can see users profile(or status update, picture etc.) or not. Twitter users have public profiles and its users to send and read text based message limit of 140 characters, it is known as "tweet".

Nowadays, those social network is good search data for marketing and finance fields. Many of company thinking about changing their marketing strategy. Billion people using social network service and those social network related to company. Then, social web-sites use pop-up ads for individual companies marketing. At the main pop-up, you see a short video

---

3) http://www.myspace.com
4) http://www.youtube.com
5) http://www.facebook.com
6) http://twitter.com

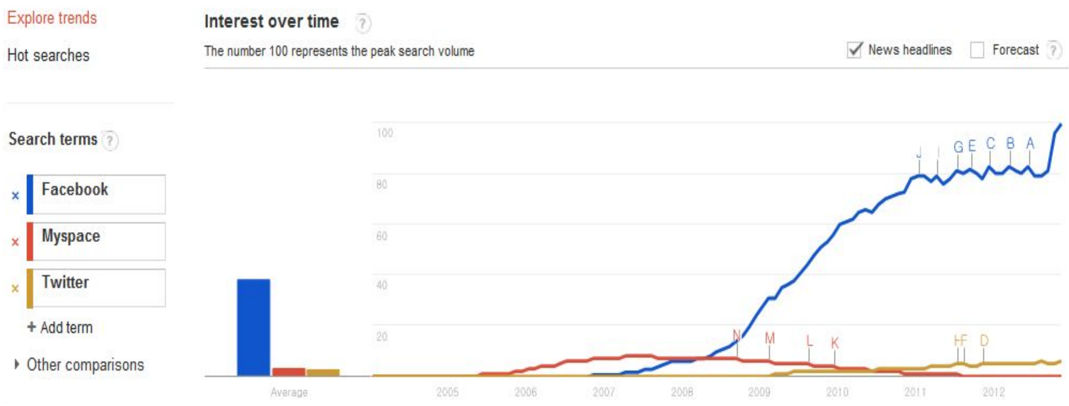
clip or pictures of what is being offered.

Moreover, we will explain about Google Trends such as on the other social network service(SNS) case. Google Trends is a public web search engine by Google. If someone want to know or buy something, they will be using internet search engine. After that, searching information and gather it. At those behavior helps decision making, also makes information aggregation. When you click on the internet search engine(For instance, Google, Yahoo etc.) makes aggregation of data (In this case, Google). It is Google Trends.

2.3 Google Trends

Google Trends[7], when people searching for information by using Google, Google Trends compute how many words or topics have been done for terms they have entered, relative to the total number of searches. Google Trends is can compare the favorite topics and enter up to five topics. Also, Google Trends shows how often they have been searched on Google over time. For example, enter up to the queries [Facebook, Myspace, Twitter] and the result shows compared to the total number of queries search volume index(The graph is search volume index graph). We were focusing on this search volume index also we called number of hitting. We thought, maybe Google Trends and KOSPI100 index's number of hitting has connection. So, we were interested in connection between Google Trends and KOSPI100 index.

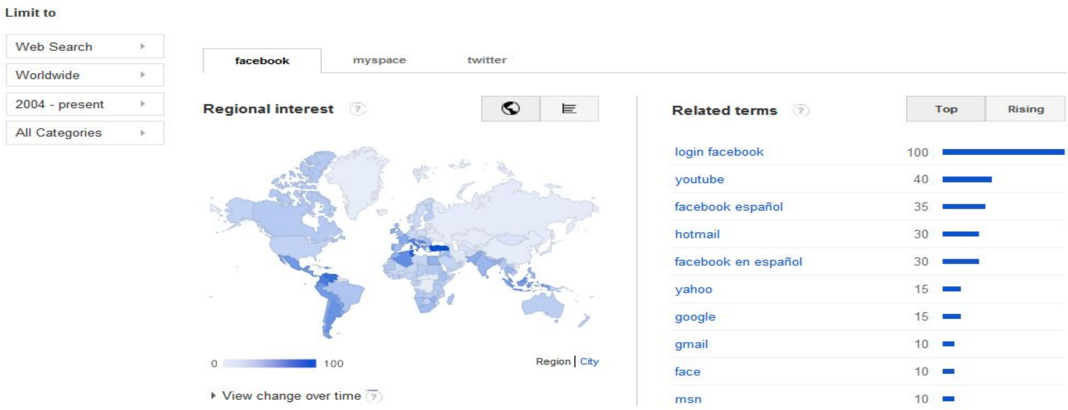
7) http://support.google.com/trends/bin/answer.py



<Figures 2-1> Google Trends output for search index for social network service (Facebook(blue), Myspace(red), Twitter(yellow) worldwide from January 2004 to September 2012. Google trends can compare the interest words and enter up to five topics. On the left side shows enter up to favorite topics and on the right side shows see how often they have been searched on Google over time.
(Source: http://www.google.com/trends/)

At the first query share examined is normalized to be zero(0) and the maximum query share is normalized to be 100. Google Trends all results are normalized and each point on the graph is divided by highest point. If Google do not have data, they will be show zero(0). The search volume index graphs for data start from January 1, 2004 to the current date(in this example, September 29, 2012).

In addition, Google Trends provide for region and related information. According to the Google Trends region map, where is the best cities of searches term. Moreover, when you entered up to search term, you can see a list of top 10 related terms. For example, facebook's related to login facebook, youtube, hotmail, and the others.



<Figures 2-2> Google Trends also shows regional interest and related terms. On the left side shows which cities have interested in facebook, myspace, and twitter. On the other side shows a list of the top 10 rising searches related to that term. (Source: http://www.google.com/trends/)

Additionally, all Google user to download the query index data as a CVS file. If we click on the "Web Search", we can download the CVS file(web, image, news) in just a matter of second. So, we gather the individual companies listed on KOSPI100’s search volume index by using Google Trends. Total KOSPI100 companies and we collected only web CVS files. Moreover, we collected empirical KOSPI100 index from FN guide. After that, compared to Google Trends and empirical KOSPI100 index.

2.4 Literatuar Review

The first Google Trends previous work focused on cancer or influenza-like diseases. Several papers giving a web search data in various fields. One of the first paper was cancer-related topic by using internet search volume [Cooper *et al*. 2005]. Google search volume data used to predict of influenza-like diseases[8] [Polgreen *et al*. 2008; Camille *et al*. 2009; Ginsberg *et al*. 2009; Herman and Elefterios 2009; Valdivis and Monge-Corella 2010].

In economic field work was how to use Google search Insights data to predict several economic field in unemployment, automobile demand, and vacation destinations [Choi and Varian (2009a,b)].

Recently, measuring consumer sentiment using by search data [Huang and Penna 2009; Preis et al. 2010] and McLaren and Shanbhoge [2011] are explain about how web search data can be used for economic nowcasting.

Finally, Tobias [2010] worked was motivate this paper's work. This paper work the smallest or largest possible scale of our economic life used search volume data provided by the Google Trends.

8) Google flu trends (http://www.google.org/flutrends/intl/en_us/)

Chapter 3

Data and Methods

In this chapter 3 we describe the data and methods. First, we will be explain about data selection, and then we calculated those data using method by autocorrelation, detrended fluctuation analysis(DFA), and hurst exponent. Before methods, we will introduce about how can we collect the KOSPI100 data from empirical price and Google Trends KOSPI100 hitting number.

3.1 Data

The first step in gathering KOSPI100 index(weekly) from January, 2004 to September, 2012 by Fn guide. This list include the exchange trading symbols and the company names(symbols name). Next, we were collecting KOSPI100 search volume index(weekly) from Google Trends, which is available for the same period of time. We use all 100 company names of the KOSPI100 components. Enter up to Korean.

We were gathering total KOSPI100 index each other(Fn guide and Google Trends) and picked up only 22 individual companies index, because we did not count hitting number zero(0). No hitting is no influential. The total 22 individual companies are 기업은행(Industrial Bank of Korea: IBK), 대한항공(Korean Air), 두산(Doosan), 삼성증권(Samsung Securities), 삼성화재(Samsung fire & Marine insurance), 신세계(Shinsegae), 외환은행(Korea Exchange Bank: KEB), 한국전력(Korea Electric Power Corporation: KEPCO), 한화(Hanwha), 현대

백화점(Hyundai Department Store), 효성(Hyosung), CJ, KCC, KT, LG, LG전자(LG Electronics), LS, NHN, OCI, POSCO, SK, and SK텔레콤(SK Telecom).

3.2 Methods

This section describes the definition of autocorrelation and use of the autocorrelation function(ACF). After that, we will describes detrended fluctuation analysis(DFA) and hurst exponent.

3.2.1 Definition of Autocorrelation

Autocorrelation[9] is the cross-correlation of a signal with itself. A mathematical representation of the degree of similarity between a given time series and a lagged version of itself over successive time intervals. It is the same as calculating the correlation between two different time series, except that the same time series is used twice - once in its original form and once lagged one or more time periods. As the results, the resulting number can range from -1 to +1. A value of -1 represents perfect negative correlation and +1 represents perfect positive correlation(i.e. -1 is increase seen in one time series results in a proportionate decrease in the other series and +1 is one time series will lead to a proportionate increase in other time series).

9) Definition from Wikipedia (http://en.wikipedia.org/wiki/Autocorrelation).

3.2.2 Statistics of Autocorrelation

The autocorrelation function [Box and Jenkins. 1976] can be used for the following two purposes:

1. To detect non-randomness in data.
2. To identify an appropriate time series model if the data are not random.

The Autocorrelation of a random process depicts the correlation between values of the two times or time difference. The autocorrelation function of a time series z for lag k is defined as:

$$\rho_k = \frac{E[(z_t - \mu)(z_{t+1} - \mu)]}{\sigma^2} \tag{1}$$

where, E is the expectation operator

z_t is value of the time series at time t

μ is mean

σ^2 is variance.

A common estimate of the autocorrelation function is:

$$r_k = \frac{\frac{1}{N}\sum_{t=1}^{N-k}[(z_t - \bar{z})(z_{t+k} - \bar{z})}{\frac{1}{N}\sum_{t=1}^{N}(z_t - \bar{z})^2} \tag{2}$$

where $\bar{z} = \frac{1}{N}\sum_{t=1}^{N} z_t$, and where the lags are $k = 0, 1, 1 \ldots K$, and K is $\leq N-1$.

3.2.3 Detrended Fluctuation Analysis(DFA) and Hurst Exponent

Detrended fluctuation analysis(DFA) is a method for determining the statistical self-similar process of a signal. Detrended fluctuation analysis (DFA) was introduced by Peng et al. [1994] and extension of the fluctuation analysis(FA). They have been established as an important tool for the detection of long-memory correlations(autocorrelation) in time series with non-stationary. Detrended fluctuation analysis(DFA) is useful for long-memory processes. In this section, we describe method of Detrended fluctuation analysis(DFA) 3 steps[10].

In the first step, accrue to subtraction of the mean with formula(3.3).

$$Y(k) = \sum_{t=1}^{N} (X_t - X_{mean}) \qquad (3)$$

where, $Y(k)$ is cumulative sum or profile

X_{mean} is mean value of time series

$t, k = 1, 2, \cdots, N$

This cumulative process is convert to self-similar process from original data.

10) 오갑진, 엄철준, 김승환(2004), "한국주식시장의 장기기억상관성: DFA방법을 중심으로", 金融工學硏究, 제3권 제2호, pp.137-138.

Next step, the integrated time series $Y(k)$ is divided into boxes of equal size, n. In each box of length n, we fit the time series by using a least squares(the local trend), $y_n(k)$. The integrated time series, $Y(k)$ by subtractiong the $y_n(k)$ in each box.

$$Y(k) = y(k) - y_n(k). \tag{4}$$

Given box size n, calculate the root mean square fluctuation,

$$F(n) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(Y(k) - y_n(k))^2} \tag{5}$$

where, $Y(k)$ is integrated time series that is removed the whole trend

$y_n(k)$ is integrated time series that is removed the local trend

The last step is we must repeat steps 1 to 2 for several time scales. Then, established the formula (6).

$$F(n) = cn^H \tag{6}$$

where, n is scale

C is constant

H is Hurst exponent.

$$\log_{10} F(n) = \log_{10} C + \log_{10^n} \tag{7}$$

where, H(Hurst exponent) is used as a measure of
long term memory of time series.

Hurst exponent(H) is a statistical measure used to classify time series. The Hurst exponent proposed by H. E. Hurst for use in fractal analysis. The Hurst exponent range between 0 and 1. H = 0.5 indicates a random series. This means that, in a random series is no correlation. $0<H<0.5$ indicates an anti-correlated(short-term memory). This means that, an increase will tend to be followed by a decrease. $0.5<H<1$ indicates a correlated(long-term memory). The time series is trending and the larger Hurst Exponent is stronger than trend.

Chapter4

Results

In this section, we analyze the fundamental mechanism of KOSPI100 individual companies(= Total 22 companies)(Table 1) and Google Trends(Table 2). This KOSPI100 index(weekly) collect date from January, 2004 to September, 2012 using by Fn Guide. Also, KOSPI100 search volume index(weekly) from Google Trends, collect date from January, 2004 to September, 2012. Table 1 show, the statistical properties of KOSPI100 of 22 individual companies mean, standard deviation, skewness, and kurtosis values. In addition, table 2 shows, the statistical properties of Google trends of 22 individual companies mean, standard deviation, skewness, and kurtosis values. Table 3 shows, the total average value of KOSPI100 individual companies and Google Trends.

<Table 1> The statistical properties of 22 KOSPI 100 individual companies.

Company	Mean	Standard deviation	Skewness	Kurtosis
Industrial Bank of Korea	13,721	4026.068	-0.3966	-1.0034
Korean Air	43,817	19227.32	0.1522	-1.0423
Doosan	92,425	60908.12	0.1446	-0.7977
Samsung Securities	54,912	19073.77	-0.2010	-0.1949
Samsung fire & Marine insurance	167,376	55761.94	-0.5724	-1.0247
Shinsegae	306,140	76449.7	-0.0983	-0.8134

Korea Exchange Bank	10,770	2911.913	-0.1369	-1.3206
Korea Electric Power Corporation	30,987	6704.44	0.0982	-0.9133
Hanwha	34,362	15809.22	0.3219	0.3172
Hyundai Department Store	95,550	41152.21	0.2881	-0.5633
Hyosung	50,474	31468.73	0.1216	-1.2051
CJ	58,692	17988.9	-0.0041	-1.1835
KCC	292,682	112767.4	0.4017	0.2853
KT	40,808	4,575	0	0
LG	50,060	23824.45	-0.1114	-1.2264
LG Electronics	82,925	22764.18	0.7001	-0.3044
LS	66,954	34033.49	-0.1448	-1.3603
NHN	146,374	73466.66	-0.2945	-1.0854
OCI	176,088	135806.8	0.4252	-0.5219
POSCO	369,510	138084.3	-0.0765	-1.1451
SK	112,557	40956.09	0.8406	0.5482
SK Telecom	184,101	25431.06	-0.0310	0.1277

<Table 2> The statistical properties of Google Trends.

Company	Mean	Standard deviation	Skewness	Kurtosis
Industrial Bank of Korea	25.6644	14.6723	2.7426	8.1474
Korean Air	46.3245	14.5460	0.6777	0.3204
Doosan	12.2543	9.6540	4.8343	34.5797
Samsung Securities	38.8486	19.4051	0.4101	0.2624
Samsung fire & Marine insurance	26.4451	17.6916	1.2355	1.9799
Shinsegae	25.6381	14.4615	1.7486	3.7706

Korea Exchange Bank	47.8464	17.7414	0.6366	-0.1453
Korea Electric Power Corporation	24.1184	14.4232	1.4541	3.3141
Hanwha	12.9254	8.5620	3.2699	24.2174
Hyundai Department Store	17.0263	10.2021	1.9257	9.2134
Hyosung	15.7609	10.3576	2.2506	11.5591
CJ	62.9276	8.1224	0.9294	1.2816
KCC	75.0504	7.8362	-0.1534	0.9872
KT	53.9714	13.6766	1.2152	0.8162
LG	58.8728	10.7061	0.4583	0.3630
LG Electronics	34.9100	16.4522	0.8324	0.0782
LS	72.6769	8.6311	0.2305	-0.8826
NHN	24.7785	7.7738	2.9667	22.2761
OCI	63.5087	13.6597	0.2073	-1.1264
POSCO	50.5416	11.6257	0.5173	1.8416
SK	72.375	11.6323	0.2000	-1.0846
SK Telecom	28.3026	14.4433	1.8572	4.1877

<Table 3> The total statistical properties of 22 KOSPI 100 individual companies and Google Trends.

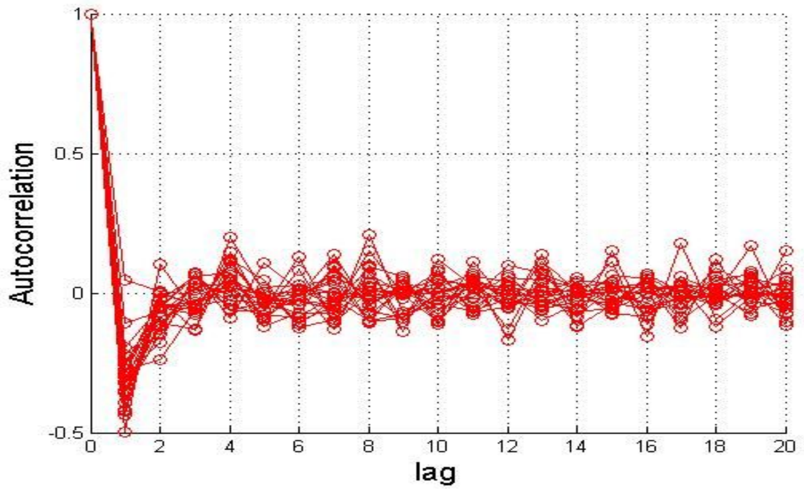
	The total statistical properties	
	KOSPI100	Google Trends
Mean	112785.7	40.4894
Standard deviation	40689.32	3.4985
Skewness	0.5263	1.2145
Kurtosis	-0.8419	3.4925

Firstly, Google Trends hitting data consists of positive number, because these data made up of the number of hitting. In addition, we investigate the effects of the individual companies Google Trends flow. In this case, Google Trends time series is taking difference and check out the movement. The reason is that, statistical means relating to the movement.

In this paper, there are two view points of the individual companies key feature. The first is Hurst Exponent. The Hurst Exponent(H value) range between 0 and 1. A H value close to 0.5 followed by a random series, H value range between 0 and 0.5 followed by anti-correlation, and 0.5 and followed by correlation. H value 0.5 means that, if before hitting number is rise, the rise and decrease probability will be each one in two(1/2). More over, a H value is smaller than one in two(1/2), if before rise, it will be decreasing probability is large and a H value is bigger than one in two(1/2), rise probability will be more large.

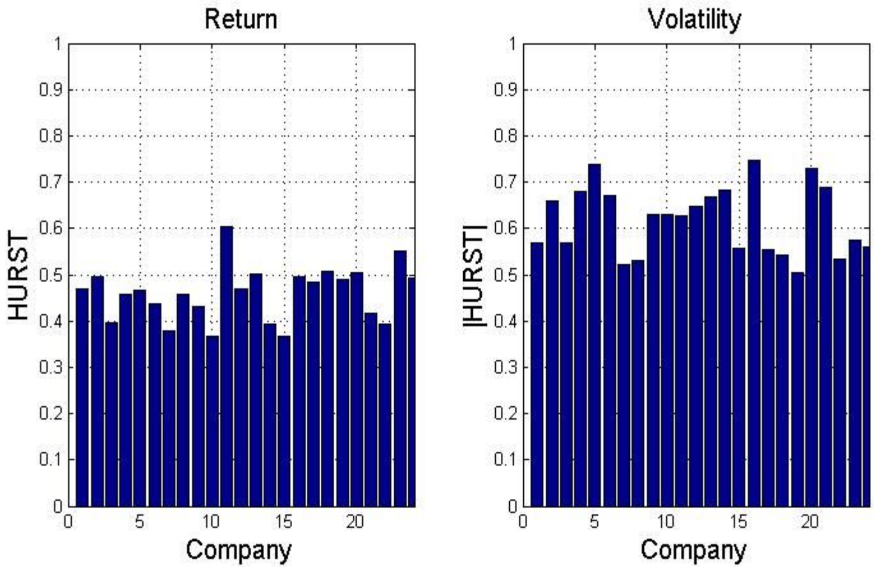
In this work, the reason that calculated of the Hurst Exponent is we looking for the pattern of the movement. Also, how to change the hitting number, going up or not(random).

The first figure shows, KOSPI100 individual companies autocorrelation. The autocorrelation of a random process depicts the correlation between values of the two times or time difference. This figures shows lag 0's autocorrelation is 1. This is a natural result, because self value relating to self value. These rules also reflect the corollary of that principle, viz. Also, this result shows, anti-correlation.



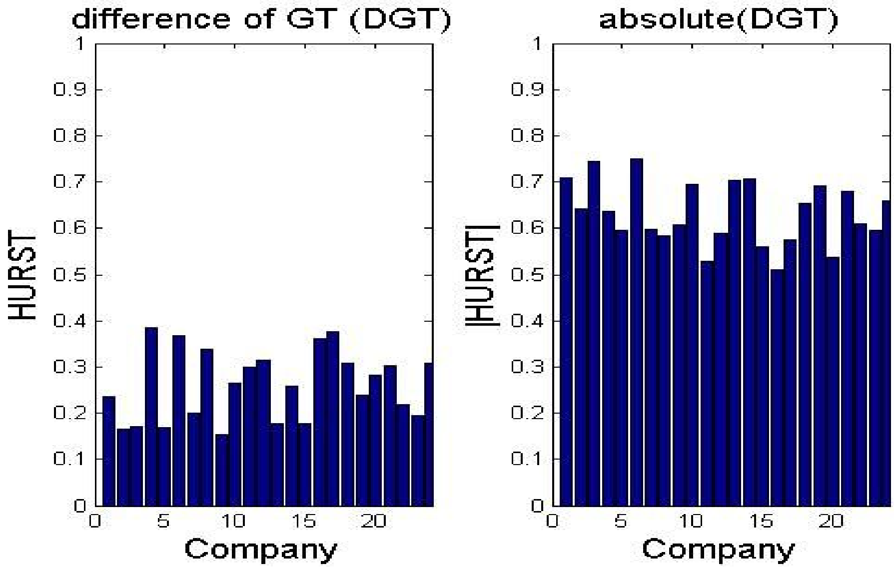
<Figures 4-1> KOSPI100 individual companies autocorrelation.

Next, we analyze Detrended fluctuation analysis(DFA) using by formula (3), (4), (5), and (6). In our work, we investigate the use of the Hurst Exponent to classify series of Google Trends and KOSPI100 index data representing different periods of time. Actually, we were gathering KOSPI100 data from Fn Guide and Google Trends. KOSPI100's company are total 100. But we use to only 22 companies, because we did not count hitting number 0. As the result, figure (4-2) shows, KOSPI100 individual companies Hurst Exponent value of return and volatility.



<Figures 4-2> KOSPI100 individual companies Hurst Exponent value of return and volatility.

Hurst Exponent of return value close to 0.5. This return value is anti-correlation(short-term memory). Also, we measure that volatility, because of clustering effect. Volatility value is upper than 0.5, less than 0.8. This result means that in the absolute value have long range correlation.



<Figures 4-3> Google Trends Hurst Exponent value and Hurst Exponent absolute value.

The figures (4-3) shows, Google Trends Hurst Exponent about difference of Google Trends and difference of absolute value. Difference of Google Trends Hurst Exponent value is less than 0.4. These means that, this Hurst Exponent value shows anti-correlation(short-term memory). Google Trends absolute value is upper than 0.5, less than 0.8. This result means that in the absolute value have long range correlation.

Chapter 5

Conclusion

In this paper, we main idea start from Social Network Service(SNS). We interested in relate to connecting between KOSPI100 index and Google Trends. Nowadays, Google Trends is one of powerful social network service. Also, our data gathering from one of part of social network in Google Trends. Next, we explain about behavioral economics, Social Network Service(SNS) and social media, and Google Trends.

More over, we were computing data to KOSPI100 and Google Trends using by autocorrelation, Detrended fluctuation analysis(DFA) and Hurst Exponent value. Normally, Hurst Exponent value is close to 0.5 or under the 0.5. This means that, Hurst Exponent value shows anti-correlation(short-term memory). In addition, Hurst Exponent absolute value is upper than 0.5, less than 0.8. This result means that in the absolute value have long range correlation.

These days, many of people using social network and many researcher analyzed this network system. These studies will continue into the future also. For example, predicting the stock market or some events are take an accurate measurement. We are hoping for good basic studies on this paper.

Reference

오갑진, 엄철준, 김승환(2004), “한국주식시장의 장기기억상관성: DFA방법을 중심으로”, 金融工學硏究, 제3권 제2호, 135-145.

About Google Trends, 2012. http://support.google.com/trends/?hl=en

About Autocorrelation. http://en.wikipedia.org/wiki/Autocorrelation

Boyd, D. M. and Ellison, N. B. (2008). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, 13(1):210-230.

Box, G. E. P., and Jenkins, G. (1976), Time Series Analysis: Forecasting and Control, Holden-Day. ISBN 0-8162-1104-3.

Box, G. E. P., G. M. Jenkins, and G. C. Reinsel. Time Series Analysis: Forecasting and Control. 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 1994.

Brooks, D. (2009). Twitter Marketing Tips. *Emereo Pty Limited*, page 200.

Choi, H. and Varian, H. (2009a). Predicting the Present with Google Trends. Technical report, Google. [Cited 1 April 2012.] Available from: http:// google.com/googleblogs/pdfs/google_predicting_the_ present.pdf.

Choi, H. and Varian, H. (2009b). Predicting Initial Claims for Unemployment Insurance Using Google Trends. Technical report, Google. [Cited 1 April 2012.] Available from: http://research.google.com/ archive/papers/initialclaimsUS.pdf.

Choi, H. and Varian, H.(2012), "Predicting the present with google trends", *THE ECONOMIC RECORD*, 88, 2-9.

Cooper, C., Mallon, K., Leadbetter, S., Pollack, L. and Peipins, L. (2005), "Cancer Internet Search Activity on a Major Search Engine, United

States 2001-2003", *Journal of Medical Internet Research*, 7. [Cited 1 April 2012.] Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1550657/.

Earle, P. (2010). Earthquake Twitter. *Nature Geoscience*, 3(4):221--222.

Ginsberg, J., Mohebbi, M.H., Patel, R.S., Brammer, L., Smolinski, M.S. and Brilliant, L. (2009), "Detecting Influenza Epidemics Using Search Engine Query Data", *Nature*, 1012-14. [Cited 1 April 2012.]

Gunn, S., Brown, M., and Bossley, K. (1997). Network performance assessment for neurofuzzy data modelling. *Data Analysis. Reasoning about Data*:, (Lecture Notes in Computer Science.):313-323.

Herman, A. C., and Eleftherios, M. (2009). Google Trends: A Web-Based Tool for Real-Time Surveillance of Disease Outbreaks. *Clin Infect Dis*. (2009) 49 (10): 1557-1564.

Hu, K., Ivanov, P. C., Chen, Z., Carpena, P., and Stanley, H. E. (2001). "Effect of trends on detrended fluctuation analysis". *Phys Rev* E 64(1): 011114.

Huang, H. and Penna, N.D. (2009). Constructing Consumer Sentiment Index for U.S. Using Google Searches, Technical report, University of Alberta. [Cited 1 April 2012.] Available from: http://econ papers.repec.org/paper/risalbaec/2009_5f026.htm.

Huberman, B., Romero, D., and Wu, F. (2009). Social networks that matter: Twitter under the microscope. *First Monday*, 14(1):8.

Hurst, H. E.(1951). "Long Term Storage Capacity of Reservoirs", *Transactions of the American Society of Civil engineers*, 116, 770-799.

Johan, B., Huina Mao, and Xiao-Jun Zeng (2011), "Twitter mood predicts the stock market", *Journal of Computational Science*, 2, 1-8.

Kantelhardt, J. W., Zschiegner, S. A., Koscielny-Bunde, E., Havlin. S., Bunde, A., and Stanley. H. E(2002), "Multifractal detrended fluctuation analysis of nonstationary time series", *Phsica A: Statistical Mechanics and its Applications*, Volume 316, Issues 1-4, Pages 87-114.

Jansen, B. J. and Zhang, M. (2009). Twitter Power : Tweets as Electronic Word of Mouth. *Journal of the American Society for Information Science*, 60(11):2169–2188.

Kahneman, Daniel and Tversky, Amos (1979). "Prospect Theory: An Analysis of Decision under Risk". *Econometrica (The Econometric Society)* 47 (2): 263-291.

Krishnamurthy, B., Gill, P., and Arlitt, M. (2008). A few chirps about twitter. *Proceedings of the first workshop on Online social networks*, pages 19–24.

Mowbray, M. (2010). The twittering machine. *Journal of Applied Statistics*, 17(2):211–217.

McLaren, N. and Shanbhoge, R. (2011). Using Internet Search Data as Economic Indicators, Bank of England Quarterly Bulletin. [Cited 1 April 2012.] Available from: June http://www.bankofengland.co.uk/ publications/quarterlybulletin/qb110206.pdf.

Nofsinger, J.R. (2005). "Social Mood and Financial Economics", *Journal of Behavioural Finance*, 6

Olson, K.R. (2006). "A literature Review of Social Mood", *Journal of Behavioral Finance*, 7

Pearson, K. (1985). "Contributions to the mathematical theory of evolution II: skew variations in homogeneous material. *Phil. Trans. R. Soc. Lond. A* 186, 343–414.(doi:10.1098/rsta.1985.0010)

Pelat, C., Turbelin, C., Bar-Hen, A., Flahault, A., and Valleron A. (2009) More Diseases Tracked by Using Google Trends. *Emerg Infect*

Dis. 2009 Aug; 15(8): 1327–1328.

Peng, C. K., S. V. Buldyrev, S. Havlin, H. E. Stanley, and A. L. Goldberger. (1994). "Mosaic organization of DNA nucleotides". *Phys Rev* E 49: 1685-1689.

Polgreen, P.M., Chen, Y., Pennock, D.M. and Nelson, F.D. (2008), "Using Internet Searches for Influenza Surveillance", Clinical Infectious Diseases, 47, 1443–8. [Cited 1 April 2012.]

Preis, T. and Stanley, H. E. (2010). Switching phenomena in a system with no switches. *J. Stat. Phys.* 138, 431446. (doi:10.1007/s10955-009-9914-y)

Ritterman, J., Osborne, M., and Klein, E. (2009). Using Prediction Markets and Twitter to Predict a Swine Flu Pandemic. *Forecast*, (2004).

Sendhil M., and Richard H.T. (2000). Behavioral Economics. *Working paper 7948* *http://www.nber.org/papers/w7948*

Valdivia, A. and Monge-Corella S. (2010). Diseases tracked bu using Google trends, Spain. *Emerg Infect Dis*. 2010 Jan;16(1):168.

Appendix

List of KOSPI100 Companies

유한양행; CJ대한통운; 두산; 대림산업; 한국타이어월드와이드; 기아차; SK하이닉스; 현대건설; 삼성화재; 삼성물산; 한화; CJ; LG상사; 동국제강; 제일모직; SK네트웍스; 오리온; KCC; 아모레G; 현대증권; 대한항공; LG; SK; 삼성정밀화학; 현대제철; 신세계; 농심; 효성; 외환은행; 롯데제과; 롯데칠성; 현대차; POSCO; 삼성전자; 우리투자증권; LS; GS건설; 삼성SDI; 대우증권; 삼성전기; 현대중공업; 현대케미칼; OCI; LS산전; 고려아연; 삼성중공업; 현대하이스코; 현대미포조선; S-Oil; LG이노텍; 호남석유; 현대상선; 현대모비스; 삼성테크원; 현대산업; 에스원; 한국전력; 삼성증권; SK텔레콤; 한라공조; 웅진코웨이; 롯데쇼핑; 기업은행; 삼성엔지니어링; STX팬오션; 삼성카드; 제일기획; KT; LG유플러스; 삼성생명; KT&G; 두산중공업; LG디스플레이; SK C&C; 강원랜드; NHN; 한국가스공사; 엔씨소프트; 미래에셋증권; 대우조선해양; 두산인프라코어; 대우건설; 대우인터내셔널; LG생활건강; LG화학; 한전기술; 우리금융; 신한지주; LG전자; 현대백화점; 한국금융지주; GS; 현대글로비스; 하나금융지주; 한화생명; 아모레퍼시픽; SK이노베이션; CJ제일제당; KB금융; BS금융지주