Scientific Research

Volatility Analysis of Web News and Public Attitude by GARCH Model

Pinrui Yu[1], Tianzhen Liu[2], Qian Ding[2*]

[1]School of Journalism and Communication, Southwest University, Chongqing, China
[2]School of Urban Design, Wuhan University, Wuhan, China
Email: *qianding2011@yeah.net

GARCH (Generalized Auto-Regressive Conditional Heteroskedasticity) model proposed by Professor Engle is successful to analyze the volatility of stock price. In this paper GARCH model is used to analyze the volatility of web news events and public attitudes by the data coming from typical news events in famous web. The results show that the volatility of web news events and public attitudes are suitable to GARCH model by some adjusting and test of parameters.

Keywords: Web News; Public Attitude; Volatility Analysis; GARCH Model

Introduction

Social hot news spread rapidly in internet with abundant information. The news may cause psychological reactions of public with some attitudes or moods: delight, moved, sympathy, angry, funny, sad, novel, sweat, and so on as **Figure 1**. The summaries of these attitudes are also called as "public attitudes". To reveal the relationship between public events and public attitudes or opinions, it is necessary to make quantitative analysis, especially volatility analysis by webometrics (Tomas et al., 1997).

There are some quantitative researches on public attitudes of social hot news in website. Researchers get the data from web forums and news website, and analyze the evolvement of public attitudes in these news examples by statistical modeling. Xie (2006) proposed ten analysis modes such as hot sports, emphasis, focus, sensitive points, frequency points, etc. Wei (2006) divided evolvement modes of crisis events into four categories by statistical distributions: exponential type, normal type, Poisson type, fluctuation type. Qi (2008) proposed an evaluation method of public attitudes by ARMA model and regression model. Lu (2010) explored the rule of spread of public attitude by variation number, correlation analysis, Tobit model. Sun (2011) built monitor and analysis model of spread of public attitude in web by agent. In fact, attitude problems in web belong to psychological problems which have been studied using mathematical models successfully (Tong, 2010; Lin, 2011).

The public attitudes occurred by emergencies have volatilities with irregular variance similar to the prices of many stocks. Professor Engle (1982) proposed Auto-Regressive Conditional Heteroskedasticity (ARCH) model to analyze the volatility of stock price successfully and obtained Nobel Prize in Economics in 2003. His student Bollerslev (1996) proposed Generalized ARCH (GARCH) model and developed the thought of Engle. Almost there is not any paper to use GARCH model to analyze the volatility of public attitudes occurred by web news. So we probe the method using GARCH model to analyze the volatility of public attitudes occurred by web news.

The Volatility Characters of Public Attitudes in Web

In general, an evolvement process of public attitudes in web has three stages: initial stage, rapid spread stage, and subsiding stage, similar to the process of the life of a man [11]. **Figure 1** shows the three stages of post amount about "Wu-Kan event" in the "Baidu Web" (**Figure 2**). In the initial stage, the post amount is small and can not distinguish from common case easily. In the second stage, the post amount increase rapidly, and if we guide the opinions in time we can control the events correctly. We notice that the subsiding process is not smooth but may have some small waves in this stage.

How about your mood after reading this news?
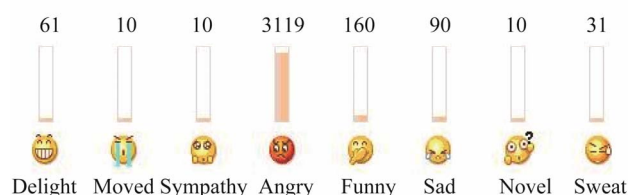
Figure 1.
Survey figure for mood of news events in Tencent Web.
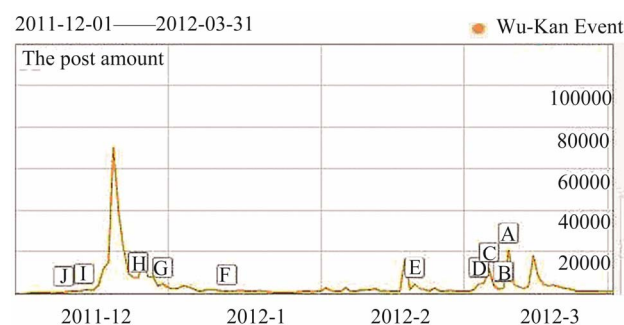
Figure 2.
Three stages of post amount about "Wu-Kan event".

---

*Corresponding author.

The post amount of news events in web is a time series, and has some characters as follows.

1) Clustering. This means the volatility of post amount is not single. A big wave is always followed many big waves, and a small wave is always followed many small waves. These phenomena show that there are herd mentalities in web hot events.

2) Sharp leptokurtic with fat tail in the density curve. We can draw the density curve of post amount, and find it does not appear as normal distribution, but has sharp leptokurtic with fat tail similar to the density curve of stock prices. In general, an emergency outbreaks suddenly and its affection dissipates slowly, so its density curve has a single fat tail on the right.

3) Leverage effect. In the spread process of news information in web, positive information or negative information may affect volatility of public attitudes greatly. Especially, the negative information may affect volatility of public attitudes more seriously.

4) Long Memory. The autocorrelation function of the series of volatility of public attitudes decreases very slowly. That is, the autocorrelation function between two sub-series with long distance is significant yet. This means a history event may affect public attitudes for a long time.

These characters can not be explained by ARMA model or Random Walk model, but can be explained by ARCH model and GARCH model easily.

Suppose y_t is a time series, ARCH(q) model can be expressed as following formula:

$$y_t = X_t\beta + \varepsilon_t \tag{1}$$

$$\varepsilon_t = \sqrt{h_t} \cdot e_t \tag{2}$$

$$h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \cdots + \alpha_q\varepsilon_{t-q}^2 \tag{3}$$

where X_t is a explanatory variable, β is a coefficient vector, ε_t is the error term with mean 0, but his variance is not constant. It is interesting the hypothesis about ε_t in ARCH model is too complex with (2) and (3), where $E(e_t) = 0$, $D(e_t) = 1$. To guarantee the second-order stability of series y_t, it is necessary to demand all characteristic roots of $h_t = 0$ must be outside the unit circle.

GARCH (p, q) model only generalizes (3) in ARCH(q) model as (4):

$$h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \cdots + \alpha_q\varepsilon_{t-q}^2 + \beta_1 h_{t-1} + \cdots + \beta_p h_{p-1} \tag{4}$$

In GARCH (p, q) model, when $p = 0$, the model becomes common ARCH (q) model. When $q = p = 0$, ε_t is a common white noise, and the model becomes common linear regression model. In general, GARCH (1, 1) or GARCH (2, 2) are sufficient to fit and forecast volatilities of public attitudes. More details about GARCH models can be found in "Developing Econometrics", Chapter 8, Multivariate and Nonstationary Time Series Models (Tong et al., 2011).

Volatility Analysis of Public Attitudes by GARCH Model

We select the hottest social event in 2011, Wu-Kan event, as research object to analyze its volatility of public attitudes by GARCH model.

Wu-Kan event is a typical conflict between local government and the people involving economic interest. In September 2011, the leader of Wu-Kan villages sold the land to developer and got a huge profit, but the people only obtained a few benefits. Economic interests lead to political conflict for three months. Serious political conflict caused attention of senior leadership. At last, the conflict is satisfactorily resolved, and the event subsides from line of sight of public gradually.

We collect the data from December 1, 2011 to March 31, 2012 from Baidu Web by keyword "Wu-Kan event", and obtain 121 numbers.

We use four models, GARCH (1, 1), GARCH (1, 2), GARCH (2, 1), GARCH (2, 2), to compare their fitting effects by the significant of test. The selected model at last has the minimum value of AIC and SC criterion and the maximum value of likelihood. The results of calculation are in **Table 1**.

Comparing with the results in **Table 1**, we select model GARCH (1, 2) as the optimal model. The parameters of GARCH (1, 2) are calculated as follows:

$$R_t = 0.7584 + 0.1334R_{t-1} + \varepsilon_t$$

Table 1.
The statistics of GARCH models.

	GARCH (1, 1)	GARCH (2, 1)	GARCH (1, 2)	GARCH (2, 2)
AIC	−.5967	−.7025	−.6062	−.4873
SC	−.5028	−.5895	−.4933	−.3557
Log	52.4909	62.1998	54.2613	45.4588
R_2	−.5907	.0233	.1623	.1858
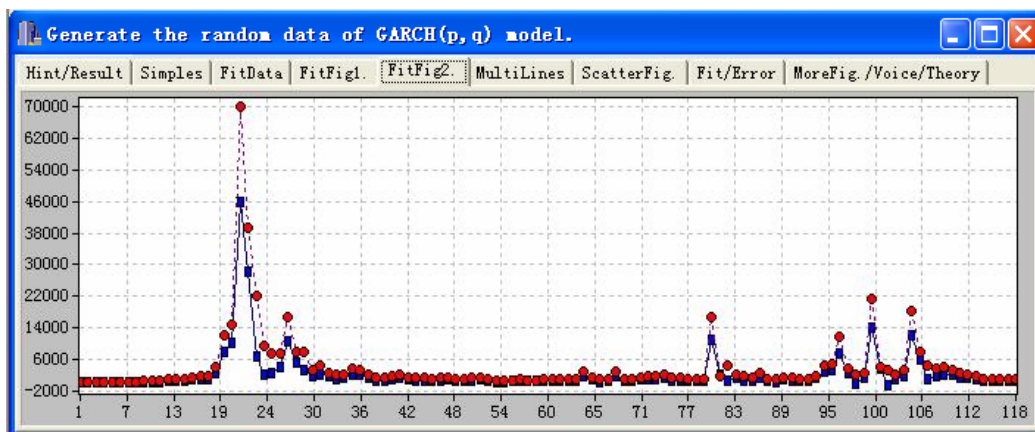
Figure 3.
The fitting effect of model GARCH (1, 2).

$$h_t = 0.0014 + 0.0012\hat{\varepsilon}_{t-1}^2 + 0.01532\hat{\varepsilon}_{t-2}^2 + 0.8123\hat{h}_{t-1}^2$$

Figure 3 shows the fitting effect of GARCH (1, 2) with the original data, which is calculated by the software DASC (Data Analysis & Statistical Computation, Tong, 2005).

It is important to capture the big volatility in analysis of the public attitude in web news events. This is good for the forecasting and pre-waning of emergency. From **Figure 3** we are satisfied with model GARCH (1, 2).

We can see the affections of the rate of change by the new information from the coefficients of model GARCH (1, 2). Besides, the coefficient (.8123) reflects long memory in system. At last, the sum of three coefficients is less than 1 which shows the system is stable.

Conclusion

Advanced statistical tools in modern network provide convenient conditions for our access to the Internet public opinion data. Our work shows that the volatilities of public attitudes in web have the character of heteroskedasticity, so it is suitable to use GARCH model. Further more, adjusting and test parameters of model is important to obtain good fitting effect, and the fitting figure is good tool to distinguish the effect of the model.

The work of this paper is only to focus on the basic data of public attitude. More deep research should focus on sensibility analysis by text information, which is more complex and interesting.

Acknowledgements

REFERENCES

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31,* 307-327. doi:10.1016/0304-4076(86)90063-1

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of variance of UK inflation. *Econometrica, 50,* 987-1008. doi:10.2307/1912773

Lin, H., Zhang, Y. B., Tong, H. Q., Tong, Q. L., & Liu, D. S. (2011). Study on psychological crisis evaluation combining factor analysis and neural networks. *Psychology, 2,* 138-142. doi:10.4236/psych.2011.22022

Lu, T. J. (2010). Research on the crisis information communication and response method in the internet. Master's thesis, Beijing: University of Science and Technology of China.

Pan, C. X. (2011). Analysis of evolvement stages of public opinion in web. *Computer and Modernization, 10,* 203-206.

Qi, H. F. (2008). Research on hot topic detection and event tracking in network public attitude. Master's thesis, Harbin: Harbin Engineering University.

Sun, B. Q., & Dong, J. W. (2011). Monitoring model on diffusion law identification of network public attitude for major public crisis. *Journal of Harbin Institute of Technology(Social Sciences Edition), 13,* 92-97.

Tong, H. Q. (2005). *Data analysis & statistical computation software (DASC)*. Beijing: Science Press.

Tong, H. Q., Krishna, K. T., & Huang, Y. X. (2011). *Developing econometrics*. Chichester: John Wiley & Sons. doi:10.1002/9781119954231

Tong, Q. L., Zou, X. C., Gong, Y., & Tong, H. Q. (2010). Analysis of psychological health and life qualities of internet addicts using structural equation model. *Psychology, 1,* 22-26. doi:10.4236/psych.2010.11004

Tomas, C. A., & Peter, I. (1997). Informetric analyses on the world wide web: Methodological approaches to "webometrics". *Journal of Documentation, 53,* 404-426.

Wei, J. C. (2006). Research on analysis and evaluation of social affects of crisis events. Doctor thesis, Beijing: University of Science and Technology of China.

Xie, H. G., & Chen, Z. R. (2006). Internet information and the pattern of public attitude analysis in depth. *Journal of China Youth University for Political Sciences, 25,* 95-100.