Memory-Forgetting Curve based on Virtual and Real Spaces for Commercial Recommendation

Yunlan Xue[#1], Lingyu Xu[#2], Jie Yu[#3], Lei Wang[#4], Gaowei Zhang[#5]
[#]School of Computer Engineering and Science, Shanghai University
[1]xueyunlan@i.shu.edu.cn
[2]xly@shu.edu.cn
[3]jieyu@shu.edu.cn
[4]wanglei727@shu.edu.cn
[5]yiqigo0215@163.com

ABSTRACT

With the social media development, semantic web and Crowdsensing blog platforms become the store of big data, which store the memory of diverse information. Memory enables past experiences to be remembered and acquired as useful knowledge to support decision making in commercial recommendations, especially when perception and computational resources are limited in Crowdsensing. However, the traditional event research lacks the view of memory and forgetting of event from the people’s cognitive psychology. On the basis of Ebbinghaus, we treat every day’s information as an independent event to calculate Memory-Forgetting Curve (MFC), then calculate MFC of episode with the fusion of events’.

Keywords

social media; Crowdsensing; event; episode; memory-forgetting curve; information influence.

1. INTRODUCTION

The general quantitative description of memory and forgetting curve is one of experimental psychology’s the oldest problems was by Ebbinghaus[1], the nonlinear function relating the probability of memory retention (R) and the forgetting between study and test (t). He found that forgetting begin immediately after the study, and the process of forgetting is not uniform. The speed of forgetting becomes more and more slowly from the beginning.

Memory plays a key role in decision making by providing past relevant event information to improve learnt knowledge[2]. For instance, the episodic memory to improve task performance in simulated environments[3]. How memory can improve the realism of virtual agents in some search [4, 5, 6, 7]. Memory as a flexible information storage can retrieve and store information. When people switch between thoughts, the incoming thought and the outgoing thought will be co-active (as the incoming thoughts activation (memory) is rising and the outgoing thought’s activation (forgetting) is falling), resulting in competition between memory and forgetting[8].

A significant amount of computational resources may still be needed to process specific items in memory to support reasoning and decision making, although memory can provide useful information about previous experiences. As the tasks and the environment become more complex, it is often impossible to make use of all the stored information necessary to make the right decision. Therefore, how to accurately find memory forgetting curve of information influence brings us a challenge.

Information (including text information, such as weibo, blog, etc.; and multimedia information, such as audio, video, images, etc.) flows and diffuses on the social network. This process is mainly by the user’s behaviors, such as information release, forward, comment and so on. In social network, the information influence not only affect the user's behavior, but also apply to the dynamic changes of the social network, its plays an important role in information flow and diffusion. In order to more accurately calculate the influence of information or event, this article introduced improved Ebbinghaus memory forgetting curve in the cycle of event. From the perspective of cognitive psychology with memory forgetting about things, we use the information of every day as a separate event, the time series events form episode. In addition, we consider an event as a life form and propose a Memory-Forgetting Curve (MFC) [9].

The rest of this paper is organized as follows: Section 2 defines and describes preliminary concepts. Sections 3 shows experimental analysis and results. Section 4 gives conclusions.

2. PRELIMINARY CONCEPTS AND ALGORITHM DESCRIPTION

2.1 The Definition of Event and Episode

An event is a snapshot of perceived experience at one moment in time which can be defined as a collection or a tuple of attributes.

Definition 1. An event e is a tuple reflecting a moment of experience such that $e = < A, M, T, S >$, where A is element of the action, and means the change process and its characteristics of event, which is the degree of movement, description of the way, method and so on. $A = < a_1, a_2, ..., a_k >$. Each attribute a_i is defined as a tuple such that $a_i = < a_1^i, a_2^i, ..., a_l^i >$ and a_j^i is normalized real value $a_j^i \in [0,1]$. M is the object element, that includes all roles participating in event. T is the time element, and it means the time from the begin to the end of event. S is the environmental element, and it means the place and its environmental characteristics of event.

MOBIMEDIA 2016, June 18-19, Xi'an, People's Republic of China
Copyright © 2016 EAI
DOI 10.4108/eai.18-6-2016.2264332

An episode can be defined as a finite list of events collected in a temporal order.

Definition 2. An episode E is a sequence of events such that $E = < e_{t_0}, e_{t_1}, ..., e_{t_n} >$, where t_i denotes the relative time point where in the event e_{t_i} occurs.

2.2 The Description of Public Opinion Data Space (OS) and Public Behavior Data Space (BS)

When an event happens, in different S the form of A is presented diversely. Someone discuss the event in forums (Public Opinion Data Space (OS)), while others change their investment strategies (Public Behavior Data Space (BS)), such as buy or sell related product or stock which are affected by the event. Therefore, we study event from a new perspective that is from OS and BS.

2.2.1 Public Opinion Data Space (OS)

When a hot event happens, a lot of views and comments about the event have diffusion fast on network to form a powerful network influence. The research and analysis of network events emerge in endlessly [10, 11]. In particularly, investors publish their views and emotions when a good or a bad event happens in finance field. The form of A is diversified in OS, such as click, reply, content and size of post, etc. in forums. The Investors response mainly by post and reply. Therefore, we build $< post, reply >$ as set of A to measure the action of event.

The number of post on Internet also reflect user attention on event, but the strength of post is larger than the strength of click to reflect the participation of event. Posting for event with their comments and analysis fully illustrate the subjective initiative of user, which reflects user hope to enhance or inhibit the spread of event influence through their own comments. In addition, the strength of reply to event embodies the role of reply who is given to the interaction of event. $A_{OS} = < post, reply >$, where $post$ is the number of user posting the comments on forums. $reply$ is the number of user reply to some post.

2.2.2 Public Behavior Data Space (BS)

In the real life of offline, it is difficult to capture the information in real world with its mass scale and few value features. Especially, capturing the group behavior after the event happening is more difficult, so the data can be used extremely limited. On the other hand, some carries such as Stock Exchange centre and electric business platform can record human behaviors of the real world. Some special carriers are good choices to observe human behavior in real world.

Firstly, $\mathcal{A}$ can reflect the degree of participation or group game mainly through $\%Turnover$, Vol, $Turnover$. One day the volume of a stock is blew up that illustrate buying and selling are more active, group $\mathcal{M}$ makes a clarity of judgment for potential stock trend. $\mathcal{A}$ illustrates an event occurs or will occur in soon, so quantitative analysing the participation degree of real social $\mathcal{M}$ is an important indicator to study event.

Secondly, with occurrence, development, decay and death of event, the price of one stock will change with event lifecycle. If it is a good event, the price will be gradually pushed up from the beginning to the end of event. The prices trend map out the development trend of events, such as $Close$, $Pre\ Close$. $A_{BS} = < Close, Pre\ Close, Vol >$, $Close$ is the closing price in the end of trading time. $Pre\ Close$ is yesterday's $Close$. Vol is the amount of trade shares.

2.3 Memory-Forgetting Curve

The memory forgetting curve as formula 1 was first described by Ebbinghaus[1]. *R* denotes the memory retention function with time elapses from the beginning of the event. *S* denotes strength of memory which controls how fast we forget about the event.

$$R = e^{-\frac{t}{s}} \qquad (1)$$

With the deepening on memory and forgetting research, there are some variants of forgetting curve models such as the power-law curve and the S-shaped curve [12]. We adapt these aforementioned forgetting curves to model the event and episode memory and forgetting curve to improve event influence accuracy. Work is energy accumulation of force moving on the space in physics, the size of work equals force multiple the distance of the force moving on space. Is hard to measured event by single element, an event is often affected by post and reply in OS, and the trend of a stock event is often measured by price and volume for judgment in BS. Inspired by work in physics, this paper puts forward the event energy of OS in the financial field that is equal to the number of post working on the reply; and the event energy of BS is equal to daily trading volume working on the price. This paper defines the *post* and the *reply* of an event in every day is a new event, as the beginning of a new memory, when the coming of *post* or *reply* information begin to decay with the improved Ebbinghaus memory forgetting curve in OS. In BS, we use the same method to calculate the memory forgetting curve about one event by the energy of *Vol, Pre close, Close.* $energy_n^{OS} = post \times reply$, $energy_n^{BS} = Vol \times (Close - Pre\ close)$. Therefore, we put forward the concept of Memory-Forgetting Curve (MFC) which is expressed as,

$$\begin{cases} energy_1 = e^{-(t_1 \times \alpha_1)}, & t_1 = 1; \\ energy_2 = e^{-(t_1 \times \alpha_2)} + w_1 e^{-(t_2 \times \alpha_1)}, & t_2 = 2; \\ energy_3 = e^{-(t_1 \times \alpha_3)} + w_2 e^{-(t_2 \times \alpha_2)} + w_1 e^{-(t_3 \times \alpha_1)}, & t_2 = 3; \\ \quad ... \\ energy_n = e^{-(t_1 \times \alpha_n)} + w_{n-1} e^{-(t_2 \times \alpha_{n-1})} + ... + w_1 e^{-(t_n \times \alpha_1)}, t_n = n. \end{cases} \qquad (2)$$

$t_n = n$, $\alpha_i \in R \geq 0$, $w_{n-1} = \frac{y_n - y_{n-1}}{x_n - x_{n-1}} * 100\%$, $0 \leq e^{-(t_n \times \alpha_1)} \leq 1$;

where α_i is a non-negative real parameter to control the memory-forgetting speed, where t_i is the time interval from the begin of an event till now, and the larger value of α_i is, the faster memory loses as time elapses. Particularly, if $\alpha_i = 0$, then MFC is degenerated to the uniform weighted model, i.e., no memory on the past events is lost. w_i is the slope of memory-forgetting curve that denotes the weight of memory. MFC characterizes how much memory of a happened event is retained by now. The more memory the event has retained, the more likely it will affect next event according to the recency effects. Besides, if we compute individual memory-forgetting speed α_i for each event, we'll get an event MFC model. The exponential representation of MFC in formula 2 can be extended to other forgetting curve models to compute the memory retention of happened event. We use this basic and widely accepted exponential form in this work.

2.4 Memory-Forgetting Curve Algorithm(MFC)

MFC algorithm of this paper need to input the relate event $\mathcal{A}$ of OS and BS, such as *post* and *reply* of OS, *vol*, *pre close* and *close* of BS. The energy time series of OS and BS about event according to the definition. Based on the definition of Ebbinghaus, we carry out on the energy 0-1 normalized processing. If *e* index function within the domain of 0-1, we calculate Memory-Forgetting Curve of event or episode.

ALGORITHM: Memory-Forgetting Curve Algorithm(MFC)

Input：every *A*;
Output: MFC of event, episode;
Scan: data sources from Eastmoney Website and Sina Website;
generate $energy_n^{OS} = post \times reply$;
generate $energy_n^{BS} = Vol \times (Close - Pre\ close)$;
normalization $energy_n^{OS}$ and $energy_n^{BS}$;
If ($0 < energy_n - w_{n-1}e^{-(t_2 \times \alpha_{n-1})} - ... - w_1 e^{-(t_n \times \alpha_1)} \le 1$)
generate $\alpha_1, \alpha_2, ..., \alpha_n$ by formula 2;
generate $e^{-\alpha_1}, e^{-2\alpha_1}, ..., e^{-n\alpha_1}, ..., e^{-\alpha_{10}}, e^{-2\alpha_{10}}, ..., e^{-n\alpha_{10}}$;
generate formula 3;
generate $M_1 = e^{-\alpha_1}, e^{-\alpha_2}, ..., e^{-\alpha_n}$ by formula 3;
generate $M_2 = e^{-2\alpha_1}, e^{-2\alpha_2}, ..., e^{-2\alpha_n}$ by formula 3;
generate $Forgetting = M_1 - M_2$ by formula 3;

3. EXPERIMENT AND ANALYSIS

3.1 Pre-processing

The data of OS from Eastmoney website, the data of BS from Chinese Stock Exchange. This article selects the Shanghai Free Trade Area Episode of 600648(WaiGaoQiao) to calculate the number of post and reply from January 1, 2013 to March 19, 2016, as show in figure 1.

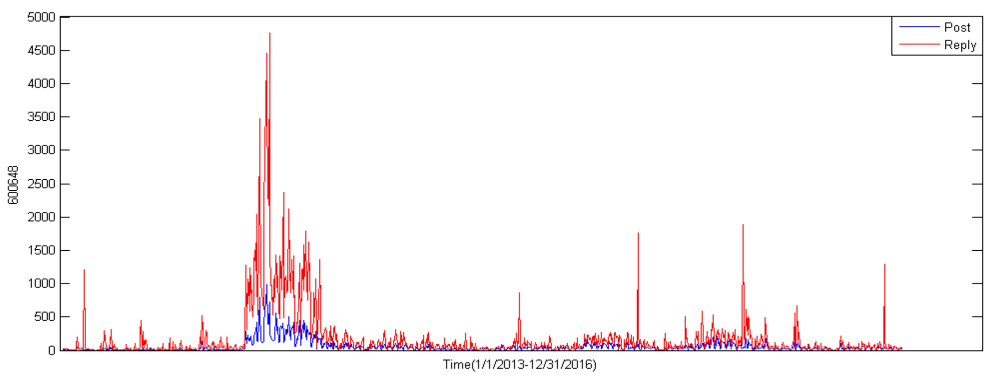
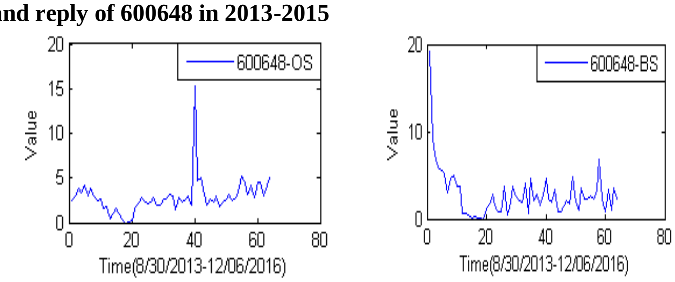


Figure 1. The number of post and reply of 600648 in 2013-2015

We find Shanghai Free Trade Area 600648 Episode from figure 1 by the TF-IDF method, the cycle is from August 30, 2013 to December 6, 2013. We select the *post* and *reply* as energy parameters of OS, as well as the *vol*, *pre close* and *close* as energy parameters in BS.

According to the formula 2 to calculate $\alpha_1, \alpha_2, ..., \alpha_n$ of Shanghai Free Trade Area 600648 Episode in the period of OS and BS.



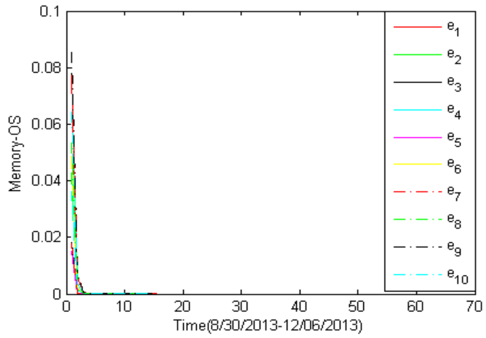
a. $\alpha_1, \alpha_2, ..., \alpha_n$ **in OS** **b.** $\alpha_1, \alpha_2, ..., \alpha_n$ **in BS**

Figure 2. $\alpha_1, \alpha_2, ..., \alpha_n$ **in OS and BS**

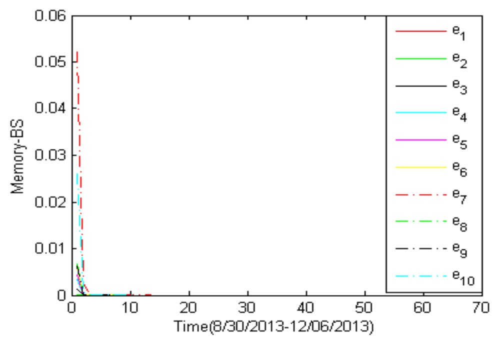
when $\alpha_{19} = 0$ in OS and $\alpha_{18} = 0$ in BS, the memory=1

3.2 Memory-Forgetting Curve of Event

According the $\alpha_1,\alpha_2,...,\alpha_n$ of OS and BS, we calculate Memory-Forgetting Curve of the top 10 events in OS and BS, $e^{-\alpha_1},e^{-2\alpha_1},...,e^{-n\alpha_1}$,…, $e^{-\alpha_{10}},e^{-2\alpha_{10}},...,e^{-n\alpha_{10}}$. We think the information of every day as a separate event to recession, the Memory-Forgetting Curves of top 10 single events are shown in figure 3.



a. Memory-Forgetting Curve in OS



b. Memory-Forgetting Curve in BS

Figure 3. 600648 Memory-Forgetting Curve in OS and BS

In the Figure 3, the time series about $e_1,e_2,...,e_{10}$ denotes $e^{-\alpha_1},e^{-2\alpha_1},...,e^{-n\alpha_1}$,…, $e^{-\alpha_{10}},e^{-2\alpha_{10}},...,e^{-n\alpha_{10}}$.

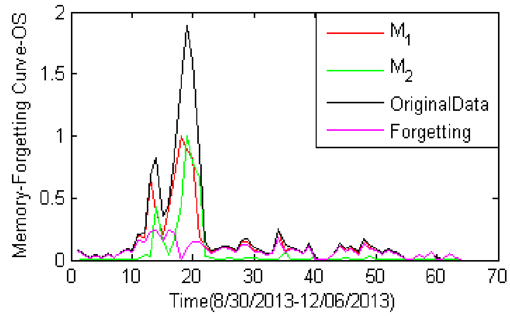
The Memory-Forgetting Curves after the third day make convergence close to zero from the figure 3. So we only need the first day and the second day data, the memory influence of the third day can be neglected. Then according to the Memory-Forgetting Curve law, we modified formula 2 to form a more simplified formula 2. We summarize a generic function as formula 3.

$$eng_n=\begin{cases} e^{-\alpha_n}, & n=1; \\ e^{-\alpha_n}+w_{n-1}e^{-2\alpha_{n-1}}, & n\geq 2. \end{cases} \quad (3)$$

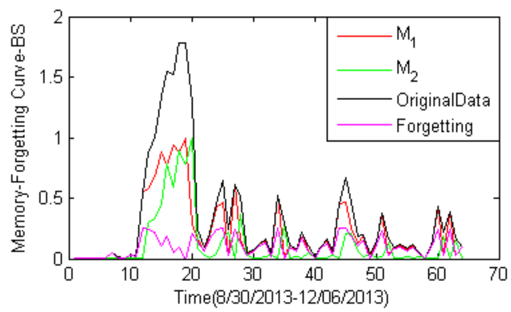
$$0<eng_n\leq 1,\ w_{n-1}=\frac{y_n-y_{n-1}}{x_n-x_{n-1}}*100\%$$

3.3 Memory-Forgetting Curve of Episode

The cycle of the Shanghai Free Trade Area 600648 Episode is from August 30, 2013 to December 6, 2013 in figure 4. We select the *post* and *reply* as energy parameters of OS, as well as the *vol*, *pre close* and *close* as energy parameters in BS. According to the $\alpha_1,\alpha_2,...,\alpha_n$, $M_1=e^{-\alpha_1},e^{-\alpha_2},...,e^{-\alpha_n}$ is the memory time series of independent event on first day. $M_2=e^{-2\alpha_1},e^{-2\alpha_2},...,e^{-2\alpha_n}$ is the memory time series of independent event on second day. $Forgetting=M_1-M_2$ is the forgetting time series of independent event.



a. Memory-Forgetting Curve in OS



b. Memory-Forgetting Curve in BS

Figure 4. Memory-Forgetting Curve of Episode in OS and BS

In the *a* of Figure 4 $\alpha_{19}=0$ and the *b* of Figure 4 $\alpha_{18}=0$, then forget=0, memory=1.

4. CONCLUSION

Semantic web and Crowdsensing blog platforms store the memory of diverse information with the social media development. This paper proposed a Memory-Forgetting Curve (MFC) from the people's cognitive psychology. We treated every day's information as an independent event to calculate MFC on the basis of Ebbinghaus, then calculate MFC of episode with the fusion of events'. MFC to be remembered and acquired as useful knowledge to support decision making in commercial recommendations.

5. REFERENCES

[1] Jun Chen, Chaokun Wang, Jianmin Wang; Modeling the Interest-Forgetting Curve for Music Recommendation, [C] *MM'14*, November 03 - 07 2014, Orlando, FL, USA.

[2] A. M. Nuxoll and J. E. Laird. Extending cognitive architecture with episodic memory. [C] *In Proceedings of the 22nd national conference on Artificial Intelligence (AAAI 2007)* Volume 2, pages 1560–1565. AAAI, Press, 2007.

[3] W. C. Ho, K. Dautenhahn, and C. L. Nehaniv. Computational memory architectures for autobiographic agents interacting in a complex virtual environment: A working model.[J] *Connection Science*, 20(1):21–65, 2008.

[4] C. Becker-Asano and I. Wachsmuth. Affective computing with primary and secondary emotions in a virtual human.[J] *Autonomous Agents and Multi-Agent Systems*, 20(1):32–49, 2010.
[5] C. Brom, T. Korenko, and J. Lukavsky. How do place and objects combine? What where memory for human-like agents.[C] *Intelligent Virtual Agents*, LNCS 5773, pages 42–48. Springer Berlin, 2009.
[6] Z. Kasap, M. B. Moussa, P. Chaudhuri, and N. Magnenat-Thalmann. Making them remember-emotional virtual characters with memory.[J] *IEEE Computer Graphics and Applications*, 29(2):20–29, 2009.
[7] M. Y. Lim, R. Aylett, W. C. Ho, S. Enz, and P. Vargas. A socially-aware memory for companion agents. [C] *Intelligent Virtual Agents*, LNCS 5773, pages 20–26. Springer Berlin, 2009.
[8] Jarrod A. Lewis-Peacock & Kenneth A. Norman, Competition between items in working memory leads to forgetting, [J] *Nature Communications*, 2014, | 5:5768 | DOI: 10.1038/ncomms6768.
[9] Chien Chin Chen, Yao-Tsung Chen, and Meng Chang Chen, An Aging Theory for Event Life-Cycle Modeling, [J] *IEEE Transactions on Systems*, Man and Cybernetics—Part a: Systems and Humans, VOL. 37, NO. 2, MARCH 2007.
[10] C. C. Aggarwaland K. Subbian. Event detection in social streams.[C] In SDM'12, pages 624-635, 2012.
[11] R. Li, K. H. Lei, R. Khadiwala, and K. C.-C. Chang. Tedas: A twitter-based event detection and analysis system. [C] *Data Engineering*, International Conference on, 0:1273-1276, 2012.
[12] L. Averell and A. Heathcote. The form of the forgetting curve and the fate of memories. [J] Journal of Mathematical Psychology, 55:25–35, 2011.