Empirical Validation of Website Timeliness Measures

Yanlong Zhang[(1, 2)] **Hong Zhu** [(1)] **and Sue Greenwood** [(1)]
(1) Dept. of Computing, Oxford Brookes University, Oxford, OX33 1HX, UK,
Email: hzhu | sgreenwood @brookes.ac.uk
(2) Dept. of Comp. & Math., Manchester Metropolitan University, Manchester M1 5GH, UK,
Email: y.zhang@mmu.ac.uk

Abstract

Information timeliness is crucial for media-based websites. Although a couple of timeliness design strategies have been developed, timeliness measurement is still in its infancy. Based on our previous research, this paper presents three timeliness measures and reports the empirical validation of them in a case study conducted during the recent Olympic Games.

Keyword: Web measurement, Web metric, Web timeliness, Web quality

1. INTRODUCTION

Timeliness is a quality attribute of media-based websites [1]. Usually, users judge the timeliness of a website by comparing the site with the real world. They believe that it would be easier to update the information online than printed documentation. Thus, websites are always expected to be kept updated in order to reflect the change of the real world in a timely manner [2].

The importance of timeliness has now been widely recognised by Web designers [3,4,5,6,7,8,9]. A number of strategies have been developed to show the timeliness of websites, such as to put date and/or time indicators on the page, to use animation associated with dynamic content to show its timeliness, or to tie content to current real-world events. However, how to measure a website's timeliness remains an open problem. Compared with the above strategies, measurement can provide quantitative analysis [10] and thus insight into the issue in terms of quality prediction and evaluation.

Based on our previous research reported in [11,12], this paper further investigates the measurement of timeliness. Two new measures are developed and validated empirically and compared with the measures proposed in our previous work. The rest of the paper is organised as follows. Section 2 reviews the related research. Section 3 defines and justifies our Web timeliness measures. Section 4 reports the results of empirical validation of the measures. Finally, section 5 concludes the paper with a discussion of future work.

2. RELATED RESEARCH

Recently, a method for the assessment of the timeliness of websites has been proposed using sets of assessment criteria in the literature. For example, in [5] the following three criteria are proposed to evaluate the Web timeliness.

- Is there an indication of when the information was created/published?
- Is the information regularly revised or updated?
- Is the information still valid for your topic?

In [3], the following five criteria were given to assist in evaluating the Website timeliness.

- When was the information created or last updated?
- Is the source appropriate for your needs with regards to the time that the source of information was published?
- How current are the links, statistical data, illustrations, etc.?
- Does the information appear to be valid and well researched, or is it questionable and unsupported by evidence?
- What is its relation to other works on the subject, especially with regards to the time?

Such criteria also provide useful design guidelines, or heuristics, for website designers. However, it is rather difficult to validate the correctness and completeness of such heuristics. Furthermore, the application of the criteria is difficult to give accurate assessment and comparison of websites' timeliness. The most common way of evaluation of a website's timeliness is by user test, which is, unfortunately, not always practical. As pointed out in [13], it is not an easy task to organise and complete such a test satisfactorily, which is time-consuming and costly especially for repetition tests. Measurement can provide a more economic and quantitative means for the issue [14]. However, timeliness measurement is still in its infancy.

In [12], we proposed four Website timeliness measures and conducted the preliminary empirical studies. In particular, we defined the measure HUF

(Homepage update frequency). It requires to monitor the change of website homepages at a set frequency, such as 1/60s, and to calculate the number of probes that detect a change of the homepage. For most websites, including the news sites, this frequency is fast enough to monitor the homepage update rate. However, in a recent case study on the websites that reports news of the Athens Olympic Games, the high frequency in updating homepages posed a challenge to this measure. We found that some web sites that reports the news on Olympic Games such as http://2004.sina.com/ changed their homepage at a frequency as high as 9 times within a minute. According to the Sampling Law [15], the sampling frequency must be at least twice the bandwidth of the sampled signal. For example, a 44.1 kHz sound sample will sample frequencies up to about 22 kHz. It is apparent that, in theory, the results of HUF cannot always reflect the true timeliness of the websites. Due to technical reasons, the HUF sampling frequency cannot be very high. Therefore, alternative ways to measure timeliness must be developed. In this paper, we propose two new timeliness measures that can overcome this drawback of the HUF measure.

3. WEBSITE TIMELINESS MEASURES

In [12], we defined Website timeliness as the ability of web-based information systems to provide and process information in a timely manner, i.e., to create, update and present information within a required time delay in order to keep the information consistent with the real world. The following measures can be derived from this definition.

3.1. Measurement 1: Mean Time Delay To Publish

According to the definition of timeliness, a direct measure of timeliness is to measure the time difference between the time when the information is published and the time when the event occurs in the real world. The Time Delay To Publish (TDTP) can be formally defined as below.

$$TDTP = T_{publish}(E) - T_{occur}(E)$$

where $T_{publish}(E)$ is the time when an event E is published online, and $T_{occur}(E)$ records the time it occurs.

Usually, it is not sufficient to assess a website's timeliness by testing only one event. A set of events will be used to obtain a more accurate measurement result. We therefore defined Mean Time Delay To Publish (MTDTP). Let $A=\{\alpha_n \mid n=1,\ldots,K\}$ be a set of events. The Mean Time Delay To Publish with regards to the set A of events can be formally defined as follows.

$$MTDTP_A = \sum_{i=1}^{K} (T_{publish}(\alpha_i) - T_{occur}(\alpha_i)) / K$$

where $T_{publish}(\alpha_i)$ is the time when an event α_i is published online, and $T_{occur}(\alpha_i)$ is the time when the event occurs.

MTDTP provides a direct measure of timeliness, but it has the following limitations.

1. The measurement relies on the availability of T_{occur}. In the current practices of web site development, T_{occur} can only be found manually in the websites or from other media types. For big events, the news content usually includes the precise time when they happened. The accuracy of the measure is dependent on the accuracy of T_{occur}. If T_{occur} cannot be found, or only a rough idea of the time is available, such as around 5 pm, the measure is not applicable.
2. The measurement relies on the set of events chosen to test a website's timeliness. Different websites may have different tastes. The same event could be regarded as a 'big' news for one site, but completely ignored by another site. The set of events chosen for testing may significantly affect the test results.
3. It is usually time-consuming to find T_{occur} and $T_{publish}$. As mentioned above, most websites put time indicators on the page. However, without an automatic mechanism, it is obvious that $T_{publish}$ involves much manual reading. Even with the aid of a software tool in this study, it was a rather tedious and time consuming task to complete.

3.2. Measurement 2: Site Evolution Speed

Site Evolution Speed (SES) calculates the number of web pages that are changed over a period of time. In the empirical study, we found that the changes of a website were often accurately reflected by the number of pages added. It is rare that pages were deleted or changed except for the homepage.

For a fast-changing website, it is always ready to publish a piece of news. Usually a headline appears in the homepage, and the detailed information is added to the site as a new page with a link from the headline. Each time a new page is added to the website, a change to the website is made. Assume that two probes of a website w are made at time moments t_0 and t_1, $AddedPages_w$ is the number of new pages added to website w during the time between t_0 and t_1 based on the state of the website obtained by the probes. The Site Evolution Speed (SES) can be formally defined as follows.

$$SES = \frac{AddedPages_w}{t_1 - t_0}$$

SES can be automatically calculated. The measure SES has the limitation that it is accurate only if the website does not delete, merge or split web pages.

3.3. Measurement 3: Homepage Update Frequency

Homepage update Frequency (HUF) was first defined in [11,12], where it was called Homepage Change Frequency (HCF). Formally,

$$HUF = \frac{C_{pr}}{N_{pr}}$$

where N_{pr} is the number of probes made to a website in a period, C_{pr} is the number of probes that detected changes to the home page. The detailed derivation can be found in [12].

In theory, if a website always changes at a frequency that between two sampling only one page is added to the website, and pages are never deleted or merged/split, then, both *SES* and *HUF* will detect all changes to a website if they are applied to the same period of time. Then, we have

$$HUF = SES / SF_r$$

where SF_r is the sampling frequency of *HUF*.

In the practical uses of HUF measure, an appropriate frequency of sampling must be carefully set so that the homepages can be downloaded between two probes. Therefore, efficient implementation of the sampling tool is the key issue of the usability of the measure, which has been discussed in [12]. The uses of *SES*, on the other hand, do not heavily rely on the efficient implementation of the measure. A question is how well the assumptions made in the uses of *SES* and *HUF* measures match the reality. Hence, an empirical case study is conducted.

4. EMPIRICAL STUDY

In this section we report an empirical study of validation of the above measures.

4.1 Experimental design

To evaluate the feasibility of the measurement method and to validate the measures, an empirical study was conducted during the recent Olympic Games. The study consisted of the selection of a set of candidate websites as the subject, the selection of a set of events in the real world, and the collection and processing of the data.

The selection of candidate website and real world events must be fair in the sense that the events should be interested to all the websites tested. To meet this requirement, we selected the events to be the gold medals won by the Chinese athletic team in the recent Athens Olympic Games. The candidate websites were selected from among the major online news media in China. These websites were amongst the top four when searching the keyword 'Olympics' (in Simplified Chinese) using Google. They are:

- Sina (http://2004.sina.com.cn)
- TOM (http://2004.sports.tom.com)
- Sohu (http://2004.sports.sohu.com/)
- Yahoo (http://cn.sports.yahoo.com/olympic/)

The empirical study was prepared in advance before the Athens Olympic Games. A software tool, consisting of small Perl scripts, was developed to collect the data and calculate the measures. During the Athens Olympic Games, the Chinese Team obtained 32 Gold Medals. For each gold medal, we collected the publicising times on these candidate websites as well as the real time when the event happened.

To obtain MTDTP measures, a software tool monitored the homepage of each website by the sampling frequency of once per minute, and downloaded it whenever a change was detected. We found that all important news items related to the selected events were reported on these websites' homepages. All these websites also provided the time indicators of the web pages using meta-data, accurate to minute, in the pages. There was a little time difference between the time indicators provided by the sites and the time we detected the change on the homepages. This was, we believed, due to the network delays or the Web writers' mistakes. To be fair to all, we used the time that we detected the changes as $T_{publish}$. Manual information processing was used to collect the time of official announcement of each Gold Medal won by the Chinese team. The accurate time of the Chinese athletic team obtaining each Gold Medal, T_{occur}, was based on the reports by the Xinhua Net (http://www.xinhuanet.com/olympic).

To obtain SES measures, a software tool is used to download the targeted websites once every hour. We only downloaded the webpages related to the news on the Olympic Games. For example, for Yahoo, we regarded the URL http://cn.sports.yahoo.com/olympic/ as the homepage and only downloaded the pages within this directory. All pages which were linked to other websites or other servers within Yahoo were ignored. It took a long time to download the site for the first time, but it was much quicker for the following downloads as it only downloaded the added pages.

We also obtained HUF for the above websites. The measuring method and tool are the same as those discussed in [12].

The software tools were run to collect data continuously from 00:00 14th to 23:59 30th August 2004.

4.2 Experimental results

The results of the experiment are listed in Table 1, 2 , 3 and illustrated in Figure 1.

Table 1 TDTP of four news websites (Minute)

Site / Medal	TOM	Yahoo	Sohu	Sina
1	1	19	10	3
2	1	20	3	3
3	1	7	4	3
4	2	8	3	7
5	1	7	13	2
6	3	5	4	4
7	2	4	3	3
8	1	5	2	2
9	2	4	4	6
10	2	5	7	4
11	1	5	7	4
12	1	3	4	2
13	1	6	4	4
14	1	2	3	3
15	3	6	3	4
16	1	3	5	2
17	1	6	6	3
18	1	4	2	2
19	1	3	3	2
20	2	6	3	4
21	2	5	6	4
22	1	6	6	2
23	1	11	8	2
24	1	3	1	4
25	1	8	1	2
26	1	2	1	2
27	1	4	2	2
28	1	7	3	2
29	2	7	6	4
30	1	9	4	2
31	1	3	1	2
32	1	18	1	3

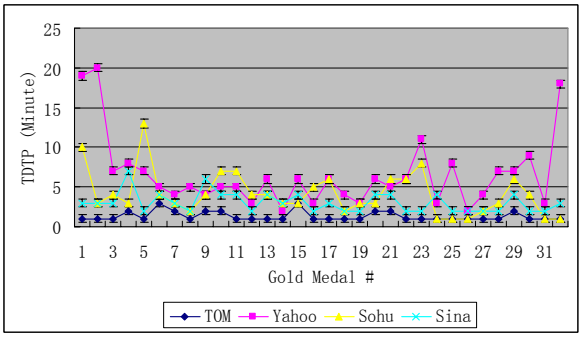

Figure 1 Pattern of TDTP of four websites

Table 2 Number of Pages Added

Site / Date	TOM	Yahoo	Sohu	Sina
14	733	267	296	433
15	745	246	270	398
16	754	342	366	495
17	739	363	298	398
18	772	439	386	516
19	853	486	160	406
20	760	397	442	446
21	711	367	310	532
22	759	329	282	473
23	635	361	297	366
24	612	287	189	406
25	508	378	390	538
26	651	306	411	269
27	632	397	283	218
28	560	259	261	177
29	672	278	432	105

Table 3 Number of Homepage Updates

Site / Date	TOM	Yahoo	Sohu	Sina
14	463	183	248	375
15	471	215	255	378
16	433	301	280	444
17	417	304	267	364
18	432	329	292	447
19	489	420	148	375
20	450	308	386	399
21	473	324	273	449
22	452	303	266	375
23	419	291	247	378
24	349	236	168	444
25	350	338	356	364
26	374	269	324	447
27	363	351	258	375
28	300	226	246	399
29	347	198	340	449

From Figure 1, we found that TOM performed the best. The other sites were similar except that Yahoo did not perform very well at the beginning and at the end. Using the data from the above tables, MTDTP, HUF and SES can be calculated (see Table 4).

Table 4 Timeliness measures of four websites

Measure / Website	MTDTP	HUF	SES
TOM	2.69	0.29	28.90
Yahoo	13.19	0.20	14.33
Sohu	7.06	0.19	13.21
Sina	6.13	0.24	16.08

From Table 4, all measures agree that TOM was the best. Results of HUF and SES seem quite consistent. Both agree that the ranking order should be: TOM > Sina > Yahoo > Sohu. However, TDTP shows the ranking order: TOM > Sina > Sohu > Yahoo.

We conducted Pearson analysis to study the correlations between the measures. The results are in Table 5 below.

Table 5 Correlations between three measures

	MTDTP	HUF	SES
MTDTP	-	-0.76	-0.72
HUF	-0.76	-	0.94
SES	-0.72	0.94	-

It can be seen that there exists strong correlation among the measures. Although the correlation coefficients between MTDTP and HUF or SES are comparatively lower, considering the difficulty of choosing MTDTP events, the results are satisfactory.

It is also interesting to note that a strong correlation exists between HUF and SES.

5. CONCLUSION AND FUTURE WORK

In the literature of website design, there exist a couple of assessment criteria to evaluate Website timeliness. However, these criteria for evaluation of websites are based on different understanding of the issue. They are difficult to validate in terms of completeness and correctness. Website quality measures can provide insight into such issue. They can be used to measure the Website quality attributes in an objective and economic way.

In this paper, we defined and validated three Website timeliness measures. Mean-Time-Delay-to-Publish (MTDTP) is derived directly from the definition of timeliness. It can provide a relatively accurate measurement of a website's timeliness with regards to a set of focused events. However, it requires manual collection of data. Homepage-Update-Frequency (HUF) measures the timeliness indirectly through the frequency that a website updates its homepage. It can be automatically computed without human interferences. However, it may be less accurate when a website updates the homepage with a frequency higher than the sampling frequency. Site-Evolution-Speed (SES) also measures the timeliness indirectly, but through the number of webpages in the site change during a period of time. Under the assumption that webpages are not deleted, merged or split once published, SES can provide a good measurement of timeliness. SES is especially useful and applicable to the websites whose update frequency is very high. The empirical study reported in this paper demonstrated their practical usability. The results showed that SES and HFC were effective measures of website timeliness.

We also developed a prototype software tool to implement the measures. Such a tool can also be used by designers to evaluate their websites' timeliness easily. As the results were obtained using the software tool and human-involved checking, they were objective and independent on the human's judgments.

This paper focused on the development and validation of timeliness measures. It leaves some interesting issues to investigate, such as 'which measure is more sensitive to small differences of websites', 'which measure is actually better for different purposes', etc. We argue that these issues can be addressed through further empirical studies.

This research focused on the website timeliness measurement. It is part of a larger project [16] whose goals are to develop measures to empirically investigate all aspects of website quality attributes, and to develop tools to help evaluate and improve the website quality.

REFERENCES

[1] Brajnik, G., Towards valid quality models of websites, *Proceedings of the 7th Conference on Human Factors and the Web*, Madison, Wisconsin, June, 2001.

[2] Navarro, A. and Khan, T., *Effective Web Design: Master the Essential*, San Francisco: SYBEX, 1996.

[3] Virtual Chase, Information timeliness, 2000,URL: http://www.virtualchase.com/govdoc/timely.html, accessed on 25 June 2004

[4] Collins Memorial Library, Criteria for Evaluating Information on the Web: Accuracy and Currency, 2004, URL: http://library.ups.edu/research/handouts/time.htm, accessed on 25 June 2004.

[5] CSUS, Evaluating Web timeliness, 2003, URL: http://library.csus.edu/services/inst/ICCS/infocomp/tutorials/module5/web/timeliness/, accessed on 25 June 2004.

[6] Powell, T.A., *Web Site Engineering: Beyond Web Page Design*, New York: Prentice Hall, 1998.

[7] Powell, T.A., *Web Design: the Complete Reference*, 2nd Edition, New York: McGraw Hill/Osborne, 2002.

[8] Lynch, P.J. and Horton, S., *Web Style Guide: Basic Design Principles for Creating Web Sites*, Yale University Press, 1999.

[9] Borges, J.A., Morales, I. And Rodrigues, J., Guidelines for designing usable World Wide Web pages, *Proceedings of ACM CHI '96 Conference on Human Factors in Computing Systems*, vol. 2, pp. 277-178, 1996.

[10] Fenton, N.E. and Pfleeger, S.L., *SoftwareMeasures: A Rigorous & Practical Approach,* 2nd edn, PWS Publishing Company, Boston, MA, 1997.

[11] Zhang, Y., Huo, Q., Zhang, L., Zhu, H., Greenwood, S., Measuring the Timeliness of Websites, *Proceedings of the Fourth European Conference on Software Measurement and ICT Control in Co-operation with DASMA*, Heidelberg, Germany, pp. 139-147, 2001.

[12] Zhang, Y., Zhu, H., Huo, Q. and Greenwood, S., Measurement of Timeliness of Web-based Information Systems, *Proceedings of the 6th World Multi-Conference on Systemic, Cybernetics and Informatics (SCI 2002)*, July 14-18, 2002.

[13] Tes, Developing a basic numeric measure for Web usability, http://www.kuro5hin.org, 14 April 2002, accessed on 2nd March 2003

[14] Ivory, M.Y., Sinha, R.R. and Hearst, M.A., Empirically validated Web page design measures, *Proc. 01'SIGCHI*, 31 March – 4 April 2001 Seattle, WA, USA, pp. 53-60.

[15] T202 Course Team, *Analogue and digital electronics*, The Open University, Milton Keynes, UK, 1990.

[16] Zhu, H., Greenwood, S., Huo, Q. and Zhang, Y., Towards Agent-Oriented Quality Management of Information Systems, *Workshop Notes of Second International Bi-Conference Workshop on Agent-Oriented Information Systems* (AOIS-2000) at AAAI'2000, Austin, USA, pp.57~64, July 30, 2000.