SOAR Project Proposal, Fall 2014

Investigating a new approach to web visitor engagement measurement

Primary Investigator: Tarah Cleveland, Faculty Mentor: Dr. Razib Iqbal

Computer Systems and Software Engineering Department
Valley City State University
[tarah.cleveland | razib.iqbal]@vcsu.edu

1. Abstract: In the United States, Internet advertising revenues totaled nearly $42.8 billion in 2013 which is a 17% increase from the year 2012 (IAB Internet Advertising Revenue Report 2014). As revenues from Internet advertising continue to grow, advertisers seek popular web pages for placing ads in an effort to maximize profits. An important measure of how well a website is doing or how attractive they are to advertisers is how engaged the web visitors are with that website. For example, Philly.com (home of two Philadelphia newspapers, *Inquirer* and *Daily News*) uses an "engagement index" which takes into consideration seven categories to measure web visitor engagement. These categories include duration index (visitors who spent at least five minutes on the site), click index (visitors who viewed at least six pages), recency index (rate at which visitors return to the site over a period of time), loyalty index (registered visitors or a minimum of three visits per week), brand index (website URL referral or visiting from a bookmark), interaction index (visitors that interact with the site's comment pages or forums), and participation index (activity on the site by posting, sharing, and uploading contents) (Beckett, 2010). Now, our preliminary study has revealed that there are some fundamental problems with current measurements for web visitor engagement. For example, session duration which tracks visit time on a web page does not take into consideration the physical away time of the web visitor from the computer or when a visitor switches to a different tab or application. Therefore, while advertisers commonly rely on a visitor's time spent looking at the website, traditional web analytics tools have some limitations to measure this accurately (Peterson and Carrabis 2008).

In order to address the need for an efficient web engagement tool, in this project we propose to investigate a method for accurately measuring session duration and web visitor engagement. Initial focus of this investigation will be on gathering and analyzing requirements to improve the measurement of "engagement index". As a proof of concept, we shall attempt to build a tool or add a feature to the existing web analytics tools such as piwik.org. The long-term objective of this

research initiative is to give guidance to the online advertisers on their selection of pages based on the actual degree of engagement of the web visitors with other similar webpages.

2. Proposed method: There are evidences of prominent research initiatives involving web engagement measurement. For example, Arapakis et al. explored user comments on web pages to devise a tool for measuring user engagement. Their initiative is focused on human curiosity in online news engagement (Arapakis et al. 2014). Attfield and colleagues adopted a framework that defines, measures, and explains user engagement through researching existing information retrieval metrics, user engagement metrics, web analytics, and measures from immersion in gaming (Attfield et al. 2011). In "Understanding the Impact of Video Quality on User Engagement," Dobrian et al. suggested that video buffering ratio has the greatest impact on user engagement across short and long video content types. They proposed a tool to measure the rate of buffering events in videos to compute the likeliness of user engagement (Dobrian et al. 1992). In "Five Tools and Technologies to Measure Engagement," Forrester, an independent technology and market research company, adopted a five-step measurement technique to profit marketers. These five steps include developing a system of records by centralizing customer data, using analytics tools to gain knowledge about customer behavior, gaining an understanding of customer attitudes through brand monitoring, gaining insights through studying customer values, and sharing information with measurement dashboards (Haven and Vittal 2008).

In this project, we shall first study the related publications, platforms, tools and technologies for measuring web visitor engagement. In the second phase, we shall investigate a reliable new measure for tracking active web visitors in a particular website. In order to achieve the desired goal, we are considering *Focus Ratio* and *Active Ratio*. Page focus ratio is the difference between the times a webpage has been loaded in a browser and the time that page was actually visible in the active tab. For web advertisers, a higher focused ratio indicates a more attractive location for placing ads on. Active ratio is the difference between the time a webpage is visible in the active tab and the user is actually considered viewing or interacting with that page. Both focused ratio and active ratio will help advertisers determine the maximum profits or earnings they may gain from a given web page.

Finally, by developing a proof of concept tool to quantify the impact of focused ratio and active ratio, our research findings will be measurable and tractable for commercial feasibility studies.

3. Timeline: We propose to spend in total 100 man-hours in this project. We have broken down our research activities into 4 major milestones. There will be overlaps among these tasks. However, for ease of understanding, we have summarized our milestones and anticipated duration in the following table:

ID	Milestones	Proposed Dates *(Duration)*	Actual Effort	Goal/Exit Criteria
1	Background study and Literature review: *User web engagement measurement tools and methodologies*	Jan. 02/15– Feb. 28/15 *(8 Weeks)*	25 hours	**Review Report**: 2-3 pages covering at least 10 related papers.
2	Investigation and Analysis: *Web user behavior and tools to capture these behaviors*	Mar. 01/15 – Apr. 25/15 *(8 Weeks)*	30 hours	**Initial Investigation Report**: 3-4 pages report covering detail analysis of the initial investigation and comparison to the reviewed literature work.
3	Proof of concept prototype/formula generation	Apr. 26/15 – Jun. 20/15 *(8 Weeks)*	25 hours	**Feature Report**: 1-2 Page report outlining the contribution/findings.
4	Evaluation/comparison of tools/formula/proposed methodology	Jun. 21/15 – Aug. 15/15 *(8 Weeks)*	20 hours	**Final Report**: 4-6 Pages final report in ACM format.

4. Budget:

ID	Category	Source	Total
1	Equipment/Software Tools/Access to scholarly articles	SOAR	$150.00
2	Supplies (e.g. Poster preparation)	SOAR	$100.00
3	Payment to undergraduate researcher (100hrs×$10/hr)	SOAR	$1000.00
4	Faculty mentor's contribution	CSSE	*In Kind*
5	Student conference registration and travel	Anticipating support from CSSE Department (*Subject to availability of funds*)	*To be determined*

5. Dissemination: Our motivation is to collaborate with the industry practitioners engaged in improving and optimizing online user engagement in order to improve success of their business. Therefore, our goal is to disseminate the final research findings at "The Workshop on User Engagement Optimization (UEO)" in 2016. The UEO workshop facilitates connections between the academic researchers and industry professionals interested in optimizing user web engagement. "The Workshop on User Engagement Optimization" would be a suitable venue to share our research findings with the researchers and professionals engaged in this domain.

We also plan to share our preliminary findings at local and regional forums, namely Valley City State University Annual Student Scholar Symposium in 2015. VCSU Annual Student Scholar Symposium is an outlet for VCSU community members to share course projects and related research with the faculty members, staff, and students.

6. Educational Objectives: As a Software Engineering major, my enthusiasm lies in exploring different aspects of Science, Technology, Engineering, and Math (STEM). My educational background in STEM includes science courses in biology, psychology, technology, and engineering. I have successfully completed courses in Software Engineering, Structured Programming, Data Structures, Calculus, College Algebra, Analytical Geometry, Plane Trigonometry, and Discrete Mathematics. By engaging in coursework covering all four areas of STEM, I have developed an eagerness to conduct research which is not limited to technology alone, but rather incorporates other domains. My intention is to enrich my knowledge in marketing and communications, human behavioral patterns, and online tools analysis. I would like to understand the behavioral patterns of web visitors and how this relates to user involvement and interactions with online applications.

7. Integrative nature of work: The proposed topic encompasses a variety of academic interests - the most prevalent interests being in Human-Computer Interaction (HCI) and Human Psychology. HCI is a school of thought that studies the relationships between humans and machines (Hewett et al. 1992) and Psychology is the discipline that studies the human behavior and the human mind (American Psychological Association 2014). Besides the technical aspects, I plan to review, investigate and analyze overall users' browsing habits and patterns, and understand the impact of each browsing pattern on the website. This process will require me to study behavioral patterns of modern day computer users as well as how they interact with the different web-based tools and

contents. I am motivated to explore the power of these two disciplines and apply my knowledge to benefit the vendors engaged in commercial aspects of the user web engagement.

8. Bibliography:

Arapakis, Ioannis, et al. “User engagement in online News: Under the scope of sentiment, interest, affect, and gaze.” *Journal of the Association for Information Science and Technology*, 2014. 65: 1988–2005. doi: 10.1002/asi.23096

Attfield, Simon, Kazai, Gabriella, Lalmas, Mounia, & Piwowarski, Benjamin. 2011. Towards a science of user engagement. WSDM Workshop on User Modelling for Web Applications, Hong Kong.

Beckett, Lois. "Getting beyond Just Pageviews: Philly.com’s Seven-part Equation for Measuring Online Engagement." *Nieman Lab*. President and Fellows of Harvard College, 26 Oct. 2010. Web. 22 Nov. 2014. <http://www.niemanlab.org/2010/10/getting-beyond-just-pageviews-philly-coms-seven-part-equation-for-measuring-online-engagement>

Dobrian, Florin, et al. “Understanding the impact of video quality on user engagement.” Proceedings of the ACM SIGCOMM conference, 2011. Pages 362-373.

Haven, Brian and Vittal, Suresh. “Five Tools and Technologies to Measure Engagement.” 2008. Web. 19 Nov. 2014. <http://www.adobe.com/engagement/pdfs/five_tools_and_tech_to_measure.pdf>

Hewett, Thomas T., et al. "CHAPTER 2: Human-Computer Interaction {p. 5}." *ACM SIGCHI Curricula for Human-Computer Interaction: 2. Definition and Overview of Human-Computer Interaction*. ACM SIGCHI, 1992. Web. 21 Nov. 2014. <http://old.sigchi.org/cdg/cdg2.html>.

"How Does the APA Define "psychology"?" American Psychological Association, 2014. Web. 19 Nov. 2014. <http://www.apa.org/support/about/apa/psychology.aspx#answer>.

IAB Internet Advertising Revenue Report. Rep. Interactive Advertising Bureau, Apr. 2014. Web. 25 Nov. 2014. <http://www.iab.net/media/file/IAB_Internet_Advertising_Revenue_Report_FY_2013.pdf>

"Marketing." Oxford Dictionaries, 2014. Web. 21 Nov. 2014. <http://www.oxforddictionaries.com/us/definition/american_english/marketing>.

Peterson, Eric T. and Carrabis, Joseph. “Measuring the Immeasurable: Visitor Engagement.” *Web Analytics Demystified*, 2008. Web. 19 Nov. 2014. <http://www.webanalyticsdemystified.com/downloads/Web_Analytics_Demystified_and_NextStage_Global_-_Measuring_the_Immeasurable_-_Visitor_Engagement.pdf>

Sharma, Saurabh. "UEO 2014: The Second Workshop on User Engagement Optimization at the 20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining." *The Journal Database*. TJDB, n.d. Web. 21 Nov. 2014. <http://www.tjdb.org/Browse-cfp-151.htm#.VG_ctI3wvVJ>.