I Can Wait a Minute: Uncovering the Optimal Delay Time for Pre-Moderated User-Generated Content on Public Displays

Miriam Greis[1], **Florian Alt**[1,2], **Niels Henze**[1], **Nemanja Memarovic**[3]

[1]University of Stuttgart
VIS
Pfaffenwaldring 5a
70569 Stuttgart, Germany
firstname.lastname@vis.uni-stuttgart.de

[2]University of Munich
Media Informatics Group
Amalienstraße 17
80333 München, Germany
florian.alt@ifi.lmu.de

[3]Universita della Svizzera
Italiana
Via G. Buffi 13
6904 Lugano, Switzerland
nemanja.memarovic@usi.ch

ABSTRACT

Public displays have advanced from isolated and non interactive "ad" displays which show images and videos to displays that are networked, interactive, and open to a wide variety of content and applications. Prior work has shown large potential of user-generated content on public displays. However, one of the problems with user-generated content on public displays is moderation as content may be explicit or troublesome for a particular location. In this work we explore the expectations of users with regard to content moderation on public displays. An online survey revealed that people not only think that display content should be moderated but also that a delay of up to 10 minutes is acceptable if display content is moderated. In a subsequent in the wild deployment we compared different moderation delays. We found that a moderation delay significantly decreases the number of user-generated posts while at the same time there is no significant effect on users' decision to repeatedly post on the display.

Author Keywords

public displays, content moderation, Twitter

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION AND BACKGROUND

Urban spaces are getting crowded with public displays [5], from small screens showing menus in bars to large ones covering an entire building facade. Although they are mainly singular installations that show Powerpoint slides and still images it is not hard to imagine that they will soon be connected over the Internet, to form a novel communication medium open to a variety of content and applications − so-called open display networks [3].

CHI'14, April 26 – May 01 2014, Toronto, ON, Canada

ACM 978-1-4503-2473-1/14/04?$15.00.
http://dx.doi.org/10.1145/2556288.2557183

Previous research has shown content creation to be one of the crucial problems for public displays and this content is often expensive, both in terms of human resources and monetary value [9]. On the other hand creating content for social networking services is considered "dirt cheap", e.g., Twitter reports staggering 340 million tweets that are posted daily. Integrating user-generated content, e.g., tweets, onto open display networks is a possible solution for content creation in open display networks and has been explored in prior work [7]. It would also allow public displays to be integrated more into users' "communicative ecology" [6]. Yet, posting user-generated content on public displays comes with the problem of *content moderation* as explicit and inappropriate content could appear. In addition, posted content might be problematic in other ways for a display's particular physical location. For example, previous work reported on inadvertently posted corporate information [2].

Prior research suggests different ways of moderating public displays content, including pre-moderation [10], post moderation [2, 4], and moderation based on audience feedback [1]. However, these works only applied a certain strategy without closely investigating the effects. In our work we aim to gather an in-depth understanding of pre-moderation as we believe this to be a central prerequisite in open display networks to encourage display owners to allow user-generated content. Prior work looked at the impact of labelling content [8] whereas we focus on the delay times caused by the review process. We believe this to be a major challenge with pre-moderation. The fact that posts do not appear instantly raises the following questions: (1) What do users expect when their content appears? (2) What effect does the delay cause? (3) How can the user be notified that content is under review and when it will appear? If a system fails to address these challenges, users will wonder where the problem occurred and either resend the content or stop using the display.

To investigate pre-moderation of content on public displays we conducted two studies. First, we distributed a survey that investigated users' expectations of optimal timing for content that is pre-moderated. We found 10 minutes to be an acceptable delay for more than 70% of the users. Within this time-frame different forms of content moderation are possible, including (a) automatically, (b) by the owner, or

(c) crowdsourcing-based approaches. We then developed an application allowing people to send tweets to a public display network. We deployed the application in the wild on five connected displays and inferred an artificial delay to investigate the effect of different pre-moderation mechanisms.

The contribution of the work is twofold. First, we report on *user preferences for content upload waiting time* on public displays (a) with and (b) without moderation. Our results show that if users are aware of moderation, more than half of them is even willing to accept delays of one hour or more. For applications that do not communicate the moderation process, a delay of up to 10 minutes is still acceptable for the majority of the potential users. Second, we provide insights about the *effect of a moderation delay on users' behavior*. Through an empirical study we found that even short waiting times of 90 seconds can confuse users. Furthermore, we show that the longer the delay time the less posts appear on a display. Finally, the delay time seems to not influence a user's decision to continuously post to a display.

STUDY I: ONLINE SURVEY

To learn about potential users' expectations towards content moderation for open display networks we distributed an online questionnaire. We were mainly interested in the expectations of people that have no previous experience with publishing user-generated content on public displays to be able to later contrast the results with the actual behavior in front of public displays.

Method

We prepared the questionnaire using our LimeSurvey server. The questionnaire consists of three pages. After an introduction to public displays and user-generated content on the first page, the second page asks about basic demographics and if the person ever published content on a public display or even operates one. The third page asks participants if content on public displays should be moderated; if participants expect that a message they post to a display would appear instantly; if they expect a delay prior to showing the message and which delay is acceptable. Since no prior in-depth investigation on delay times exists (prior research that comes closest to investigating delay times simply classified them as "fast" or "slow" [11]), we came up with concrete values that we felt would best reflect a user's expectations.

We distributed the questionnaire through our university's mailing list, our social network, and our research projects' mailing lists. In total, 114 participants completed the questionnaire. 50 participants are female and 64 are male. On average they are 24.10 years old (SD=4.54). 89 participants are students with a variety of majors including computer science, biology, and medicine.

Results

101 participants never posted content on a public display or operated one. 10 participants used research prototypes or installations in museums to post content on public displays and 7 participants operate, maintain, or work with public displays. We excluded the 13 participants that had previous experience with public displays from further analysis.

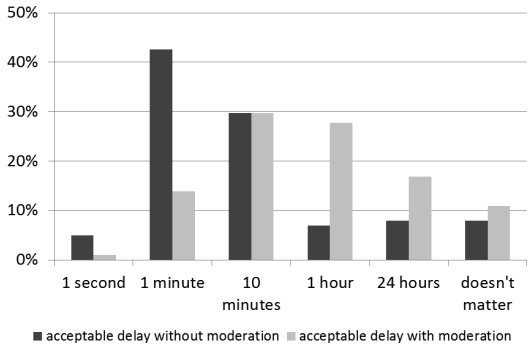

Figure 1. The acceptable delay for messages to appear when posting messages on a public display without or with moderation.

83% of the participants agree or strongly agree with the statement that a moderation or control process for content on public displays is necessary to avoid misuse. The average level of agreement is 4.22 (SD=1.05) on the 5-point Likert scale (1=don't agree at all, 5=totally agree). With an average of 3.89 (SD=1.08) on the 5-point Likert scale most participants agree with the statement that they expect posted messages to appear on a display instantly (81% agreed or strongly agreed). Asked about the acceptable delay (see Figure 1), 48% state that only a delay of 1 minute is acceptable if there is no content moderation. Another 30% think that a delay time of up to 10 minutes is still acceptable. If there is content moderation, 59.4% of the participants would accept a longer delay. 55.5% would even accept a delay of one hour or more.

Discussion

The results indicate, that communicating the fact that a display is moderated overall leads to an increase of the acceptable delay. Yet, a challenge in open display networks may be the fact that moderation is handled differently across displays, even if running the same application. For example, a digital bulletin board may be globally available for use on public displays but then, supermarkets running the application may require moderation whereas universities may not. This raises the need to understand, how to employ content moderation without the user being aware of it.

We see different opportunities to enable pre-moderation of open display content within a time frame of 10 minutes. First, posts could be checked *automatically* (less than one second) by searching for explicit keywords and potentially banning the content from publication. Second, display owners may employ dedicated moderators to instantly approve or reject content by *manually* checking incoming items (i.e., within 30 seconds). Third, display or space owners may decide to outsource the reviewing task via *crowdsourcing* platforms such as Mechanical Turk. We believe that in such cases a response time of 90 seconds is feasible. In the following we explore these options by means of a real-world deployment.

STUDY II: IN THE WILD DEPLOYMENT

We implemented a web-based app displaying tweets which mention the display network's Twitter profile. In this way people can post text and images to the display by adding the network's Twitter id to a tweet.

Figure 2. The five display locations with the running display client. Locations included entrance areas, cafeterias, and coffee kitchens.

The display client is depicted in Figure 2. It runs in a full screen browser and shows the 12 most recent tweets in a 3x4 grid in portrait and a 4x3 grid in landscape mode. The Twitter id to tweet to is displayed below the grid. If a tweet only contains text we just render the text and if it contains a link to a photo only the photo is shown. Each tweet is logged in a database together with a timestamp and the user's Twitter ID. If a tweet contains a link to a photo we also store the photo in the database. Hence we cannot only easily exclude explicit tweets but we can later contact the poster via his or her ID, e.g., to send an online survey.

The display client polls new tweets from the database every 1000ms. We do not provide any other interaction than showing the messages posted on Twitter to keep the application simple and concentrate on users' reactions to delays caused by moderation.

Deployment Setting

We deployed the web-client on five displays across our campus where it ran 24/7 for the duration of 2 months (June–July 2013). All content appeared simultaneously on all five displays. Two displays were installed in the entrance area of faculty buildings and in close proximity to lecture theaters. A third display was deployed in the vicinity of a coffee kitchen shared by two research groups in one of the university building. The fourth display was deployed in a university cafeteria. The display was mounted on the wall in close proximity to tables but was visible from almost any location inside the cafeteria. The last display was located in the main canteen building of the university with a throughput of several thousand people per day. The display stood at the intersection of two aisles with tables in the vicinity. Passersby for all displays were both university employees as well as students attending lectures and courses. Figure 2 depicts the five locations.

According to the previously mentioned options for pre-moderation, we configured our display application to support a 0, 30, and 90 seconds delay. The time intervals were changed every 2 hours. Lectures on campus follow a centralized schedule which allowed us to always switch the time intervals while lectures were in progress. The reason for this is that most people were present shortly before or after the lecture as they waited for fellow students. We did not communicate the delay times to users simulating pre-moderation without informing users. During the time of deployment 518 messages were posted by 116 different users.

Questionnaires

Each time a user posted to the display for the first time, we sent a personalized link to an online questionnaire via Twitter encoding the user ID, the tweet, and the condition in the link. To ensure that the correct user answered the questionnaire, participants had to enter their Twitter ID as they filled in the questionnaire. In the questionnaire we first asked participants about demographics, previous experience with public displays, whether they were in front of the display when posting, and whether they waited until their message appeared. In addition, we were interested in their overall experience.

Out of the 116 people who posted to the display and were asked to fill in the questionnaires, 2 female and 29 male participants replied. They were students or employees at our university in different engineering fields (avg. age = 23.71 years, SD=4.70). 9 participants interacted with a public display before using our application. 23 participants posted while being in front of the display out of which 17 waited in front of the display for their post to appear (48.6 %). 14 participants posted in the 0-seconds condition, 11 in the 30-seconds condition, and 6 in the 90-seconds condition.

Participants were asked to rate their experience with the application. Overall, participants give an average rating of 4.16 (SD=0.90) on a 5-point Likert scale (1=very low, 5=very high). The average rating after posting in the 0-seconds condition is 4.00 (SD=0.96), 4.55 (SD=0.69) after posting in the 30-seconds condition, and 3.83 (SD=0.98) after posting in the 90-seconds condition. With an average of 3.35 (SD=1.56) on the 5-point Likert scale most participants agree with the statement that they expect their message to instantly be displayed.

Data Analysis

To understand the effect of moderation times on content behavior we analyzed the number of posts as well as how often participants posted on the display. We considered multiple consecutive posts by a user that depended on each other (e.g., a large picture) as a single post. This resulted in 349 posts that remained for the analysis. First, we found an increase in the overall *number of posts* as moderation time decreases. During the 0-seconds condition, 150 messages were posted, compared to 127 in the 30-seconds condition and only 72 in the 90-seconds condition. The difference is statistically significant ($\chi^2(2) = 27.62, p < 0.001$) and indicates a strong effect of moderation delay on number of posts. There was no significant association between the conditions and whether or

not users posted again ($\chi^2(2) = 0.73, p = 0.76$). In the 0-seconds condition, 43,2% posted once, in the 30-seconds condition 51,7%, and in the 90-seconds condition 43,1%. The small difference suggests that delay times do not impact on whether people post again or not.

Observations and Interviews

We observed users of our system during the deployment and invited seven people who filled in a questionnaire to semi-structured interviews. We were particularly interested in the first contact of the people with the display, their first message, their expectations, and why they decided to post again.

Six of the seven interviewees stated that they did not post on the first encounter but, instead, first observed the display and its mechanisms before making the decision to post themselves (*"My friends posted something first. They took a photo of themselves and waited in front of the display until the message was shown."*–P4). Interviews and observations also revealed that users often were not sure whether the display really works. They discussed with others if a moderation process was in progress, particularly because they felt 90 seconds were quite a long time for a message to be displayed. As a result, many of them spent a significant amount of time in the display vicinity. We observed that people who decided to post even sat down on the floor in front of the display to be able to post and observe the display at the same time.

Five interviewees stated that after they verified that the display is actually working they started to post remotely quite extensively, i.e. while they were not in the display vicinity. It seems that people that later frequently post 'give credit' to the display the first time, hence forgiving even longer delays. This may be an explanation why the delay time does not have any impact on whether people post or not.

Interviewees also had quite different ideas as to when they expected the posted content to appear. Whereas two did not expect any moderation, two others expected moderation and did not think that their posts would appear instantly.

We also observed a *digital honeypot effect* where posts coming in on the display as people were watching motivated them to post themselves, sometimes directly replying to the new post. This effect seems to also impact on expectations towards content moderation as new posts not only communicate what is appropriate but also that moderation is currently in progress and new posts can be expected to appear shortly.

DISCUSSION & CONCLUSION

In this paper, we investigate the moderation of user-generated content on open public displays. We assessed users' expectation about content moderation through an online survey and used an in the wild deployment of a public display network to compare different moderation delays.

We learned that, in general, most users expect content to be moderated. Content should appear within one minute without content moderation and within 10 minutes with content moderation. The in the wild deployment revealed that delay times caused by content moderation significantly influence the number of user-generated posts on a display. Hence, for display applications that aim to maximize the number of people posting, post-moderation or automated moderation, for example through simple keyword spotting, is advisable.

At the same time we discovered no effect on the users' decision to post – neither for cases where people posted just a single piece of content nor where they posted multiple times. We believe the reason for this to be the fact that people try to compensate for the lack of information on the moderation process by observing the display prior to posting. Once they found out about moderation they posted without reservations.

In the future we plan to build upon our findings and investigate different mechanisms that convey (a) how the display is moderated and (b) what the expected delay time is.

Acknowledgments: This work is partly funded by the German Research Foundation within the SimTech Cluster of Excellence (EXC 310/1).

REFERENCES

1. Alt, F., Kubitza, T., Bial, D., Zaidan, F., Ortel, M., Zurmaar, B., Lewen, T., Shirazi, A. S., and Schmidt, A. Digifieds: insights into deploying digital public notice areas in the wild. In *Proc. MUM'11*, ACM (2011).
2. Churchill, E., Nelson, L., Denoue, L., Murphy, P., and Helfman, J. The plasma poster network. In *Public and Situated Displays*. Kluwer, 2003.
3. Davies, N., Langheinrich, M., Jose, R., and Schmidt, A. Open display networks: A 21st century communications medium. *IEEE Computer* (2012).
4. Elhart, I., Memarovic, N., Langheinrich, M., and Rubegni, E. Control and scheduling interface for public displays. In *Adjunct Proc. UbiComp'13* (2013).
5. Kostakos, V., and Ojala, T. Public displays invade urban spaces. *Pervasive Computing* (2013).
6. Memarovic, N., Langheinrich, M., Rubegni, E., David, A., and Elhart, I. Designing "interacting places" for a student community using a communicative ecology approach. In *Proc. MUM'12*, ACM (2012).
7. Munson, S. A., Rosengren, E., and Resnick, P. Thanks and tweets: comparing two public displays. In *Proc. CSCW '11*, ACM (2011).
8. Schroeter, R. Engaging new digital locals with interactive urban screens to collaboratively improve the city. In *Proc. CSCW '12*, ACM (2012).
9. Storz, O., Friday, A., Davies, N., Finney, J., Sas, C., and Sheridan, J. Public ubiquitous computing systems: Lessons from the e-campus display deployments. *Pervasive Computing* (2006).
10. Taylor, N., Cheverst, K., Fitton, D., Race, N. J., Rouncefield, M., and Graham, C. Probing communities: study of a village photo display. In *Proc. OzCHI* (2007).
11. Wise, K., Hamman, B., and Thorson, K. Moderation, response rate, and message interactivity: Features of online communities and their effects on intent to participate. *Comp.-Med. Comm. 12*, 1 (2006).