A System Designed to Collect Users' TV-Watching Data Using a Smart TV, Smartphones, and Smart Watches

Jehwan Seo
Core Tech Lab, VD Division
Samsung Electronics
jehwan.seo@samsung.com

Hyun-kyu Yun
Core Tech Lab, VD Division
Samsung Electronics
hyunkyu.yun@samsung.com

Hyunchul Lim
Human Centered Computing Lab
Seoul National University
hyunchul@snu.ac.kr

Bongwon Suh
Human Centered Computing Lab
Seoul National University
bongwon@snu.ac.kr

Changhoon Oh
Human Centered Computing Lab
Seoul National University
yurial@snu.ac.kr

Joongseek Lee
User Experience Lab
Seoul National University
joonlee8@snu.ac.kr

TVX'16, June 22-24, 2016, Chicago, IL, USA
ACM 978-1-4503-4067-0/16/06.
http://dx.doi.org/10.1145/2932206.2933562

Abstract

In this study, we suggest an enhanced smart TV logging system composed of a smart TV, smartphone, and smart watch. It can be used to research the audience's complicated and segmented behaviors while watching TV. We designed a prototype of the system, which can not only detect whether viewers are located in the TV-viewing area but also measure their movements and activities by analyzing beacon signals and sensor data from the smart watch. We conducted a technical evaluation to verify its fidelity and measure its performance, and a user study identified what factors affect the users' level of engagement with the TV content. The experiment results showed that the system accurately detected and measured users' location and engagement levels while watching TV. We found that smartphone usage while watching TV is important in understanding users' TV-viewing behavior.

Authors Keywords

TV-watching behavior; beacon; smart TV; smart watch; TV rating; machine learning.

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI), miscellaneous

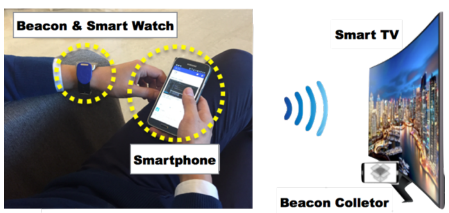

Figure 1: The research prototype system comprises a smart watch with a beacon, a smartphone, and a smart TV with a beacon collector.

Introduction and Related Work

TV-viewing behaviors are not simple. Viewers do many other activities after turning on the TV, including moving around the house, using a smartphone, and doing chores, which may distract viewers from watching TV [2, 9]. Hence, advertisers and TV networks want to know not only if the TV is turned on but also if viewers are paying close attention to TV programs. Understanding the viewers' TV-watching context could help to measure their engagement level. Recognizing a viewer's engagement level could benefit both viewers and TV-rating firms: viewers could receive a personalized service, while TV-rating firms could build enhanced business models by using viewers' engagement level.

To address the issue, we suggested the smart TV logging system [9] that utilized a beacon and smartphones to investigate the viewer's TV-watching behavior. However, the system has a number of limitations: (1) viewers do not necessarily carry their smartphone with them all the time, (2) the system does not capture viewers' activities not involving digital devices (e.g., house chores), and (3) multiple beacons incur the high cost of installation and maintenance.

In order to address these issues, we propose an enhanced smart TV logging system by adding smart watches with a beacon to the previous system. Using smart watches has many benefits. First, TV viewers can wear a smart watch all the time, which can help avoid data loss. Second, the inertial sensors embedded in smart watches can track the motion of viewers and infer their activities [3, 4] while they watch TV. Lastly, our system uses a few beacons for detecting viewers' location because it requires only one beacon per a viewer.

To verify the feasibility of our system, we installed a prototype in two types of apartments and conducted experiments, consisting of a technical evaluation and user study. The results of the technical evaluation show that (1) both beacon signals and inertial sensors from the smart watches can accurately detect whether viewers are in front of TV, and (2) they provide information to infer viewers' activities while watching TV. Moreover, through the user study, we found that the TV-viewing engagement level could be measured based on the system logs.

System Design

We designed an enhanced smart TV logging system that comprises smart watches with a beacon, smartphones, and a smart TV with a beacon collector (Figure 1). Our goal is to collect viewers' TV-watching data from the system for understanding their TV-viewing behavior. To attain this goal, we used the RSSI (Received signal strength indiccation) [8], accelerometer, and gyroscope data from smart watches to keep track of users' activities and then detect whether users are located in TV-watching zone. Along with these users' information, smartphone app-usage data were employed to figure out what users actually did while watching TV.

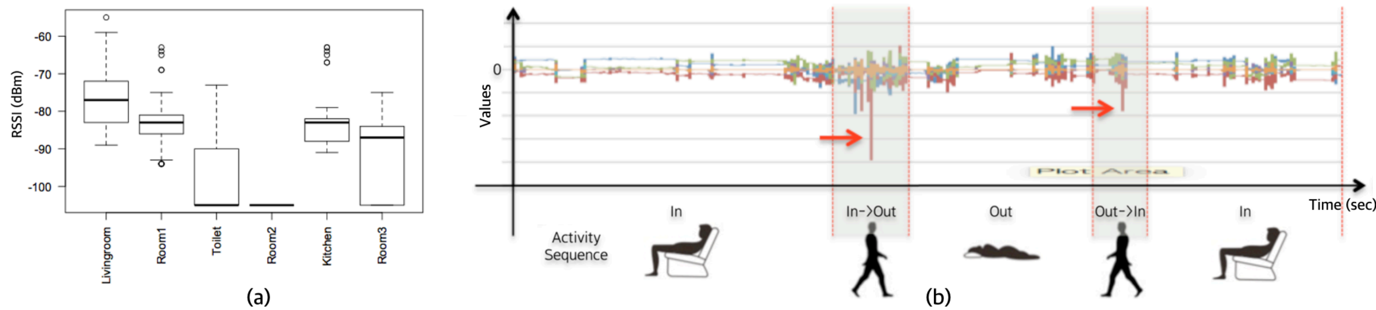

Figure 3: (a) Box plot of RSSI from a smart watch when a participant moves around the house. (b) The six values from an accelerometer and a gyroscope show typical patterns when a person moves from one place to another.

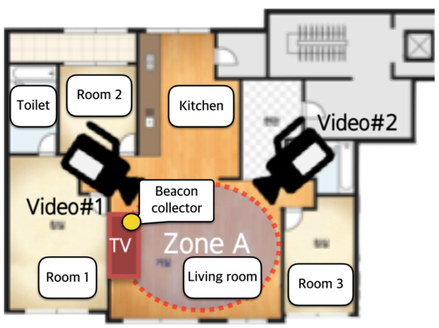

Figure 2: We assumed that the TV with the beacon collector is located in the living room (called Zone A), where the person usually watches TV. For the user study, we recorded the TV screen and viewers with two cameras.

To assess the effectiveness of data metioned above as features for detecting viewers' location, we performed a pilot study with two participants who we asked to move around a house while wearing a smart watch. We collected the data at 100 Hz from a beacon (Gimbal) and a smart watch (Samsung Galaxy Gear Live). As shown in Figure 3, when participants moved from one place to another, (a) there are differences in RSSI values among locations in a house. Also, (b) the accelerometer and gyroscope values had typical patterns when participants were moving or staying in the places. From our initial insights, we select RSSI and sensors' values to detect whether viewers are located in a certain area as Zone A, the user's TV-watching zone (Figure 2).

In order to determine if a person is not watching TV when he or she is located in Zone A, we collected smartphone usage logs using the App Usage Tracker [1]. We analyzed the usage logs and the activitiy logs to identify the relationship among smartphone use, viewer's activitiy and TV viewing.

Technical Evaluation

In the first part of the evaluation, we conducted a technical evaluation of the prototype. We estimated how the system accurately finds the user's location and how much it has been enhanced according to the each system component.

Method

We recruited two households for the experiments and installed the prototype in these two houses. Both houses were typical Korean apartments. One is an 80-m^2 apartment with three bedrooms, one bathroom, a living room, and a kitchen. The other is a 159-m^2 house with four bedrooms, two bathrooms, a living room, and a kitchen. Two family members from each household took part in the experiment, and we gathered data from a total of four participants (two males and two females aged between 29 and 38). Participants wore a smart watch with a beacon on the left or right wrist. A beacon collector was installed in front of the TV. This setup allowed us to receive the beacon's signal from the smart watch and then identify if viewers were in

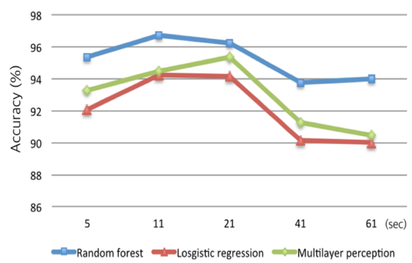

Figure 4: The accuracy of RSSI and sensor values combined with different size of window and classifiers (ten-fold cross validation).

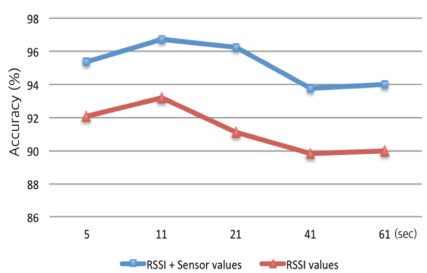

Figure 5: The accuracy of RSSI values alone vs. The accuracy of RSSI and sensor values combined.

Zone A (Figure 2). We simultaneously collected three types of data: (1) we recorded the TV screen and viewers' TV-viewing activities with two video cameras (Figure 2); (2) we collected beacon signal logs and sensor values from the smart watches at 100 Hz; and (3) we gathered smartphone usage logs with the App Usage Tracker application on the participants' smartphones. Each experiment lasted about two hours.

After the experiment, we analyzed the beacon signal logs and measured how the signals detected if viewers were in the TV-watching zone using machine-learning approaches, such as random forest, logistic regression, and multilayer perceptron.

Result 1. The Effect of Sensor Values, Window Size and Majority-Voting Method
First, we examined the effect of the window size because it could affect the performance of classifiers [6]. We examined five types of window sizes: 5, 11, 21, 41, and 61 seconds. We used RSSI and sensor values combined to validate the performance of random forest, logistic regression, and multilayer perceptron. The results (Figure 4) show that accuracy improves until the 11-second window. The random forest approach performed better than the other two classifier methods. Thus, we decided to use the random forest technique with the window size of 11 for the following evaluation.

Next, we investigated the effect of the sensor values from the smart watches. We compared the accuracy of RSSI values alone with the accuracy of RSSI and sensor values combined. As shown in Figure 5, using both beacon data and sensor values achieved high accuracy of 96.58% with small standard deviation.

Lastly, since RSSI values fluctuate considerably due to environmental factors such as reflections and wall damping [5], we used the majority-voting technique to improve the accuracy of the system. The result shows that the accuracy of using RSSI values alone increased from 91.16% to 94.00.%, and the accuracy of using beacon and sensor values combined increased from 93.95% to 96.58%.

Result 2. Per-User Classifiers vs. Per-Location Classifiers
Inspired by Harrison et al. [7], we evaluated our system in three different conditions: ten-fold cross validation using all data, per-user classifiers, and general classifiers.

	System Accuracy (%)			
	BO	**BO (MV)**	**BO+SV**	**BO+SV (MV)**
10-fold cross validation	91.16%	94.00%	93.95%	96.58%
Per-user classifiers	90.90% (1.97%)	94.00% (1.17%)	94.20% (0.79%)	96.80% (0.72%)
Per-location classifiers	90.04% (1.62%)	94.36% (1.48%)	94.67% (0.11%)	97.11% (0.77%)

Table 1: Accuracy (SD in parentheses) of our system: ten-fold cross validation with all data, per-user classifiers, and per-location classifiers. BO = beacon data only; SV = beacon and sensor values combined; MV = majority-voting technique used in the data.

Ten-Fold Cross Validation with All Data: We ran a conventional ten-fold cross validation using all data from the four participants. As shown in Table 1, the 10-

Response	Count
Watched	62
Don't know	6
Didn't watch	22
Total	90

Table 2: Video clip survey responses

.

	Video Clip Survey	
	Watched	Didn't Watch
TV Watching Zone In/Out (times)	In: 62	In: 7
	Out: 0	Out: 15
Smartphone Use(times)	8	5

Table 3: Video clip survey results classified by TV watching zone in/out and smartphone usage.

fold cross validation showed high accuracy of over 96% when using combined beacon and sensor values and majority-voting technique. This result gives us a basic evaluation where the users' data is known a priori, and we can train and adjust a model to a particular group of users. Two further evaluations and analyses examined our techniques in more realistic situations.

Per-User Classifiers: It is important to understand how the features perform at a per-user level [7]. We divided the data between the participants and conducted a ten-fold cross validation for each participant's data. The features are not related to specific users (see Table 1; there is a small standard deviation).

Per-Location Classifiers: The experiment locations had different dimensions. Therefore, we performed an analysis to find out if spatial dimensions affect our system and if we can use it as a general model regardless of house size. Table 1 shows that house size does not affect system accuracy.

User Study

Participants completed a survey to identify in detail what actually happens while they watch TV. In other words, we aimed to identify users' level of engagement and whether they focused on TV content, even when the system log showed they were in front of the TV.

Methodology

To measure the users' engagement level with TV content, we provided the participants with a simple survey about the TV content and conducted a post hoc interview. The survey consisted of video clips and questions asking whether they remember the content. We extracted the video clips from the content the participants had watched and edited them into one-minute units. We randomly showed the clips to the participants and asked them to choose an answer from three options: (1) 'watched', (2) 'don't know', and (3) 'didn't watch.' We collected a total of 90 responses from the participants. After the survey we interviewed the participants about what they thought of the system and what factors could affect their TV-watching level of engagement.

Result

From 90 answers, the number of 'watched' was 63, the number of 'don't know' was 6, and the number of 'didn't watch' was 22 (Table 2). The result shows that the participants watched the TV more than half of the time. However, even if the participants were located in front of the TV, they sometimes lost concentration and did other things.

We classified each response according to the system log information, such as users' location. As shown in Table 3, among the responses that were classified in Zone A (69), 62 responses were the cases where the participants actually watched the content. However, in the remaining seven cases (10.1%), they didn't concentrate on the TV even if they were in front of the TV. To investigate what factors affected their distraction, we checked the users' smartphone app usage from the system logs and their behavior video data. In four out of seven cases, the participants showed various non-TV-watching behaviors, such as doing the laundry or talking to other people. In the other three cases, the participants used their smartphones and did not concentrate on the TV content. Using a web browser, they sometimes checked portal sites and read news articles. One participant looked at old pictures in his Gallery folder for more than three minutes. Noting the importance of the smartphone, we also checked smartphone usage in the cases where users answered that they watched the video clips. We identified eight cases, and users often browsed messengers apps or Facebook, or they sometimes searched the Internet for information related to the TV content they were watching. These results show that it is important to integrate users' smartphone usage in order to comprehensively understand viewers' TV-watching behavior and measure their level of engagement with TV content.

We briefly present the result of the post hoc interview, which might be helpful to understand the participants' thoughts about the experiments and their engagement with the TV screen. P01 said, "Remember the first 30 seconds, but I do not remember the rest!" Because we had split the video clips into one-minute units, users could not entirely remember the content if they had moved or been distracted when watching the unedited TV content during the experiment. Results from the technical evaluation logs support these explanations. The participants frequently moved in and out of Zone A during the experiment (62 times). Meanwhile, the interview let us discover the effect of smartphone usage on the viewers' TV immersion. We asked P02 to describe the video content played when he was using his smartphone, and he said, "This scene looks as though I haven't been paying much attention." Lastly, the issue about the boundary of TV-watching Zone A was also raised. P03 said, "I was not in the living room, but I was watching once in a while in the room while doing my chores." This was possible because the door of the room was open, and the participant could watch TV without entering the living room. In the experiment, we made the TV-watching zone a fixed area, such as the living room. However, depending on the TV's location or the layout of the rooms, the boundary of the TV-watching zone could be enlarged into marginal areas.

Conclusions and Future Work

In this study, we proposed a system that can be used to accurately measure and analyze the diverse and subdivided patterns of users' TV-watching behavior. Above all, by integrating users' smart watches into the system, we could overcome the disadvantages of previous user location tracking systems. In order to evaluate our system's feasibility, we installed prototypes and conducted a technical evaluation and user study. The result of the experiment showed that the accuracy and reliability of the proposed system were significantly acceptable. Moreover, from the results of the user study, we could find out users' behaviors while watching TV and what factors affect their level of engagement with the TV content.

Although we obtained statistically significant results from the technical evaluation of the proposed system, we could not gather enough data to analyze users' TV-watching patterns. Obtaining clearer data and training the model will lead to more accurate study results. In future work, we plan to conduct a large-scale user study to overcome this drawback and to introduce various sensor technologies in the system design. In addition, a pattern recognition method could be used to monitor the TV-watching behavior of users in Zone A. Information about viewers' smartphone application usage while watching TV, their search queries, and their application usage time could be used to calculate the specific level TV viewers' immersion.

References

1. App Usage Tracker https://play.google.com/store/apps/details?id=com.agrvaibhav.AppUsageTracking&hl=ko
2. Abreu, Jorge, et al. "Viewer behaviors and practices in the (new) Television Environment." In Proc. on Interactive TV and video. ACM, 2013
3. Bieber, Gerald, Marian Haescher, and Matthias Vahl. "Sensor requirements for activity recognition on smart watches." *Proceedings of the 6th International Conference on PErvasive Technologies Related to Assistive Environments*. ACM, 2013.
4. Case, Meredith A., et al. "Accuracy of smartphone applications and wearable devices for tracking physical activity data." *Jama* 313.6 (2015): 625-626.
5. Feldmann, S., Kyamakya, K., Zapater, A., & Lue, Z. (2003, June). An Indoor Bluetooth-Based Positioning System: Concept, Implementation and Experimental Evaluation. In *International Conference on Wireless Networks* (pp. 109-113).
6. Fujinami, Kaori, and Satoshi Kouchi. "Recognizing a Mobile Phone's Storing Position as a Context of a Device and a User." *Mobile and ubiquitous systems: computing, networking, and services*. Springer Berlin Heidelberg, 2012. 76-88.
7. Harrison, Chris, Julia Schwarz, and Scott E. Hudson. "TapSense: enhancing finger interaction on touch surfaces." Proceedings of the 24th annual ACM symposium on User interface software and technology. ACM, 2011.
8. RSSI, Wikipedia https://en.wikipedia.org/wiki/Received_signal_strength_indication
9. Seo, J., Kim, D., Suh, B., & Lee, J. (2015, April). Design of a Smart TV Logging System Using Beacons and Smartphones. In Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (pp. 2157-2162). ACM.