Smart Signage: A Draggable Cyber-physical Broadcast/Multicast Media System

James She and Jon Crowcroft
Computer Lab
University of Cambridge
Cambridge CB2 1TN, UK
Email: james.she@cl.cam.ac.uk,
jon.crowcroft@cl.cam.ac.uk

Hao Fu
Department of Electronic and Computer Engineering
The Hong Kong University of Science and Technology
Hong Kong
Email: hfu@ust.hk

Pin-Han Ho
Department of Electrical and Computer Engineering
University of Waterloo
Waterloo, ON, Canada
Email: p4ho@uwaterloo.ca

***Abstract*—Digital signages are increasingly common for out-of-home advertising. Latest advancements in smartphones, wireless communication and displays make it possible to design smarter interactive signage systems for more effective advertising. Although existing research have started a trend of cyber-physical interactions, they are generally not scalable for multiple users and not intuitive to interact with. Smart Signage - a "draggable" cyber-physical broadcast/multicast (B/M) media system is therefore proposed here. With a novel cyber-physical B/M protocol that enables a display concurrently interacting with physical actions of multiple user smartphones, a large number of users can simultaneously acquire any running content on a display by simply using an intuitive "dragging" hand gesture with their smartphones. Analytical formulations are derived to identify the key parameters and their dynamics in the system, which provide the condition for achieving the average response time of a user "dragging" gesture within 1 second limit. Implementations are demonstrated for possible real-world deployments. Experiments with 30 primary students are conducted to prove the system offering scalable and intuitive interactions for the first-time users even when they are moving.**

***Keywords*-Interactive display, cyber-physical system, draggable broadcast/multicast media, smartphones, advertising.**

I. Introduction

Traditional static signages at public/semi-public areas have been used as a convenient way to broadcast advertisements to a large number of targeted audiences who pass by the signages. As one of the early interactive version, tear-off advertisements are commonly adopted with creative designs as shown in Fig. 1 [1]. A tear-off advertisement allows individuals to physically tear off a part of the signage to keep specific information (e.g., contacts and websites) for follow-up advertising actions.

Digital signages, as an attractive replacement of traditional static signages for out-of-home advertisements, have gained increasing popularity due to the appealing multimedia presentation. Whereas, it lost the physical advantages of tear-off advertising signages for users to acquire and keep the follow-up advertising information. On the other hand, digital signages generally do not provide the interactivity to multiple users concurrently. However, the recent advances in wireless networking and sensing capabilities in smartphones makes it possible now to introduce a smart signage system in this work. A novel cyber-physical interactive modality is therefore proposed to provide multiple users concurrently with scalable and intuitive interactions. A user could feel like physically "dragging" a tear-off advertisement through their smartphones along with cyber functionalities for more effective advertising.

Figure 1. Examples of tear-off advertisement: (a) conventional tear-off advertisement; (b) tear-off advertisement with creative design [1].

Generic interactive signage systems, as shown in Fig. 2, can be abstracted into three major components, which are the *display*, *system* and *interactive modality*. The *display* represents a visualization equipment, which could be a flat-panel LCD unit, a screen with projector, a digital billboard, a LED matrix display, a 3D display, a PC monitor, or any emerging flexible and transparent display. The *system* is an embedded media playback and control system with/without network connectivity that coordinates the display and interactive modality. The *interactive modality* is the technical method, physical scale and social style such that users can interact with the displayed content.

Although many signage systems today have equipped with modern networking and wireless communication capabilities, most of their interaction modalities so far are unfriendly and limited to one-to-one interaction. Bluetooth [2] [3] [4] is

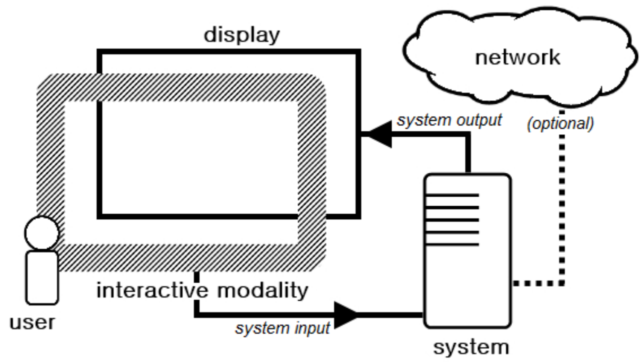

Figure 2. General interactive display system.

Figure 3. Examples of: (a) QR code advertisement [9]; (b) U-tie [10].

a common technology being used in the interactive modality. However, its master-slave pairing architecture is not intuitive to users, in which its scalability is limited to 7 pairings only. Using an camera with image processing is an alternative interactive modality, where the camera can be either on signage-side [5] [6] or user device-side [7] [8]. QR code is a case of user device-side widely adopted in recent years, which provides the user a scalable interaction and follow-up actions after scanning the code by the user smartphone. For a distance that allows a large number of users scanning the code concurrently, a large and unappealing QR code is required to compromise the precious advertising area as showed in Fig. 3(a) [9]. With the powerful capability of smartphones today, novel image recognition techniques on smartphones are employed in some new commercial cases like U-tie in Fig. 3(b) [10]. The attractive images in the advertisement could be thus preserved in a precious signage space, while a large number of multiple user interactions and information retrievals could be achieved by using user device-side camera and computing resources. Unfortunately, this approach requires a good Internet access and relatively a longer processing time without the user moving. Near-field communication (NFC) [11] [12] is an emerging technology for one-to-one interaction that allows users to quickly collect information, but the user smartphones must be within a range of several centimeters. Wi-Fi is another widely accepted technology today for the interactive modality, which has a higher throughput and more flexible interactive range. [13] [14] [15] are the latest attempts using Wi-Fi together with gestures detection by user smartphones to offer intuitive interactive modalities, but they all are limited to one-to-one interaction.

Motivated by the above observations, Smart Signage is proposed here with the key contributions summarized below:

1) Smart Signage: a new cyber-physical "draggable" broadcast/multicast (B/M) media system is proposed, which allows a large number of users to simultaneously interact with a display through an intuitive "dragging" hand gesture. A novel cyber-physical "draggable" B/M protocol is developed to synchronize the displaying content with user gestures on their smartphones. It is easy to integrate into any signage device and user smartphone for scalable interactions.
2) Performance analysis and evaluation: The system performance is characterized by the average user "dragging" response, which is meaningfully modeled with some key system parameters (e.g., the number of advertisements per signage) and their dynamics for the system achieving instant responses within a second. Comprehensive evaluations are conducted to show that satisfying user experiences are achievable in real-life deployments.
3) Implementation: The proposed Smart Signage system is practically implemented using a VxWorks-based signage device and various Google Android- and Apple iOS-based smartphones. The proposed B/M protocol is developed using latest Wi-Fi standards, which successfully demonstrated all expected functionalities.
4) Experimentations: By focusing on the stringent requirement of user experience (i.e., fast "dragging" response within 1 second), the prototype is tested over various aged users and scenarios to prove its scalability by providing intuitive interactivity to multiple users simultaneously even the users are walking around.

To the best of our knowledge, Smart Signage has advantages over other existing approaches in terms of the scalability to engage a large number of users, more flexible interaction range, more intuitive interaction, less visual compromises on the advertisement, higher throughput and supporting interaction with mobile users. These contributions are significant when compared with conventional interactive modalities used in current interactive signage systems such as Bluetooth, cameras with image processing, QR code, and NFC, as summarized in Table I.

The rest of the paper starts with the architecture of the proposed Smart Signage system in Section II. Section III describes the performance analysis of the system. Section IV shows the implementation and experimental results, and the paper is concluded with future works in Section V.

Table I
ADVANTAGES OF DRAGGABLE CYBER-PHYSICAL B/M MEDIA

Metrics \ Technologies	Bluetooth	System-side camera	QR code	Device-side camera	NFC	Smart Signage
Simultaneous user interaction	×	×	×	×	×	✓
More flexible distance	×	×	✓	✓	×	✓
More intuitive interaction	×	✓	×	×	✓	✓
Less visual compromise	✓	×	×	✓	✓	✓
Higher throughput	×	×	×	×	×	✓
Mobility support	×	×	×	×	×	✓

II. PROPOSED SMART SIGNAGE: A CYBER-PHYSICAL DRAGGABLE BROADCAST/MULTICAST MEDIA SYSTEM

Fig. 4 shows the architecture of the proposed Smart Signage system, which mainly involves a digital display, a signage device and smartphones. The signage device is equipped with a wireless router to enable a new interactive modality with user smartphones by adopting the proposed B/M protocol below. Although the current system design uses Wi-Fi in the implementation, any wireless standard supporting B/M radio signals is also applicable. The signage device downloads updated content for the display from a content provider through the Internet access.

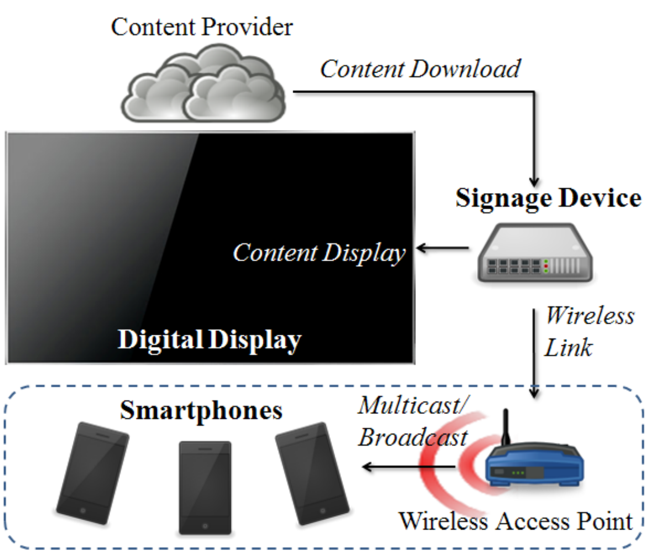

Figure 4. System architecture of Smart Signage.

A. Cyber-physical Broadcast/Multicast Protocol

Different from the conventional B/M protocols (e.g.,[16]) simply only for the content transfers to multiple user devices, the proposed protocol must let multiple user smartphones aware of the existence of a display seamlessly in a physical space, and aware what content on the display a user could physically "drag" at any moment.

A "dragging" event by a user is an unpredictable request for the "draggable" media of a content, which is indeed the version to appear on a user smartphone of the corresponding content showed on the display at the "dragging" moment.

Such "dragging" event by a user is difficult to predict, but the limited number of all "draggable" media of a signage are possible to be all transferred and made available in any user smartphone upfront.

The protocol indicates the updated information about what content is being showed on the display at any moment. A user smartphone therefore is aware of which content on the display is possibly being "dragged", and the corresponding "draggable" media of that content may or may not be already available inside the user smartphone. Fig. 5 shows the sequence diagram of the protocol.

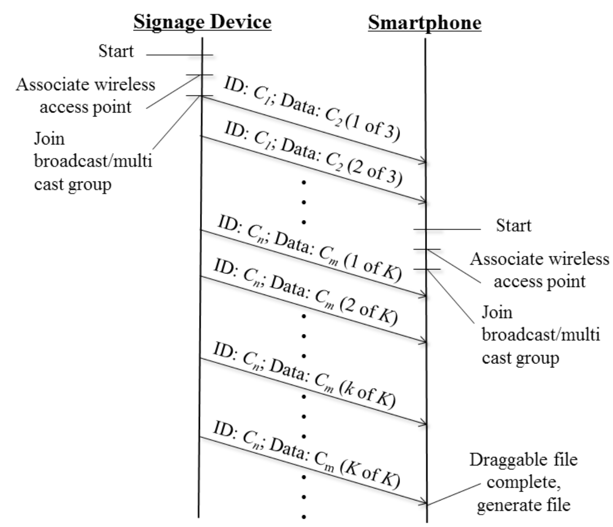

Figure 5. Sequence diagram of the proposed protocol.

Signage device continuously sends B/M packets containing data fragments of a "draggable" media that is corresponding to a content showed on the display in a round-robin fashion. As shown in Fig. 6, each B/M packet generated by the signage device contains: 1) a display ID that indicates the display identity; 2) a showing ID, C_n, (where $1 \leq n \leq N$ and N is the total number of contents), that indicates which content is being shown on the display at that moment; 3) a sending ID, C_m, (where $1 \leq m \leq N$) that indicates which content that the data fragment in this packet belongs to (Note that n is not necessarily equal to m); 4) a data header that indicates the size of the "draggable" file, total number of packets K and current packet sequence number k; 5) data payload that is a data fragment of the "draggable" media file C_m indexed by k (where $k \leq K$).

Display ID (4B)	Showing ID (4B)	Sending ID (4B)	Data Header (4B+4B+4B)
Data Payload (Maximum 1024B)			

Figure 6. Data packet structure.

Smartphones can start and join the same B/M group asynchronously to receive the proposed B/M packets. Once it has received all the data fragments of a "draggable" media file, the "draggable" media of a corresponding content is reconstructed and stored in the buffer as long as the smartphone is still associating with the B/M group. Once a "draggable" event is detected on a user smartphone, it first checks C_n from the received B/M packet for the showing ID, and then checks if the corresponding "draggable" media file is already in the buffer. If the "draggable" file is ready, it will be showed on the smartphone display and copied into another permanent buffer for other application purposes. Otherwise, it will continue receiving the remaining data fragments of C_n. It is advantageous to buffer all the "draggable" media files temporarily for better system responses, as a user "draggable" event is hardly predictable. Any "draggable" file that is not "dragged" and copied to the permanent buffer will be discarded when the smartphone is out of the interaction range.

B. Software Designs of Signage Device and Smartphone

The software designs of the signage device and smartphones is shown in Fig. 7.

Signage device: once it starts running, the program splits into two subroutines: Subroutine 1 checks content changes pushed by the content provider and updates the contents stored in the signage device; Subroutine 2 first joins a specific B/M group and continuously send the proposed B/M packets to the smartphones, and updates the showing ID in the sending packets when the content on the display changes.

It is important to aware that there is always a "draggable" media file associated with a corresponding display file for each advertisement content on a signage. The *display file* is the content shown on the display, whereas the corresponding *"draggable" file*, is the content delivered to and showed on the smartphones. The size of the "draggable" file is generally much smaller than the one for the display.

Smartphones: first joins the same B/M group, and the program splits into two subroutines: Subroutine 1 continuously detects the "dragging" event (i.e., user performing a "dragging" gesture) and displays the "dragged" media on the smartphone;

Subroutine 2 receives the B/M packets transmitted by the signage device and buffers them to reconstruct the "draggable" files.

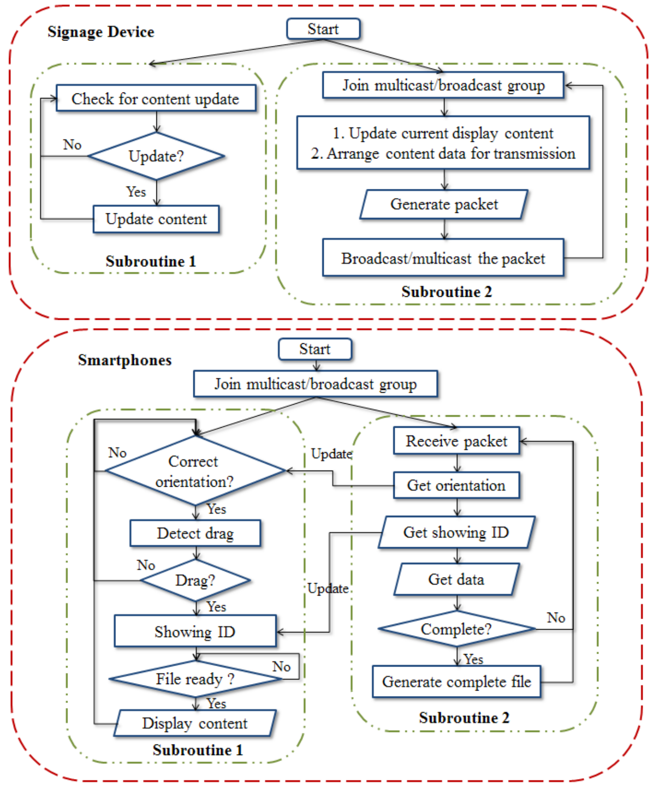

Figure 7. Software flow charts of both signage device (top) and smartphones (bottom).

The smartphone uses built-in accelerometer to track a "dragging" event, which only detects one specific but simple "dragging" hand gesture as shown in Fig. 8. Conventional hand gesture recognition requires gesture spotting and gesture segmentation [17]. Since only one simple gesture is required here makes it possible to detect a "dragging" event simply by checking the total acceleration value, $\overline{A} = \sqrt{A_x^2 + A_y^2 + A_z^2}$ (where A_x, A_y and A_z are the acceleration values of X-, Y- and Z-axis respectively), reaching its maximum.

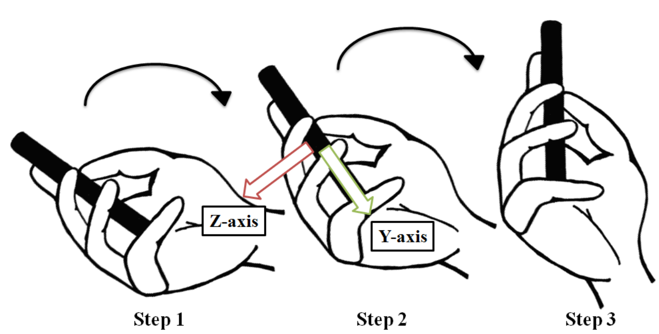

Figure 8. Illustration of the "dragging" hand gesture.

III. Performance Analysis

The interactivity of the proposed system involves both the hardly predictable human processes in the physical world and also the multi-user data transfers in a wireless B/M channel. It is difficult to evaluate the performance in such cyber-physical media system [18] using conventional approaches. In the context of interactive signage, the unpredictable user behaviors must be accounted into the system performance in order to evaluate and engineer the optimal user experience (i.e., the perception after a “dragging” event). To quantitatively characterize such user experience perceived in the interaction process, the *response time* T_r after a “dragging” event perceived by the user is identified as an important evaluation metric.

The timing profile of the system is shown in Fig. 9. The signage device transmits “draggable” files in a round-robin fashion and the smartphone goes through a process characterized by a series of time intervals. *Association time* T_a is the duration of joining a B/M group. *Decision time* T_d is the time interval between smartphone successfully joins the B/M group and the user performs a “dragging” gesture. *Gesture detection time* T_g is the time required for the smartphone to recognize a successful “dragging” gesture. *Transmission time* T_x is the time interval between a “dragging” gesture is confirmed and the intended content is successfully showed on a smartphone.

As discussed, the performance of a Smart Signage is characterized by *response time* T_r, which is the time perceived by the user after a “dragging” event. T_a is excluded from T_r as the association process is prior to an interaction, and the duration is unnoticeable because of some skillful implementation which will be discussed in the next section. Hence, T_r is defined as:

$$T_r = T_g + T_x. \quad (1)$$

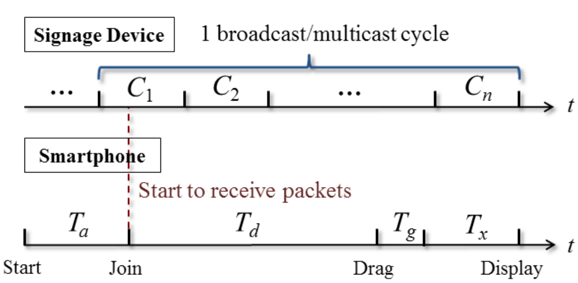

Figure 9. Timing profile of signage device and smartphone.

In Eq. (1), T_g depends on the complexity of the gestures, which is small as only one simple gesture is in this system, and T_x depends on T_d. To be more specific, if T_d is larger than a B/M cycle, then T_x is 0 as the intended “draggable” file has already been buffered, otherwise, T_x is a function of T_d. To compute T_x, some assumptions are made to simplify the problem:

1) An ideal channel is assumed, which means that there will be no packet loss;
2) Each “draggable” file has the same size and is uniformly divided into K datagrams to be transmitted as B/M packets, and the time needed to send out one complete “draggable” file (*file transfer time*) is denoted as T_f;
3) The B/M packets of each “draggable” file will be repeatedly transmitted again in a round-robin fashion after the turns of sending the packets of other“draggable” files are done;
4) $T_c = N \times T_f$ is the time of one complete B/M cycle and the analysis will only focus on T_c after the user has successfully joined the B/M group;
5) Only T_d that falls into the range of $0 \leq T_d < T_c$ would be considered;
6) The user successfully joining a B/M group is a random process, which is uniformly distributed over one complete B/M cycle T_c.

Fig. 10 shows the diagram of a simple example to calculate T_x, where $N = 3$, $T_f < T_d < 2T_f$ and the user is interested in C_3. The formulation is divided into 3 cases according to t, defined as the time interval between the time point when a complete B/M cycle starts and the time point when the smartphone has successfully joined the B/M group, where $0 \leq t < T_c$:

1) *Case 1:* $0 \leq t < T_f$, as only part of C_3 is received, $T_g + T_x = T_c - T_d$;
2) *Case 2:* $T_f \leq t < T_c + T_f - T_d$, where the user has to wait until C_3 has been received, so $T_g + T_x = T_c - T_d - (t - T_f)$;
3) *Case 3:* $T_c + T_f - T_d \leq t < T_c$, as all the packets of C_3 are received before the “dragging” event, $T_x = 0$.

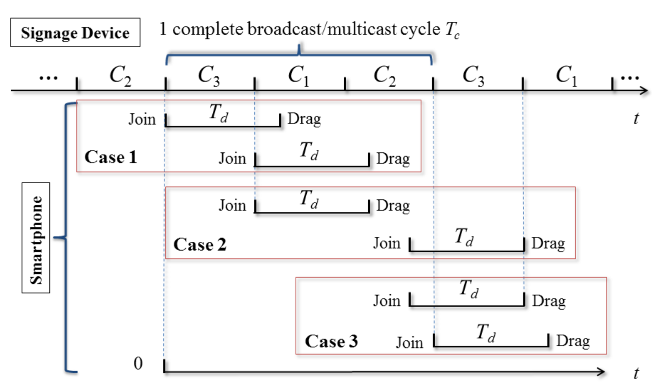

Figure 10. An example to calculate T_x with $N = 3$, $T_f < T_d < 2T_f$ and C_3 is interested.

One thing to notice in this example is that although the user is interested in C_3, the results will be the same for any content, C_n, as one complete B/M cycle is considered. As the gesture detection time T_g is constant and negligibly small because of some skillful implementation discussed before, $T_g = 0$ is assumed in the following analysis further simplify the formulation problem. In this case, the transmission time T_x can be expressed as:

$$T_x(t) = \begin{cases} T_c - T_d & \text{for} \quad 0 \leq t < T_f; \\ T_c + T_f - T_d - t & \text{for} \quad T_f \leq t < T_c + T_f - T_d; \\ 0 & \text{for} \quad T_c + T_f - T_d \leq t < T_c. \end{cases} \tag{2}$$

Eq. (2) is valid for $N \geq 2$ and $T_f \leq T_d < T_c$, which can be verified in a similar manner. In terms of $N \geq 2$ and $0 \leq T_d \leq T_f$,

$$T_x(t) = \begin{cases} T_c - T_d & \text{for} \quad 0 \leq t < T_f; \\ T_c + T_f - T_d - t & \text{for} \quad T_f \leq t < T_c. \end{cases} \tag{3}$$

Eq. (2) and (3) are for the situation that $N \geq 2$. However, there could be only one content on the signage display (i.e., $N = 1$), hence the transmission time is:

$$T_x(t) = T_f - T_d \text{ for } N = 1. \tag{4}$$

As assumed earlier that a user successfully joining the B/M group, or equivalently t, is uniformly distributed over one complete B/M cycle T_c, taking the expectation of T_x with respect to t will give an *expected transmission time* $\overline{T_x}$ experienced by the user as:

$$\overline{T_x} = \begin{cases} \left[(T_c - T_d)T_f + \frac{1}{2}(T_c - T_f)(T_c + T_f - 2T_d)\right]/T_c \\ \text{for } 0 \leq T_d < T_f \text{ and } N \geq 2; \\ \left[(T_c - T_d)T_f + \frac{1}{2}(T_c - T_d)^2\right]/T_c \\ \text{for } T_f \leq T_d < T_c \text{ and } N \geq 2; \\ T_f - T_d \text{ for } N = 1. \end{cases} \tag{5}$$

Hence, the average response time $\overline{T_r}$ is:

$$\overline{T_r} = T_g + \overline{T_x}. \tag{6}$$

Eq. (5) shows that $\overline{T_x}$ is a function of N, T_f, and T_d. Note that $T_f = S_f/B$, where S_f is the size of the "draggable" file and B is the bit rate of the channel. So the average response time of the system $\overline{T_r}$ is a function of T_g, N, S_f, B and T_d. As T_g is small, it is fixed to be $0.01s$ in the following numerical analysis.

Fig. 11(a) shows the plot of $\overline{T_r}$ against T_d. As larger T_d means a longer time for the smartphone to buffer the "draggable" files, $\overline{T_r}$ will decrease as T_d grows. $\overline{T_r}$ is plotted against T_d under the condition of $N = 1, 3$ and 5. Note that $\overline{T_r} = T_g$ after $T_d = T_c$, as all the "draggable" files are buffered after $T_d = T_c$, such that $\overline{T_x}$ will be $0s$ afterwards. T_f is fixed at $0.06s$ as the "draggable" file size S_f is set as 30KB (240 × 320 JPEG image with the resolution quality sufficient for the application) and the bit rate B is assumed to be 4Mbps. The plot shows that the decision time T_d required to achieve small $\overline{T_r}$ is also small. When a user discovers an interesting content and performs a "dragging" hand gesture, the user will experience almost no delay before the content is showed on the smartphone. Even for the worst case, the average response time, $\overline{T_r}$, will be less than 1 second.

Fig. 11(b) shows the plot of $\overline{T_r}$ against N. A range of values of T_d were used in this numerical analysis, which are $T_d = 0s$, $T_d = T_f/2$, $T_d = T_f$, $T_d = T_c/2$ and $T_d = T_c$. These are boundary cases and turning points as well as two representative points in Eq. (5), such that the boundary cases of $\overline{T_r}$ can be shown in the plot as well as some representative curves in between. T_f is fixed at $0.06s$. As observed from the plot, for each value of N from 1 to 10, the lower bound of $\overline{T_r}$ is T_g, and the upper bound is approximately proportional to N. In overall, this system could give an reasonable average response time, $\overline{T_r}$, for this range of N. Requiring a large number of contents in the system is not realistically needed in the real-world deployments for advertising signages. If $N = 5$, the implication is that the user can freely "drag" up to 5 different contents, while the system can still maintain $\overline{T_r}$ below 1 second.

Fig. 11(c) shows the plot of $\overline{T_r}$ against S_f. The same values of T_d used in Fig. 11(b) are used in this plot. N is fixed at 5 and B is fixed at 4Mbps. As observed from the plot, the lower bound of $\overline{T_r}$ is T_g, and the upper bound is linear with S_f. It is not ideal to use a large and non-constant file size, S_f, if $\overline{T_r}$ is needed to be a constantly small. On the other hand, it is not necessary to use a large S_f as the resolution of the smartphone display is limited. However, the size of the "draggable" file cannot be reduced too much, otherwise the user experience will be compromised. As a result, a 240 × 320 JPEG image (i.e., approximately 30KB in size) allows the user to "drag" a content with good quality, while still experiences a small response time below 1 second.

Fig. 11(d) shows the plot of $\overline{T_r}$ against B. The same values of T_d used in Fig. 11(b) are used in this plot. N is fixed at 5 and S_f is fixed at 30KB. It is observed from the plot that the lower bound of $\overline{T_r}$ is T_a and upper bound of $\overline{T_r}$ is inversely proportional to B. Although further increasing B will reduce $\overline{T_r}$, the gain will diminish as B becomes larger. Hence, $B = 4$Mbps is adequate to evaluate the real implementations for achieving relatively small $\overline{T_r}$. Considering a non-ideal channel with effective bit rate B, Fig. 11(d) can be interpreted as a larger error probability leading to a lower effective bit rate and thus a larger $\overline{T_r}$. As observed from the plot, $B = 4$Mbps is close to an effective bit rate of Wi-Fi in common situations, it is sufficient to achieve small $\overline{T_r}$ below 1 second in the real implementations using Wi-Fi.

IV. Implementation and Experimentation

Fig. 12 shows the implementation of Smart Signage system, which consists of a LCD TV as the signage display, a programmed VxWorks-based embedded media playback system as the signage device, and an IP-multicast enabled Wi-Fi wireless router. The wireless router with the Internet access is used by the signage device to multicast packets to

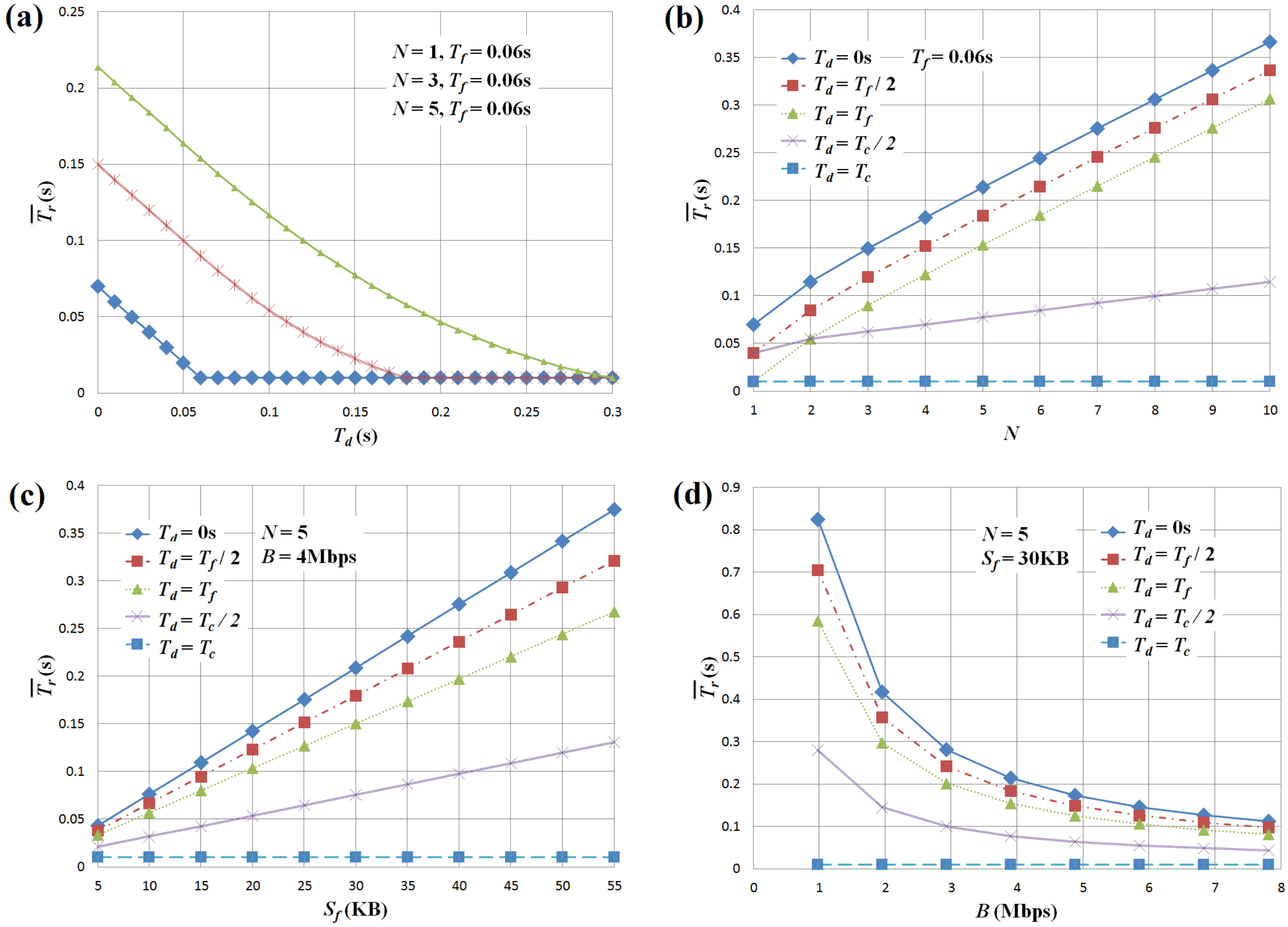

Figure 11. Numerical analysis of the average response time $\overline{T_r}$: (a) against T_d; (b) against N; (c) against S_f; (d) against B.

smartphones. The client-side software design is implemented on Android- and iOS-based smartphones. For the Android platform, the process of associating to an SSID and joining a multicast group is automated by skillful software implementation. The resulting association time T_a is within a few seconds, and this automatic process only requires once when the mobile application program is started. Comparing with the initial launches of other common mobile applications, users do not perceive such one-time duration of T_a as a response delay at all. Fig. 12(b) shows the Android-based smartphone to "drag" a content.

Fig. 13 shows the photographs of the 3 steps of a "dragging" gesture in the real implementation. A vertical levitation of the smartphone is observed during the process. Evoked by the "dragging" event, the corresponding visual responses to these 3 steps are implemented accordingly: Step 1: Before a user performs a "dragging" gesture, the display on the smartphone shows "Drag what you like"; Step 2: After a "dragging" event is detected, a successful "dragged" media is displayed on the smartphone; Step 3: Finally, the

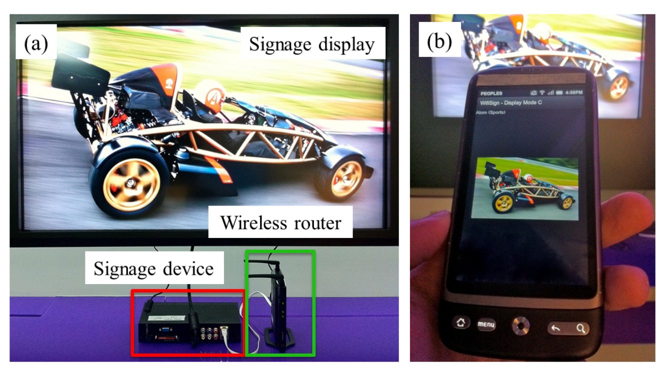

Figure 12. Implementation of the system.

user performs follow-up actions to store and use or even discard the "dragged" content.

In order to test the system implementation, a series of experiments are done. In one of representative experiments, 30 primary school students were invited to the laboratory to experiment with the proposed cyber-physical system.

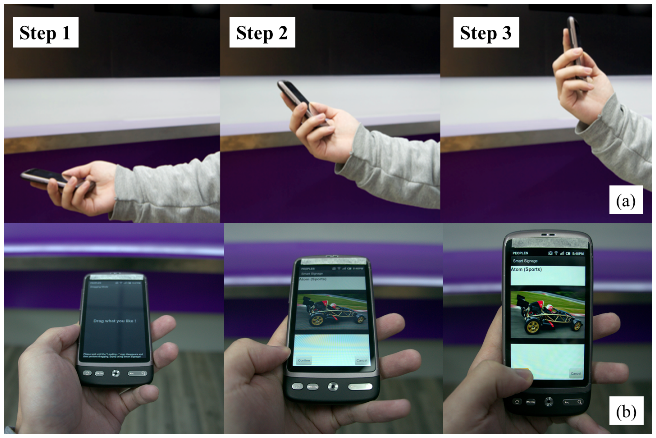

Figure 13. Implementation: a) the 3 steps of the "dragging" gesture; b) the corresponding visual responses.

Smartphones with different brands and OSs are installed with the proposed client-side mobile applications, which are distributed to these students and let them to "drag" the content on a display while they are walking around in the space of the laboratory. Without any prior knowledge about the interaction modality, these students easily got themselves familiar with the system after seeing a quick demo of "dragging" a content from the display into the smartphone showed by a researcher.

The students in this experiment reported neither any response delay nor any interaction glitch at all. Fig. 14 shows the students with the smartphones who successfully "dragged" the intended content into their smartphones simultaneously. Smart Signage system is therefore proven to provide an intuitive and scalable cyber-physical interactive modality to a large number of users simultaneously even they are walking around, while still maintain good user experiences (i.e., fast response time).

Figure 14. Smart Signage used by multiple primary school students.

V. Conclusion and Future Work

In this paper, Smart Signage - a "draggable" cyber-physical B/M media system is proposed. With a novel cyber-physical B/M protocol that enables a display concurrently interacting with physical actions of multiple user smartphones, a large number of users can simultaneously acquire any running content on a display by simply using an intuitive "dragging" hand gesture with their smartphones. Analytical formulations are derived to identify the key parameters and their dynamics in the system, which provide the condition for achieving the average response time of a user "dragging" gesture within 1 second limit. Implementations are demonstrated for possible real-world deployments. Experiments with 30 primary students are conducted to prove the system offering scalable and intuitive interactions for the first-time users even when they are moving.

Compared with existing interactive display systems, Smart Signage has the advantages of scalability to engage a large number of users, more flexible interaction range, more intuitive way of interaction, less visual compromise on the precious signage space, higher throughput of data transmission and supporting interactions with users on-the-move. Future research on cyber-physical interactions with multiple signages in a physical space is in process.

References

[1] WeightWatchers, "Weight watchers advertisement," 2009. [Online]. Available: http://gcms.weightwatchers.com

[2] T. Vajk, P. Coulton, W. Bamford, and R. Edwards, "Using a mobile phone as a "wii-like" controller for playing games on a large public display," *Int. J. Comput. Games Technol.*, vol. 2008, pp. 4:1–4:6, Jan. 2008.

[3] N. Davies, A. Friday, P. Newman, S. Rutlidge, and O. Storz, "Using bluetooth device names to support interaction in smart environments," in *Proc. of the 7th Int. Conf. on Mobile Systems, Applications, and Services*, ser. MobiSys '09. New York, NY, USA: ACM, 2009, pp. 151–164.

[4] K. Cheverst, A. Dix, D. Fitton, C. Kray, M. Rouncefield, C. Sas, G. Saslis-Lagoudakis, and J. G. Sheridan, "Exploring bluetooth based mobile phone interaction with the hermes photo display," in *Proc. of the 7th Int. Conf. on Human computer Interaction with Mobile Devices & Services*, ser. MobileHCI '05. New York, NY, USA: ACM, 2005, pp. 47–54.

[5] A. S. Shirazi, C. Winkler, and A. Schmidt, "Flashlight interaction: a study on mobile phone interaction techniques with large displays," in *Proc. of the 11th Int. Conf. on Human-Computer Interaction with Mobile Devices and Services*, ser. MobileHCI '09. New York, NY, USA: ACM, 2009, pp. 93:1–93:2.

[6] D. Schmidt, F. Chehimi, E. Rukzio, and H. Gellersen, "Phonetouch: a technique for direct phone interaction on surfaces," in *Proc. of the 23nd Annual ACM Symposium on User interface software and Technology*, ser. UIST '10. New York, NY, USA: ACM, 2010, pp. 13–16.

[7] R. Ballagas, M. Rohs, and J. G. Sheridan, “Sweep and point and shoot: phonecam-based interactions for large public displays,” in *CHI ’05 Extended Abstracts on Human factors in Computing Systems*, ser. CHI EA ’05. New York, NY, USA: ACM, 2005, pp. 1200–1203.

[8] N. Pears, D. Jackson, and P. Olivier, “Smart phone interaction with registered displays,” *Pervasive Computing, IEEE*, vol. 8, no. 2, pp. 14 –21, April-June 2009.

[9] Fox, “Fox quick response code advertisement,” 2 2008. [Online]. Available: http://www.amplify.com.au/blog/tag/coding/

[10] JCDecaux, “U-tie,” 2 2012. [Online]. Available: http://www.u-tie.com.hk/front/en/home

[11] K. Seewoonauth, E. Rukzio, R. Hardy, and P. Holleis, “Touch & connect and touch & select: interacting with a computer by touching it with a mobile phone,” in *Proc. of the 11th Int. Conf. on Human-Computer Interaction with Mobile Devices and Services*, ser. MobileHCI ’09. New York, NY, USA: ACM, 2009, pp. 36:1–36:9.

[12] R. Hardy, E. Rukzio, M. Wagner, and M. Paolucci, “Exploring expressive nfc-based mobile phone interaction with large dynamic displays,” in *Proc. of the 2009 First Int. Workshop on Near Field Communication*, ser. NFC ’09. Washington, DC, USA: IEEE Computer Society, 2009, pp. 36–41.

[13] R. Dachselt and R. Buchholz, “Natural throw and tilt interaction between mobile phones and distant displays,” in *Proc. of the 27th Int. Conf. Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA ’09. New York, NY, USA: ACM, 2009, pp. 3253–3258.

[14] J. Scheible, T. Ojala, and P. Coulton, “Mobitoss: a novel gesture based interface for creating and sharing mobile multimedia art on large public displays,” in *Proc. of the 16th ACM Int. Conf. on Multimedia*, ser. MM ’08. New York, NY, USA: ACM, 2008, pp. 957–960.

[15] RPA-Technology, “Mobile mouse for the iphone, ipod, & ipad,” 4 2011. [Online]. Available: http://www.mobilemouse.com/overview.html

[16] A. Waluyo, B. Srinivasan, and D. Taniar, “Optimal broadcast channel for data dissemination in mobile database environment,” in *Advanced Parallel Processing Technologies*, ser. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2003, vol. 2834, pp. 655–664.

[17] S. Mitra and T. Acharya, “Gesture recognition: A survey,” *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Trans. on*, vol. 37, no. 3, pp. 311 –324, may 2007.

[18] E. Lee, “Cyber physical systems: Design challenges,” in *Object Oriented Real-Time Distributed Computing (ISORC), 2008 11th IEEE Int. Symposium on*, may 2008, pp. 363 –369.