

Investigating the Use of Voice and Text Chat in a Social Television System

Joe Tullio¹, Gunnar Harboe¹, Noel Massey¹

¹ Motorola Labs
Schaumburg, IL, USA
{joe.tullio, gunnar.harboe, noel.massey}@motorola.com

Abstract. We tested a social television system with both voice and text communication in the field for one week, in four households. Participants expressed a preference for the freeform communication features over an earlier system with restricted communication options. Use of the communication features was nevertheless sporadic. Text was used more often than voice communication. However, both were usually used in combination. The in-home context had marked effects on the usage patterns, and we observed issues of privacy and identity verification.

Keywords: social television, computer-mediated communication, field study

1 Introduction

Social television systems promise to enhance the experience of watching television with the simultaneous benefit of socialization with remote parties. These parties could include friends, family, or even strangers with similar tastes in television programming. The communication facilities in current research systems vary, incorporating among other things voice chat capabilities, text chat, simple emoticons, lightweight, predetermined messages, or some combination thereof. Each of these facilities seems to have inherent benefits and drawbacks that may be dependent on factors such as the context of use, user preference, user demographics and relationships, and perhaps even the content being viewed. To this point, there are few studies in the literature that give any guidance as to the use of different communication modalities in a real-world setting for social television systems.

In this paper, we present results from a preliminary study of an experimental social television system called STV3 that incorporates voice and text chat facilities as well as lightweight graphical messages (emoticons). This study builds on earlier work examining a previous version of the system (STV2) that used lightweight messages only [4], and serves as preparation for an upcoming, larger study. We report on how a group of close friends and family used our system in the field for a period of one week. Based on interviews and use logs, we suggest that text chat may work better than voice in the real-world context, but that the combination of both is better still. In

addition, conversations seem to be only occasionally related to the TV content, and that identity verification remains a challenge.

2 Related Work

While there has been plenty of research on different electronic communication modalities, studies looking at social television specifically are scarce. Geerts compared text and voice communication in a within-subjects lab trial of two different systems [2], and found that voice was considered more natural, although young people preferred text. Baillie *et al.* compared voice communication to emoticon-based communication in a similar experimental setup, concluding that enjoyment and social presence was higher with voice [1]. Harboe *et al.* reported from group interviews reacting to a system concept and from field trials of an audio-only Social TV system that participants expressed a preference for voice over text [3].

Previous work, then, points to voice communication as the most promising modality for social television. However, there is no evidence from the field to support or contradict this finding. There may be important factors of the real-world context that do not arise in the lab or occur to interviewees discussing hypothetical scenarios.

3 STV3

STV3 is the result of several major enhancements to its previous incarnation, STV2. These enhancements address issues voiced by participants in our 2-week field study of STV2. In that study, participants felt that the limited communication options (initially just three emoticons, later also a selection of 20 predetermined text messages) were too restrictive, and wanted the ability to express themselves more freely [4]. STV3 contains both instant messaging-like text chat and multi-party voice call functionality, thus allowing users to communicate in their own words using the modality of their choice.

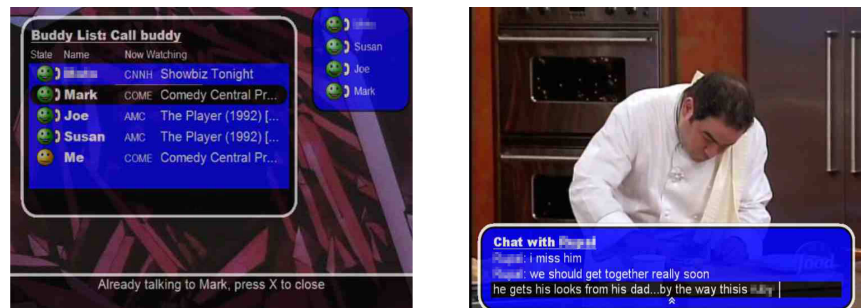


Fig. 1. Screenshots of buddy list and voice call presence (left), and a text chat window (right).

STV3 runs on a laptop PC, is displayed on a regular TV, and is controlled with a TV remote control. An on-screen buddy list allows users to contact friends and family, and also displays their presence, current TV show, and whether they are in a call (Figure 1). The prototype also includes a color-changing light to signal when a call is ongoing, and to indicate others' presence when the TV is off.

Voice calls are supported through the use of an echo-canceling room microphone, while text chats are supported using a wireless keyboard. When a call is attempted, the recipient can choose whether or not to answer. If the call is accepted, the two parties are connected and can begin talking to one another. Any participant in a call can invite others to join. Anyone can leave at any time, and a call ends when there is only one participant left.

For text chats, text messages appear immediately on the recipient's television in a small window near the bottom of the screen. At this point, either participant can type messages using the keyboard and send them. Each text conversation only supports two parties, users can have multiple simultaneous text chats with different buddies.

3 Method

For this preliminary study, we recruited four participants external to Motorola, all of whom lived in different households (A1, A2, A5 and A6). Three of the participants (A1, A2 and A5) took part in the earlier study of STV2. This decision was intended to minimize the novelty effects of using the system, since most participants were already familiar with its basic operation, and focus on the newly added communication features instead. The participants were between the ages of 25 and 35. A2, A5 and A6 were sisters, while A1 was a best friend of A2. All except A5 were married, and all watched TV frequently and regularly discussed certain TV shows.

The system was deployed in their households for a period of one week in mid-December of 2007. An initial interview was used to gather background data, a phone interview was conducted mid-week to investigate participants' experiences with the system during the first half of the week, and a final interview was conducted at the end of the week to examine their experiences in-depth. All interactions with the system, including text conversations, were logged, and voice conversations were recorded. Interview clips were transcribed and grouped along general themes to arrive at common insights regarding the experiences of our participants as a group.

4 Results

During the study period, participants had 4 voice calls and 12 text chats. Three attempts to make calls went unanswered, and 9 text messages (or series of messages) received no response. The first call (A1-A2 on the night of the setup) was a simple 1-minute system test which led into a joking text chat. This was followed by a 40-minute long call later that evening, with all participants joining in. 4 of the text chats took place within or around this conversation. The last two calls, A1-A2 (10 minutes) and A2-A4 (1 minute), took place on the fifth night, and although the calls were

separate, all the participants were communicating with each other in 3 simultaneous text chats.

The three participants from the STV2 study all agreed that STV3 was a significant improvement from the previous version, describing it as “a good concept”, “more interactive”, and “a lot more fun.” While there were some complaints that the system, had too many wires and too many peripheral devices, participants understood that the system was still only a research prototype and not yet a commercial product.

4.1 Feature use and choice of modality

The amount of communication through the system was lower than we had hoped. After the first day, voice calls were used on only one occasion, and text chats on four separate occasions. This may be due to busy holiday schedules, lack of interesting programming that week, or it may reflect the natural frequency of use of the system.

While previous research has presented text and voice as contrasting alternatives, we found that participants combined the two. All of the calls in our study took place just before, during or after a text chat between the same participants. Indeed, most of the text chats were associated with voice calls. Participants generally stated that they preferred text chat to voice calls. Their reasons for this choice varied. One participant did it out of concern for the other households in the study, stating “there’s no sound or anything, so it’s not disruptive,” especially at night after children’s bed time. Another participant cited its familiarity: “I guess that’s a technology I’m used to, because I text-message a lot, and that’s all I do at work.”

This is not to say that they had no use for voice calls. One participant said “I would prefer the voice if it was a more lengthy or detailed conversation... it’s a little bit more time consuming to sit and type it.” Another cited her sister’s personal preference: “she would much rather talk and get the personal interaction.”

4.2 Communication and topics of conversation

All of our participants communicated daily prior to the study, usually by phone during their afternoon commutes and in some cases via email during the day. Evening, as reported by one participants, was “our time with our significant others that we live with”. The addition of STV3, however, encouraged additional communication outside of their daily routines. One participant said “in the evening it’s a bonus chat...just to say ‘hi, what are you doing?’” Husbands who had not previously been part of the conversation could join in, as one stated jokingly: “Usually I wouldn’t know what they talk about. Here I get more involved, probably more than I wanted to.”

Conversation topics mostly did not relate to the programming being viewed during those conversations. One participant described them as “almost like a side conversation.” Topics were typically about “things in general” and “more casual” than conversations conducted over the phone earlier in the day. However, there were two text chats and one voice call about the soap opera *General Hospital*.

4.3 Transparency and privacy

Participants recounted a number of experiences concerning privacy, from lighthearted jokes to more disruptive misunderstandings. Twice, the husbands of participants attempted to pose as their wives in text chats, though they were quickly discovered because as one participant put it, “it was the way he said ‘hey dude’, and my sister doesn’t talk to me that way.” In a more disruptive case, one participant accidentally mentioned a surprise Christmas present in a text chat, not realizing that the intended gift recipient was in the living room watching the television, causing awkwardness for both the participants involved.

Not knowing exactly who was on the other side of the conversation was a common issue for all participants. One participant likened the chat feature to her IM client at home: “at first I wasn’t conscientious [*sic*] of the fact that...I thought it was just a private conversation. In my mindset I’m comparing it to my Sametime chat I have at work which to me is private.” Another participant was wary of using the voice feature, stating “you feel like it’s kind of an invasion of your privacy, because they can be listening to you the whole time.”

5 Conclusions and Future Work

While our results are only preliminary, they highlight issues that designers of social television systems should be aware of, and that warrant further attention from researchers. Most of these issues have received little if any consideration in previous work, perhaps because observations such as the reluctance to intrude with a voice call, and other household members impersonating the main participant are not likely to be made in a lab. In fact, due in part to such factors, our findings seem to contradict previous studies, by favoring text chat over voice calls.

With our method validated by this small study, we are currently conducting a larger trial where we hope to gather more substantial data. Our prototype has already benefited from a design iteration based on the input collected in this round. We encourage others to test our findings independently to see how well they generalize to other groups and contexts.

The authors would like to thank Frank Bentley, Elaine Huang, Larry Marturano, Crysta Metcalf, and Guy Romano for their contributions to this work.

References

1. Baillie, L., Fröhlic, P., Schatz, R.: Exploring Social TV. In: ITI 2007 29th Int. Conf. on Information Technology Interfaces, pp. 215–220. Cavtat, Croatia (2007)
2. Geerts, D.: Comparing Voice Chat and Text Chat in a Communication Tool for Interactive Television. In: NordiCHI 2006, pp. 461–464. ACM Press, Oslo, Norway (2006)
3. Harboe, G., Massey, N., Metcalf, C.J., Wheatley, D., Romano, G.: The Uses of Social Television. In: Computers in Entertainment, vol. 6 (1). ACM Press (2008)

4. Harboe, G., Metcalf, C., Bentley, F., Tullio, J., Massey, N., and Romano, G. (2008)
Ambient Social TV: Drawing People into a Shared Experience. In: Proc. CHI 2008, pp.1-10.