ExciTube
Video Player for Sharing Viewer's Excitement

Takumi Shirokura, Nagisa Munekata and Tetsuo Ono
Graduate School of Information Science and Technology, Hokkaido University N14-W9, Kita-ku, Sapporo, Japan

Keywords: Video Player, Virtual Reality, Physiological Computing, Entertainment Computing, User Experience.

Abstract: We can share non-verbal emotional experiences, such as excitement and pleasure, by watching movies and sports events with others, like our friends and family. These shared experiences are thought to enhance excitement and pleasure compared to when watching videos alone. Our research provides this shared experience on the internet by sharing the viewer's excitement with others while watching videos that are on the web. We studied the relationship between users' excitement while watching videos on the web and their impressions of those videos. Here, we introduce a video player called ExciTube that allows users to share their excitement and view other users' excitement as visual information alongside the video they are watching. The user's excitement is expressed and shared by using avatars. We carried out user-involved demonstrations of ExciTube at our laboratory and at a Japanese domestic Computer Entertainment Developers Conference, and confirmed that people did enjoy using the system and felt other people's sense of excitement.

1 INTRODUCTION

We have all watched movies, sports events, and videos with family or friends. This experience is stronger than when watching videos alone because we can share our excitement and impressions. We can engage with those we are watching videos with through sharing emotions. However, it is difficult for people who live in remote places to do this because viewers need to gather at one location.

However, we can easily communicate on the subject of videos with friends through comments and positive/negative icons on online video sites, such as YouTube (http://www.youtube.com/), Vimeo (https://vimeo.com/) and NicoNico (http://www.niconico.com/). These services only have conscious information, such as text and icons, to represent users' impressions. However, there are many reactions such as unconscious non-verbal information caused by watching videos together with friends that is spread among viewers in the real world. Comments and positive/negative icons are important content on most online video websites that represent users' evaluations of videos. These comments, though, may contain untrue information because viewers leave them consciously, or they are left multiple times by one viewer. We feel and share

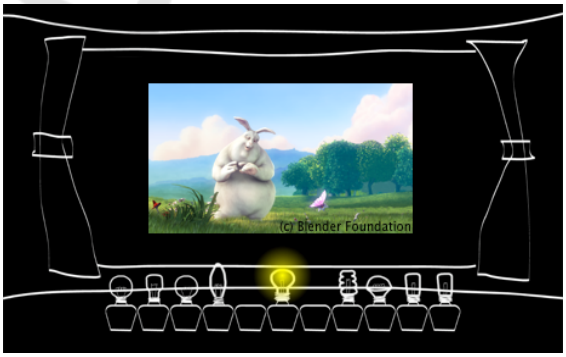


Figure 1: Overview of ExciTube. Center avatar represents user currently watching video. Avatars on both sides represent other users who have watched video that user selected. These avatars reflected each user's SCR.

evaluations of videos in the real world not only through each other's verbal comments, but also through each other's unconscious non-verbal information, such as gestures, facial expressions and eye movements. Viewers' truthful reactions are particularly contained in much unconscious non-verbal information. However, the previously mentioned webvideo services do not support the representation of unconscious non-verbal information. We believe that unconscious non-verbal information, and not only comments and

Shirokura T., Munekata N. and Ono T..
ExciTube - Video Player for Sharing Viewer's Excitement.
DOI: 10.5220/0004726503150322
In *Proceedings of the International Conference on Physiological Computing Systems* (PhyCS-2014), pages 315-322
ISBN: 978-989-758-006-2

positive/negative icons, are necessary for webvideos to enhance the watching experience. Viewers' excitement in reactions to videos, which is contained in unconscious non-verbal information, is important because people who live in different places can experience a similar sense of excitement with them by having that information provided to other viewers.

We carried out an experiment to achieve this goal on the relationship between a webvideo of short duration (3 minutes) on Youtube, and the emotional excitement of a user when watching it. Along with this experiment, we also developed a video player called ExciTube (Figure 1) that allowed users to share their excitement with that of others as visual information alongside the video they were watching. The excitement of video viewers was measured using their skin conductance response (SCR). This system provided the experience as though it occurred in the real world where we can notice the excitement of other video viewers, which in this case was then expressed as non-verbal information through SCR. We conducted user-involved demonstrations of ExciTube at our laboratory and at a Japanese domestic conference, and confirmed that people did enjoy using the system and feeling other people's sense of excitement.

The remainder of the paper is organised as follows. Section 2 describes advanced research on social/interactive television and the sharing of physiological signals. Section 3 describes an experiment on the relationship between webvideos and emotional excitement of viewers and presents the results. Section 4 describes the ExciTube system, highlighting distinctive features that appreciably affect usage. Section 5 describes user-involved demonstrations of ExciTube. Finally, Section 6 presents a discussion and the conclusion.

2 RELATED WORK

2.1 Sharing Viewing Experiences

Many researchers have proposed systems that provide experiences like those of watching videos together with family or friends even though they live in different locations (Regan et al., 2004) (Shamma et al., 2008). These systems are divided into two types that can be used synchronously or nonsynchronously with other people.

Coppens developed a system that had to be used synchronously with other people, called AmigoTV (Coppens et al., 2004) and they focused on broadcast television. This system consisted of a video screen and personal avatars, which represented the video viewers' faces. Users could express their emotions through the avatars, and use voice chats to communicate directly in relation to the television programme. Users felt closer to others and connected with them through voice chats and facial expressions while sharing their time together. This system led to other similar systems such as Social TV (Harboe et al., 2008) and Abreu's 2BeOn (Abreu et al., 2002). Both systems had voice chat functions while viewers were watching videos. Other systems not only had voice chats but also text chats and instant messaging. Geerts (Geerts, 2006) and Baillie (Baillie et al., 2007) explored how we communicate with other people who live in different locations while watching videos. Even though these systems were used in different locations, users had to share them at the same time. In our research, we focused on a nonsynchronous system because users may not be able to share videos at the same time.

Harrison developed CollaboraTV (Harrison, 2007) that provided nonsynchronous communication to users. It consisted of a video screen and avatars, which represented users who had watched a video and left comments and positive/negative evaluations. The user's comments and positive/negative evaluations when watching the video were shared through the avatars with others who watched the same video. Viewers who used this system did need not worry about time when watching videos because the shared comments and evaluations were synchronized with the video timeline. Users of NicoNico, which is a webvideo service, can leave comments on a video screen. The comments flow over the video screen from right to left. Users nonsynchronously communicate with others through the comments. We did not use comments and positive/negative expressions, but video viewers' sense of excitement, which is unconscious non-verbal information, to enhance the webvideo experience.

Most research has focused on how to place the expressions of other people on the video screen to communicate with others, but Harboe's research presented an ambient display like a color-changing lamp, which expressed someone's presence. This system displayed the current number of people watching television through colours.

We aimed at creating viewing experiences, like those in the real world, on the web in this study, and we propose using unconscious non-verbal information in addition to communications that have been used in previous research to enhance webvideo

experiences. Therefore, we employed the concept of avatars displayed on the video screen because they have most frequently been employed by many researchers.

2.2 Sharing Physiological Signals

There has been much advanced research on sharing physiological signals that include the emotional excitement of people (Fairclough, 2009) (Janssen et al., 2010) (Mandryk et al., 2006). Slovak explored sharing heart rates when in the laboratory and at home (Slovak et al. 2012). Shared heart rates were displayed on a screen or converted into background music that was played in a room. They found through the study that participants used the shared heart rates to communicate with others and they then felt strong connectedness to others by understanding their condition through the heart rates. Werner proposed a United-pulse (Werner et al., 2008), which is a ring type device that could measure pulse waves and vibrate, to share heartbeats. Lotan developed an Impulse system (Lotan et al., 2007) that measured pulse waves and vibrated, and allowed users to wirelessly transmit their heartbeat rhythms to an audio speaker and a lamp. Although these studies used pulse waves, it was difficult to detect timing when users became excited because the physiological signals changed slowly. Consequently, we used SCR, which is also known as the galvanic skin response. SCR changed quickly and could be used to capture the moment the emotional response to a video occurred.

3 WEBVIDEO WITH EMOTINOAL EXCITEMENT

We explored the relation between webvideos and the emotional excitement of users while they were watching them. We compared emotional excitement, which was measured from SCR, and the user's tastes/preferences in an experiment.

3.1 Measuring Emotional Excitement

We used electrical signals detected on the human body to measure the emotional excitement of viewers. These signals were objective, and the quantitative data reflected psychological states and physiological functions. Such signals have been used for diagnosis and treatment in medical care and in lie detectors used during police interrogations. One of the psychological signals that a lie detector recognizes is SCR, which occurs when mental states, such as agitation, surprise, or excitement, induce changes in conductance on the skin's surface. We have little awareness of the physiological functioning of our own bodies because most physiological functions are involuntary and therefore uncontrollable. SCR is a typical example. No one is aware of the minute amounts of perspiration during mental agitation unless there is an unusually large amount of mental stress. Therefore, observing one's own SCR produces a strange feeling that this is not a feature of one's own body but rather that of another person. People generally believe that inner agitation or excitement during communication in daily life can be concealed. However, SCR can reveal concealed agitation despite a person's best intentions to conceal it.

Figure 2 has a photograph of an SCR sensor and shows a typical SCR signal trace. We developed the SCR sensor in our previous work (Munekata et al., 2006) using an improved circuit recommended by Fowles (Fowles et al., 1981) that could measure the magnitude of a user's SCR. The SCR magnitude was converted into an SCR value from 0 to 255. Note that the SCR response is an integrated SCR value in this paper.

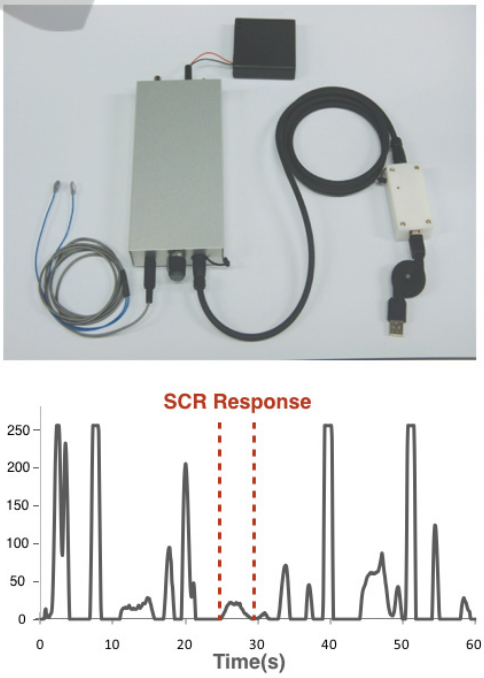


Figure 2: SCR sensor and typical SCR signal trace.

3.2 Experiment

We conducted an experiment to examine the relation between users' tastes in relation to a webvideo and their emotional excitement while watching it. The emotional excitement was evaluated by using SCR values.

We recruited 10 participants (A-J) from Hokkaido University. Their ages ranged from 21 to 43 with an average age of 25.3 (σ =33.0). Each participant watched three out of five official movie trailers on YouTube chosen by each of them, and their SCR values were measured for the duration of each trailer. After that, they ranked the three videos according to their preferences (1st, 2nd, and 3rd). All the trailers were 2 minutes long. The genres for the trailers were two fantasies, two action trailers, and one drama. All participants said they watched TV and webvideos daily.

3.3 Results

Table 1 summarizes averaged SCR responses while participants were watching the 1st and 3rd videos. We compared participants' SCRs and the participants' ranking through paired t-tests. As a result of this experiment, we found that the average SCR response while watching the 1st-ranked video was larger than that while watching the 3rd-ranked video (p=0.04607*). Therefore, viewers' preferences for the videos had a strong relationship with SCR. Therefore, we think SCR reflected viewer's preferences for videos as unconscious non-verbal information.

Table 1: Average of excitement reaction value while watching videos that are rated 1st and 3rd by participants. (*p < 0.05).

Participants	1st	3rd
A	398	330
B	190	181
C	874	380
D	276	199
E	90	41
F	327	312
G	394	279
H	318	282
I	255	176
J	146	55
Paired T-Test		
t	2.3122	
df	9	
p	0.04607*	

However, participants B and F did not have a strong relationship between their favourite rankings of videos and the average SCR response. There could be a couple of reasons for this. For example, if both videos had many sound stimuli, they could make the users just as excited. Again, the participants may not have had a strong interest in any of the videos. Consequently, these cases revealed that these users' SCR responses differed from those of other users. Note here that users choose favourite videos when using webvideo services.

We developed a video player from this experiment that enabled users to nonsynchronously share emotional excitement (SCR) through the internet while watching webvideos with others who lived in different locations.

4 VIDEO PLAYER

We developed a video player called ExciTube (Figure 1) that enabled a viewer's SCR to be shared with others. Users could utilize ExciTube by just attaching a SCR sensor to their own hand. This system automatically shared SCRs while participants were watching a video. The shared SCRs were represented by three avatar designs. A user could feel others just by watching a webvideo because other users' SCRs were visualized on the video screen as avatars. This system could be used for watching videos that had been uploaded onto YouTube, because the system used YouTube's API.

4.1 Visualization of Emotional Arousal

Shared SCRs were represented by avatars in this system because avatars are suitable for representing people on computers, and users can easily sense other users (Harrison et al., 2010).

ExciTube had two kinds of avatars (Figure 3). The user avatar was displayed by default, but users could select the visibility of user avatar (true/false).

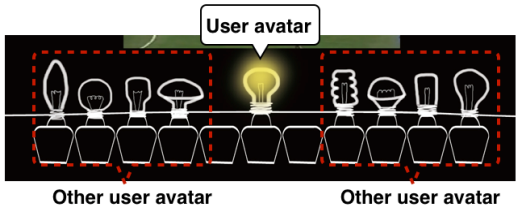
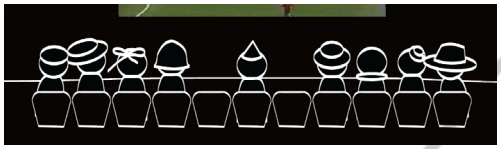
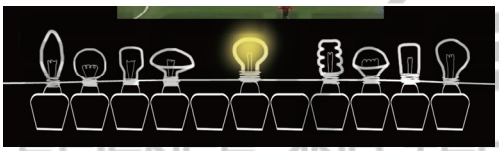


Figure 3: User avatar, and other user avatars, which represent other users who watched video that user selected.

Other user avatars reflected the shared SCRs that were measured while other users were watching

Table 2: Avatar design and behaviour adaptation. Top: Person avatar, middle: cartoon avatar, bottom: bulb avatar.

Avatar designs	Default	Behavior	SCR magnitude	Response duration
			Vibration	None
			Change size	Change time
			Brightness	Lighting Time

videos that the user had selected. The system provided three designs for the avatars - a person design, a cartoon design, and a bulb design. Because avatar designs and actions have a huge effect on users, every avatar design had unique actions that reflected emotional excitement. This system displayed eight other user avatars at maximum because it became difficult to watch the videos and all the other user avatars when there were too many of them. If this system had more than eight other user avatars, they would be shown to be randomly selected within this system.

4.1.1 Person Avatar

The person avatar imitated a person's silhouette to provide an experience like that at a movie theatre (Table 2: Top) to the user. The person avatar had a vibrating behaviour. When the SCR value exceeded a certain threshold, the person avatar vibrated once.

4.1.2 Cartoon Avatar

The cartoon avatar was inspired by cartoon animation. This avatar could change size just like cartoon characters. If there was an SCR response occurring, the avatar changed according to the SCR value.

4.1.3 Bulb Avatar

The bulb avatar represented the SCR value as brightness. Therefore, this avatar was the most intuitive of the three avatar designs because many people express brightness as Hi/Low. It was easy for users to understand their own emotional excitement and that of others through this avatar.

4.2 Interface

Figure 4 shows the ExciTube interface. Avatars with the design the user selected are at the bottom of the application window. The video screen is in the middle of the window. When users move the mouse cursor to the top of the window, a search tool bar

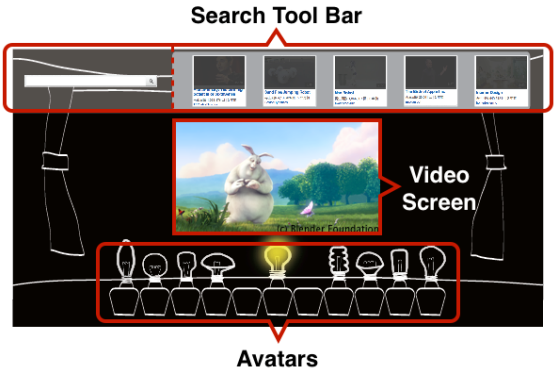


Figure 4: Screenshot of Interface.

appears. Users could search videos that had been uploaded onto YouTube, and they then watched the videos. The search results had the videos' thumbnails, durations, titles and numbers of shared SCRs.

4.3 System Architecture

ExciTube consisted of a client application, server application, and an SCR sensor. The client application was built using Adobe AIR and Java. This application had three functions; the first was to save the SCRs of users made available by the SCR sensor, the second was to display avatars, and the third was to play YouTube videos through the YouTube API. The server application was built using the Ruby on Rails framework. The web server used WEBrick and MySQL. The information from the videos and SCRs of users was recorded in MySQL.

5 DEMONSTRATION

We demonstrated ExciTube at our laboratory and at a Japanese domestic Computer Entertainment Developers Conference (CEDEC) as a pilot study. Users could watch three videos in these demonstrations because ExciTube contained shared SCR data. We prepared the SCRs of four test users for each video to demonstrate ExciTube. Therefore, ExciTube had a user avatar and four other user avatars from the beginning.

This section introduces comments made by participants and their observations.

5.1 Laboratory Demonstration

The user avatar was disabled in this demonstration to observe the effect of other users' avatars. The participants were 10 university students (10 men: 20-27 years old). All participants used webvideo services such as Youtube, Vimeo and NicoNico on a daily basis. We presented an introduction to this system before the demonstration.

The positive comments provided were "*As the behaviours of the avatars were synchronised with the sound effects, the video and avatars were interesting.*", "*The moment when all the avatars did an action at the same time was interesting.*" and "*When I was surprised, the avatar also did an action.*" Therefore, participants felt the presence of others through the avatars. ExciTube can help participants get excited even when they are alone similarly to when they watch videos with others. Additionally, ExciTube can provide new experiences through avatar actions as observed from comments, such as "*I looked for interesting points in the video when the avatar acted*". The participants also made negative comments, such as "*I was curious about the avatar and was not able to concentrate on the video.*" and "*A small avatar did not have a large enough presence about it. I was only viewing the video without the avatars.*" As this system was used to convey a similar sense of excitement to when they watched videos with others, we thought these negative comments by participants were not problematic.

This demonstration revealed that ExciTube could help participants become excited even when they were alone similarly to when they watched videos with others, and also that it could create a new experience through avatar actions that made watching the videos more interesting.

5.2 Conference Demonstration

We demonstrated ExciTube at CEDEC2011. The audience consisted not only of game developers and researchers but also of non-human-computer interaction attendees. Most of the audience did not have knowledge of physiological computing. More than 50 people used our system and more than 100 people watched the demonstration, where the user avatar and other user avatars were displayed.

Most participants observed their own avatar (user avatar) and then they understood the behaviour of the other avatars through their own experience of this system. After this initial observation, they enjoyed watching the videos and other users' avatars. Some participants had particular interest in the other users' avatars' reaction points because their own avatar's reaction point was different to theirs. Additionally, participants left comments about this system, such as "*It was intuitive and very comprehensible that excitement was expressed by the brightness of the bulb.*", "*I want to use this technology to evaluate a system we have developed.*" and "*I was able to observe other users' sense of excitement*". We could easily notice users' excitement points in the videos with this system. Therefore, we found this system has the possibility of being applied to games and their evaluation.

We observed that participants did enjoy ExciTube through the two demonstrations, and we found that ExciTube had the potential for providing new experiences to video viewers.

6 DISCUSSION AND CONCLUSIONS

Because this system used the sense of excitement of video viewers, the avatars did not demonstrate behaviours when users did not feel excited while watching videos. Therefore, sharing emotional excitement made no sense in this case. We do not become excited in the real world when we watch boring videos with friends or family. Therefore, this case was not a problem because our goal was to offer experiences like those in the real world. ExciTube was useful for users who wanted to watch videos with others.

This system had a problem in that users may get excited by stimulation that is not related to the videos. This was confirmed in the demonstration. However, it is thought that a user watching a webvideo alone receives almost no outside stimulation. Additionally, we thought that outside stimulation would not excite users because they were concentrating on watching a video they chose themselves.

We provided experiences on the web in this study as those found in the real world such as when watching movies with friends or family members. We conducted an experiment to achieve this by using the emotional excitement of users when they were watching webvideos. As a result of the experiment, a correlation was found in the relation between emotional excitement and webvideos. We developed ExciTube that could feel the presence of others and provide evidence that a user could feel the emotional excitement of others from SCR values. This system used avatars that represented the user and other users, which were displayed on a video screen. People felt the presence of others through the avatars. The avatars' behaviours reflected the emotional excitement of users and this was shared when they watched the videos on their own. The users' emotional excitement was measured from their SCRs. We demonstrated ExciTube at our laboratory and at CEDEC to investigate its usefulness. The demonstration revealed the positive effects of participants sharing their emotional excitement in webvideo experiences.

REFERENCES

Abreu, J., Almeida, P., & Branco, V., 2002. *2BeOn: interactive television supporting interpersonal communication*, pp. 199–208.

Baillie, L., Frohlich, P., & Schatz, R., 2007. Exploring Social TV. In *2007 29th International Conference on Information Technology Interfaces*. IEEE, pp. 215–220.

Chagas, A. B. & Ferraz, C. A. G., 2012. ConnecTV. In *Proceedings of the 18th Brazilian Symposium on Multimedia and the Web - WebMedia '12*. New York, USA: ACM Press, pp. 83–90.

Coppens, T., Trappeniers, L., & Godon, M., 2004. Amigotv: towards a social TV experience. In *Proceedings from the Second European Conference on Interactive Television*.

Ducheneaut, N. et al., 2008. Social TV: Designing for Distributed, Sociable Television Viewing. *International Journal of Human-Computer Interaction*, 24(2), pp. 136–154.

Fairclough, S., 2009, Fundamentals of physiological computing. *Interacting with Computers*, 21(1), pp. 133–145.

Fowles, D.C. et al., 1981. Publication Recommendations for Electrodermal Measurements. *Psychophysiology*, 18(3), pp. 232–239.

Geerts, D., 2006. Comparing voice chat and text chat in a communication tool for interactive television. In *Proceedings of the third Nordic conference on human-computer interaction - NordiCHI '04*. New York, USA: ACM Press, pp. 461–464.

Harboe, G. et al., 2008. Ambient social TV. In *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*. New York, USA: ACM Press, pp. 1–10.

Harrison, C. and Amento, B., 2007. Collaboratv: Using asynchronous communication to make TV social again. *Adjunct Proceedings of EuroITV*, pp. 218–222.

Janssen, J. H. et al., 2010. Intimate Heartbeats: Opportunities for Affective Communication Technology. *IEEE Transactions on Affective Computing*, 1(2), pp. 72–80.

Lotan, G. and Croft, C., 2007. Impulse; In *CHI '07 extended abstracts on human factors in computing systems*, pp. 1983–1988.

Luyten, K. et al., 2006. Telebuddies. In *CHI '06 extended abstracts on human factors in computing systems - CHI EA '06*. New York, USA: ACM Press, pp. 1049–1054.

Mandryk, R. L., Inkpen, K. M., & Calvert, T. W., 2006. Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour & Information Technology*, 25(2), pp. 141–158.

Munekata, N. et al., 2006. Design of positive biofeedback using a robot's behaviors as motion media. *In Proceedings of the 5th international conference on Entertainment Computing -ICEC'06*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 340–349.

Regan, T. & Todd, I., 2004. Media center buddies. In *Proceedings of the third Nordic conference on human-computer interaction - NordiCHI '04*. New York, USA: ACM Press, pp. 141–144.

Shamma, D. A. et al., 2008. Enhancing online personal connections through the synchronized sharing of

online video. In *CHI '08 extended abstracts on human factors in computing systems - CHI '08*. New York, USA: ACM Press, pp. 2931–2936.

Slovák, P., Janssen, J., & Fitzpatrick, G., 2012. Understanding heart rate sharing. In *Proceedings of the 2012 ACM annual conference on human factors in computing systems - CHI '12*. New York, USA: ACM Press, pp. 859–868.

Werner, J., Wettach, R., & Hornecker, E., 2008. United-pulse. In Proceedings of the 10th international conference on human computer interaction with mobile devices and services - MobileHCI '08. New York, USA: ACM Press, pp. 535–538.