Master 2 in Computer Science - Interaction Specialty

Situated Breakdowns in Video Mediated Communication

Author:
Yi Zhang

Supervisor:
Wendy Mackay

Hosting lab/enterprise:
Inria ExSitu Research Group

March, 2018 – August, 2018

Secrétariat - tel: 01 69 15 66 36 Fax: 01 69 15 42 72
email: Murielle.Benard@u-psud.fr

Contents

Summary

In an ideal world, VMC systems would work perfectly, with fully synced audio and video signals on both sides. In real life, however, user do not necessarily have a perfect network signal, hardware, or physical environment, which might lead to breakdowns of the communication. In this thesis, we asked the research question: what is the user experience of video-mediated communication **under real-world conditions**.

To answer the question, we conducted two iterations of user studies. In study one, four members of the research team explicitly recorded data about every VMC call they conducted over a period of one month. The results suggest that VMC systems often result in technical breakdowns that disrupt the user's conversations, leading to unsatisfying quality of experience (QoE). Besides, a pattern of "breakdown-confusions-workarounds" is revealed. We proposed a seven-dimension breakdown space to describe a breakdown from multiple perspectives, introduced two types of confusions that hinder users from solving breakdowns, and categorized users' workarounds into three steps.

We then conducted a diary study with external VMC users to test the generality of the patterns we found in the 1st iteration. The results support the previous findings.

In the end, we proposed design suggestions to the VMC techniques to mitigate poor service and better support for users' natural approach to workarounds. Moreover, we put forward a method to calculate the cost of poor QoS. These two will be the main future work for this research.

Keywords

Video-Mediated Communication, VMC, Breakdown, Workaround, User experience

CHAPTER 1

Introduction

Video-Mediated Communication (VMC) has become increasingly accessible and popular with the rapid development of the information and communication technologies (ICT). In 2017, Facebook Messenger alone had logged a total of 17 billion video calls [26], and WhatsApp users are reported to make 55 million video calls per day along with a total of over 340 million video calling minutes per day. Video-Mediated Communication has permeated every aspect of our life. It has gained widespread acceptance by the general public to support their lives from professional work tasks to informal everyday communication. [10]

While VMC technologies are broadly adopted by the general public to facilitate their work and daily life, is the quality of experience (QoE) in VMC always satisfying? In an ideal world, VMC systems would work perfectly, with fully synced audio and video signals on both sides. However, in real life, it is possible that users might not have good settings of network and devices, etc., which might influence their user experience in VMCs. We ask the following specific research questions:

1. Under real-world conditions, how often do VMC systems result in technical breakdowns that disrupt the user's conversations?

2. How helpful are the system's error messages in helping user's identify what is wrong and how to fix it?

3. What kinds of breakdowns, from the user's perspective, actually occur?

4. What kinds of workarounds do users try, and how successful are they?

In order to address these questions, we conducted two studies. Study one used a diary approach in which four members of the research team explicitly recorded data about every VMC call they conducted over a period of one month. For each call, we recorded a pre-determined set of factual data, including number of participants, number of sites, relationship of communication participants, purpose, topic, the one who initiate the call, date, time, duration, context of the participants, city, network settings, primary hardware, extra hardware, primary software, extra software, media involved, as well as an overall assessment of the level of disruption, if any.

We also recorded a qualitative description of any audio, video or other technical problems that arose, and whether and how they were fixed. We analyzed both the quantitative and qualitative data to identify 13 breakdown symptoms and how users reacted to them.

In order to test the generality of these breakdowns and workarounds, we conducted a second study, using the same diary-study protocol, with ten external participants.

We required the user to fill in a self-report sheet each time they had a video call, to collect 1) raw data of the call, e.g., nature of the call, settings, etc.; 2) qualitative description of the VMC story; 3) report on disruptions, confusions, and workarounds.

This thesis first reviews the related work, then describes the first study and the resulting framework describing breakdown symptoms, breakdown space, confusion types, and workaround category. I then describe the results of study two and conclude with directions for future research.

CHAPTER 2

Related Work

A goal of this thesis is to study technology breakdowns that cause poor user experience. To achieve the goal, a comprehensive understanding of technology breakdowns is essential. In the first section, we outlined the three main technical factors that can lead to poor VMC quality. These factors could help us better understand the source of a breakdown. Also, we summarized the current user studies in VMC regarding its topic and methods. To our knowledge, little research has addressed the influence of technology breakdowns on the user experience of VMC under real-world conditions.

2.1 Technical factors of VMC quality and stability

There is a rich body of work from computer scientists and telecommunication researchers that studied the quality of video calls and its influencing factors from the technical perspective. Factors that affect the perceptual quality of video and audio streams can be seen in three categories: network factors, equipment factors, and physical environment factors.

Network factors

Video-mediated communication relies heavily on the network traffic; therefore the impact of the network on the VMC quality is highly noticeable. Packet loss, latency, and jitter are the most common network factors that reduce the quality of audio and video. If the packets carrying the audio or video information are lost, late, or uneven in the network transmission, the problem of packet loss, delay and jitter shows up. They may cause frozen screens, delays, artifacts in the image, stuck voice, lagged audio or the worst situation- call outage [1, 5, 23, 8], which severely affects the quality and intelligibility of the call. On the public Internet, especially in the case of long-distance communications, the effect of these factors is even more noticeable.

Equipment factors

Equipment factors include device's video and audio codec technology, device's acoustic design, device's adaptability to the harsh network environment. Poor acoustic design can lead to severe microphone speaker coupling, nonlinear distortion, and microphone noise, and cause some common audio algorithms (echo cancellation, noise reduction) not to work properly. Poor VMC quality caused by equipment factors can be strong

reverberation and howl. Much work seeks to improve the quality of video and audio transmission by improving algorithms such as Acoustic echo cancellation [24, 12], Background noise cancellation, Video compression [25, 28].

Physical environment factors

Physical environment factors include closed environment, noise, howling, etc. The impact of the physical environment on audio is less noticeable than the other factors, but in many cases, it can disrupt the acoustic device from working well. Physical environment factors may cause a positive feedback loop between the microphone and the speaker and the background non-white noise [6]. The research [14] provides an efficient method of echo cancellation and background noise suppression.

2.2 User study of VMC systems

Topics

Simulating face-to-face conversation has always been considered as a main goal of VMC. A large number of studies have addressed the effects of video systems on a conversation, trying to explore the factors that enable the VMC system to simulate face-to-face communication better. For instance [20] concluded that the advent of high-speed multimedia networking improved the conversation as an interpersonal communications tool.

Other work has addressed the value of the video channel. There has always been controversy towards the role of visual information in conversation. In the context of VMC, the primary focus has been on whether the visual information improves conversation or not. [27] Some argue that non-verbal information such as head nods, gestures, facial expressions and the postural information is an integral part of normal face-to-face communication. [4, 15] While some [19]found evidence that the presence of a video channel does not affect communication behavior. A task-centered view that believes the effect of the visual information depends on the tasks has gained its popularity. It is believed by [4] that tasks relying on social cues tend to be affected by visual channel.

There are also user studies to test usability issues of novel VMC systems. [Ω_?, 18]. For instance [18] conducted a study of media space CAVECAT to find potential problems and implications for the next generation design.

Methods

A variety of methods has been employed in VMC research: laboratory studies, field studies, interviews, surveys, diary studies and ethnographic studies.

Laboratory study is the most common method for VMC user study. Typically users are required to perform predefined tasks in specific settings. Both quantitative data (e.g., complete task rate, time) and qualitative data (e.g.users' opinion) are collected. [2] is an example of laboratory studies where the video-mediated contexts (VMCs) were compared with face-to-face and audio-only interactions regarding the success of the task outcome, the process of communication and decision making and user satisfaction.

It worth noting conversation characteristics are frequently used in VMC laboratory study as criteria to evaluate a conversational medium. Frequency and duration of turns, length of pauses, and number of interruptions are conversational characteristics that have been shown to be important in face-to-face interaction. [3, 9, 22] It is frequently used to assess the effect of a conversational medium, especially in the comparison of VMC with face-to-face conversation. For instance, O'Conaill [20] used the conversational characteristics as the criterion to compare the degree of approximation between face-to-face conversation and two VMC systems.

Compared to laboratory study, field studies offer less control, but have better external validity, since they take place in usually occurring social settings. In some researches, both field study and laboratory studies were used. For instance, in a study conducted by Kies [16] on desktop video conferencing systems, a two-part study involving a controlled laboratory experiment and a field study evaluation was conducted. It is interesting that field study got a slightly different result from employing similar dependent measures.

Some VMC researches [21, 17, 13] have employed dairy study. There are two types of diary study: One is elicitation studies, where the media captured by participants are used as prompts for discussion in interviews followed. Elicitation study is seen as a way to trigger the participant's memory; another is feedback studies, where participants answer predefined questions about events so that immediate answers can be collected from the participants. [7]

Diary study allows the participants to report events and experiences in context and in-the-moment thereby minimizing the potential for post rationalization.[1] Although the advantage of diary study is distinct, it has some drawbacks. Firstly, the use of memory triggers might cause inaccurate recalls. Secondly, researchers have low control over the study, and there exist the risks of disruption. Also, participants may find it troubling and disturbing to write everything down. [11]

[1] http://uxpamagazine.org/dear-diary-using-diaries-to-study-user-experience/

CHAPTER 3

User study first iteration

3.1 Method

In the previous section, we hypothesized that there were mismatches between engineer's and users' perspectives in current video-calling software, resulting in imperfect user experience in existing VMC. We assumed that it was common for users to encounter breakdowns in video-calling practices. We were interested in these breakdowns that interrupted the communication and users' workarounds, if any, to mitigate the breakdowns. Therefore, we conducted the 1st iteration user study to a) test the hypothesis by exploring the user experience of current video-mediated communications (regardless of the specific technologies); b) collect practices of breakdowns and workarounds from real-life stories of the usage of VMC technologies.

The 1st iteration user study was introspection, where the researchers of this study reported their own stories of video calls using pre-designed forms. The introspection study lasted for around one month, during which the researchers kept records of every video call they had, preferably right after the call to make sure that the researchers can clearly remember the details of the VMC. The form consists of two parts: one for the respondents to write down the whole story of the VMC, with focuses on the details of any breakdowns and workarounds; another to record raw data of the VMC, including the setting data of not only the respondents themselves but also their communication partners. Definitions of the data collected can be seen in Table 3.1.

3.2 Results

In total, seven researchers have reported their recent VMC stories; 43 real-life VMC stories were collected. The average call length for video conversations was 58.1 mins. Note that it was reported in 13 stories that the conversations ended earlier than desired because of technology breakdowns. As for breakdowns and workarounds, ten of the conversations were reported as "flawless" or "reasonably good," while the rest (33) all had at least one breakdown. In some cases, the respondents had workarounds to mitigate the breakdowns, some not. We will thoroughly analyze these breakdowns and workarounds in the next section.

Among the 43 stories, 29 were work-related, with the purposes of rehearsing, group meeting, researching, interviewing, and so on; the rest 14 stories were related to social contact, in which the communication partners were friends, families, game teammates, or romantic partners, with the purpose of pure chatting, decision making,

Table 3.1: Definition of the collected raw data for the 1st iteration user study

Data	**Definition**	**Examples**
Story	Each video call has story number	S1
Participant	Each participant has a unique code within the relevant story	P2
Site	Co-located participants have the same site number	3
Tot-P	Total number of participants in the call	3
Disrupted	Was the call significantly disrupted?	yes, no
ID	Initials of participants, if known.	YZ
Relationship	The relationship of respondents and communication partners	Family
Purpose	What was the nature of the call?	Interview
Topic	What was the main topic of the call?	research
Initiated?	Did this participant initiate the call?	yes, no
Date	Date the call occurred	15/06/2018
Start	Starting time for the call (when first tried to initiate)	13:07
End	Time when the call ended	13:27
Duration	Duration of the call	20 min
Too early?	Did the call end earlier than desired?	yes, no
City	What city did this participant call from?	Paris
Context	In what context did this participant call?	Sitting in office
Network	Network setting of the device	Wired
Primary HW	Main hardware this participant used for the call	laptop
Extra HW?	Extra hardware to facilitate the call	Microphone
Primary SW	Main software for the video call	Skype
Extra SW?	Supplementary software	Google Doc
Video?	Did this participant display video?	yes, no
Audio?	Did this participant use audio?	yes, no
Text?	Did this participant use text within the VMC software?	yes, no
Breakdown?	Any disruptions where the tool didn't work as intended?	yes, no, n/a
Workaround?	Any actions taken to mitigate the disruptions?	yes, no, n/a

and gaming. The context of the video communication varies as well: most of the communication participants locate at home or office, some at video conference room, and some more challenging, for example from a moving vehicle.

Concerning the communication modalities used, in 3 out of the 43 cases, users only had audio on throughout the entire conversation. In another 3 cases, users turned off the video in the middle of the conversation out of the same reason: they had inferior VMC quality, and by turning off video, they hoped to improve the communication quality by allocating more network from video to audio.

The software used to facilitate the video communication were Skype (26/43 conversations), Appear. In (10/43 conversations), WeChat (5/43 conversations), FaceTime (2/43 conversations), Scopia (1/43 conversations), Zoom (1/43 conversations), Overwatch (1/43 conversations), and Discord (1/43 conversations). In three cases, users switched the VMC software during the conversation in order to achieve better communication quality, in one of which the respondent had changed software twice. In two cases, users used two different VMC software on a different device at the same time in order to communicate with two (sub)group of people.

Twenty-four of the conversations had used supplementary communication media to facilitate the video communication. These media include the text message, e-mail, shared document, shared screen, slideshow, file transfer, YouTube video, and shared game space, among which text message was the most frequent one: in 19 conversations the respondents sent text messages to each other for communication, either inside or outside the primary VMC software.

Of the 43 video calls, 32 involved only two sites. Note that this does not mean that there are only two participants because sometimes there are multiple participants in one site. The remaining 11 video calls are many-to-many, which had increased the complexity of the breakdowns.

3.3 Findings

We defined technology breakdown as the situation where the communication gets interrupted because the tool does not work as intended. By analyzing the 43 self-report VMC stories, we found that 33 stories had at least one breakdown, and there were 90 breakdown symptoms in total among these 33 video conversations. We summarized the symptoms of these breakdowns, the source of the breakdowns (if known), and also users' workarounds (if any), from which some interesting patterns of VMC use were revealed. breakdown-confusion-workaround

Breakdowns

Breakdown symptoms

A symptom is the most intuitive aspect of a breakdown. We characterized these breakdown symptoms into 13 types, as shown in Table 3.2. Note that we only focus on technology breakdowns. Below we will explain them one by one.

T1 poor audio quality was the most frequent symptom in these conversations, which had appeared for 14 times. "Poor audio" is a quite general description that can include varieties of situations, for instance, noise, echo, miss word, or garbled audio. The reason for the poor audio, as reported by the respondents, was the poor network for most of the time. Sometimes it was because of the noisy environment. In one case

Table 3.2: Summary of breakdown symptoms from the reported VMC stories

Type	Symptoms	Freq.	Stories involved
T1	poor audio quality (e.g.noise, garbled)	14	2,4,6,27,30,31,33,34,35,36,37,39,40
T2	video freeze	14	3,5,6,7,9,13,14,28,31,35,29,31,37
T3	have trouble initiating call	11	2,3,4,11,13,15,17,42
T4	call drop off	11	2,6,8,11,13,17,42
T5	unable to hear others	10	3,7,9,16,18,21,27,28,37,40
T6	audio outage	5	2,5,17,24,37
T7	nonsimultaneousness in voice and images	5	14,16,23,30,35
T8	weird/ wrong view	4	2,10
T9	poor image quality (e.g.fuzzy, pixilated)	3	9,27,31
T10	conflict with other program	3	12,14,37
T11	unable to be heard	2	7,27
T12	audio cannot get connected	1	3
T13	fail to join existing call	1	5

the user reported that *"delay has each of us speaking over the other"*(Story30), which lead to the poor audio. The symptom of "poor audio" ranges from mild to severe: in Story39, the conversation was *"tiny bit garbled a few times, but easy recovery"*, while in Story37,40, the conversation was so disrupted that users had to switch to another VMC software.

T2 video freeze is when frames of content repeatedly show on the screen—"freezing" the action. As reported by the respondents, most of the times video freeze came along with poor audio or audio outage. In 7 cases the respondents shut down the call because of freeze and poor audio, among which 2 tried the alternative software, 3 reconnected the call within the currently-used software, and 2 ended the conversation. In two cases, the reporters reboot the Internet or used alternative network options.

T3 having trouble initiating calls was a common symptom. The source of the problem was heterogeneous: It could be the unclear user interface (Story2,4) that hindered the users from accepting or setting up the call; Sometimes it was the problem with the other site, for example, the communication partner had broken WiFi (Story1), or dead phone (Story2); A respondent (in Story11) found that if both sites initiated the call to another at the same time, the call invitations would get suspended. They tried three times before finally understanding what was wrong. In four cases (Story 13,15,17,42) the user was not aware of the source of the breakdown, which made them confused: *"My friend appears connected with green light. When I tried to call, it did not even ring and closed the call mode by itself."* **T13 fail to join existing call** is a branch of having trouble initiating calls. One respondent reported

this problem. In this case, it was since the app did not support the function.

Another common breakdown symptom was **T4 call drop off**, which was considered as VERY disruptive. Mostly it was due to the network problem such as broken WiFi, or hardware problem such as device out of power. In one case(Story8) the respondent got a phone call from others, although he did not pick it up, the ongoing video call got cut off. Respondents' reaction toward call drop-off were surprisingly consistent: all of them reconnected the video call right after the drop-off.

T5 unable to hear others refers to the situation where the communication partners seem to be muted for the user, while **T11 unable to be heard** the opposite. T5 was reported ten times. When a user could not hear others, the problem could lie in the user's site (e.g., the user have a poor network – Story4, 9), or his communication partner's (e.g., the user's communication partner had wrong microphone setting – Story 27). T10 was reported 2 times. In both cases where the respondents could not be heard (Story 7, 27), they were not aware of the problem until their communication participants informed them, which is an interesting phenomenon that will be discussed in the following section.

In three cases where more than two sites were involved in the VMC, it happened that only some of the communication participants had trouble hearing or being heard. As reported in Story18, user *" had the experience of not being able to see or hear one of the call participants, even though another participant could see and hear that participant"*. This situation made the breakdown more complicated and confusing for users to understand, and thus increased the difficulty for users to have workarounds.

T6 audio outage means the voice channel is wholly cut off. An audio outage can be seen as an extreme case of T1 "poor audio quality," and a part of T4 "call drop off." It was reported in five cases. Respondents had various workarounds towards the symptom: one changed the network setting (Story5); one reconnected the video call (Story24); one used alternative software and kept the audio only (Story37); one kept repeating the contents when the conversation recovered (Story2). A similar symptom is **T12 audio cannot get connected**, which appeared once.

T7 non-simultaneousness in voice and images refers to the situation where the audio goes slow or fast than video. This symptom was found in 5 stories, among which four reported that it was repetitive. One solved the problem by rebooting the Internet and trying another VMC software.

T8 weird/wrong view means the view of the user in video call is different from what he expected. A respondent complained that she had two possible cameras from the laptop and a big screen and did not know how to choose the camera as wish.

T9 poor image quality can be a fuzzy image, pixilated image, etc. The reason that we do not consider video freeze as a branch of "poor image quality" is because video freeze is usually a problem with the timeline, which doesn't have to be related to image quality. A video can have high image quality but gets frozen. The problem of poor image quality arose in three stories, all of which were reported to be repetitive

or continuous.

T10 conflict with another program were found in three stories. Sometimes video calls can affect other software, by slowing down the network speed (Story 12), or taking over space (Story 14), etc.; sometimes they can be affected as well. In one case (Story 37), the user got one video invitation during the video call, resulting in the conflict between the two video calls: the incoming caller got a busy response.

Breakdown Space

In the previous section, we characterize breakdowns into 13 types based on their symptoms. However, we find that there are many drawbacks to this classification. First of all, some symptoms overlap each other, for example, "audio outage"(T6) and "call drop off"(T4). Most of the video calls involved two modalities, i.e., video and audio. In that case, we can consider calling drop-off as the combination of an audio outage and video outage. That is to say, T6 "audio outage" can be seen as a part of T4 "call drop off." However, If we do not differentiate these two but classify T6 into T4, we risk losing details of the breakdown.

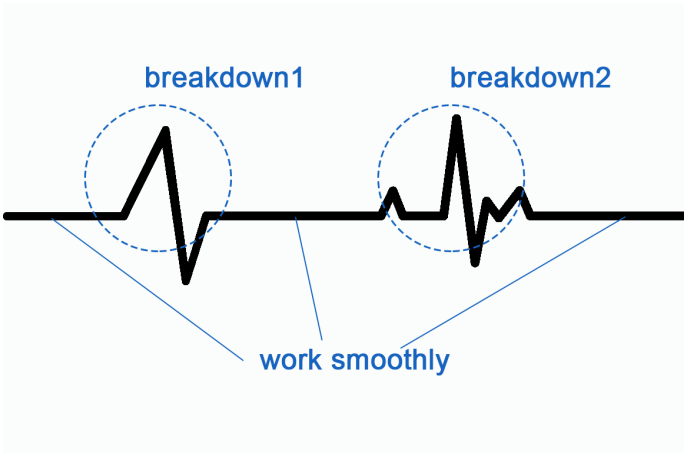

Figure 3.1: Illustration of breakdown using ECG: The segments of the straight line represent the states where the tool works smoothly, while the waves represent breakdowns.

Secondly, it is hard to decide the scale and the boundary of a breakdown: what is ONE breakdown? Can a breakdown have several symptoms? For instance, as mentioned before, video freeze always comes along with audio outage. If a user has a video freeze and audio outage at the same time, should we consider it as one breakdown or two? To answer the question, let's have a look at our definition for "breakdown": the state where the tool does not work as intended. In Figure 3.1, the ECG is used to illustrate our understanding towards breakdown. The segments of

the straight line represent the states where the tool works smoothly, while the waves represent breakdowns. The period from the moment when the smooth state begins to change, to the next smooth state, can be considered as the state of breakdown. Regardless of how many breakdown symptoms appear during this period, we attribute them to one breakdown.

Thirdly, a breakdown involves multiple aspects besides the symptoms. One aspect could be persistence– whether the symptom is repetitive or continuous, and how long it last. Of the 90 breakdown symptom cases, 56 were reported to be repetitive or continuous. An example of repetitive symptom can be seen in Story17, where the respondent had *" several brief audio outages (10-20 sec), until the total audio outage"*. Another aspect could be the level of severeness – how disruptive the breakdown is. In Story39, the conversation was *"tiny bit garbled a few times, but easy recovery"*, while in Story31, the conversation was so disrupted that users chose to end the conversation: *"there were enough freezes that we ultimately ended the call"*.

In summary, the previous classification of symptoms is far from enough to depict the overall picture of a technology breakdown thoroughly and clearly. Therefore, we propose a breakdown space (as shown in Figure 3.2), where the technology breakdowns are described not based on the single aspect of symptom but from multiple perspectives. Seven dimensions of the breakdown space are introduced below:

When?	initiating the call	during the call	at the end (forced disconnect)			
Audio Problems?	none	poor quality (noise, echo)	no voice	garbled	slowdown or speedup	nonsimultaneousness
Video Problems?	none	poor quality (foggy, pixelated)	no picture	freeze	slowdown or speedup	nonsimultaneousness
Persistence?	once	repetitive	constant			
Sites?	one site	some of the sites	all sites	unknown		
Passiveness?	trouble seeing/hearing	trouble being seen/heard	both	unknown		
Severeness?	not disruptive	easy recovery	seriously affected	totally disruptive		

Figure 3.2: Seven dimensions of Breakdown space

- When did the breakdown occur, relative to the video call? The same symptoms can have different influence at different stages. For instance, "no audio" when

initiating the call means the voice channel does not get connected, while "no audio" during the call means typically there's audio outage.

- What was the audio problem, if any? It could be poor audio quality (e.g., noise, echo), no voice (e.g., audio outage, audio could not get connected), garbled audio, slow down or speed up, or non-simultaneousness with video.
- What was the video problem, if any? It could be poor video quality (e.g., noise, echo), no picture, frozen video, slow down or speed up, or non-simultaneousness with audio.
- Was the breakdown persistent? It might appear only once, or it was repetitive, or it was constant. The criterion for judging whether a breakdown is constant is whether it persists without human intervention. If so, it is constant. If the same problem happened again and again, it is seen as repetitive.
- How many sites were influenced? One site? All sites? Or some of the sites? Note that sometimes it is hard to figure out how many sites are affected, especially when the conversation involves multiple sites. Users need to communicate to figure out what happened, but the breakdown causes their communication channel to be blocked.
- Was the problem related to the difficulty seeing/hearing others, or the difficulty being seen/heard? The difference is, the former makes the user's breakdown clear at a glance, while the latter can be implicit for the user and always arise confuses.
- How disruptive was the breakdown? Regarding this dimension, we mainly consider two factors. First, whether the existence of this breakdown made the conversation impossible to be continued; second, whether the problem disappeared by itself or was easily solved by the user?

Here's an example (Story3) to explain the breakdown space: it was a conference call among P1 (reporter), P2, P3, and P4, who located in three sites. P1 and P2 are on the same site. The story goes as below.

> *"P3's Wi-Fi went down. He tried his phone, connected to P4, but he could not hear, and we could not hear him. So he dropped out of the call altogether. The three remaining people continued with skype with video...P4's image kept freezing, along with the sound. However, hard to tell if he could still hear us...his voice was very muddled and we strained to here...Eventually, we turned off the video, but it did not help much...the lack of video made it harder to tell when he stopped talking or if it was frozen."*

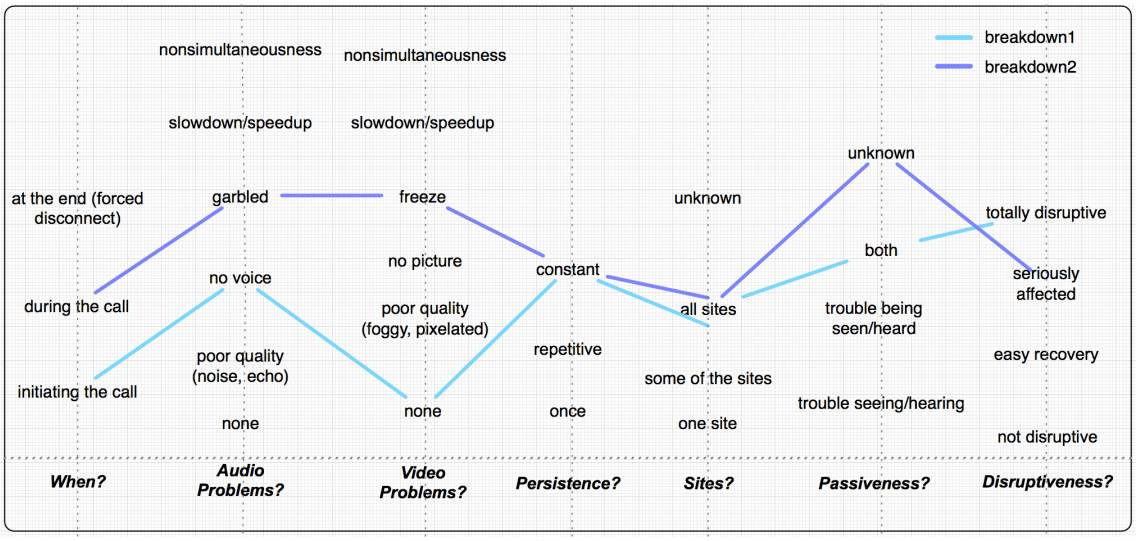

Figure 3.3: Breakdown space example of Story3

There were two breakdowns in this story, labeled with blue and purple respectively in Figure 3.3. The first one was the problem with P3 when initiating the call. The audio problem was "no voice," and there was no video problem. Since P3 could not hear the others and the others could not hear P3 as well, all three sites were affected by this breakdown. The breakdown can be seen as constant because it persisted until the P3 exited the conversation. Users, in this case, have difficulty both hearing others and being heard. Finally, the breakdown made the dialogue impossible to continue, and it neither self-healed or was solved by users; therefore, we considered it as totally disruptive.

The second one was the problem with P4 during the call, which involved both a video problem, i.e.freeze, and an audio problem, i.e., garbled audio. The breakdown was constant because it lasted for a while and didn't recover by itself. Both remaining sites were affected by this breakdown. Regarding the passiveness, for the reporter P1, it should be unknown, because P1 was sure that she has difficulty hearing and seeing P4. However, she was not sure whether P4 could hear or see her. Regarding the disruptiveness, the conversation could continue but severely affected.

Through the above example, we can see that the breakdown space can illustrate multiple aspects of a breakdown more clearly and comprehensively than the breakdown symptoms. Another advantage of the breakdown space is that it can display all the breakdowns of a VMC in one diagram, allowing people to understand all the problems in a VMC at a glance.

Confusions

In the process of studying the collected stories, we found that many users fell into a confused phase when the breakdown occurs. There were two main reasons for them to be confused: a) they didn't know the status of all the sites, so they couldn't figure out what the breakdown was from the overall perspective; b) even if they understood the full picture of the breakdown, it was difficult for them to diagnose the source of the breakdown accurately. These confusions have severely affected the user experience and limited users' ability to resolve breakdowns. In this section, we will analyze the users' confusions in detail.

Confusions about the breakdown status

To understand a breakdown, the user needs first to be aware of the breakdown. In some cases, for instance, Story7, the user was unconscious of the problem until being informed by the communication partners.

> *"When I was talking, he said:' wait, now you are frozen, and I cannot hear you.' However, for my part, the picture and sound of both of us were very smooth."*

As can be seen from the above case, when a user has difficulty being seen or heard by the other party but has no difficulty seeing or hearing, it is difficult for him to realize the existence of the problem consciously. We can call this situation a **stealth symptom**. The opposite is the **dominant symptom**, where the user can spontaneously realize that there is a breakdown. Confusions about the breakdown status do not just exist on the site with stealth symptoms. In another case (Story 37), persons that have dominant symptoms also have confusions about the breakdown status.

> *"Neither P2 nor P3 could hear me, but they could hear each other, and they could still see me. I could hear them and they were talking about it, wondering if I could hear them."*

In this case, the reporter(P1)'s communication partners P2 and P3 both have difficulty hearing P1. They were aware of the breakdown, but they were still confused about whether P1 could hear them or not. Another type of confusion is about the latest state of the breakdown, for example in Story7, when P2 could not hear P1, P1 did not know if the breakdown had disappeared.

> *"He kept talking to inform me of the latest state:" it is still stuck, wait a minute...", "wait wait...", "okay, now you can talk." Otherwise, I wouldn't know when can I continue talking."*

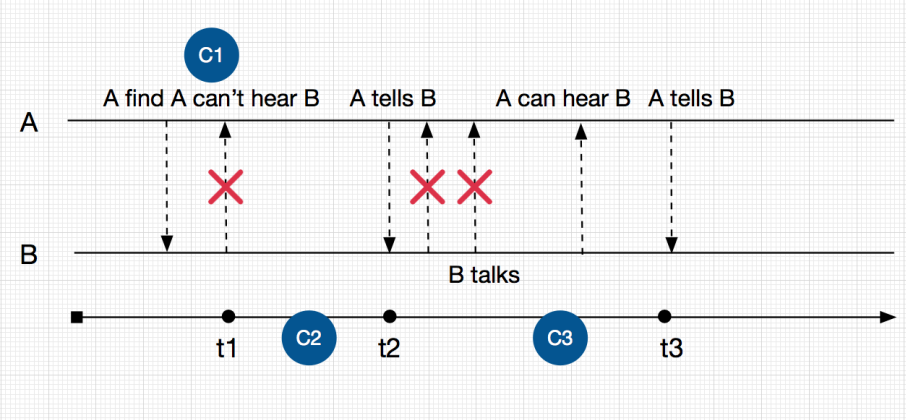

Figure 3.4: Illustration of a use case to show the possibilities of confusions

Below we use a simple but typical scenario to discuss the possibilities of confusions. When B is talking, A finds he could not hear B. Figure 3.4 shows the dialogue between A and B. The bottom line is the timeline, and the dotted line between A and B indicates their voice output. Red crosses represent failed signal transmissions. At t1, A realizes that A can't hear B, here comes the first confusion: A doesn't know whether B can hear A or not. A tells B about the breakdown. If we assume that B can hear A, then at t2, B realizes that he could not be heard. Here comes second confusion: during the period from t1 to t2, B is entirely unaware of this breakdown. Therefore, B does not know when this situation started, which makes it harder for him to resume the conversation later. Another confusion is, starting from t2, B never knows when he could be heard again. He could only judge by talking and waiting for A's response. Finally, A hears B and informs B at t3.

The above is already the most simplified situation. If B cannot hear A as well at first place, they can fall into a loop where both keep asking "can you hear me?" while no one can get an answer, until one of them shut down the call or the breakdown disappears by itself. Alternatively, if they keep video off, it is possible that they can not even tell if there is a breakdown, as commented by the respondent in Story3: *"The lack of video made it harder to tell when he stopped talking or if it was frozen."*

If more than two sites are involved in this conversation, things can get even more complicated, for instance in Story 40: *" at the beginning was a tiny bit confusing because some people could see everyone, but others could not see one other person."* In this case, the user needs to communicate a lot with each other to figure out what is happening. *"...we were using the built-in chat to figure out who could hear who."*

Confusions about the breakdown source

In order to solve a problem, one has first to diagnose the source of the problem. However, even when users have acquired a comprehensive understanding of the breakdown status, they might still be confused with the source of the breakdown. Explicitly knowing the source of the problem always make it easier for users to solve it. In one case (Story31), the user had a poor wifi network and knew the sweet spot to get the best signal. It was a rare case where a user knew exactly what the issue was and was able to address it.

> *"A friend...uses the wifi network from the house across the street, which means that the signal often comes in and out, and he knows that he should sit in a particular part of the house to have a better connection...The call was so bad at the beginning that he had to move to a better spot and then we could discuss his network etc."*

However, this is not always the case. Mostly the user does not know the source from the beginning and has to explore to figure it out, which may or may not work. Story17 shows an interesting case of user's confusion towards the breakdown source. The same problem (i.e.call drop off) happened three times. The first time the user did not explore and therefore did not know the source of the problem. The second time the user explored and found that her WiFi was the problem. The third time the user explored the source of the problem again, but could not find it.

> *"Call completely dropped...Not sure the cause, could have been my wifi (didn't think to look). I re-entered the call no problem...Call dropped again, shortly after. This time it was my wifi (I paid attention this time)...Then dropped again, again shortly after. This time it was not my wifi (could access other sites no problem). "*

Most of the times, there are no system cues to inform the user of the problem or the source of the problem. However, sometimes, the system does give notifications, but still can't provide enough information for the user to get to the cause. Below is an example as reported in Story13 where a system notification was sent to the user, but not entirely useful.

> *"...try calling back...I get a message that says the other person is unavailable, but it is not clear if they are calling me or if there's a trouble with the network, or what. Finally, after two tries to reconnect on my end, we end up reconnecting."*

A comprehensive understanding of breakdown status can help users find the source of breakdown. For example, in a VMC case involves three people A, B, and

C, when A couldn't hear B, the possibilities could be: 1) it is a problem by A's side, e.g., A wasn't receiving audio signals. In this case, B and C should be able to communicate with each other flawlessly; 2) it is a problem by B's side, e.g., B was not sending audio signals. In this case, both A and C could not hear B; 3) it is a problem with the connection among all involved sides; namely, none of them could hear each other. If the user can figure out the states of all communication partners, then there is a higher possibility for them to know where the source of the problem lies.

Workarounds

Workaround refers to the actions the user takes to mitigate the breakdown. Users can have several workarounds for one breakdown. Users' ultimate goal of taking workaround is to solve the breakdown and let VMC continue smoothly. In order to achieve this ultimate goal, users sometimes have intermediate workarounds a) to understand the breakdown, and b) to diagnose the source of the breakdown. We divide users' workarounds into three steps.

1. Workarounds to understand the breakdown

   As discussed in the previous section, when a breakdown occurs, users do not always have a comprehensive understanding of what is happening, which can make them confused. Therefore, some users take workarounds in order to figure out the whole picture of the breakdown. The most common workaround in this stage is communicating with VMC partners, either through the ongoing video call or by sending messages. An example can be seen in Story40, where some of the conversation participants could not see or hear. In order to understand what was going on, all conversation participants sent messages in the group to figure out who could hear who and to test whether all of them were connected.

2. Workarounds to diagnose the source of the breakdown.

   In order to solve the problem, merely understanding the breakdown is not enough. The user needs to be able to figure out the source of the breakdown. Sometimes there are system cues to notify of the problem reason, for instance, "poor network." Under the circumstances where the user does not receive system cue, or the system cue is not useful, users have their workarounds. Story9 shows an example of user's workaround to diagnose the source of the problem:

   > *"...when she's frozen again, I shut down the video and texted her asking the reason. I told her I have full signal, and she said she only got two of five. So we realize it must be her signal problem."*

In the above case, the user figured out the source of breakdown by checking the network of both sites. In another case (Story16), the user diagnosed the source with the knowledge of all sites' breakdown symptoms:

> *"One participant could not see my video, yet all the other participants could see it, so not an issue with my computer not sending the video signal"*

3. Workarounds to solve the breakdown.

   We observed that users generally have several ways to solve breakdowns: 1) using alternative or supplementary hardware, e.g.landline, laptop; 2)using alternative VMC software, e.g., Skype, WhatsApp; 3) using alternative or supplementary media, e.g., text messages, file transfer; 4) change network settings, e.g. from 4g to WiFi; 5) change the involved modalities, e.g., turning off video to save audio; 6) reconnect the call. It is an interesting phenomenon that in most cases when the user cannot diagnose the cause for a breakdown, they tend to reconnect the video call, expecting that the problem disappears automatically. However, in many cases, it does not work.

A summary of the users' workaround can be seen in Table 3.3. Users do not always have workarounds. When the user does not have any workarounds, it may be the following situation: 1) user gives up the conversation, i.e., turns off the call and ends the conversation; 2) the user ignores the breakdown because it is not so disruptive; 3) the problem disappeared by itself before the user takes any action.

Table 3.3: Summary of workaround categories

Purpose	**Actions**	**Options**
To understand the breakdown	Talk about the breakdowns?	no; a little; a lot
To diagnose the source	User seeks?	no; yes, e.g. network setting
To solve the breakdown	Modify the involved modalities	no; yes, e.g. turn off video
	Use alternative hardware?	no; yes, e.g landline
	Use alternative media	no; yes, e.g. SMS
	Use alternative VMC software?	no; yes, e.g. Skype
	Change network setting	no; yes, e.g. from wifi to 4g
	Reconnect the call?	no; yes
	Success?	no; poor; okay; fixed

3.4 Discussions & Limitations

The user study above provides a deep understanding of how people currently use VMC technology. We found a pattern of "breakdown-confusion-workaround" where the user encounters breakdown, get confused, and then take workarounds to address the confusion and solve the breakdown. We proposed a seven-dimension breakdown space to describe a breakdown from multiple perspectives. We introduced two types of confusions that hinder users from solving the problems, i.e., confusions about the breakdown status and confusions about the breakdown source. The confusions can be seen as the pain points in user experience, which can be used to guide the design improvements for the future VMC software. Finally, we summarized users workarounds and categorized them into three steps according to the purposes.

There are some limitations to this study. Firstly, all the results and finding are based on self-reports of the researchers themselves, which could be biased. Other users could have different patterns or habit of VMC use. Secondly, only 43 VMC stories were collected, which may not be enough to cover all situations. Thirdly, the self-report forms can hardly cover all the details needed. For some stories, it is hard for us to understand the whole picture of all sites.

CHAPTER 4

User study second iteration

4.1 Introduction

In the first iteration, we conducted introspection with the researchers of this study and captured 43 cases of video-mediated communication both at work and in daily life. We analyzed these VMC stories and found interesting patterns of "breakdown-confusion-workaround." However, although only factual data was collected in the 1st iteration, the homogeneous group of respondents may result in biased results. Therefore, in the second iteration, we ran a diary study with ordinary users of VMC technologies, to test whether the findings from the previous study still held water, to gain a better understanding of existing video-mediated communication, and to impact the design of VMC that could better support users.

4.2 Method

Diary study

Our research participants took part in diary studies where they recorded entries about their VMC use in every day lives in a log over a period. This allowed us to study their regular VMC practices in real-world scenarios. It is worth noting that we collected all stories of the video calls during the period of diary study, not just the problematic ones. The goal was to collect three cases of VMC from each participant.

Participants

Facebook advertisement and word-of-mouth recruited 10 participants (5 female and five male). The participants of the diary study were ordinary VMC users from all around the world (China, France, Canada, the US, and India).

Before taking part in the diary study, participants were required to fill in a pre-questionnaire. We preferred the participants that used more than one VMC technique, because we were interested in whether user acted differently using different software, and we did not want to limit the scope of this research to one specific VMC software. All of the ten participants had ever used at least two VMC techniques, including frequently used ones and infrequently used ones. Also, we asked the participants' frequency of VMC use but did not screen them based on it. Six of the ten participants reported to have video calls a few times a week; three had video

calls respectively about once a day, once a week, and a couple of times a month; and one rarely had video calls.

Procedure

Before the study, the participant filled in a short questionnaire asking basic personal information including occupation, gender, email address, and information related to their habits in VMC use including the frequency of video call, and the VMC software they ever used.

Our goal was to collect three VMC stories from each participant, therefore, the duration of the study depended on their frequency of video calls. During the study, the participant kept a log of video calls by filling in a set of self-report forms, until they had three video calls. We strongly suggested the participant did the self-report right after each video call when their memories towards the communication were still fresh. The set of forms consists of three parts:

1. Raw data of the call;

   This includes date, time and duration, participant's network (e.g. WiFi, 4g), participant's device (e.g. smartphone), software used (e.g. Skype, WhatsApp), participant's context (e.g. in a moving car, sitting in the office), nature of the call (e.g. social, work), general purpose (e.g. interview, decision-making, pure chat), relationship between communication participants (e.g. friends, colleagues), media involved (e.g. video, audio, file transfer, text messages), number of communication participants involved, number of sites involved, and intensity level.

   Here low-intensity level can be the situation where the user takes the video call as a side channel to parallel activities, while high-intensity level indicates aggressive conversation. The question of intensity level was inspired from our observations in the 1st iteration, where we noticed that user tended to pay less attention to breakdowns in deep, intense conversations than in high intense ones. Here we would like to test whether our hypothesis was right.

2. Story report;

   We would like the participant to describe the story of the video calls briefly. We were interested primarily in the breakdowns (if any) that interrupted the participant from having a smooth conversation, and the reactions of the participants and their communication partners towards the breakdowns.

3. Breakdowns and workarounds questionnaire;

   Here we would like the participant to report the details of the breakdowns and workarounds in the questionnaire. We required the user to fill out one

questionnaire for each breakdown they encountered. If it was a perfect call, they had to leave the questionnaire blank. Repeated breakdowns of the same type in one video call can be put in one sheet.

Inspired by the findings from the 1st iteration, we designed the questionnaire into three sections, with topics of breakdowns, confusions, and workarounds respectively. In the breakdown section, we asked questions about the relative time, symptoms, persistence, number of sites affected, and disruptiveness. In the confusion section, we explored whether the user knew the source of the breakdown, and how. Also, we asked whether there were system cues to notify of the breakdown or the source of the breakdown. In the workaround section, we asked the participant to report if they took any actions to mitigate the breakdown, and if yes, to what extent did they think the workaround succeeded in solving the problem.

4.3 Results & findings

Verification of 1st iteration

Did the participants encounter breakdowns?

Up to now, we have received 22 self-reports from the 10 participants. Seven were work-related calls, and fifteen were social calls with friends or families. It worth noting that only two of these stories were reported flawless, while all of the others were problematic. That is to say, the frequency of user encountering breakdowns was surprisingly high. 34 technology breakdowns were reported in total, among which ten were repetitive (meaning that the breakdown happened for several times in one VMC) and nine were constant (meaning that the breakdown persisted until the user took actions or the call ended).

Were they confused about the breakdowns?

In the 1st iteration, we found that some users got confused about the breakdown or the source of the breakdown. Evidence to support this finding can also be found in stories of the second iteration. In some cases, the user mistook the reason for a breakdown, which caused them to take the wrong workaround. For instance, in a case where the participant kept hearing noises and thus couldn't hear her communication partner (her mother) clearly, the participant thought it was because of her noisy environment, so she held the speaker in her hand to avoid noise, which didn't work. Talking with her mom about the breakdown, then she realized that actually, it was her mom who did not turn on the head free mode that leads to the breakdown.

In 10 of the 34 problematic cases, the participants did not know the source of the breakdown throughout the conversation. In 12 cases, the participants did not know the source at the beginning but fortunately managed to figure it out, mostly by talking with their communication partner (either orally or through text messages), and by checking their network.

In the other 12 cases, the participants claimed that they knew the source of the breakdown immediately when it occurred. We dug into the reason for which they were able to diagnose the source immediately. In most of the cases, it was because of their previous experience, for instance, a participant always knew that his parents had poor wifi signal at home, so when his call with his parents got distorted, he immediately knew that it was because of the poor network in the other site. Sometimes the source of the breakdown was evident, for example, one participant easily recognized that a room was comfortable to produce an echo, which leads to poor audio quality. In some cases, the breakdown was triggered by the users themselves. For instance, a participant locked his iPhone screen during the video call, resulting in an outage. In this case, although the user did not expect the call to be terminated, he was clear about what caused this problem.

Did they take workarounds?

In the 1st user study, we categorized users' workarounds into three types, i.e., workaround to understand the breakdown, workaround to diagnose the reason, and workaround to solve the problem. All the three types of breakdown could be found in the stories collected in the 2nd iteration.

We found it difficult to distinguish the boundaries between the workarounds to understand the breakdown and workarounds to diagnose the source. In most cases, the user did these two at the same time. As mentioned before, workarounds in these two stages were communicating with other sites or check own network.

In 25 of cases, participants took workarounds to solve the breakdown. These workarounds included (in parentheses is the frequency): using alternative/supplementary hardware (2), turning off video to save audio (3), using alternative software (1), reconnecting (8), changing physical environment e.g. moving to better location (2), changing network settings (3), changing device setting e.g.turning speakers off(1).

Other findings

More than half breakdowns were reported as totally disruptive

We asked the participant to rank each breakdown from 1 to 5 based on it disruptiveness. The purpose was to figure out participants' subjective feelings about how much a breakdown had affected the communication. One referred to the situation where the conversation was not affected at all, and five means the conversation was too

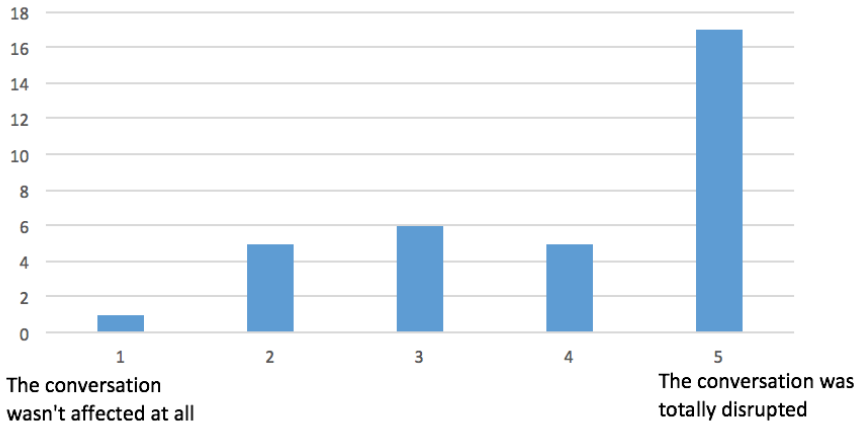

Figure 4.1: Participants' perception of the disruptiveness of breakdowns

disrupted to be continued. As shown in Figure 4.1, more than half (17) breakdowns were reported as totally disruptive, where the conversation could not continue unless the user took actions to solve it. We found that participants took actions more actively toward those highly disruptive breakdowns. While for those less disruptive ones (with score 2), in four of the six cases, the participants chose to ignore it, as commented by a participant: *"I did not do anything to solve it, because it was not a big problem. Moreover, I was in the middle of the interview, so better not to interrupt the interviewee."*

Visual problems occur less frequently than auditory problems

Another interesting phenomenon we found was that the participants reported much more auditory problems than video ones. Of the 34 problematic cases, 32 had symptoms of the audio problem (e.g., distorted audio, noise), while only 17 had symptoms of video problem (e.g., frozen video) The reason for this phenomenon is to be explored in the future.

Cost of breakdowns

In the previous section, we used "disruptiveness" to quantify the degree to which a breakdown affects the VMC conversation. However, the score of disruptiveness relies heavily on the subjective judgment of the user. Here we explored an objective way to quantify the costs of the poor Quality of Service (QoS) in VMC by calculating the ratio of the time user spent off the conversation content to the length of the video call.

Figure 4.2 showed an example of a problematic VMC that has three breakdowns. At the bottom is the timeline for the video call: solid lines depict the time user spent on the conversation content itself, while the dashed lines (i.e.t1 and t2) represent the time they wasted because of poor QoS. Breakdown1 and Breakdown3 are both highly disruptive that prevent the conversation to be continued. The user takes time to work them out before getting back to the conversation. Breakdown2 is not a big problem and therefore be ignored by the user. In this case, the cost of poor QoS can be (t1+t2)/T: the length of the VMC divided by the time spent off the conversation content.

In a flawless VMC, ideally user do not need to spend any time on the quality of the conversation; therefore, the cost would be 0; while in the worst case where the user spends all time on the conversation quality rather than the content, the cost is 1.

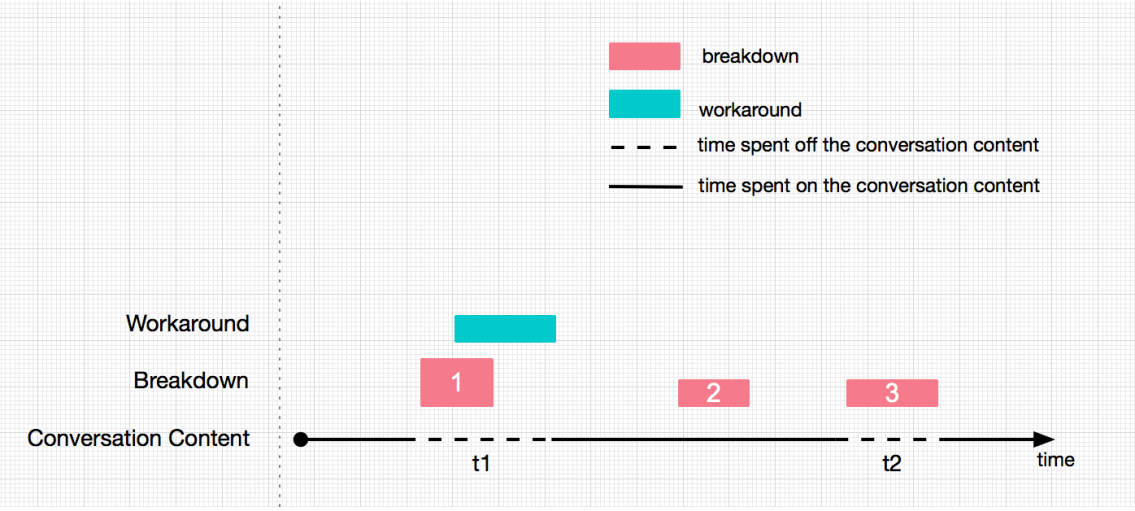

Figure 4.2: Timeline to illustrate the cost of breakdown

4.4 Design suggestions

Based on the findings from the two user studies, we proposed design suggestions for the future VMC software to better support user in case of breakdowns.

First, when a breakdown occurs, provide timely feedback of what is the problem, and what could be the source of the problem to help user go through the "confusion" period. VMC system could take into account the situation of each site, diagnose the problem from the technical perspective, and then notify the user. The feedback should contain richer information than just "poor connection."

Second, provide interactive feedforward to guide the user to solve the problem. For instance, give trade-off options for media channels (i.e., video and audio) and communication quality. When limited bandwidth causes low VMC quality, ask the user if they are willing to sacrifice video quality for better audio quality, or the other way round.

CHAPTER 5

Conclusion and perspectives

In this thesis, we aimed to understand the user experience of video-mediated communication under real-world conditions. We conducted two studies in which 64 real-life use cases of VMCs were captured and analyzed. The results suggest that VMC systems often result in technical breakdowns that disrupt the user's conversations, leading to unsatisfying quality of experience (QoE).

We summarize the breakdowns from collected use cases and identify 13 breakdown symptoms (see Table 3.2). We then propose a seven-dimension breakdown space that characterizes a breakdown from multiple perspectives, i.e., 1) When did the breakdown occur; 2) What was the audio problem; 3) What was the video problem; 4) Was the breakdown persistent or repetitive; 5) How many sites were influenced; 6) Was the problem related to the difficulty of seeing/hearing others or the difficulty being seen/heard; and 7) How disruptive was the breakdown.

We find that in most cases of breakdown, the current system's error messages aren't useful in helping the user identify what is wrong and how to fix it. The user easily gets confused towards a breakdown. We recognize two main types of confusions that hindered users from understanding and solving breakdowns: a) users don't always know the status of all the sites, so it is difficult for them to figure out what the breakdown is from the overall perspective, for example, when A couldn't hear B, A doesn't know if B could hear him; b) even if the user understands the full picture of the breakdown, it is difficult for them to accurately diagnose the source.

To address the confusions and to solve the breakdowns, users have tried various kinds of workarounds, which can be categorized into three steps: workarounds to understand the breakdown; workarounds to diagnose the source of the breakdown, and workarounds to solve the breakdown. A summary of breakdown categories can be seen in Table 3.3.

Based on the findings above, we propose implications for the design of VMC techniques to mitigate poor service and better support for users' natural approach to workarounds. We suggest that timely feedback and interactive feedforward should be provided in VMC software when a breakdown is detected. Future work of this research is to implement the design proposal and to test whether it improves the user experience.

Finally, we present a method to calculate the cost of poor QoS in VMC. The length of the conversation divides the time the user spends off the conversation content. 0 indicates that the user spends all the time on the conversation content itself, while one means that the conversation quality captures the user's full attention, suggesting a low efficiency of conversation. Again, the method needs to be further studied in future work.

Acknowledgments

I want to express my sincere gratitude to Wendy Mackay and Joanna McGrenere, the supervisors of the project, for their constant advice, guidance, and engagement through the process of this master thesis.

Furthermore, I would like to thank my teammate John MacCallum for his support and valuable suggestions, and my colleague Nicolas Taffin, who is always an excellent source of inspiration.

Also, I like to thank all the participants in my study for their time, effort and candor.

Finally, I would like to thank my parents for supporting me throughout the entire process.

Bibliography

[1] OYEKANLU Emmanuel Adebomi and JOHN Samson Mwela. Impact of packet losses on the quality of video streaming. Master's thesis, , School of Computing, 2010.

[2] A. H. Anderson, A. Newlands, J. Mullin, A. Marie Fleming, G. Doherty-Sneddon, and J. Van der Velden. Impact of video-mediated communication on simulated service encounters. *Interacting with Computers*, 8(2):193–206, June 1996.

[3] Michael Argyle, Mansur Lalljee, and Mark Cook. The effects of visibility on interaction in a dyad. *Human Relations*, 21(1):3–17, 1968.

[4] Jonathan Benthall, Michael Argyle, and Mark Cook. Gaze and mutual gaze. *Rain*, (12):7, 1976.

[5] Vasanthi Dwaraka Bhamidipati and Swetha Kilari. Effect of delay/delay variation on qoe in video streaming. Master's thesis, , School of Computing, 2010.

[6] S. Boll. Suppression of acoustic noise in speech using spectral subtraction. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 27(2):113–120, April 1979.

[7] Scott Carter and Jennifer Mankoff. When participants do the capturing: The role of media in diary studies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '05, pages 899–908, New York, NY, USA, 2005. ACM.

[8] Mark Claypool and Jonathan Tanner. The effects of jitter on the peceptual quality of video. In *Proceedings of the Seventh ACM International Conference on Multimedia (Part 2)*, MULTIMEDIA '99, pages 115–118, New York, NY, USA, 1999. ACM.

[9] K.M. Cohen. Speaker interaction: Video teleconferences versus face-to-face meetings. *Teleconferencing and Electronic Communications*, pages 189–199, 1982.

[10] Kathleen E. Finn. *Video-Mediated Communication*. L. Erlbaum Associates Inc., Hillsdale, NJ, USA, 1997.

[11] E. Goodman, M. Kuniavsky, and A. Moed. Observing the user experience: A practitioner's guide to user research (second edition). *IEEE Transactions on Professional Communication*, 56(3):260–261, Sept 2013.

[12] Jack Arthur Gowan and James Michael Downs. Video conferencing human-machine interface: A field study. *Information Management*, 27(6):341 – 356, 1994.

[13] Jeffrey T. Hancock, Jennifer Thom-Santelli, and Thompson Ritchie. Deception and design: The impact of communication technology on lying behavior. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04, pages 129–134, New York, NY, USA, 2004. ACM.

[14] W. Harrison, Jae Lim, and E. Singer. A new application of adaptive noise cancellation. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 34(1):21–27, February 1986.

[15] Adam Kendon. Some functions of gaze-direction in social interaction. *Acta Psychologica*, 26:22 – 63, 1967.

[16] Jonathan K. Kies, Robert C. Williges, and Mary Beth Rosson. Evaluating desktop video conferencing for distance learning. *Computers Education*, 28(2):79 – 91, 1997.

[17] David S. Kirk, Abigail Sellen, and Xiang Cao. Home video communication: Mediating 'closeness'. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*, CSCW '10, pages 135–144, New York, NY, USA, 2010. ACM.

[18] Marilyn M. Mantei, Ronald M. Baecker, Abigail J. Sellen, William A. S. Buxton, Thomas Milligan, and Barry Wellman. Experiences in the use of a media space. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '91, pages 203–208, New York, NY, USA, 1991. ACM.

[19] Robert B. Ochsman and Alphonse Chapanis. The effects of 10 communication modes on the behavior of teams during co-operative problem-solving. *International Journal of Man-Machine Studies*, 6(5):579 – 619, 1974.

[20] Brid O'Conaill, Steve Whittaker, and Sylvia Wilbur. Conversations over video conferences: An evaluation of the spoken aspects of video-mediated communication. *Hum.-Comput. Interact.*, 8(4):389–428, December 1993.

[21] Kenton O'Hara, Alison Black, and Matthew Lipson. Everyday practices with mobile video telephony. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '06, pages 871–880, New York, NY, USA, 2006. ACM.

[22] D. R. Rutter and G. M. Stephenson. The role of visual communication in synchronising conversation. *European Journal of Social Psychology*, 7(1):29–37.

[23] Y. M. Saputra and Hendrawan. The effect of packet loss and delay jitter on the video streaming performance using h.264/mpeg-4 scalable video coding. In

2016 10th International Conference on Telecommunication Systems Services and Applications (TSSA), pages 1–6, Oct 2016.

[24] M. M. Sondhi, D. R. Morgan, and J. L. Hall. Stereophonic acoustic echo cancellation-an overview of the fundamental problem. *IEEE Signal Processing Letters*, 2(8):148–151, Aug 1995.

[25] Wai-Tian Tan and A. Zakhor. Real-time internet video using error resilient scalable compression and tcp-friendly transport protocol. *IEEE Transactions on Multimedia*, 1(2):172–186, June 1999.

[26] Chris Welch. Facebook messenger doubles number of video chats to 17 billion in 2017, Dec 2017.

[27] Steve Whittaker. Rethinking video as a technology for interpersonal communications: theory and design implications. *International Journal of Human-Computer Studies*, 42(5):501 – 529, 1995.

[28] Dapeng Wu, Y. T. Hou, Wenwu Zhu, Ya-Qin Zhang, and J. M. Peha. Streaming video over the internet: approaches and directions. *IEEE Transactions on Circuits and Systems for Video Technology*, 11(3):282–300, March 2001.