Managing Quality of Experience for Wireless VOIP Using Noncooperative Games

Jahan A. Hassan, Member, IEEE, Mahbub Hassan, Senior Member, IEEE, Sajal K. Das, Senior Member, IEEE, Arthur Ramer, Senior Member, IEEE

Abstract—We model the user’s quality of experience (QoE) in a wireless voice over IP (VoIP) service as a function of the amount of effort the user has to put to continue her conversation. We assume that users would quit or terminate an ongoing call if they have to put more efforts than they could tolerate. Not knowing the tolerance threshold of each individual user, the service provider faces a decision dilemma of whether to fix the network problem immediately whenever he detects a user effort in the VoIP system, or ignore it with the hope that the user may still continue the call anyway. In this paper, we formulate the provider’s dilemma as a non-cooperative game between the provider and the VoIP user experiencing a deteriorating QoE. We demonstrate that providers implementing the equilibrium solutions can expect to not only increase their revenues, but also reduce the number of cases when users quit out of frustration thus minimizing potential churning. We also discuss conditions under which a sophisticated user may or may not benefit from faking unwarranted efforts with a goal of receiving a better service from the provider. Finally, we conduct a subjective experiment of VoIP over WiFi, which verifies the key model assumption that perceptual quality is negatively correlated to the amount of effort the user has to put to continue the call.

Index Terms—Quality of Experience, Wireless Multimedia, Wireless Voice over IP, Game Theory.

I. INTRODUCTION

Wireless voice over IP (VoIP) extends the traditional Internet-based VoIP services to smart mobile devices using WiFi or cellular networks. Market forecasts suggest that wireless VoIP will continue to enjoy an exponential growth over the coming years [1]. One of the factors driving this growth is the need for reduced cost of delivering voice services on a mobile phone. However, with wireless links at the edge, a major issue facing wireless VoIP is the prospect of occasional link quality problems caused by a multitude of possible sources including interference from nearby WiFi access points, sudden increases in delay due to data overload in the base station, and so on. Irrespective of the actual source of the link problem, users would naturally put some human-level efforts to restore the conversation. These efforts are basically user attempts to retransmit lost speech at the natural language level using words or phrases such as “I can’t hear you, can you repeat please?” With such human efforts, technically it may be possible for users to continue a conversation even with occasional link problems. However, each time a user has to put in such effort, it would cause irritation and eventual frustration if the link quality does not improve during conversation. Given that a user has an expanding list of options for communication, e.g., SMS, e-mail, or even making the call through landline or another VoIP provider, frustration may lead to user quitting the on-going call prematurely, which may also result in eventual churning from the service provider.

Call quitting probability (CQP), i.e., the probability that a call admitted to the system quits prematurely, therefore becomes as important as the call blocking probability (CBP), which merely defines the probability of a call request rejected by the admission control mechanism due to lack of resources. Clearly, if a provider wants to reduce the CQP, he would need to dynamically monitor and manage user irritation for all admitted calls. User irritation is inherently linked to the so-called quality of experience (QoE) or “overall acceptability of a service” [2], [3]. Unfortunately, it is now well established [4] that network parameters such as packet loss rate, alone cannot give a true account of user experience for multimedia services. User experience is rather subjective and context dependent, thus making QoE-based network adaptation an extremely challenging problem. While researchers have traditionally focused on modeling subjective QoE from objective audio-visual data [5], there is a growing evidence in recent years that user experience with a given computing or communication service can be modeled directly from live human biometric data such as galvanic skin response (GSR) and pupil dynamics [6], or even body gestures [7] and facial expressions [8]. Given that many of these biometric data can be gathered with the help of various sensors embedded in today’s smartphones, it is only a matter of time before such human-level QoE modeling become widely available for the wireless VoIP.

Modeling QoE directly from human data is expected to enhance the provider’s capability to obtain a more authentic account of how a particular user is experiencing a service at any instance. Based on this trend in QoE modeling, we assume that, in future, human efforts in a VoIP session may be detected automatically in real-time. This will enable the providers to reliably detect user irritation in wireless VoIP networks and take necessary resource management actions dynamically to improve the link quality. While the details of a particular resource managemenget technique may vary across different kinds of wireless networks, in general the user irritation management through dynamic resource allocation to users not
having good connection experience at the time would imply additional overhead. Given that radio resource is limited, the provider faces the challenge of managing user irritation in a way that does not negatively affect his revenue generation. In particular, when an irritation alarm is generated, the provider must decide whether to allocate additional resource and fix the link quality immediately, or ignore it for the moment hoping that the user will not quit.

Although rate adaptation at the VoIP codec has been studied extensively to degrade the voice quality more gracefully during periods of adverse network conditions [9], the prospect of dynamic network resource allocation to directly improve the link quality in response to a deteriorating user QoE has not been addressed sufficiently in the current literature. In particular, the provider’s dilemma of whether to allocate network resource in an attempt to further improve the link quality or to ignore a ‘QoE alarm’ remains to be formally analyzed. This motivates our work.

The resource allocation decision on the provider’s part and the user’s decision of whether to keep on putting additional effort or to actually quit the call leads to an important decision-interaction problem that can be best analyzed using game theory. It is important to highlight that we have more than one independent agent each trying to optimize its own outcome, but each agent’s outcome depends on the decision taken by the other party. We are more interested in finding out the behaviour that will emerge from such interactions in the equilibrium, rather than optimizing the outcome of a single agent. As such, we believe that game theory is the most natural choice to model this problem compared to other optimization tools. In this paper, we make three major novel contributions:

- We propose the concept of human effort to model user QoE and predict the (frustrated) user state that may lead to call quitting. We validate this concept and characterize human efforts in wireless VoIP through practical experiments involving real subjects.
- We formulate the provider’s dilemma as a non-cooperative game between the provider and the VoIP user experiencing deteriorating QoE, and derive Nash equilibrium solutions. Using simulation experiments, we demonstrate that providers implementing the equilibrium solutions can expect to not only increase their revenues, but also reduce the number of users who quit out of frustration.
- We define and formally analyze the case of a ‘sophisticated’ user, who may decide to generate fake efforts with a goal of receiving better service from the provider. We derive conditions under which a sophisticated user may or may not benefit from faking unwarranted efforts.

The remainder of the paper is structured as follows. Section II positions our work with respect to related work in the area of QoE in wireless networks. Subjective experiments to study user effort in wireless VoIP and its influence on user QoE are presented in Section III. Section IV describes the non-cooperative game between the user and the provider and derives the game’s Nash equilibria. Simulation-based performance evaluation of the Nash strategies are presented in Section V. We conclude the paper in Section VI with direction of future research.

II. RELATED WORK

QoE in wireless networks has been a subject of immense interest in the recent years. One key issue is to find a suitable method to dynamically obtain a subjective measure of user QoE in an on-going service from the objective measures. The International Telecommunication Union (ITU) has released the E-model for VoIP [10] and video telephony [11], which attempts to derive mean opinion score (MOS) from the usual network quality of service (QoS) parameters as well as some other parameters such as noise level, screen size, etc. These models provide a quick and easy way to estimate QoE from objective parameters, but are still inherently limited as no direct feedback is obtained from the users during network operation. Some researchers have proposed the use of artificial intelligence or neural network based approach [12] to more accurately learn and predict the mapping between the network QoS and MOS.

Chen et al. [13], [14] proposed two new methods for capturing user QoE in a VoIP call. In the OneClick work [14], they propose a method to capture real-time user perceptions by requiring the user to press a button whenever she feels unhappy about the quality of the call. In the other work [13], they propose a user satisfaction index which can be calculated on-line by monitoring packet traces as the call progresses. Our work has a similar spirit to [13] in the sense that we also attempt to gauge user reactions in real-time during a call, but we differ in the underlying approach. Specifically, our approach is to model QoE as a function of efforts put in by the users in a VoIP call.

Another key issue is how to adapt application rates or network resources dynamically in response to fluctuations in user QoE. Mohamed et al. [15] proposed a codec rate adaptation algorithm for VoIP that uses the realtime MOS scores obtained from a trained neural network to optimize user satisfaction and bandwidth utilization. Other researchers demonstrated that the use of real-time MOS scores in connection admission control [16] and network selection [17] can improve the overall user satisfaction in the network. In our work, we dynamically make resource allocation decisions at the network level based on human efforts in a VoIP call.

Finally, use of game theory for dynamic resource allocation in wireless networks is becoming a popular approach to simultaneously optimize both user satisfaction and network utilization. In [18], the authors have shown that a coalition based cooperative allocation strategy can maximize network resources while satisfying user performance in heterogeneous networks. Chatterjee et. al. [19] studied integrated admission control and rate control in CDMA systems by modeling the conflicting interest between the provider and users as noncooperative games. The objective here is to provide a balanced optimization between the provider’s revenue and user satisfaction as well as QoS. Beyond resource management for wireless voice calls, game theory has also been applied to study various other interesting problems in wireless networks supporting voice calls. For example, the work presented in [20] focuses on how to compensate a user during the period of total outage in interworked wireless-LAN-3G networks through a noncooperative game-theory-based pricing mechanism. In
[21], the authors model the dynamics of resource sharing where \( N \) users dynamically share a wireless LAN access point and select a bit rate selfishly to improve the quality of their own voice calls.

The game we present in this paper is different from all of the above mentioned work in the following aspects: (i) it studies the management of additional resource allocation after a call has been admitted, in the hope to reduce irritation of a VoIP user (which is different from the call admission control problem), (ii) it uses instances of user ‘effort’ during a VoIP session as the triggering event when the game is played, and finally (iii) it analyzes user behavior in terms of prematurely abandoning an ongoing VoIP session. Although Ref. [22] considers premature ‘call drops’ by the users, they are in the context of assessing the TCP friendliness of VoIP traffic.

As compared with our preliminary conference draft that was presented in [23], this paper has considerable extensions and new material summarized as follows. First, we have conducted additional analyses of the experimental data. In particular, we have included a scatter plot and computed the Pearson correlation between QoE and effort data sets, which demonstrate that QoE is negatively correlated to effort. Second, we have significantly extended and generalized the preliminary game model that was presented in [23], which assumed that ‘complete knowledge’ was available to both players. In this paper we have added necessary theory (Theorem 3) and analysis for more realistic scenarios of ‘incomplete knowledge’, which accommodates the concept of quitting probability. We have presented and discussed several alternatives for a provider to estimate these quitting probabilities (Theorems 4-6). Using curve fitting to experimental data, we have shown that the quitting probability could be approximated as a quadratic function of the accumulated efforts in a VoIP call. We further explain and analyze the case of a ‘sophisticated’ user, who may decide to generate fake efforts with a goal of receiving better service from the provider. We have also derived conditions under which a sophisticated user may or may not benefit from faking unwarranted efforts. Last but not the least, we have significantly extended the related work providing a more comprehensive comparison of our work with others published in the literature.

III. CORRELATION BETWEEN QoE AND EFFORT

Our game theoretic model for provider-user interactions is founded on the assumption that human effort is a factor with a major influence on the user QoE or the perceived quality of the voice call. In this section, we set out to verify this assumption through a small-scale subjective experiment involving real users making calls over an wireless VoIP test-bed.

A. User Experiments and Effort Data Collection

Our wireless VoIP test-bed (see Figure 1) was built and installed in a graduate research laboratory space at the University of New South Wales using (i) a laptop equipped with a 2.4 GHz WiFi card configured as an access point and connected to the Internet using Ethernet enabling direct access to the public Skype server, and (ii) two HP iPQAs [24] with Skype clients [25] and AudioNotes [26] installed, which are used by two users, a caller and a callee, at any one time. In the vicinity of the test-bed, there were microwave ovens and departmental WiFi APs operating in the same frequency band with a potential to affect the quality of the two wireless links of the test-bed. We have not controlled the wireless link qualities in any way, i.e., we have neither introduced any packet loss, jitter, etc., nor have we attempted to improve the link quality through any specific resource management techniques (Skype may have some proprietary rate adaptation mechanism at the application layer [27] to gracefully degrade the voice quality during any prolonged periods of link problems, but it does nothing to improve the actual link quality). Because the caller and the callee are located at different rooms and have separate wireless links connecting their iPQAs to the AP, they are not expected to experience identical link qualities. Therefore, we examine the experiences of the caller and the callee separately by recording each user’s audio stream on the respective iPQAs using AudioNotes. Consequently, we generate two audio streams, referred to as ‘VoIP sessions’ for each single ‘VoIP call’ between a caller and a callee. Recorded audio streams are later transferred to a central repository for post processing, i.e., for effort data collection from the semantic information in the audio streams.

For subjective experiments, we recruited seven graduate students (users). They made a total of 29 calls between each other at different times over a period of 2 weeks, generating 58 distinct audio recordings for post processing and effort data collection. All users were instructed to have a conversation with the calling partner on a topic of their choice and continue the call for five to ten minutes. However, they were given the liberty to prematurely terminate (abandon or quit) the call if they found that the call quality was not good enough to continue. Whether a call was terminated naturally or prematurely, each user was instructed to provide a rating or opinion score (OS) of the call on a scale between 1 and 5 (where 5 is the best [28]) after the call was terminated. Users were told that their calls would be recorded for research purposes without mentioning anything specific about collecting ‘effort’ data in their conversations.
For each VoIP session, three types of effort data were collected, the total number of times the user exerted human effort, the timing of these efforts within the session, and whether the session ended naturally or prematurely. All effort data were collected by listening and manually post-processing each of the 58 recorded audio streams. In this post-processing, the instance of a user effort was identified by the speaker’s attempt to recover lost speech using keywords or phrases like “sorry?”, “hello?”, “Can you repeat it?” and so on, with respect to the context in which those words were spoken (for example, the word “hello?” could be used as part of natural conversation to greet the person on the other end of the call, as well as when the user is attempting to recover lost speech). The identification of a session as a naturally ending call (NEC) or prematurely ending call (PEC) was determined by scrutinizing the terminating words or phrases of the respective user.

B. QoE-Effort Correlation Analysis

We represent the total effort (TE) of a VoIP session as the total number of efforts detected during the post processing of the recorded speech of that session. Since each user provided an OS for every session, we obtained a total of 58 pairs of (OS, TE) data. Figure 2 shows a scatter plot of these 58 pairs, where the OS is plotted along the y-axis and the TE along the x-axis. Although sometimes there are different OS for the same TE, which is quite normal for any subjective experiments, in general a negative correlation between OS and TE can be visually detected. The Pearson correlation coefficient (PCC) [29] of the 58 pairs was found to be $-0.76$. Since PCC yields a number between $-1.0$ and $+1.0$, with $-1.0$ being perfectly negatively correlated, we can say that the OS is strongly negatively correlated with TE for the data set considered. If we accept OS and TE as measures of QoE and effort, respectively, this outcome verifies that QoE is indeed negatively correlated with human effort. This means that a user spending a lot of effort would have a poor QoE and vice versa.

It would be interesting to do curve fitting and obtain a more precise picture of the relationship between QoE and effort (is the dependency logarithmic or exponential, for example), but the current data set is too small to obtain a good fit. We leave for future work the large scale experiments and the investigation of a function that precisely models the relationship between QoE and effort. Instead, in the following section, we focus on developing a game theoretic model for the user-provider interaction, which only assumes a negative correlation between QoE and effort, but do not require the precise nature of the relationship.

C. Effort Distribution

We now know that OS is negatively correlated to number of efforts. This is a useful information for managing QoE, but it would be also useful to know how many efforts a user would tolerate before she decides to quit. However, since different users may have different tolerance levels to human effort, it makes sense to consider the distribution of efforts before quitting a call for a given user base. This would require systematic experiments and effort monitoring (see more discussions on this in Section IV-D), but as an illustration, let us have a look at this distribution using the data from our small-scale experiment. In our experiment, we found that out of 58 sessions, 31 ended prematurely. Figure 3 shows the distribution of efforts (we did not record any PEC with efforts 1, 13-15, and 17-20) for these 31 sessions. It is clear that the effort distribution in prematurely ending calls has a large spread. While some calls were abandoned only after experiencing 2 efforts, others lasted for 21 efforts before the user decided to quit. While we did not have a large amount of data to obtain precise frequencies or probabilities for these effort numbers, it appears that some effort numbers are clearly more likely than others. For example, there were 6 sessions abandoned after 7 efforts, but only 3 ended after 2, 8, or 9 efforts. In Section IV-E, we will discuss how such effort distributions may influence some properties of a CQP function.

IV. GAME THEORETIC ANALYSIS

In this section we formulate the game between the service provider and a user experiencing a deteriorating QoE. The
A. Game Formulation

We define the interactions between the provider and a user as a non-cooperative, two-player, non-zero-sum game. The game is non-cooperative because the VoIP user and the provider make their decisions independently without any agreed or even implicit cooperation. It is a non-zero-sum game because one player’s loss is not offset by the other player’s gain. In fact, the two players use different value scales - the user her QoE utilities, while the provider his revenue gains or losses. These are nontransferable utilities, thus the zero-sum concept is not applicable. Let us now formalize the game formulation. The strategic form (also called game table or game matrix) of the game is shown in Table I, which contains three types of entities, players, strategies, and payoffs.

- **Players.** The two players of this game are the wireless VoIP user or customer (say C), and the service provider (say S).
- **Strategies.** The user chooses from two possible options \( \{C_1: \text{quit}, C_2: \text{stay}\} \). The quit decision means that the user will quit unless the provider fixes the link quality problem soon, whereas stay means that the user has decided to stay no matter what decision is taken by the provider. The provider selects from two strategic choices \( \{S_1: \text{fix}, S_2: \text{ignore}\} \). The game is a simultaneous move game [30], where the provider and the user make their decisions (or select their strategies) simultaneously as soon as a QoE event occurs without being able to know what move is being made by the other player.
- **Payoffs.** The payoffs in the game formulation are shown in Table II. The provider gains revenue \( R \) when a user makes a VoIP call, but loses revenue \( r < R \) if the user quits prematurely. The provider incurs a cost \( L \) if he decides to allocate additional resource (note that, if the provider chooses to fix the problem, i.e., improve link quality, the user would stay even if she selected the quit option earlier). We will assume that the user QoE will decrease as she puts more efforts in the call. She may not perceive any value in continuing the call further if the required effort exceeds her tolerance threshold \( \sigma \), at which point she would decide to quit because her utility would be worse if she stayed back. The value of \( \sigma \) may be (i) her private knowledge, in which case the provider knows only the probability distribution of its values, or (ii) it might be public information, in which case the provider knows it exactly for every user. We model such user QoE by a general cost function \( f(x, \sigma) \). We expect \( f(x, \sigma) \) to be monotonically decreasing in \( x \) for any given value of \( \sigma \), but increasing in \( \sigma \) if \( x \) is kept constant — for example, a more patient user is prepared to endure longer. Since by quitting the call the user is unable to finish her conversation, we can assume that she gets zero utility when she quits. Then, as will be shown later, the equilibria of this game depend only on the sign of \( f(x, \sigma) \) instead of its absolute value. In our conference paper [23], we used \( f(x, \sigma) = -\ln \left( \frac{x}{\sigma} \right) \), although other functions that monotonically decrease in the first argument \( x \) and increase in the second \( \sigma \), can also be used.

B. Equilibria for Complete Information Game

Let us first assume that the provider has complete knowledge of the user’s payoff parameters including the value of \( \sigma \). For this scenario, we find that the game equilibria depend on the sign of \( f(x, \sigma) \) as well as on the relative magnitude of the provider’s costs, \( L \) and \( r \). Here we consider two cases:

**Case 1.** \( L > r \), which means that the system is overloaded and the cost of resources needed to improve QoE is greater than the cost of losing the user.

**Case 2.** \( L < r \), which means that the system is not overloaded and the cost of resources needed to improve QoE is less than the cost of losing the user.

In both cases, we look at two scenarios: \( x < \sigma \) which represents an effort level below the user tolerance threshold, and \( x > \sigma \) which represents a frustrated user situation.

We derive the following two theorems giving the Nash equilibrium strategies of the provider.

**Theorem 1.** When the system is overloaded (\( L > r \)), the Nash strategy for the provider is to ignore all irritation alarms irrespective of the value of \( x \).

**Proof:** When the system is overloaded (\( L > r \)), we have \( S_2 > S_1 \), which means the ‘ignore’ strategy dominates the ‘fix’ strategy. Therefore, the provider ignores all irritation alarms irrespective of the value of \( x \). In turn, the user considers the ‘quit’ strategy over ‘stay’, i.e., \( C_1 > C_2 \) if \( x > \sigma \), and \( C_2 > C_1 \) otherwise.

**Theorem 2.** When the system is not overloaded (\( L < r \)), the Nash strategy for the provider is \( S_2 \) for \( x < \sigma \), and \( S_1 \) for \( x > \sigma \).

**Proof:** When the system is not overloaded (\( L < r \)), let us look first at the user behavior. She chooses strategy \( C_1 \) (‘quit’) or \( C_2 \) (‘stay’) as before - for her the columns \( S_1 \) and \( S_2 \) provide matching choices. If \( x < \sigma \), then \( C_2 \) is dominant; then, by iterated domination, the provider chooses strategy \( S_2 \): ignore. If \( x > \sigma \), the choice \( C_1 \): Quit dominates and, by iterated domination, the provider chooses strategy \( S_1 \): fix.

These Nash strategies give us an understanding of how a typical user and a rational provider would behave in equilibrium. These are called equilibrium solutions [30] because none of the players would gain by changing its strategy, subject to the other player’s current decision.

---

\(^1\)A strategy is a dominant strategy when it is the best choice for a player regardless of what the others might be doing.
C. Equilibria for Probabilistic Knowledge

The basic model of Nash equilibrium discussed in the preceding section assumes that both players are fully informed about all the payoffs - their own as well as their counterparts - under all possible choices of strategies by both players. In particular, the Nash equilibrium derived in Theorem 2 would depend on the provider’s knowledge of the value of \( \sigma \) for each user and each call made by the user. In practical systems, however, such knowledge is probabilistic - the provider may only know the probability, \( p \), that the user will quit when an effort is detected. This knowledge would be based on the experimental study of a (sufficiently large) group of users; we discuss later how to compute this probability from the call termination data available at the provider side.

A generalization of the basic model, known as (static) Bayesian equilibrium [31], [32], permits for such payoffs (to the counterparts) to be known only probabilistically. We discuss later how to compute this probability from the call termination data available at the provider side.

We have collected data from a specially designed experimental study or from a long observation of a representative group of users, for which all efforts were monitored, but no attempt was made to improve QoE. Let there be \( k_1 \) users who quit after the first effort, \( k_2 \) users who quit after two efforts, and so on, where \( n \) is the maximum number of efforts that any user would tolerate. Thus a population of \( \sum_{i=1}^{n} k_i \) users is monitored. Let us assume that the provider wants to monitor the situation in the system at a particular QoE event point. For obvious reasons, the defining events are the effort moments of a particular user. There are two distinct scenarios to consider.

One is when we ‘jump in’ and observe an effort randomly selected without having monitored that user from the beginning of her call. The other scenario arises when we monitor an entire call from its inception, observe an effort number \( x \) and wish to estimate the likelihood that the user will terminate after this effort.

Monitoring Random Effort

There are two natural ways to proceed. The provider can simply monitor the entire system for the efforts made by users, without paying attention to the specific user who made a particular effort. Thus we select a random effort, uniformly from all the efforts in the system. This results in the following theorem.

**Theorem 4.** If the provider monitors randomly selected efforts in the entire system, the probability of user quitting after that effort is given by

\[
p = \frac{\sum_{i=1}^{n} k_i}{\sum_{i=1}^{n} ik_i}
\]

*Proof:* In the data described above, \( k_i \) users made \( i \) efforts each (before quitting) for the total of \( \sum_{i=1}^{n} ik_i \) efforts, of which \( \sum_{i=1}^{n} k_i \) were terminating efforts - one for each user. Therefore, the probability that a randomly observed effort will be the last one for this user is derived as

\[
p = \frac{\sum_{i=1}^{n} k_i}{\sum_{i=1}^{n} ik_i}
\]

Another and perhaps more practical method is to select a user at random, uniformly from all the active users in the system, and then randomly select one of her efforts. For this random monitoring method, we have

**Theorem 5.** If the provider monitors randomly selected users and chooses to manage a random effort of that user, the probability of quitting after that effort is given by

\[
p = \frac{\sum_{i=1}^{n} \frac{1}{i} k_i}{\sum_{i=1}^{n} k_i}
\]

*Proof:* A user of type \( i \) that quits after \( i \) efforts will be selected with probability \( \frac{k_i}{\sum_{i=1}^{n} k_i} \). The probability that a specific effort of this user is her terminating one is given by \( \frac{1}{i} \) as she makes \( i \) efforts in total. Therefore, the probability that an effort selected through this two-stage process is a terminating one is obtained as

\[
\sum_{i=1}^{n} \frac{1}{i} \times \frac{k_i}{k_1 + \cdots + k_n} = \frac{\sum_{i=1}^{n} k_i/i}{\sum_{i=1}^{n} k_i}.
\]

---

**TABLE II**

**PAYOFF NOTATIONS**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R )</td>
<td>Revenue earning if user ends the call naturally</td>
</tr>
<tr>
<td>( r )</td>
<td>Revenue loss if user quits prematurely (( r \leq R ))</td>
</tr>
<tr>
<td>( L )</td>
<td>Revenue loss for fixing a problem with extra resource</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>User utility function</td>
</tr>
<tr>
<td>( f(x, \sigma) )</td>
<td>User tolerance threshold in number of efforts</td>
</tr>
</tbody>
</table>

---

### PAYOFF NOTATIONS

- \( R \): Revenue earning if user ends the call naturally
- \( r \): Revenue loss if user quits prematurely (\( r \leq R \))
- \( L \): Revenue loss for fixing a problem with extra resource
- \( \sigma \): User utility function
- \( f(x, \sigma) \): User tolerance threshold in number of efforts
Monitoring Entire Call

For various reasons the provider may monitor the entire call of a specific user and want to evaluate the likelihood of the x'th effort being a terminating effort. This is a conditional probability, where the conditioning is done on the fact that this user is willing to make at least x efforts before quitting. Thus, we have

Theorem 6. If the provider monitors the entire call of a user, then the accumulated effort number x is a terminating one with probability

\[ p(x) = \frac{f(k_x)}{f(k_x) + 1 - F(k_x)} \]

where \( f(k_x) \) and \( F(k_x) \) are the probability distribution function (PDF) and cumulative distribution function (CDF), respectively, of \( k_x \).

Proof: Only the users who could tolerate x or more efforts would attempt to make x efforts, while the others would quit earlier. There are a total of \( \sum_k^x k_i \) users with tolerance threshold x or higher. Of those users, \( k_x \) will quit after this x'th effort, thus yielding

\[ p(x) = \frac{x}{\sum_k^x k_i} = \frac{k_x}{k_x + \sum_k^x k_i - \sum_k^x k_i} \]

Dividing both numerator and denominator by \( \sum_k^x k_i \), we have

\[ p(x) = \frac{x}{\sum_k^x k_i + 1 - \sum_k^x k_i} = \frac{f(k_x)}{f(k_x) + 1 - F(k_x)} \]

From Theorem 6, we can see that the properties of \( p(x) \), e.g., whether it is an increasing or decreasing function of the effort number x in a given call, would depend on the distribution of \( k_x \). We discuss these issues and their implications in the following subsection.

E. Sophisticated Users

We look for an opportunity for a strategic behaviour on the user’s part. The user may be familiar with the fact that the calls are monitored, perhaps through the announcement by the provider. It could be to her advantage to create the impression that she has made more genuine efforts than the actual case. In other words, she may have made x genuine efforts but would like it to appear as being, say, her (x + 1)’st effort. To accomplish it, she may make spurious efforts, not actually warranted by the QoE. We can evaluate whether such a strategy would benefit her.

For this strategy to be advantageous for the user, it is required that

\[ p(x) < p(x + 1) < p(x + 2) < \cdots < p(n), \]

i.e., \( p(x) \) is an increasing function of \( x \). It can be proved (see Appendix) that a simple condition on concavity of sequence \( k_x \) implies this property. In particular, this property holds for a sequence \( k_x \) that can be approximated by normal distribution (see proof in the Appendix).

The quitting probability \( p(x) = \frac{k_x}{\sum_k^x k_i} \) computed from the experimental \( k_x \) data is plotted in Figure 4. From a choice of several typical regression types, including linear, exponential, power, logarithmic, and polynomial, we found that a second degree polynomial function, \( y = ax^2 + bx + c \), best fits the data (here \( R^2 = 0.907 \) denotes the value of the coefficient of determination, which is a statistical measure of goodness of fit, and should approach 1 for a perfect match). With positive values for all three coefficients, \( a, b, \) and \( c \), the quadratic function would yield a quitting probability function that is increasing in the cumulative effort \( x \). Now let us visit the game equilibrium for the probabilistic knowledge, which says that the provider would attempt to fix the QoE problem only if \( p > \frac{1}{R} \) (Theorem 3). This means that the provider would ignore \( \omega \) efforts of a user before attempting to fix the problem, where \( p(x = \omega) = \frac{1}{R} \). The value of \( \omega \) for a given \( \frac{1}{R} \) can be directly obtained from the quadratic \( p(x) \) curve in Figure 4. For example, for \( \frac{1}{R} = 0.1 \), we have \( \omega = 4 \), meaning the provider would always ignore the first 4 efforts of any user and would only attempt to fix the problem at the fifth effort. Clearly, in this case, the sophisticated user would gain from faking efforts.

However, the very same user may disadvantage herself if \( p(x) \) is modelled as a decreasing function of \( x \). For example, it can be proved (see Theorem 9 in Appendix) that if \( k_x \) is Cauchy distributed, which is also ‘bell shaped’ like the normal distribution, then \( p(x) \) is a decreasing function for \( x \geq 1 \). In such cases, the provider may well decide that the conditional probability of call terminating is actually decreasing with the number of efforts. Thus the longer-lasting calls could be given a less favourable treatment.

Next, we simulate a VoIP provider with limited wireless resources and compare the CQP, CBP, and the provider revenue for different schemes of QoE management.

V. Simulation Study

Broadly speaking, the provider has three different options or schemes to choose from when it comes to real-time management of user QoE within an active VoIP call. For example, the provider may decide not to implement any QoE monitoring at
all and hence ignore all efforts by the user. We will refer to this default option as the ‘baseline’ strategy. On the other hand, if the provider chooses to implement QoE monitoring, he can either respond to each user effort detected by the system, or he can decide to ignore the first few efforts before attending to them. The latter is the equilibrium strategy derived from the game formulation in Section IV. In this section, we compare these three schemes via simulation study.

A. Simulation Model

We have simulated a wireless network provider with a limited amount of available resources (wireless bandwidth) to support VoIP calls. Since the focus of the paper is on the game theoretic modeling, the results can be generalized to different wireless technologies and even other systems that require resource management to improve service quality. As such, in our game models, we have not strongly linked wireless technology aspects with user experience. For the same reasons, we have not simulated the lower layers, such as PHY and MAC, of the wireless network, but assumed that resources can be dynamically allocated and deallocated to a VoIP call. New call requests arrive to the provider following a Poisson process with an arrival rate of \( \lambda \) calls per second. Each arriving call is allocated a minimum amount of resource, say 1 unit of bandwidth, if there is 1 or more units of bandwidth available at the time, otherwise the call request is rejected. If the network provider decides to allocate extra resource to an ongoing call when an effort within the call is detected, he allocates a 0.8 (say) additional units of bandwidth provided there is 0.8 or more units of bandwidth available at that time, otherwise the provider ignores the effort.

The properties of all accepted calls in terms of their durations, effort arrival timings within the calls, and whether they are supposed to end prematurely or not, are guided by the traces collected from our experimental test-bed as described in Section III. The game is simulated (played) each time an effort instance arrives within a call. If the call selected from the pool of 58 traces was a naturally ending call (NEC), it always ends naturally (user never quits). For a prematurely ending call (PEC), the user’s decision to quit the call if the provider does not allocate additional resource is completely guided by the traces. If the provider does allocate additional resource, the call is converted to a NEC, implying the user does not quit. Therefore, some of the PEC calls may eventually be NEC in our simulation depending on whether the call received additional resource from the provider before the user quits. These functionalities of the simulation model was implemented by writing a C program. Table III shows the values of all simulation parameters used in our experiments.

In the simulation, the provider ignores the first \( \omega \) efforts and attempts to allocate additional resource at effort number \( \omega + 1 \). This enables us to simulate all three different schemes by simply manipulating the value of \( \omega \). Two extreme values of \( \omega \) represent two extreme schemes for the provider; \( \omega = 0 \) represents the case when the provider never ignores any effort, while \( \omega = 21 \) is the default or ‘baseline’ case when the provider does not implement any QoE monitoring system, i.e., all efforts are always ignored. The values of \( \omega \) between these two extremes represent the game equilibrium discussed in the preceding section. We run a separate simulation for each values of \( \omega \) between 1-21. To ensure that simulation results have a high statistical confidence, we run each simulation for a long time until each of the performance metrics reaches steady state, where ‘steady state’ is determined when the last 100 observations have a 95% confidence level with a relative precision less than or equal to 5%.

B. Performance Metrics

For each simulation experiment, we measure the following three metrics:

- **Call Quitting Probability (CQP):** It represents the fraction of total admitted calls that end prematurely due to user quitting the call (because the provider did not attend to her efforts soon enough). From our experimental trace, this fraction is 53% (recall that out of 58 total user sessions, 31 sessions ended prematurely). Therefore, for the ‘baseline’ strategy, we also expect it to be 53%, as all our simulations are driven by these traces. However, this should decrease if the provider implements effort-based QoE management. Any reduction in CQP would have a positive effect on the provider revenue.

- **Call Blocking Probability (CBP):** This defines the fraction of total call requests that are not admitted due to lack of available resources. Since the total available resource in the system is limited, allocating any additional resource to a user will increase the CBP.

- **Provider’s Revenue:** It is calculated as \((mR - zr)\) normalized to the entire simulation time, where \(m\) is the total number of calls accepted in the system and \(z\) is the total number of calls terminated prematurely. Therefore, the revenue is affected by both CQP and CBP, but in opposite directions.

C. Simulation Results

Figure 5 plots the CBP and CQP as a function of \( \omega \). As expected intuitively, for decreasing \( \omega \), the CQP decreases while the CBP increases. This is because, with smaller \( \omega \), the provider allocates additional resource much earlier in the call, thus preventing many calls from ending prematurely, but inevitably increasing the probability of a new call not finding enough resource in the system to be admitted. What is interesting (and not so intuitive) is that the CQP decreases at a much faster rate than the increase of CBP. This non-linear relationship between CQP and CBP is shown more clearly in Figure 5(b). For example, if the provider selected a threshold
of 4 efforts, the Nash strategy of resource allocation would reduce the CQP from 53% to 14%. This dramatic reduction in CQP would increase the CBP from 4% to 7.6%.

Next, we examine how the provider revenue compares across the three schemes. Figure 6, which plots revenue as a function of effort threshold ($\omega$), clearly shows that the provider would enjoy a small-scale wireless VoIP test-bed involving real subjects.

Using game theoretic analysis, we have shown that in equilibrium, the provider may ignore some of the efforts before attending to efforts after ignoring the first few. The fact that there is an optimum $\omega$ for which the revenue is maximized can be explained as follows. The revenue drops if the provider never takes an action to repair the link, because many users quit prematurely (loss of $r$ revenue per user). The revenue also drops if the provider always takes action as soon as the first QoE alarm is raised, because additional resources are allocated to a call too soon, which immediately increases the probability of blocking a new user (loss of $R$ revenue).

Therefore, from the viewpoint of revenue, the optimal policy for the provider appears to be the one where the provider does respond to QoE events (efforts in this case), but not attend to the very first event. This is also the strategy obtained from the game equilibrium.

It is interesting to note that, managing user QoE by allocating additional resource to certain efforts is actually more profitable for the provider than ignoring it altogether. The magnitude of the economic reward is, of course, directly dependent on the relative values of $r$ and $R$. Nevertheless, as shown in Table IV, even with a very small value of $r$, and an additional resource allocation ($\delta$) as large as 80% of the basic allocation, the provider can still expect an increase in its net revenue as compared to the baseline case where all efforts are ignored. For example, for $r = 4$, which is only 40% of $R = 10$, a 10% increase in revenue is possible by implementing a QoE management scheme.

Figure 6 also shows the CQP as a function of $\omega$. Of course, CQP could be totally eliminated by responding to every single effort, but from the revenue point of view, that would not be very satisfactory as it would significantly reduce new customers in the system. However, with the scheme of attending to efforts after ignoring the first few, CQP can be reduced dramatically from the ‘baseline’ value of 53%.

VI. CONCLUSION AND FUTURE WORK

In this paper, we have proposed a new model for quality of experience (QoE) for mobile VoIP services, based on the amount of effort a user has to put to continue her conversation. The dependency of the QoE, in terms of opinion score, on the amount of effort a user has to put to continue her conversation. The dependency of the QoE, in terms of opinion score, on the amount of effort a user has to put to continue her conversation. The dependency of the QoE, in terms of opinion score, on the amount of effort a user has to put to continue her conversation. The dependency of the QoE, in terms of opinion score, on the amount of effort a user has to put to continue her conversation. The dependency of the QoE, in terms of opinion score, on the amount of effort a user has to put to continue her conversation.

The dependency of the QoE, in terms of opinion score, on the amount of effort a user has to put to continue her conversation. The dependency of the QoE, in terms of opinion score, on the amount of effort a user has to put to continue her conversation. The dependency of the QoE, in terms of opinion score, on the amount of effort a user has to put to continue her conversation.
receiving a better service from the provider. We have shown also that the user would benefit from faking efforts (which is kind of threat) only if the provider uses the effort distribution from all users to predict the probability of a user quitting after a given effort and that the distribution is logarithmically concave giving a quitting probability that increases with the number of efforts.

The current work provides some good insights to the issue of human effort in wireless VoIP, but it can be extended in several ways for further investigation. We only conducted a small-scale user experiment to verify the existence of (negative) correlation between QoE and human effort. Although we observed a strong negative correlation, we did not have enough data to fit a function with good accuracy. It would be an interesting future work to carry out larger scale experiments and use the resulting data to obtain a function that more precisely represents the nature of relationship between QoE, say in terms of MOS, and the amount of effort spent by the user. Inclusion of additional metrics at signal and packet layer could also be beneficial for fine tuning the function.

In this work, we studied interactions between a provider and a single user. An interesting future direction would be to address the scenarios of multiple users interacting with a single or multiple providers. The situation of multiple users ‘facing’ a single provider is fundamental. Its analysis should provide several deeper insights into game-theoretic aspects of social interaction among the users. In particular, our analysis of conditional quitting probability can be continued along the lines of the ‘Tragedy of the Commons’ [33]. It addresses the phenomenon when it is individually optimal to exploit more resources, even though it leads to the overexploitation and eventual collapse of the system. Such patterns have been noted in ecological systems, overgrazing of pastures, or overfishing the seas. Modelling it for a large number of selfish VoIP users, however, can be quite complex; we leave it for future investigations.

We used simulation experiments to demonstrate that best revenue is obtained when some efforts are ignored. To show this analytically, we need to derive expressions for CBP and CQP. Note that CBP in traditional networks, where providers do not allocate additional resources in the middle of a voice call, can be derived easily by formulating the problem as a Markov Chain. However, the construction of the Markov Chain becomes non-trivial with additional resource allocation. On the other hand, although the data from our small-scale subjective experiment suggest that the CQP can be approximated by a quadratic function of the accumulated effort, larger scale experiments are required to verify this. We leave derivation of accurate expressions for CBP and CQP as part of our on-going work.

ACKNOWLEDGEMENTS

This research was funded by the Australian Research Council Discovery Project DP0881553. We thank the graduate students who took part in the subjective experiments. Finally, we are greatful to the anonymous reviewers whose comments have helped us significantly improving the quality of the final version of this paper.

APPENDIX

In this appendix, we derive the conditions for call quitting probability $p(x)$ as an increasing function of the accumulated effort number $x$. We first consider the discrete data case, followed by the continuous case.

The Case of Discrete data

Conditional probabilities of user quitting are an increasing sequence means

$$p(x+1) = \frac{k_{x+1}}{k_{x+1} + k_{x+2} + \cdots + k_n} > \frac{k_x}{k_x + k_{x+1} + \cdots + k_n}$$

$$p(x) = \frac{k_x}{k_x + k_{x+1} + \cdots + k_n} \text{ for } x \geq 1$$

Here $p(x)$ represents the probability of quitting after exactly $x$ efforts given that the user have endured until this effort. In other words, $p(x)$ is the ratio of the number of users that quit exactly after the $x$'th effort to the number of the users who last for $x, x + 1, \ldots$ efforts.

There is a very simple and practical sufficient condition for these inequalities

$$\frac{k_{x+1}}{k_{x+2}} > \frac{k_x}{k_{x+1}}$$

It is better visualized using the logarithms of the numbers of users of given types $l_i = \ln k_i, \text{ for } i = 1, \ldots, n$. Then

$$l_{x+1} - l_{x+2} > l_x - l_{x+1} \text{ or } \frac{l_x + l_{x+2}}{2} < l_{x+1}.$$ 

This means that

Theorem 7. $p(x)$ is an increasing function of $x$ if $k_x$ is logarithmically concave.

Proof: First let us note that the concavity for three consecutive terms $k_x, k_{x+1}, k_{x+2}$ implies

$$l_x - l_{x+z} < l_y - l_{y+z}, \quad y > x.$$ 

This is because

$$l_x - l_{x+z} = \sum_{i=x}^{x+z-1} l_i - l_{i+1}, \quad l_y - l_{y+z} = \sum_{i=x}^{y+z-1} l_i - l_{i+1}$$

and every term in the first sum is less than the corresponding term in the second sum. With change of subscripts, we get

$$l_x + l_y < l_{x+z} + l_{y-z}, \quad x < y, \quad z \leq y - x.$$ 

Reverting to $k_x$ we have

$$k_x k_y < k_{x+z} k_{y-z}, \quad x < y, \quad z \leq y - x.$$ 

Now we can tackle the fractions representing the consecutive conditional probabilities $p(x) < p(x+1)$. Multiplying out their denominators gives an equivalent inequality

$$k_{x+1} k_x + k_{x+1}^2 + \cdots + k_{x+1} k_n <$$

$$k_x k_{x+1} + k_x k_{x+2} + \cdots + k_x k_{n-1} + k_x k_n$$

and as the first term on either side is identical, we need to verify that

$$k_{x+1}^2 + \cdots + k_{x+1} k_n < k_x k_{x+2} + \cdots + k_x k_{n-1} + k_x k_n.$$
We note that every term on the left is less than the corresponding term on the right which follows from the concavity. Furthermore, there is one more term on the right, namely the last term $k_x k_n$, which is not matched.

The Case of Continuous data

In practice, the reasoning about possible strategies would use smoothed data. Most likely one would fit a continuous probability distribution $f(x)$, $1 \leq x < \infty$ to the raw data. If $f(x)$ is differentiable (always the case in practice), there is a particularly simple test of the required concavity: $(\ln f(x))'' < 0$.

Theorem 8. Normal distribution is logarithmically concave.

Proof: In its normalized form, the distribution has frequency $\propto e^{-x^2}$ whose logarithm is $-x^2$, which is negative.

Although we proved for the standard normal distribution, the same analysis applies to the general form of normal function. Change of mean and/or standard deviation simply modifies its graph linearly - it is then shifted and stretched or compressed horizontally. Concavity is a geometric property invariant with respect to such changes.

Now that we have proved that normal distribution is logarithmically concave, $p(x)$ should be an increasing function of $x$ if $k_x$ is normal distributed (Theorem 7). For (standard) normal distributed $k_x$, we have a conditional quitting density $\frac{f(k_x)}{1-F(k_x)}$. This represents the failure rate (of a call) as used in reliability theory [34]. Our objective is to verify that this density is monotonically increasing. This is equivalent to verify that

$$p(x) = \frac{f(k_x)}{f(k_x) + 1 - F(k_x)} = \frac{e^{-x^2/2}}{\sqrt{2\pi}} + \frac{1}{2} \text{erfc}(\frac{x}{\sqrt{2}})$$

is monotonically increasing. There is an intuitive advantage in referring to function $p(x)$. It is always bound between 0 and 1 as visualized in Figure 7, which also clearly shows that $p(x)$ is an increasing function of $x$.

Theorem 9. Cauchy distribution is not logarithmically concave.

Proof:

$$g(x) = \frac{1}{\pi(1 + x^2)}$$

$$(\ln g(x))'' = -2 \frac{(1 + x^2) - x \cdot 2x}{(1 + x^2)^2}$$

The latter simplifies to $2 \left( \frac{x^2 - 1}{x^2 + 1} \right)$ positive for $x > 1$.

Similar to the Normal distribution, the use of a general form of Cauchy distribution changes the shape of graph of the function in a linear manner; therefore, it does not affect the conclusion.

Now that we have proved that Cauchy distribution is not logarithmically concave, $p(x)$ should be a decreasing function of $x$ if $k_x$ is Cauchy distributed. For (standard) Cauchy distributed $k_x$, we have

$$p(x) = \frac{f(k_x)}{f(k_x) + 1 - F(k_x)} = \frac{1}{\pi(1 + x^2)} + \frac{1}{2} - \frac{\arctan(x)}{\pi}$$

The plot of the above equation is shown in Figure 8. It is clearly seen that $p(x)$ is a decreasing function of $x$ for $x \geq 1$.

REFERENCES


Jahan A. Hassan is an adjunct academic member and a former research fellow at the School of Computer Science and Engineering, University of New South Wales, Sydney, Australia. Prior to joining the University of New South Wales, she worked as a research fellow at the School of Information Technologies, University of Sydney, Australia. She received her Ph.D. in 2004 from the University of New South Wales, and her Bachelors degree in 1995 from Monash University, Melbourne, both in computer science. Her Ph.D. thesis was published as a book by VDM Verlag, Germany. She served as Guest Editor for IEEE Communications Magazine and IEEE Network. She was a member of the Technical Program Committee of IEEE LCN 2006, IEEE ICC 2007, IWQOS 2008, IEEE WCNC 2010, and IEEE GLOBECOM 2010.

Mahbub Hassan is a full Professor in the School of Computer Science and Engineering, University of New South Wales, Sydney, Australia. He earned his PhD in Computer Science from Monash University, Melbourne, Australia (1997), MSc in Computer Science from University of Victoria, Canada (1991), and BSc in Computer Engineering (with High Honor) from Middle East Technical University, Turkey (1989). He has co-authored three books, one US patent, and more than 150 technical papers. He serves in the Editorial Advisory Board of Computer Communications (Elsevier Science) and previously served as Guest Editor for IEEE Communications Magazine and IEEE Network. Professor Hassan has written several successful Australian Research Council (ARC) grants, worked on collaborative R&D projects with large industrial research laboratories, and developed industry short courses on leading edge networking topics. He held visiting appointments at Osaka University, Japan (2011), University of Nantes, France (2005) and the National ICT Australia (2006 and 2012).

Sajal K. Das is a University Distinguished Scholar Professor of Computer Science and Engineering at the University of Texas at Arlington. During 2008-2011 he was a Program Director at the National Science Foundation in Computer Networks and Systems. He has published over 500 papers and 47 book chapters, and holds five US patents. He co-authored three books: “Smart Environments: Technology, Protocols, and Applications” (2005), “Handbook on Securing Cyber-Physical Critical Infrastructure: Foundations and Challenges” (2012), and “Mobile Agents in Distributed Computing and Networking” (2012). He is a recipient of seven best paper awards, IEEE Computer Society Technical Achievement Award (2009) and IEEE Region 5 Outstanding Engineering Educator Award. At UTA, he is also a recipient of Graduate Dean’s Award for Excellence for Doctoral Students Mentoring, Lockheed Martin Teaching Excellence Award, Academy of Distinguished Scholars Award, University Award for Distinguished Record of Research, and College of Engineering Research Excellence Award. A founder of IEEE WoWMoM and PerCom conferences, he serves as the founding Editor-in-Chief of Elsevier’s Pervasive and Mobile Computing (PMC) journal.

Arthur Ramer is Associate Professor in computer science at the University of New South Wales, which he joined in 1992. He worked as mathematician in 1968–72, life and health actuary in 1973–81, and computer scientist since 1982. He holds PhD in computer science from SUNY-Binghamton in 1984, following his earlier PhD work in mathematics at the University of Chicago in 1970–72. He has published over 60 journal articles and over 120 international conference papers. He received nine ARC grants, graduated four PhD students, and is currently supervising a PhD in quantum games. His current work is in computational decision methods and in quantum computing. He has significant expertise in fuzzy information theory. He had consulted to a wide range of industries, including insurance, gaming companies, and IBM.