THE ROLE OF PSYCHOPHYSICS TO VOIP RATE ADAPTATION: A STUDY ON SKYPE CALLS

Anonymous ICME submission

Paper ID: 663

ABSTRACT

Effective end-to-end transport of delay-sensitive voice data has long been a problem in multimedia networking. One of the major issues is determining the sending rate of the real-time VoIP streams such that the user experience per unit network resource consumed is maximized. A complication we find particularly interesting and yet to be addressed is that available bandwidth is often dynamic. It is not clear whether a marginal increase warrants better user experience. Furthermore, if one naively tunes the sending rate to the optimum at any given opportunities, the user experience could fluctuate.

To understand the effect of magnitude and frequency of rate changes to user experience, we recruit 100+ human subjects to score emulated Skype calls with different combinations of rate changes — systematically varying magnitude and frequency of rate changes. We are able to (1) verify magnitude and frequency of rate changes affect the user experience at the logarithmic scale, echoing the Weber-Fechner’s Law [1], and (2) derive a closed-form model of user perception to rate changes with 97%+ goodness of fit for Skype calls.

Index Terms—VoIP, Rate Adaptation, User Perception, QoE, Psychophysics

1. INTRODUCTION

Effective end-to-end transport of delay-sensitive voice data has long been a subject of study in multimedia networking. In recent years, a number of pioneer works have proposed to approach this issue from a user-centric view, i.e., adapting the sending rate of voice calls based on user satisfaction [2][3]. In that, the rate adaptation mechanisms ramp up the sending rate quickly when the available bandwidth is sufficient, and carefully tune up or down the sending rate when the network starts to congest. Note that although users prefer calls with a higher bit rate, sending voice data with an unnecessarily high bit rate could be a waste of network resources or result in congestion. This in turn could compromise the quality of experience.

The user-centric approach is shown promising and relevant. However, much of the prior work has focused on identifying a sending rate for optimal user experience. An issue that has largely been overlooked is: how do users perceive rate changes. Available bandwidth of an end-to-end connection is often dynamic. A (naive) increase of sending rate might not result in better user experience in that: (1) the change might not be detectable by the user and (2) the change might be disturbing, shall the available bandwidth fluctuates. Yet, increasing sending rate costs the system more network resource.

What we are advocating uniquely is that designing the rate adaptation mechanism is a significantly more subtle task. It is not just to determine the optimal sending rate, but to determine, for what it is worth, the optimal magnitude and frequency of changes in the sending rate.

How human beings perceive the quality of voice, images, or motion pictures has long been a subject of study in Psychology. Weber and Fechner have proposed back in 1834 [1] that the human ‘notice-ability’ of a change is relative to the current experience, i.e., the degree of notice-ability tends to the log of the stimulus’s intensity. To put the subject of our interest, VoIP calls, in context, we ask: (1) whether the user experience is logarithmic to the sending rate, and extendedly (2) whether user experience is logarithmic to the time interval of rate changes. In short, does the Weber-Fechner’s Law apply to the streaming of VoIP calls?

The objective of the work is two-fold. The first is to address the above-posed questions and the second is to address a fundamental problem towards user-centric rate adaptation mechanisms — modeling the relationship of user perception and magnitude/frequency of rate changes in VoIP calls. Such a relationship is codec and environment sensitive. We select Skype, the most popular VoIP service and recently acquired by Microsoft for 8.5B US dollars, as the first target. By that, the experiments in this work are geared to emulate the specifics of Skype conversations.

Our methodology is as follows. We record a variety of human speeches. These raw audio tracks are encoded at different rates using the SILK codec [4], an open source toolkit made available by Skype. A wide range of test tracks, i.e., with different degree and different frequency of rate changes, are synthesized. 100+ human subjects are invited to score the synthesized tracks. Three data sets, different speech contents and different sets of human subjects, are collected independently. These data are first examined by the ANOVA [5] tests and then used to model and verify the
relationship of user perception and rate changes. We have achieved in (1) verifying the user experience-magnitude of rate change relationship exhibits the log-like behavior, echoing the Weber’s theory, (2) discovering that experience-frequency of rate change relationship also exhibits the log-like behavior, and (3) deriving the closed form model of user experience to rate changes with 97%+ goodness of fit. These findings provide as a foundation of rate adaptation mechanism for voice data. In that, the mechanism can be designed to optimize for user experience in the presence of network dynamics.

The rest of the work is organized as follows. Related work on multimedia networking and psychophysics are detailed in Section 2. Findings from the preliminary experiments in Section 3 allow us to construct plausible models formulating the relationship of user experience and rate changes in Section 4. With large-scale experiments in Section 5, we derive the specific models. The models are validated via the preliminary, the large-scale, and yet a third set of experiments in Section 6.

2. BACKGROUND AND FUNDAMENTALS

2.1. Quality of Service

Traditionally, whether a network application is providing a ‘good enough’ service is measured by throughput, loss rate, delay, and delay jitter. These measurements are referred to as the Quality of Service (QoS). For years, measured QoS stands for the performance of a network application. In particular, the quality of the best-effort data, transported via TCP, is measured mainly by throughput, whereas the quality of real-time streaming media, transported via UDP, is measured by additional metrics such as loss rate, delay and delay jitter.

The rate control mechanisms in early years aim at improving QoS for the data they transport. TCP, for example, employs an additive-increase-multiplicative-decrease (AIMD) policy in the control of send-window size. This allows exploration of maximum available bandwidth and avoids data losses in the presence of network congestion. AVoIP [6], designed to transport voice data, employs a similar AIMD policy controlling the sending rates of VoIP calls.

2.2. Quality of Experience

The rising trend is to measure the network services by the Quality of Experience (QoE). As defined by the International Telecommunication Union (ITU), QoE is “the overall acceptability of an application or service, as perceived subjectively by the end-user” which “includes the complete end-to-end system effects” and “those that may be influenced by user expectations and context” [7]. Metrics of QoE, such as responsiveness and the Mean Opinion Score (MOS), reflect directly user’s experience of the network services.

QoE measurements are, however, difficult to acquire without application-level support. Responsiveness might require data content analysis and MOS requires user feedback. Rate adaptation mechanisms based on such measurements are inherently challenging to implement. Furthermore, the delay of acquiring the measurement might be longer than the granularity of network dynamics, rendering the approach impractical.

2.3. QoS-QoE Synergy

Despite the differences, QoS and QoE are not in competition of each other. Rather, they are complementary pieces of the long-standing puzzle – assessing quality of multimedia services in real time. Recent works have been dedicated to map the objective, network-centric QoS to the direct, user-centric QoE, thus enabling practical implementation of user-centric rate adaptation mechanism. In [8], the authors proposed a formula mapping bit rate, loss, delay, and jitter to user satisfaction of Skype calls. Further in [9] and [10], QoS and QoE of a wide range of network services, including webpage browsing, photo sharing, file downloading, and VoIP, are investigated. Logarithmic relationships are commonly observed in the QoS-to-QoE mappings.

The proposed QoS and QoE mappings so far have provided insights to how users perceive multimedia streams in steady states. They, however, are not quite sufficient deriving the adaptation strategies delivering real-time multimedia streams, that are under the influence of frequent network dynamics. Our work aims at deriving the relationship of rate changes to human perception. This is an essential issue that has not yet been thoroughly investigated before.

2.4. Psychophysics

The Weber-Fechner’s Law [1] in psychophysics provides a plausible explanation to the logarithmic relationships observed between various QoS and QoE metrics. Studies of relationships between stimulus and human perception date back to 1834, when Ernst Heinrich Weber published his insight upon the human sensory system. Quantitatively speaking, Weber found that the ratio of the noticeable threshold of stimulus intensity change to the intensity of original stimulus is a constant, i.e.,

$$\frac{\Delta I}{I} = K$$

(1)

where $\Delta I$ is the amount of intensity difference being just noticeable, I is the original intensity, and the constant K is called the Weber fraction. Take weight lifting as an example. Assume that one starts from carrying a 25 kg object and the carrier does not notice the difference until we increase the weight by 5 kg. The just-noticeable increment will then be 1
kg, if the carrier starts from 5 kg. Here, the K for weight lifting is 1/5.

Shortly after the publishing of Weber’s Laws, Gustav Theodor Fechner stated a mathematic model as an extension known as the Weber-Fechner’s Law. Given a stimulus S and its responding quantitated perception P, the relationship between S and P is as follows.

\[ dP = k \times \frac{dS}{S} \]  

where \( dP \) and \( dS \) are the differences of stimulus and perception, respectively, and \( k \) is a constant scale factor. By integrating both sides, one obtains the following.

\[ P = k \times \ln S + c \]  

where \( c \) is the constant introduced by integration. The Weber-Fechner’s Law has been shown plausible in a wide range of human perceptions including hearing, vision, taste, sense of touch and heat, and even temporal, spatial, and numerical cognitions.

3. PRELIMINARY EXPERIMENTS

The purpose of our preliminary experiment is to examine through user tests whether changes in sending rates indeed worsen the user experience and at what scale the MOS is influenced by the magnitude and frequency of changes. Here, for the sake of clarity, we define magnitude and frequency, the two key aspects, of a rate change as follows. As shown in Fig. 1, magnitude is decided by a pair of bitrates, namely the high rate (hr) and the low rate (lr); and frequency is decided by the time interval (\( \Delta T \)) between two adjacent rate changes.

3.1. Methodology

**Audio Source:** Following recommendations of ITU-T P.830 [7], our source material consists of a number of simple, short, meaningful sentences which do not have obvious contextual connections. Two female and two male speakers are recruited to produce the voice of the audio source. Length of the source material is 30 seconds and the recording quality is 44.1kbps.

**Fixed-Rate Tracks:** We take Skype, the state-of-the-art commercial VoIP application, as our experimental subject. Skype uses SILK [4] as the audio codec for its PC-to-PC service in the latest version. Thus, the audio source is encoded by SILK for all the experiments. Limited by the SILK freeware’s functionality, we choose to encode the audio sources in 10 different bitrates, uniformly from 40.6 to 5.6 kbps.

**Variable-Rate Tracks:** By combining two audio tracks of different bitrates into one with different time intervals, we synthesize test tracks with varying qualities. The high and low bitrates are chosen among 40.6, 28.9, 17.2, and 5.6 kbps. This introduces 6 pairs of high-low rate changes. The time interval between rate changes is chosen from 1, 2, 3, 5, and 10 seconds. There are, in total, 30 test tracks generated to form the variable-rate test group for the experiment.

**Number of Participants:** Each test track is rated by 14 non-expert participants using 5-point MOS, 5 for the most desirable quality and 1 for the least. The original audio source of 44.1kbps bitrate is presented to the participants at the beginning of the experiment for reference. Other than the reference track, the rest of the 40 test tracks are randomly ordered to avoid time-dependent bias. Each participant rates 41 audio tracks in total, which takes slightly more than 20 minutes to complete.

3.2. Fixed-Rate Results

Raising the sending rate does not improve the user experience in proportion. Fig. 2 shows the MOS of the fixed-rate tracks. The x-axis indicates the sending rate and the y-axis the corresponding MOS. One can observe that the opinion score increases as the sending rate increases. In particular, the degree of MOS increase is not proportional to that of the rate increase. I.e., the MOS-bitrate relationship tends sub-linear. Having applied the regression test on the data set, we find a logarithmic fit to the convexity with a substantially high R-square value – 0.9607. This is a strong indication that the MOS (P) and sending rate (S) relationship exhibits the Weber-Fechner’s Law.

In the preliminary experiments, the sending rates tested are evenly distributed, approximately 11 kbps apart. This is in fact a shortcoming in support of the Weber-Fechner’s Law in MOS-bitrate relationship. Note the region between 7 and 10 kbps. It is where the regression hints on a rapid increase in MOS. However, there lack data points to affirm the relationship is indeed logarithmic. We will take this shortcoming into consideration and address the issue in the large-scale experiments.
3.3. Variable-Rate Results

Fluctuation in sending rates does play a significant role in user experience. Keeping the rate low and steady might be significantly better than maximizing it. Fig. 3 plots the MOS of a variable-rate track, where \((hr, lr) = (40.6, 17.2)\) kbps. The MOS increases with \(\Delta T\), which is not surprising. What we find particularly interesting is the comparison of MOS between the variable-rate and fixed-rate tracks. MOS of the fixed-rate 17.2, 28.9, and 40.6 kbps tracks are 4.0, 4.7, and 4.8 respectively. They are indicated in the figure to highlight the following findings. (1) The user experience at the fixed low rate (17.2 kbps) can be better than a dynamic one that runs no lower than the low rate throughout the track (17.2, 40.6 kbps). The phenomenon is particularly distinct when the change frequency is high. (2) The MOS of the fixed average-rate track (28.9 kbps) is higher than the variable-rate track (17.2, 40.6 kbps). This suggests that user experience of a dynamic audio track is not quantitatively comparable to the average scores of the two steady tracks. Modeling the user experience of the fixed-rate audio streams will not be sufficient to derive that of the variable-rate ones.

3.4. Effects of Rate Change Magnitude

The variable-rate tracks with the same average rate do not result in the same level of user experience either. For variable-rate tracks with the same average sending rate, participants prefer the one with a smaller magnitude rate change. Shown in Fig. 4 are the MOS of two variable-rate tracks, where \((hr, lr) = (28.9, 17.2)\) kbps and \((40.6, 5.6)\) kbps. The two tracks share the same 20.6 kbps average rate. They, however, have rather different MOS. Regardless of the \(\Delta T\), the tracks that are smaller in change magnitude \((28.9, 17.2)\) kbps are consistently better than those that are large in change magnitude \((40.6, 5.6)\) kbps. This observation hints also on the dependency of MOS to the magnitude of rate changes, i.e., a function of \(hr\) and \(lr\), and sets the stage for the modeling task next.

3.5. Effects of Rate Change Frequency

Frequent changes in sending rate do frustrate users. Results of the variable-rate track tests indicate a negative correlation between rate change frequency and MOS. Shown in Fig. 5 is the resulting MOS-\(\Delta T\) plot for three of the variable-rate cases, where the \((hr, lr) = (40.6, 28.9)\) kbps, \((40.6, 17.2)\) kbps, and \((40.6, 5.6)\) kbps. We can see that as the interval of change increases, i.e., the frequency of change drops, the MOS increases. Furthermore, the logarithmic trend indicates that the MOS-\(\Delta T\) relationship likely obeys also the Weber-Fechner’s Law.

Regression tests are carried out on each of the variable-rate track results. The R-square values are all higher than 0.9 except for the \((28.9, 17.2)\) kbps case. This may be attributed to the similarity of the two bitrates. According to the post-experiment feedbacks, the participants are indifferent or unable to notice quality changes between two similar rates. The MOS-\(\Delta T\) relationship, although shows a logarithmic behavior in general, is more subtle and depends on the magnitude of rate changes.

4. PROPOSED MODEL

We take the findings in Section 3 one step further and propose the models that quantify user experience for the fixed-rate and variable-rate Skype calls.

4.1. Fixed-Rate Model

The fixed-rate model is straightforward. Based on the logarithmic relationship observed in MOS-bitrate in Section 3.2, we propose a closed form formula to predict MOS from the sending rate as follows.

\[
f_{FIX}(br) = \gamma \times \ln(br - \alpha) + \beta
\]

where \(br\) is the bitrate, and \(\alpha, \beta, \gamma\) are coefficients to be determined by the characteristic of SILK. The bitrate shift \((\alpha)\) is due to the limit of human perception. When the quality drops below a certain threshold, users are unable to notice any difference. In Section 5, we will provide through a large-scale experiment the specifics of the coefficients.

4.2. Variable-Rate Model

As for the variable-rate model, we base the formula on the finding in Section 3.4 that the MOS of variable-rate tracks
depends on $hr$ and $lr$, and the finding in Section 3.5, the MOS is logarithmic to $\Delta T$. To capture these effects, the closed-form formula is as follows.

$$f_{FLUC}(hr, lr, \Delta T) = SCALE(hr, lr) \times \ln(\Delta T) + SHIFT(hr, lr)$$ (5)

where the effect of change frequency is distributed to the logarithm term; and magnitude is distributed to the two subroutines, $SCALE()$ and $SHIFT()$.

-ln($\Delta T$) is the term capturing the MOS-$\Delta T$ relationship. In that, the MOS increases logarithmically to $\Delta T$.

-$SCALE()$ represents the influence of rate change magnitude to $\ln(\Delta T)$. The finding in Section 3.4 suggests that the larger the difference between $hr$ and $lr$, the higher the influence. $SCALE()$ shall be a function increases with the rate difference. Another caution to take with $SCALE()$ is that it shall be able to deal with the small rate difference cases. The (28.9, 17.2 kbps) case in the preliminary experiments, for example, does not fit well to a logarithmic formula. In particular, the MOS does not vary significantly with $\Delta T$. That is, the influence of $\ln(\Delta T)$ component to MOS needs to be restricted to a small amount. Thus, $SCALE()$ is supposed also a function returning a small value when the difference of $hr$ and $lr$ is low.

-$SHIFT()$ is the remaining portion of MOS that is not influenced by $\Delta T$. This value can be derived by projecting the offset of MOS when the $\Delta T$ approaches the duration of audio track. Although the term does not depend on $\Delta T$, we do find the value varies from case to case in the preliminary results. Therefore, this $SHIFT()$ subroutine is formulated with association with $hr$ and $lr$ as well.

5. LARGE-SCALE EXPERIMENTS

We continue in this section with large-scale experiments for two purposes. The first is to re-examine the MOS-bitrate and MOS-$\Delta T$ relationships via ANOVA tests, which provides as an affirmative support of the proposed models. The second is to derive the unknown coefficients and the exact forms of $SCALE()$ and $SHIFT()$ in the proposed formulas.

5.1. Methodology

**Audio Source:** The audio source is the same as that used in the preliminary experiments.

**Fixed-Rate Tracks:** 9 rates are selected in this set of experiments. Different from the preliminary experiments, the chosen rates are separated evenly by their expected MOS, not by the bitrates, which are estimated by the logarithmic fit we have obtained from the preliminary experiments. The sending rates are (r1, r2, r3, r4, r5, r6, r7, r8, r9) = (40.552, 27.694, 19.422, 14.100, 10.676, 8.473, 7.056, 6.145, 5.558 kbps).

**Variable-Rate Tracks:** Each variable-rate track contains a high rate and a low rate, which are selected from the 9 bitrates used in the fixed-rate group. There are 36 ($hr$, $lr$) combinations. Frequencies used here are the same as those in the preliminary experiments, namely, 1, 2, 3, 5, and 10 seconds.

**Number of Participants:** There are 189 30-second tracks to be rated, 180 variable-rate tracks and 9 fixed-rate tracks. Base on the recommendation of ITU-T P.911[7], in which each track shall be rated by 6 to 40 participants and each experiment shall not exceed 30 minutes, we recruit 128 human subjects and each of them rates for 45 randomly chosen tracks (5760 scores acquired). The experiment duration is approximately 25 minutes, including the time listening to the reference tracks inserted in the experiment for score calibration. Each track is rated by at least 30 participants.

**Score Calibration:** 45 tracks are quite a few to rate. To calibrate scores that might be biased due to fatigue, we take the ACR-HR approach [7]. In that, the differential quality scores (DMOS) are computed between each track and its corresponding reference track.

5.2. ANOVA Tests

**MOS-bitrate:** The one-way ANOVA test is performed to examine the significance of bitrate’s influence to MOS. The p-value for MOS-bitrate test is 1.38e-58, indicating the influence is significant. The R-square value of a logarithmic fit to MOS-bitrate relationship is 0.9645. This affirms the Weber-Fechner’s Law observed in the preliminaries.

**Interaction between $\Delta T$ and ($hr$, $lr$):** To confirm there is indeed interaction between the two factors, a two-way ANOVA test is conducted on $\Delta T$ and the difference of $hr$ and $lr$. The p-value of interaction is 7.94e-14, which strongly support the significance of an interactive term. This affirms the multiplication of $SCALE(hr, lr)$ and $\ln(\Delta T)$ in the proposed variable-rate model.

**MOS-$\Delta T$:** For each given pair of $hr$ and $lr$, the one-way ANOVA test is conducted with respect to $\Delta T$ of the rate changes. A p-value less than 0.05 indicates a significant degree of influence to MOS by $\Delta T$. Otherwise, it suggests that participants are unable to notice the change of bitrates with varying $\Delta T$.

<table>
<thead>
<tr>
<th>Test</th>
<th>p-value</th>
<th>Test</th>
<th>p-value</th>
<th>Test</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>r1r2</td>
<td>0.31</td>
<td>r6r7</td>
<td>0.31</td>
<td>r7r8</td>
<td>0.26</td>
</tr>
<tr>
<td>r3r4</td>
<td>0.42</td>
<td>r6r8</td>
<td>0.11</td>
<td>r7r9</td>
<td>0.34</td>
</tr>
<tr>
<td>r4r5</td>
<td>0.31</td>
<td>r6r9</td>
<td>0.09</td>
<td>r8r9</td>
<td>0.32</td>
</tr>
</tbody>
</table>

Table 1 shows the variable-rate data sets with p-value larger than 0.05. The $hr$ and $lr$ pairs are all similar to each other in these data sets. This echoes the odd case identified in Section 3.5 – the $\Delta T$ would not matter if the rate change is not even discernible. This effect will be captured by the $SCALE()$ subroutine, which tends to a very small value when the difference between the two bitrates is small.

5.3. Model Specifies
Goodness of fit

The data is fitted by the proposed fixed-rate formula with $\alpha=4.091$, $\beta=1.515$, and $\gamma=1.000$. Thus, the relationship between QoS(bitrate) and QoE(MOS) of a fixed-rate Skype VoIP service can be described by the following formula:

$$f_{fix}(br) = \ln(br - 4.091) + 1.515$$  \hspace{1cm} (6)

Subroutines of Variable-rate Formula: Same as our observation in the preliminary experiments, the collected data reveals that $SCALE()$ and the magnitude of rate changes are positively correlated while $SHIFT()$ is determined by the non-$\Delta T$-dependent component and fitted to the $hr$, $lr$ and measured MOS.

$$ SCALE(hr,lr) = 0.1633 \ln(hr - lr) + 0.0598 \hspace{1cm} (7)$$

$$ SHIFT(hr,lr) = D(hr,lr) - SCALE(hr,lr) \times \ln(30) \hspace{1cm} (8)$$

where $D(hr,lr)$ is:

$$\begin{cases} 
MOS_h & \text{if } hr < 14.1 \\
(1.5332 - 0.371 \times MOS_{diff}) \times MOS_{diff} + MOS_i & \text{else}
\end{cases}$$

$$MOS_h = f_{fix}(hr) \hspace{1cm} MOS_i = f_{fix}(lr)$$

$$MOS_{diff} = f_{fix}(hr) - f_{fix}(br)$$  \hspace{1cm} (9)

6. VALIDATION AND DISCUSSION

Evaluation of the proposed model is twofold. First, we compare the predictions of our model to the actual scores given by human participants in both preliminary and large-scale experiments. Second, in order to verify the generality of our model, we conducted a verifying experiment with different audio sources and human participants. This audio source comes from a conversation between two male speakers recorded in 44.1 kbps quality. The length of the sound track is 1 minute. We created rate change in the track by combining two of three different bitrates (44.1, 11.8, 6.4 kbps) with three frequencies (1, 5, 10 seconds). Each track is rated by 13 participants.

Results are shown in Table 2, where the goodness of fit is defined as the percentage of predicted MOS which reside in the interval of one standard deviation away from the average MOS rated by participants. As we can see, the models give significantly accurate predictions.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Large-Scale Exp.</th>
<th>Prelim. Exp.</th>
<th>Verifying Exp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goodness of fit</td>
<td>97%</td>
<td>97%</td>
<td>100%</td>
</tr>
</tbody>
</table>

It does not appear surprising that the validation result is good. After all, a good portion of the data is also used during the derivation of our model. Nonetheless, the 97%-like goodness of fit indicates that the data do not deviate much from the statistical fit. Meaning, the proposed $SCALE()$ and $SHIFT()$ functions and the logarithmic approximation do capture the characteristics of the data.

Two major issues to address next are: (1) in-depth exploration of the terms revealed after numerical fit and (2) whether the Weber-Fechner’s Law exists in other VoIP services. First of all, the log-like relationship between the rate change magnitude and $SCALE()$ is not clear in the preliminary experiments. The term $D()$ has a rather complex structure. A closer look reveals a certain connection between the fixed-rate and variable-rate models, which might lead to one concise model for all. Secondly, although the near-perfect goodness of fit over 3 sets of experiments supports the proposed models to a certain extent, it is not clear for a different codec or experimental setting (ex. mobile or lossy VoIP calls) whether the relationships of MOS to magnitude/frequency of rate changes remain the same. These are the subjects we are particularly interested in pursuing next.

7. CONCLUSION

Our findings in this work provide as a foundation of rate adaptation mechanism design for real-time voice data delivery. In particular, we have achieved in (1) verifying the user experience versus bitrate relationship exhibits a log-like behavior; and (3) deriving the closed form model of user experience to rate changes with 97%+ fit.

REFERENCES