A framework for understanding and defining Quality of Musicians' Experience in Network Music Performance environments

Konstantinos Tsioutas*, Ioannis Doumanis† and George Xylomenos*
*Athens University of Economics and Business, Department of Informatics, Greece
†University of Central Lancashire, School of Physical Science and Computing, United Kingdom

***Abstract*—Considerable research has been conducted on network and system level metrics related to Network Music Performance (NMP). However, little empirical evidence is currently available for assessing the actual Quality of Musician's Experience (QoME) over NMP. We propose a research framework that integrates both subjective and objective aspects of musicians' experience, by explicitly considering the psychological state and profile of each musician, the environment acoustic variables, the music performance variables and the Quality of Service of the network as the key dimensions that impact QoME. We will use the proposed framework to drive empirical studies designed to explore the QoME of musicians performing over the Internet; this paper is a first step in this direction.**

I. Introduction

Network Music Performance (NMP), where two or more musicians perform together through their Internet connections, ideally as if they were placed in the same room, is an extremely demanding service. The critical differentiation of NMP from standard teleconferencing is the need for very low audio latency between the musicians, which places strict limits on the underlying network and coding latencies. The *Mouth to Ear* (M2E) delay, that is, the delay between a microphone at one end and a speaker/headphone at the other end, must be kept below 24 ms [1][2][3] for musicians to be able to synchronize. In contrast, teleconferencing works well even with delays of over 100 ms.

As a result, NMP is not currently feasible for plain Internet users located behind residential (ADSL) links, requiring instead the very fast connections of high speed research networks. Even in such networks, keeping latency low requires optimizing the network, avoiding servers and using low delay codecs or, even, no coding at all. Of course, assessing the suitability of a specific setup for NMP requires more than measuring latency. If we want to evaluate the *Quality of Experience* (QoE) of NMP participants, we have to take into account all the perspectives related to it. A strategy to accomplish this, is to design a research framework which describes the experience through all the correlated variables that have to be taken into account. The main goal of this paper is to understand and define the concepts related to the QoE of NMP, as well as the relations between those concepts. We define this particular instantiation of QoE as *Quality of Musicians' Experience* (QoME).

In Section II, we briefly present related work on NMP services and on assessing musicians' experience. Section III describes a framework for an NMP ecosystem, via which we propose subjective and objective parameters that are parts of the ecosystem puzzle. In Section IV we describe a first set of experiments for measuring and finding correlations between various parameters of the framework, while in Section V we discuss our plans for future work. We summarize our work in Section VI.

II. Related work

A large amount of research touches upon QoME evaluation, looking at it from different perspectives, but without taking a holistic view of its aspects. When discussing live music and concerts where a rock band or an orchestra performs music for the audience, everyone focuses on the experience of the audience. "Was the concert good?" "Did you like the music?" "Did you have fun?" are some of the questions individuals can answer after such an experience. Through these questions the experience can be evaluated subjectively, since the audience is a set of individuals. There is existing research related to music's emotional expression which focuses on the experience of the audience and what feelings are generated to individuals when listening to music. For example, [4] is a study on the connection of emotion in music performance with emotional intelligence, where 24 students were asked to complete listening tests, trying to identify the intended emotions in performances of classical piano music.

Nevertheless, when talking about NMP the subject of interest is the musician himself; there may not even be an audience. Hence, the experience of the musician during the performance needs to be explored. In the past, research has been conducted regarding how network latency affects the musicians behavior and, more specifically, his tempo, concluding that as the latency increases a musician slows down his tempo [5]. Chafe [2] reached the same conclusion, experimenting with musicians who clapped their hands, indicating that when the latency was below 11 ms, the tempo was accelerating. Olmos et al [6] experimented with two opera singers and a conductor over a network and evaluated two bio-metric measures, the Galvanic Skin Response (GSR) and the number of Skin Conductance Responses (SCR) using software for behavior recording alongside with questionnaires. Furthermore, [5] offers a wide understanding of jazz musicians experience in the creation of live performances, using the method of interviews, concluding that their live performance is an experience created by the product itself. Finally, [7] investigates the famous basist's Jeremy Kelshaw's performance experience via an interview with him.

Extensive research has also focused on *Music Performance Anxiety* (MPA). [8] is a study on the MPA of classical music performers in professional performances where interviews were used. Research reported in [9] explored the inter-relationships between occupational stress, perfectionism, aspiration, and MPA in a group of elite operatic chorus artists employed full-time by a national opera company. [10] examines the theoretical adequacy of establishing MPA as a subtype of social phobia. But is MPA the only aspect of music performance experience, and if other factors affect music performance, which are they? Furthermore, is the musician's mood a piece of the puzzle for his overall performance? Even if the session is a rock bands studio rehearsal, is music performance affected by the musicians psychological state, the room and its characteristics? If, on the other hand, the session is a duet of highly trained classical musicians in a rehearsal room, do all the above parameters affect their performance?

The diverse types of work found in the literature indicate the different aspects of QoME that can be studied. Our research goal is to take a holistic view of the field, assessing multiple objective and subjective factors and their influence on QoME, through a comprehensive measurement campaign. Such a framework is proposed by [11] to describe QoE for communications ecosystems. Similarly, [12] explores the QoE for e-health ecosystems. Finally, [13] explores the ecosystem for users of video streaming services.

III. The NMP framework

We propose a research framework consisting of four components for assessing the QoME in NMP sessions. The first component is the physical space where the musician is performing, which is described by its acoustic characteristics, or *Environment Acoustic Variables* (EAV), such as reverberation time, resonance, and the room's impulse response for each range of sound frequencies. The second component is the musician/user himself and his state, which includes both transient characteristics, such as anger or sadness, and long term personality traits, such as enthusiasm or ambition. The third component is the technical equipment used for the performance, which includes the user interface, computers, networks, equipment and all the technical aspects involved. The fourth component is the musical context of the performance, including aspects such as genre, tempo, instruments and musical scales involved, as well as performance characteristics related to the user, including experience and expectations. Figure 1 shows some of the factors affecting a musical performance. Each component includes the context and variables that interact with the user and affect QoME.

The proposed NMP framework is outlined in Figure 2, which shows the ecosystem of two musicians performing while placed in separate rooms connected via the Internet, computers and an NMP server; this is the technical part of the framework. As shown, QoS can be evaluated for the technical part in many ways, using objective metrics, such as latency and audio quality.

Unlike QoS, the quality of each musician's experience (QoME1 and QoME2) is a function of multiple variables, including QoS. With the term *Music Performance Variables* (MPV) we refer to aspects that are related to characteristics of the music performed. The term *Psychological State* (PS) refers to transient aspects of the performer like the mood of the day and happiness or sadness at the time of the performance, while the *Psychological Profile* (PP) refers to the musician's overall personality and his fixed traits. We will discuss all these classes of variables in the following subsections.

A. QoS variables

The technical components of the ecosystem start with the *User Interface* (UI) offered by the NMP system, include the computer, the network and the server and end with the other user's UI. This component, considered as a system, has variables which affect QoS. Network latency and audio quality and network jitter are three parameters which are strongly correlated to the QoS. There is existing research related to how network latency affects QoS for NMP systems and, furthermore, how the musician's performance is affected by the audio latency, which concludes that as audio latency increases the musicians slow down their tempo. The audio quality on the other hand is a parameter that can be configured in such a way that latency can be reduced. Changing, for example, the audio sample rate to a value less than 44100 samples per sec, the audio bandwidth is getting narrower, which is perceived as poor audio quality by the musician and, of course, affects his performance. Finally since the UDP protocol is used and there is always a percentage of packet loss, network jitter is being perceived with noticeable clicks by the musician something also undesirable. Hence, these three variables, network latency, audio quality and network jitter are strongly correlated to the performance and will be taken into account in our study.

B. Music Performance Variables

Music Performance Variables are aspects related to the musical context. The first variable in this section is the musician's performance experience. Experience is an aspect that affects performance in musical sessions: a more experienced musician can perform with more confidence than a less experienced one, adapting to more difficult performance conditions. On the other hand, a more experienced one has greater expectations from the NMP system (user expectations are discussed in the following sub-section). Another critical variable in this section is the performance of the other participating musician: each musician's performance is directly affected by his peer, especially in musical genres where joint improvisation takes place. Additionally, perceptual hearing and music tempo are aspects that have strong correlation to the performance.

C. User variables

In this section, aspects related to musician's personality are discussed. A key piece of the puzzle in the discussed

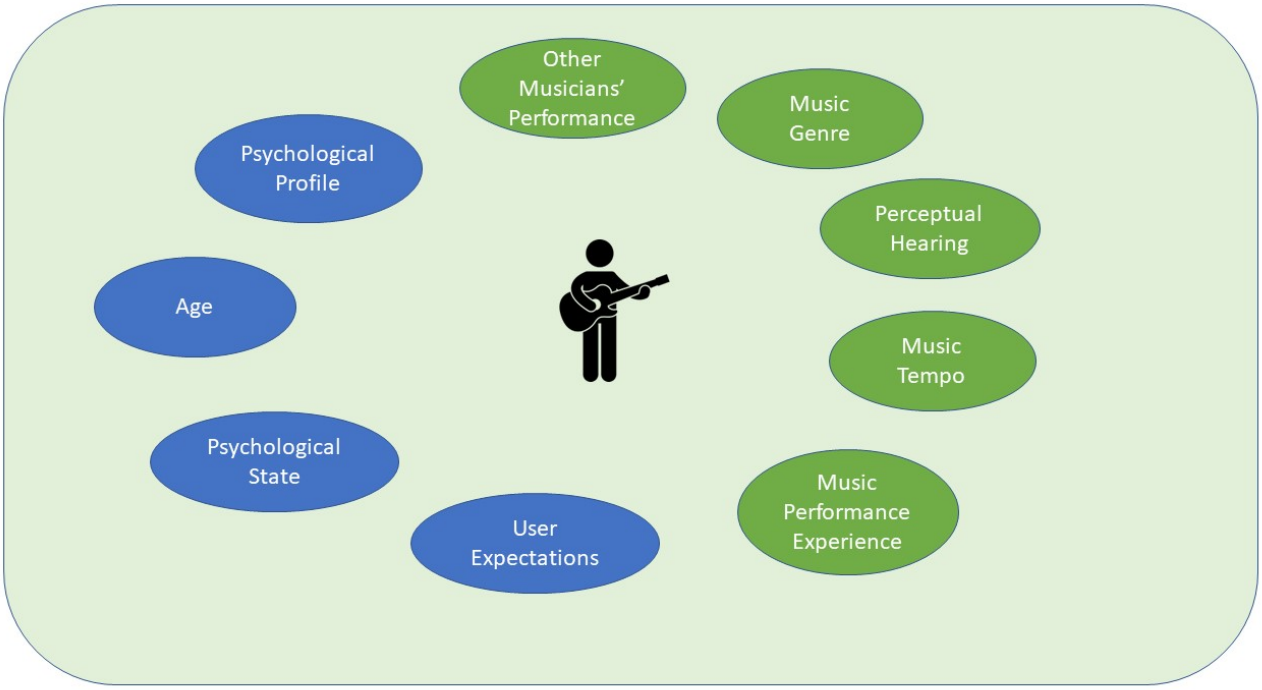

Fig. 1: Factors affecting a musical performance.

ecosystem is the musician's current psychological state. For example, a happy (in life) musician is expected to perform with high energy and enthusiasm, unlike an unhappy one who would probably perform with lower energy due to personal problems and the things that make him unhappy. There are many aspects that could describe the psychological state of a musician like anger, happiness, sadness, depression, boredom and many others, which may have a bearing on performance. Our goal is to evaluate these parameters in real time by using emotional recognition captured through video recordings of the sessions.

The musician's personality is another factor that plays an important role in performance. Personality aspects like aggressiveness, passivity, enthusiasm, patience, greediness could affect performance in general. For example, a soloist tends to play the central role in an orchestra or a band, something crucial in our case, as it tends to affect the performance of other musicians, too. We will use the Big Five Personality Test[1] to class the personality of each musician according to five factors of personality: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. The Big Five personality traits are the best accepted and most commonly used model of personality in academic psychology.

D. Environment Acoustic Variables

A subjective parameter that affects performance and overall experience is the way musicians perceive sound. As is well known, any individual perceives sound in a very specific psychoacoustic way, different from others. For example, elder people perceive a narrower band of frequencies than younger ones. In addition, the acoustic profile of the room, such as reverberation time, resonance, and the room's impulse response for each range of sound frequencies, change the audio that the performers experience. These variables are related to the construction of the room, whether it is a home studio, a professional studio or any other type of room.

[1]https://openpsychometrics.org/tests/IPIP-BFFM/

E. Quality of Experience as a function

Based on the above analysis of the parameters of the ecosystem, we come to the conclusion that QoME is correlated to all of them, as well as the experience of the peer, that is, QoME can be expressed as the function

$$QoME_1 = f(QoME_2, PS, PP, EAV, MPV, QoS)$$

Where $QoME_1$ stands for Quality of Experience of the first musician, $QoME_2$ refers to the Quality of Experience of the second one, PS stands for the psychological state, PP stands for the psychological profile, EAV includes the environment acoustic variables, MPV includes the music performance experience introduced above and QoS includes the metrics for the technological aspects of the NMP system.

IV. Pilot study

We have developed a prototype software named Aretousa for real-time audio streaming for NMP experimental purposes. Aretousa supports the initialization, configuration, control and mix of multiple outgoing and incoming audio streams using the ALSA audio drives in a Linux machine Aretousa supports both peer-to-peer and server-based architectures, uncompressed and compressed (via Opus) audio. Aretousa allows

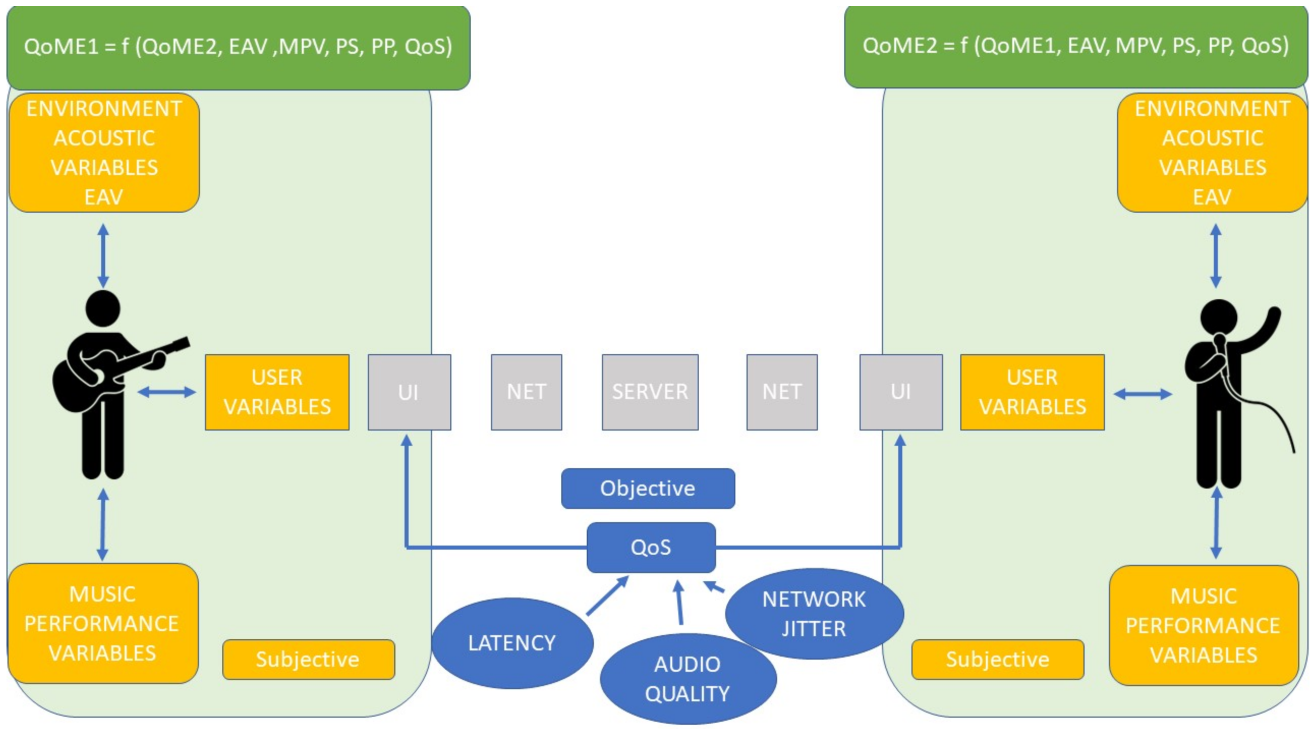

Fig. 2: Network Music Performance Framework.

configuring the audio buffer size, the Ethernet packet size and the parameters of the Opus codec. The musician can listen to his direct sound, his sound coming back from the peer (echo), and the peer's sound, controlling the sound level for each case using volume sliders.

In our experiments, we used the topology shown in Figure 3, with two endpoints (and corresponding musicians) connected via an NMP server; the server did not perform any processing, it only forwarded packets. Each endpoint was connected to a small mixing console, a condenser microphone for the instrument and closed-type headphones for the musician. We used the GRNET[2] infrastructure, to which our laboratory is connected through fiber optic links, as the test network. The computers used for the experiments ran Ubuntu 16.04 with i7 processors and 12 GB of RAM; we used the onboard sound card of each machine for audio capture and playback.

We first ran experiments to determine the latency of the system, and found that with PCM audio we could satisfy the upper limit in M2E delay of 24 ms, as long as the audio capture buffer did not exceed 10 ms worth of audio. We then conducted sessions with real musicians performing over the system and used questionnaires to evaluate subjectively the *Mean Opinion Score* (MOS) for each question.

In the first session, two musicians playing acoustic guitar and bouzouki (a traditional instrument), performed ten folk songs in various tempos. For each song, we used uncompressed PCM audio and increased the audio buffer size starting from 5 ms. The musicians were then asked to answer the following questions, using a 5 point Likert scale (1 is very bad and 5 is very good):

[2]https://grnet.gr/en/

Fig. 3: Topology through NMP server.

1) Evaluate the sound quality during the last musical performance.
2) Evaluate the degree of sync during the last musical performance.
3) Evaluate the degree of delay of the sound you experienced during the last musical performance
4) Evaluate the degree of your musical and emotional expression during the last musical performance
5) Evaluate the degree of interruptions in the sound during the last musical performance
6) Evaluate your degree of satisfaction during the last performance

The results are shown in Figure 4, with each group of columns showing the scores given by one musician in each of the five questions; odd column groups are for musician 1, and even column groups are for musician 2. In the second session, two other musicians participated, playing again guitar and bouzouki, using the same topology and settings. The results for this session are shown in Figure 5.

It is interesting to note that interruptions (question 5) were evaluated as nonexistent in all cases, while the perception of synchronization (question 2) did not show a clear correlation to the buffer-size changes. The level of delay perceived (question

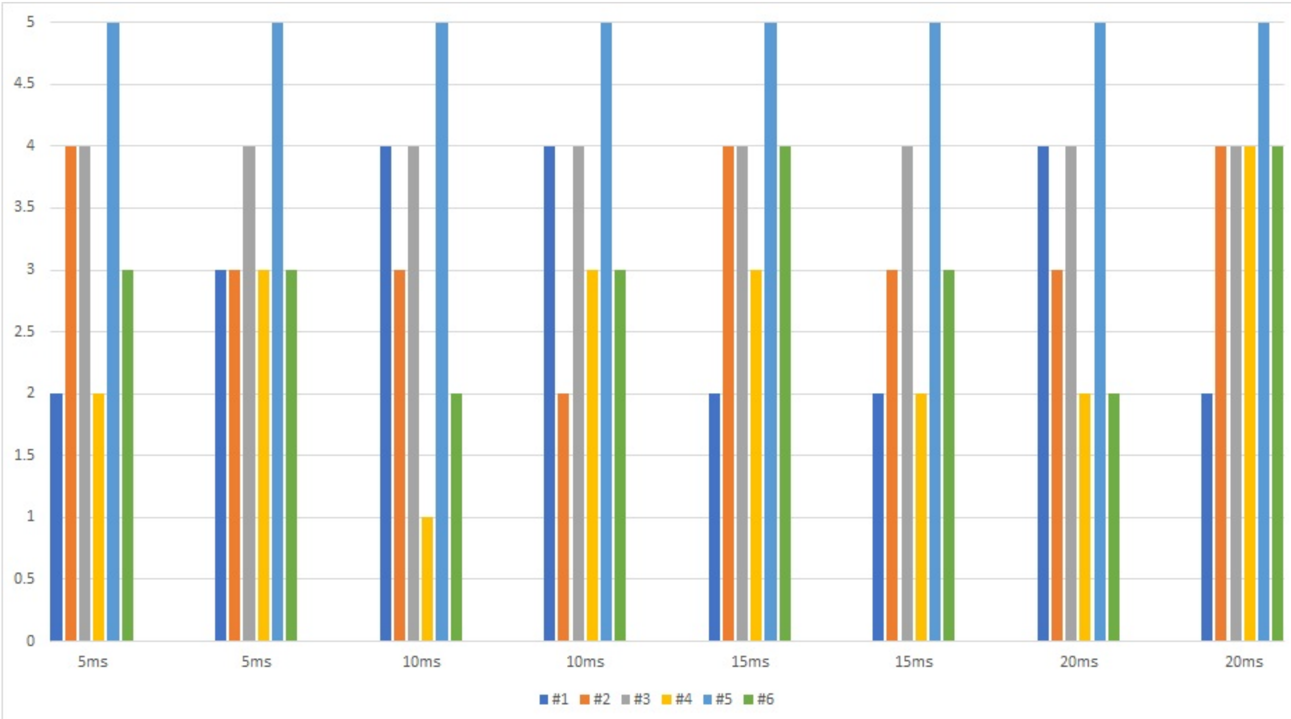

Fig. 4: Mean Opinion Score for Session 1 (one bar per question, odd column groups for musician 1 and even column groups for musician 2).

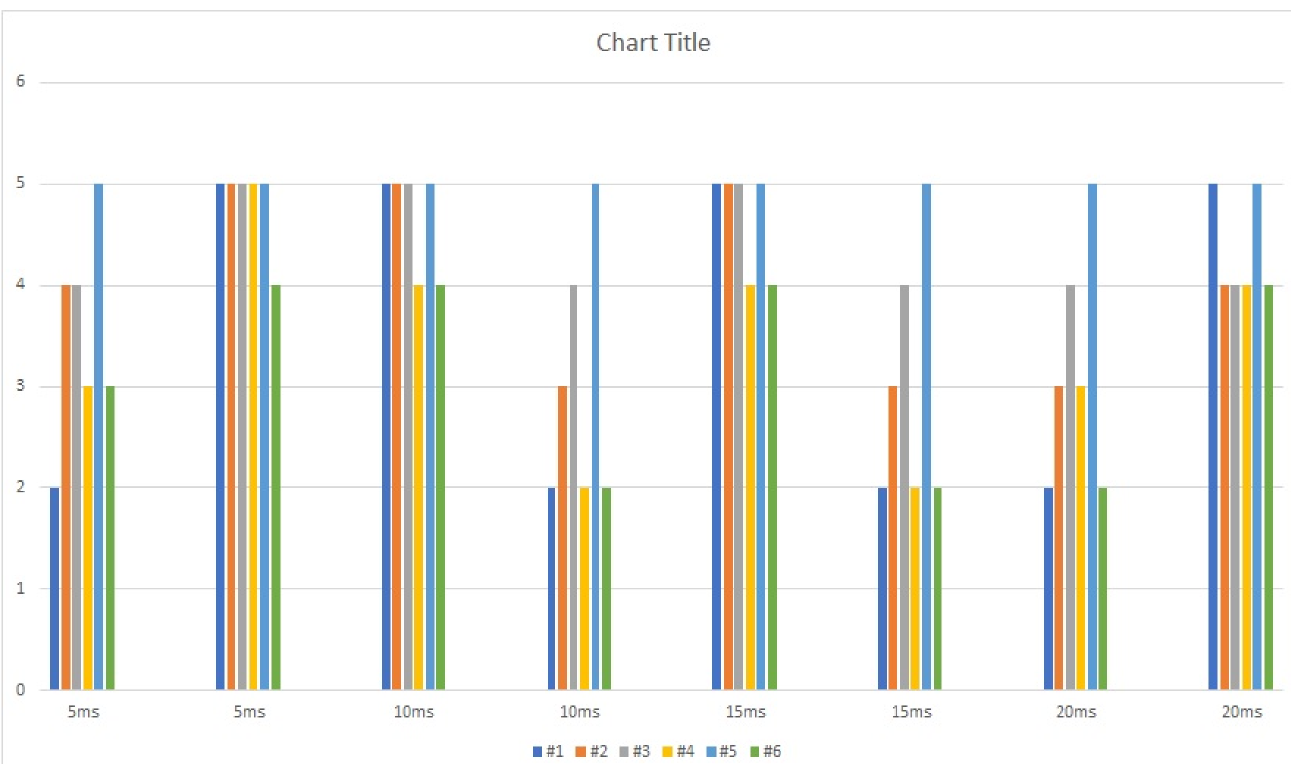

Fig. 5: Mean Opinion Score for Session 2 (one bar per question, odd column groups for musician 1 and even column groups for musician 2).

3) was found to be quite good (4 or more) throughput, despite the changes in buffer size. While the sample size was too small to draw statistically valid conclusions, these results indicate that the perception of QoME can be quite different from what a simple QoS analysis shows.

V. Future work

Creating a framework like the one presented above, and running experiments to assess QoME, is the beginning of a far more extensive study that could take many steps towards an integrated evaluation perspective of the NMP experience. The disciplines of psychology and psychoacoustics are strongly involved with such a study and theories from these areas have to be taken into account to accomplish an as much as more clear approach. Our goal is to investigate correlations among the variables introduced and the aspects discussed above. Building the QoME framework helps to analyze and explore theoretically the context of the research framework.

Our next step is to use biometrics for the extraction of useful data for analysis during the NMP sessions, by first exploiting face responses using cameras and proprietary software and then using an electroencephalography (EEG) headset and appropriate software to measure the response of musicians to events during their performance. EEG headsets can measure, for example, levels of frustration, which can be correlated with changes in the environment, for example, increased network delays.

VI. Summary

We have introduced a framework to explore the correlations between multiple aspects which affect musician's quality of experience in NMP. These aspects are components of an NMP ecosystem which is described by our framework. The main goal of our approach is to evaluate the quality of experience for musicians who participate in NMP sessions, by taking into account a large set of objective and subjective variables. As a first step, we have measure a few objective and subjective aspects in a small study conducted using our own open source NMP software.

References

[1] C. Bartlette, D. Headlam, M. Bocko, and G. Velikic, "Effect of network latency on interactive musical performance," *Music Perception: An Interdisciplinary Journal*, vol. 24, no. 1, pp. 49–62, 2006. [Online]. Available: http://www.jstor.org/stable/10.1525/mp.2006.24.1.49

[2] C. Chafe and M. Gurevich, "Network time delay and ensemble accuracy: Effects of latency, asymmetry," in *Procedings of the Audio Engineering Society Convention 117*, Oct 2004. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=12865

[3] P. F. Driessen, T. E. Darcie, and B. Pillay, "The effects of network delay on tempo in musical performance," *Computer Music Journal*, vol. 35, no. 1, pp. 76–89, Mar. 2011. [Online]. Available: http://dx.doi.org/10.1162/COMJ_a_00041

[4] J. E. Resnicow, P. Salovey, and B. H. Repp, "Is recognition of emotion in music performance an aspect of emotional intelligence?" *Music Perception: An Interdisciplinary Journal*, vol. 22, no. 1, pp. 145–158, 2004. [Online]. Available: http://mp.ucpress.edu/content/22/1/145

[5] K. Kubacki, "Jazz musicians: creating service experience in live performance," *International Journal of Contemporary Hospitality Management*, vol. 20, no. 4, pp. 303–313, 2008.

[6] A. Olmos, M. Brulé, N. Bouillot, M. Benovoy, J. Blum, H. Sun, N. W. Lund, and J. R. Cooperstock, "Exploring the role of latency and orchestra placement on the networked performance of a distributed opera," in *Proceedings of the 12th Annual International Workshop on Presence*, 2009.

[7] A. Geeves, D. Mcllwain, J. Sutton, and W. Christensen, "Expanding expertise: Investigating a musician's experience of music performance," in *Proceedings of the 9th Conference of the Australasian Society for Cognitive Science*, 2010, pp. 106–113.

[8] R. Matei and J. Ginsborg, "Music performance anxiety in classical musicians: what we know about what works," *British Journal of Psychiatry International*, vol. 14, pp. 33–35, 05 2017.

[9] D. Kenny, P. Davis, and J. Oates, "Music performance anxiety and occupational stress amongst opera chorus artists and their relationship with state and trait anxiety and perfectionism," *Journal of anxiety disorders*, vol. 18, pp. 757–77, 02 2004.

[10] M. Osborne and J. Franklin, "Cognitive processes in music performance anxiety," *Australian Journal of Psychology*, vol. 54, pp. 86 – 93, 08 2002.

[11] K. Kilkki, "Quality of experience in communications ecosystem." *Journal of Universal Computer Science*, vol. 14, pp. 615–624, 01 2008.

[12] V. Rojas-Mendizabal, A. Serrano-Santoyo, R. Conte, and A. Gomez-Gonzalez, "Toward a model for quality of experience and quality of service in e-health ecosystems," *Procedia Technology*, vol. 9, pp. 968–974, 12 2013.

[13] K. U. R. Laghari and K. Connelly, "Toward total quality of experience: A qoe model in a communication ecosystem," *Communications Magazine, IEEE*, vol. 50, pp. 58–65, 04 2012.