

# A Game Assessment Metric for the Online Gamer

David DENIEFFE<sup>1</sup>, Brian CARRIG<sup>2</sup>, Damien MARSHALL<sup>3</sup>, Dorel PICOVICI<sup>4</sup>

<sup>1),2),3),4)</sup>Institute of Technology Carlow

david.denieffe@itcarlow.ie, brian.carrig@itcarlow.ie, damian.marshall@itcarlow.ie

<sup>4)</sup>University of Limerick

dorel.picovici@ul.ie

**Abstract**—This paper describes a new game assessment metric for the online gamer. The metric is based on a mathematical model currently used for network planning assessment. Beside the traditional network-based parameters such as delay, jitter and packet loss, new parameters based on online players' game experience/knowledge are introduced. The metric aims to estimate game quality as perceived by an online player. Measurements can be achieved in real-time or near real-time and could be useful to both online game players and game service providers. In order to validate and calibrate the proposed metric a subjective game quality assessment is also proposed. Two 5-point scales are introduced: a game-quality scale and a game playing-effort scale. The mean average of each scales termed, as Mean Opinion Score (MOS), will indicate the game quality ( $MOS_{GQE}$ ) and the playing-effort required ( $MOS_{GPE}$ ). The results obtained using subjective tests are divided in two parts. The first part is used during the development and calibration of the proposed objective algorithm. The second part is used to validate the proposed algorithm. The algorithm's performance can be measured using Pearson correlation between the subjective and objective  $MOS_{GQE}$  scores.

**Index Terms**—Subjective game quality assessment, objective game quality assessment, end-user opinion estimation.

## I. INTRODUCTION

This paper introduces a new assessment metric capable of estimating end-user/players perception and performance for online games. The metric is inspired from an ITU-T recommended network planning model and the estimation of the end-user/player perception is termed as the Mean Opinion Score Game Quality Estimation ( $MOS_{GQE}$ ). The assessment can be achieved in real or non-real time for various categories of games such as: games of skill, games of chance, games of strategy, and simulation games.

Traditionally the end-user perception for online games has been measured using subjective game quality assessment only. This can be seen in the work of Armitage [1] and Armitage and Stewart [2].

The proposed approach, which represents and includes an objective alternative, based on network impairments, to the subjective game assessment, could have several applications:

a) *Game service providers*: the impact caused by network impairments that may arise during a multiplayer game session will be quantified by the  $MOS_{GQE}$ . Using the  $MOS_{GQE}$ , the game service providers could enhance the online games quality and the overall Quality of Service

(QoS).

b) *Game players*: the  $MOS_{GQE}$  can be used to inform the game player about the game quality at a point of time before starting and for the duration of playing the game. If the players' choice is to start/join an online game, the  $MOS_{GQE}$  will provide the game player with a continuous feedback about the network state and game quality.

c) *Game Developers*: game developers could take into account the  $MOS_{GQE}$  in developing and or adapting an aspect of the online game which may be affected by network impairments.

d) *Network Service Providers*: those providers who aspire to the provision of large-scale distributed interactive applications as part of their network service could use the  $MOS_{GQE}$  to assist the policing of traffic and control congestion particularly at the access points or in a wireless environment.

The research motivation is based on the continuous growth in the popularity of online games. The availability of Internet has provided an extensive infrastructure with global connectivity for the games industry to develop and deploy online games. Over the last five years, existing research reports revealed that online games have more severe requirements that are not fulfilled by the Internet's best effort model when compared with bi-dimensional Internet interfaces such as World Wide Web (WWW). Monitoring and measuring the end-user requirements is thus of great importance for both game service providers and game players. This is particularly relevant with the launch of 'network-ready' consoles such as the Playstation 3 and the provision of network based services to consoles such as Xbox LIVE.

Research carried out by US Entertainment Software Association (ESA) revealed that the percent of online game players is as high as 44% from total game players [4]. The statistics indicate that the average game player has been playing games for 12 years and the average game player is 33 years old. Between 1999 and 2005 there was an increase from 9 to 25 percent of game players with the average age over 50 years old. These statistics demonstrate a high level of online game players and in average each player provides revenue for game service industry for almost 30 years.

Researchers from University College London also published results of a questionnaire targeted at the online game community [5]. The aim of this research was to provide an insight into "What Online Gamers Really Think

of the Internet". The main results show the followings:

a) For 75 % of the questionnaire respondents the average weekly time expenditure with online gaming was 5-10 hours.

b) 73% of the respondents consider themselves advanced players and 68% of them determine the purchase of their game equipment according the game requirements.

c) The network condition has a huge impact for online game players. 86% consider network problems as "annoying", 73% consider these problems as the main attribute the majority of game disruptions and 60% of players will abandon the game.

d) The network delay has been identified as the major component of network problems. 69% of respondents are using simple metrics such as "ping" to identify the network delay. A high proportion (60%) consider individual network delay to be a disadvantage and a majority of 85% would prefer game servers where such networks effects are equalized for all players.

e) A significant percentage of players (26%) are not able to adjust the game play in the presence of network problems and a majority of 85% wants to break the user isolation from the network and consider beneficial any form of feedback about the network state.

In order to maintain a large number of online players it is becoming essential for the game service providers to estimate game player perception and performance for their games. On the other hand, advanced information about game playing conditions would allow game players to select different online games, different networks and ultimately different tariffs.

Following this introduction, this paper is broken into 4 further sections. Section 2 outlines the current subjective and objective models and their weaknesses. Section 3 proposes a new model based on the concepts of a "Game Rating Factor" and leading to the  $MOS_{GQE}$ . A Subjective Game Assessment mechanism for comparison to the  $MOS_{GQE}$  is outlined in Section 4 while future work and conclusions are discussed in Section 5.

## II. CURRENT MODELS

Much work has been done at examining, at many different layers, features and metrics to quantify and compensate at in an effort to improve player reaction in the online environment. The focus on distributed interactive applications and the adaptation of the application at both client and server has led to changes in the way games are designed and developed for online purposes. The work of Mauve [6] and the proposal for a Real-Time Transport Layer Protocol with interactive focus shows the work carried out at the Application and Transport Layer level.

Cognizance must be given to the Network Layer and the impairments that occur or the improvements which can be made. With the move to wireless media and the growth of high-speed access networks, the changing nature of the provision of online interactive applications such as games has meant that distributed bandwidth-intensive features now require network awareness. In terms of network awareness, user perception in terms of player tolerance is crucial. There are, as outlined by Armitage, two distinct approaches for discovering player tolerance to network disruptions. The

first is to build a controlled lab environment in which to test small groups of players under selected conditions or secondly to monitor player behavior on public servers over thousands of games. In conducting any experiment requiring assessment of psycho-perceptual factors, there are possible problems with the parameters defined and examined and any assumptions made.

Ubicom's [7] Quake-3 G-model defines R as an impairment factor given by the equation:

$$R = (W_L \times L + W_J \times J)(1+E) \quad (1)$$

where:

$W_L$  is the Latency Weighting Factor set to 1 for this case;

$W_J$  is the Jitter Weighting factor;

E is the packet loss;

L is the latency or one-way delay in ms;

J is the Jitter as defined in RFC 1889 in ms.

A successive approach proposed by Wattimena [3] uses a similar model named the Quake IV G-Model to predict the perceived quality of a First-Person Shooter. The perception is given by the following mapping function:

$$MOS_{model} = -0.00000587 X^3 + 0.00139 X^2 - 0.114 X + 4.37 \quad (2)$$

where:

X – the network impairment -is given by the following equation:

$$X = 0.104 * ping_{average} + jitter_{average}. \quad (3)$$

The approaches described above have at least one of the following weaknesses:

- Although they claim to estimate a user perception, the parameters used are based on network components which are limited in fully predicting the user satisfaction.
- Both models are only defined for one specific game.
- Using simple network measures (Latency, ping, jitter) which assume the network is symmetrical.
- Implementing limited subjective testing for a short period of time with a small user pool
- While a high level of correlation between the subjective and the proposed models is shown, this could be justified only for one defined game in restricted testing conditions.

Our proposed model overcomes these weaknesses by:

- Extending the network based model to include all parameters-packet loss must be considered in the bandwidth-intensive application environment
- Including new parameters not proposed or used previously such as the user-based experience/knowledge factors and distortions introduced by user equipment.
- Moving from a game-specific model to a wide range of existing online games (including console based games)
- Using a number of users for subjective testing in line with accepted test models used in telecommunications based scenarios recommended by ITU-T
- Expecting a level of correlation in excess of 95%.

### III. PROPOSED MODEL

Most of existent game quality assessment techniques take into consideration only network impairments therefore the measured games quality is only correlated with the network impairments (delay, jitter and to a limited extent packet loss) [8]. To estimate the player’s overall perception of games quality the proposed objective game quality assessment extends the traditional objective game quality methods by introducing the end-user experience/knowledge. As shown in Figure 1 the proposed objective game quality assessment takes into consideration the following parameters:

- a) end-user experience
- b) distortions introduced by game client equipment (memory, graphic card) and I/O devices (screen, keyboard, and joystick)
- c) distortions introduced by the network (end-to-end delay, jitter, packet loss)
- d) distortions introduced by game server (number of users, game type, game capability to adapt to network distortions)

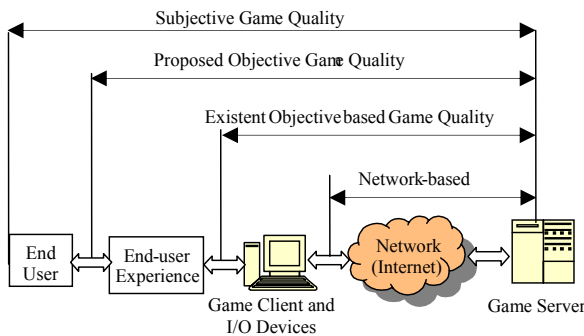


Figure 1. Game Quality Assessment

Using the above-mentioned parameters a “Game Rating Factor” (*GRF*) is proposed. The *GRF* is inspired from an International Telecommunication Union; Telecommunication Standardization Sector (ITU-T) recommended computational model (E-model) [9]. The model is used to assess the combined effects of variation in several parameters that may affect end-user perception of speech quality. The computation of the *GRF* can be described as follows: a maximum value that reflects the highest level of game quality will be reduced in proportion with the distortions caused by various impairment parameters. Mathematically, *GRF* can be calculated using the following equation:

$$GRF = GRF_{MAX} - IGCD - IN - IGS + A \quad (4)$$

where:

*GRF<sub>MAX</sub>* is the maximum Game Rating Factor (90)

*IGCD*: impairment factor representing all impairments due to Game Client and I/O device

*IN*: impairment factor representing all impairments due to network connection between the game server and game client

*IGS*: impairment factor representing all impairments due to Game Server

*A*: represents the end-user experience with online games (max 10)

The *GRF* can lie in the range from 0 to 100, where *GRF*=0 represents an extremely bad game quality and *GRF*=100 represents a very high game quality. The maximum value of 100 is in line with the E-model. The above-mentioned parameters (excepting *GRF<sub>MAX</sub>*) will be individually calibrated and their effect on user perception will be measured directly using a subjective assessment. The subjective test used for this process is detailed in Section 4.

An estimated game quality *MOS<sub>GQE</sub>* for an online game situation in the scale 1-5 can be deducted from the *GRF* using the following mapping function:

For *GRF* < 0: *MOS<sub>GQE</sub>*=1

For 0 < *GRF* <=100:

$$MOS_{GQE} = 1 + 0.035GRF + GRF(GRF - 60)(100 - GRF)^{-7} * 10^{-6}$$

For *GRF* > 100: *MOS<sub>GQE</sub>*=5 (5)

Table 1 shows the provisional guide for the relation between the *GRF*, *MOS<sub>GQE</sub>* and user satisfaction.

TABLE 1. RELATION BETWEEN THE OBTAINED SCORE AND USER SATISFACTION

<i>GRF</i> (lower limit)	<i>MOS<sub>GQE</sub></i> (lower limit)	User satisfaction
90	4.34	Very satisfied
80	4.03	Satisfied
70	3.60	Some users dissatisfied
60	3.10	Many users dissatisfied
50	2.58	Nearly all users dissatisfied

### IV. SUBJECTIVE GAME ASSESSMENT

To date various five-point category-judgment scales have been proposed by several researchers. An example of such scale is given in [10]. However, such a scale was designed to take into account only a limited number of network impairments. The scales proposed in this paper are in line with the ITU-T Recommendation P.800 [11] and takes into consideration all impairments in order to provide an indication of the overall game quality. Two 5-point scales are proposed: a game-quality scale and a game playing-effort scale. The mean average of each scales termed, as Mean Opinion Score (MOS), will indicate the game quality (*MOS<sub>GQE</sub>*) and the playing-effort required (*MOS<sub>GPE</sub>*).

#### 1) Game-quality scale

The game-quality scale is a scale from 1 to 5. Details of wording and score of this scale are given in Table 2.

TABLE 2. GAME QUALITY SCALE

Excellent	5
Good	4
Fair	3
Poor	2
Bad	1

The quantity evaluated from the scores is represented by the symbol  $MOS_{GQE}$  and represents the overall game quality.

### 2) Game playing-effort scale

As shown in Table 3, Game playing-effort scale in a scale from 1 to 5 that reflects the effort required playing the game.

TABLE 3. GAME PLAYING-EFFORT SCALE

Complete relaxation possible; no effort required	5
Attention necessary; no appreciable effort required	4
Moderate effort required	3
Considerable effort required	2
Not possible to play the game	1

The quantity evaluated from the scores is represented by the symbol  $MOS_{GPE}$  and will give an indication about the effort required to play the game. This indication will be especially useful to calibrate the individual parameters used for the proposed objective assessment as detailed in Section 3.

The results obtained using subjective tests will be divided in two parts. The first part will be used during the development and calibration of the proposed objective algorithm. The second part will be used to validate the proposed algorithm. The algorithm's performance will be measured using Pearson correlation between the subjective and objective  $MOS_{GQE}$  scores.

## V. CONCLUSIONS AND FUTURE WORK

This model functions in a role similar to the ITU-T E-model for telecommunications and can be used to predict the subjective effect of combined network impairments (relative/absolute delay, jitter, loss etc.) on a variety of online games. In validating this model, future work will initially concentrate on only FPS games but the model, we believe, will be general enough to adapt easily for other gaming categories. In order to accurately predict and validate the effects of network impairments, other influences on a user's subjective experience will need to be isolated and studied. Such external factors as highlighted above may include issues as varied as user experience/ability, opponent's experience/ability, equipment such as peripheral devices and graphics cards, server processing capacity, number of players, nature of the game tournament/friendly etc.

Going forward, this will be broken into four major tasks:

1) Isolation and subjective testing of external factors and their influence.

2) Subjective testing of isolated network impairments.

3) Measuring the effect of combined network impairments. Currently we have assumed that (like the E-model) combined network impairments will have an additive effect in reducing MOS. However this will need to be validated. Also consideration will be given to the non-additive effect of these parameters on user perception.

4) Games employ (with varying degrees of success) a number of different compensation techniques to mask network impairments from the user. The reductive effect of compensation techniques employed on network impairments will need to be accounted for in the model. Again this will involve subjective testing of the effectiveness of such techniques.

The completion of these four tasks will allow us to prove our model. Network impairments can then be passively or actively measured on any given network and the likely subjective effect on various games calculated.

## REFERENCES

- [1] G. Armitage, "Sensitivity of Quake 3 Players to Network Latency", Poster Session, SIGCOMM Internet Measurement Workshop, San Francisco, November 2001
- [2] G. Armitage, L. Stewart, "Thoughts on Emulating Jitter for User Experience Trials", Proceedings of NetGames 2004 Workshop, ACM SIGCOMM2004, Portland, Oregon, USA, August 2004.
- [3] A.F. Wattimena, R.E. Kooij, J.M. Van Vugt, O.K. Ahmed, "Predicting the perceived quality of a First Person Shooter: the Quake IV G-model", Proceedings of NetGames 2006, Singapore. October 2006
- [4] 2006 Essential Facts about the Computer and Video Game Industry, Entertainment Software Association, E<sup>3</sup>, May 2006, <http://www.theesa.com/archives/files/Essential%20Facts%202006.pdf>
- [5] M. Oliveira, T. Henderson, "What Online Gamers Really Think of the Internet", Proceedings of NetGames 2003, Redwood City (California), USA, May 2003.
- [6] Martin Mauve. "How to Keep a Dead Man from Shooting." Proc. of the 7th International Workshop on Interactive Distributed Multimedia Systems and Telecommunication Services (IDMS) 2000, pp. 199-204, Enschede, The Netherlands, October 2000.
- [7] UbiCom White Paper, "OPScore or Online Playability Score: A Metric for Playability of Online Games with Network Impairments", 2005, <http://www.ubicom.com/news/whitepapers/whitepapers.html>
- [8] M. Dick, O. Wellnitz, and L. Wolf "Analysis of Factors Affecting Players' Performance and Perception in Multiplayer Games", 4<sup>th</sup> Workshop on Network & System Support for Games (NetGames'05), pp:1-7, New York.
- [9] ITU-T Recommendation G. 107 "The E-model, a Computational Model for use in Transmission Planning", Geneva, March 2003.
- [10] P. Quax, P. Monsieurs, W. Lamotte, D. D. Vleeschauwer and N. Degrande "Objective and Subjective Evaluation of the Influence of Small Amounts of Delay and Jitter on a Recent First Person Shooter Game" Network and System Support for Games, SIGCOMM 2004, pp: 152-156, Portland, Oregon.
- [11] ITU-T Recommendation P.800 "Methods for Subjective Determination of Transmission Quality", Geneva, August 1996.