Analysis of Factors Affecting Players' Performance and Perception in Multiplayer Games

Matthias Dick Oliver Wellnitz Lars Wolf

Institut für Betriebssysteme und Rechnerverbund
Technische Universität Braunschweig
Mühlenpfordtstrasse 23, 38106 Braunschweig, Germany

{dick|wellnitz|wolf}@ibr.cs.tu-bs.de

ABSTRACT

In this paper we analyse different factors affecting players' perception and performance in multiplayer games. We introduce a mean opinion score metric borrowed from the subjective analysis of audio and video content to classify the player's perceived game quality. With a survey we investigated the player's view on network latency for a variety of games. Furthermore, we examine four different games under different network conditions and collect data on game score and subjective perception from every player. During the evaluation, we investigate the effect of latency and jitter on multiplayer games. We use regression analysis methods to identify coherences between delay, jitter, skill, game score, and the subjective impression of the player. Finally, we compare the results from the survey and the experiments.

1. INTRODUCTION

In 2004, the computer and video game market in the US exceeded 7.3 billion dollars and multiplayer games are an established factor in this market. This year 42% of most frequent game players are playing online. This is roughly the same percentage as in 2004 and up from 31% in 2001 [2].

At the same time the number of Internet connections is still on the rise. For digital subscriber lines (DSL), there are as of March 2005 now over 107 million customers worldwide. Together with this growing proliferation of residential broadband Internet access, an ever increasing number of households are in a position to participate in all different kinds of realtime multiplayer games.

The main important characteristics of a network connection are throughput, end-to-end delay (latency), variation of latency (jitter), and packet loss. The majority of Internet connections has sufficient bandwidth available to play on the Internet as today's games require only a small total amount of bandwidth, e.g. approx. 34kbit/s for Counter Strike [4] or 5kbit/s for Warcraft III [9]. Hence, the main problems for games lie with network latency, jitter, and packet loss. In this paper we analyse games with different network settings for end-to-end latency and jitter while we leave the analysis of packet loss for future work.

Besides card, board, or trivia games which are quite modest from a networking point of view, many online games require a more or less constant and low end-to-end delay between the players or between a dedicated game server and the players. Games like first person shooters or sports games demand for the lowest latency because the player has direct control over his avatar. Therefore, he notices even low delays between his action and the subsequent action of his avatar. In contrast to directly controlled games, realtime strategy games (RTS) do not have such strict delay requirements. Here, the latency may be higher without interfering with the enjoyment of the player since he just controls his units indirectly. His actions are translated into orders for a single unit or a specific group of units which execute that order on their own without any further need for control. For example, upon arrival of a new destination command for a unit in a strategy game it will calculate a path to the destination and start moving on its own. This unit will usually bypass any obstacles on its way, even those which were not known at the time the command was given by the player. Because of this indirect control of his units, a player can adjust to the network conditions easily [9].

In this work, we present the results of a survey to study users' view on latency requirements for a variety of games. We afterwards discuss the correlation between different objective and subjective parameters which influence the game score and users' perception of multiplayer games. The remainder of this paper is structured as follows: In Section 2 we will discuss related work. Section 3 explains the motivation behind this work. Section 4 describes the methodology used throughout the paper, while Section 5 shows the results of the survey. The experiments and their results follow in Section 6. Finally, in Section 7 we conclude this paper.

2. RELATED WORK

A lot of research work has been done recently on analysing game traffic characteristics and the effects of network latency on multiplayer games.

Sheldon et al. [9] showed that network latency up to several seconds has a minimal effect on the overall users' performance in Warcraft III, a realtime strategy (RTS) game.

While Nichols et al. [6] achieved similar results with Online

NetGames'05, October 10–11, 2005, Hawthorne, New York, USA.

Madden NFL Football with delays below 500 ms, latency above that limit can degrade performance by almost 30 %. They also found that its latency compensation technique failed for asymmetric latencies.

Pantel and Wolf [7] analysed another sports game, a car racing simulation, and found that latency above 50 ms affects game results and above 100 ms degrades a realistic driving behaviour.

Beigbeder et al. [1] examined the affects of loss and latency on users' performance in Unreal Tournament 2003. They explain that players were able to notice latencies as low as 75 ms and found game play less enjoyable at latencies over 100 ms. However, the game scores did not reflect any additional delay in the tests. The authors suggest to keep the delay below 150 ms and the packet loss below 3 %.

In [8], Quax et al. concluded from their experiments with Unreal Tournament 2003 that players feel impaired with delays above 60 ms and that jitter does not play a prominent role. They also notice two groups of players, the "complainers" and the "optimists" which perceive the game quality differently.

From all of these papers, we conclude that network impairment in general has a negative effect on how the players perceive a game, while the exact threshold varies between games of different categories. On the other hand, it has little impact on the outcome of the game, although the player may perceive it differently.

Most other work does not focus on subjective users' perception but try to determine the effects of latency or packet loss on the outcome of the game while trying to keep the game experience itself enjoyable. Additionally, few research has been done about the impact of jitter on multiplayer games or the effect of player's skill on the game. In this paper we focus on differences between multiple games and perform a survey in addition to game experiments.

3. MOTIVATION

Games are played for very different reasons. Game players want to relax during a game, they use it as a diversion from their normal day-to-day routine or just want to get some action. Without taking players who participate in matches or tournaments into account, the majority of players simply enjoys the game.

A game with other fellow players can even be fun if you lose the game. Also, when the network connection of a strong player has a larger delay and so from an objective point of view the game itself gets unfair, because the strong player is now at a disadvantage, this may help a weaker player so that the result of the game is in fact more enjoyable than without the additional delay. In this case, the delay can create an advantage for the weaker player.

This admittedly artificial scenario shows that objective measurements, such as the result or score of a game or the delay of a player does not necessarily reflect the subjective impression of the player.

4. METHODOLOGY

The goal of this study is to understand the factors which have an impact on the subjective users' perception in online games due to network impairment of game traffic. The factors selected for the analysis include player's experience (skill), latency and jitter. In order to achieve good results we have performed a survey with 309 participants. In addition, extensive tests in a controlled environment were accomplished in order to confirm the results obtained from the survey. Finally, a regression analysis of the obtained data was performed.

MOS	Description
1	unacceptable environment impossible to play game
2	very annoying environment server change necessary
3	clearly impaired environment although still acceptable
4	minor impairment noticeable very good environment
5	no noticeable impairments perfect environment

Table 1: The Mean Opinion Score (MOS) values

To identify the impact of network specific parameters and player's skill on users' perception, we evaluated three different kinds of games: two first person shooters, a car racing simulation, and a realtime strategy game. To achieve comparable results we introduce the Game Outcome Score (GOS) and the Mean Opinion Score (MOS) [3] for multiplayer games. GOS is an objective, normalised and therefore game-independent value which represents the game outcome for a certain player in comparison to other participants and is measured in percentage. GOS is the relative contribution of a player's score to the summarized score of all players. It is calculated as follows:

$$\mathrm{GOS}_i = \frac{\mathrm{score}_i}{\sum_{j=1}^{n} \mathrm{score}_j} \cdot 100$$

In this formula n is the total number of players in a game, while score_i represents the score of the i-th player. For example, if we consider a game with four players ($n = 4$) and the score of each player is 4, 16, 8, and 12, the corresponding GOS values would be 10, 40, 20, and 30 respectively. Please note that GOS is not comparable across different games or for a different group of players.

MOS is a common metric used for subjective quality measurements of audio and video contents. It is based on human perception and can take values from 1 to 5. In our case a MOS value of 1 means that the game environment is unacceptable, while a MOS value of 5 represents a perfect environment with no noticeable impairments. The MOS values for multiplayer games are shown in Table 1 and were given as guidelines to all players in the experiment and participants in the survey. For the players' experience, we asked each player to rate his skill for every game on a scale from 1 (never played) to 5 (professional player).

We investigated ten different games from the main genres during our online survey. In the practical evaluation we examined four games. The games selected for the survey and for the practical evaluation in the controlled environment are listed in Table 2.

5. SURVEY

For the survey we created a questionnaire containing questions on the users' experience, their computer hardware, and

Game	Type	Survey	Eval.
Quake 3	FPS	X	
Battlefield 1492	FPS	X	
Battlefield Vietnam	FPS	X	
Unreal Tournament 2003	FPS	X	
Call of Duty	FPS	X	
HALO	FPS	X	
Diablo 2	RPG	X	
FIFA 2004	Sport	X	
Counter Strike	FPS	X	X
Unreal Tournament 2004	FPS	X	X
Warcraft 3	RTS	X	X
Need for Speed Und. 2	CRS		X

Table 2: Games considered in survey and evaluation

their Internet connection. We asked them for the maximum latency with which a game could be played in a perfect environment (MOS = 5), up to which value they believed the game is very good playable, however, they might notice some minor network impairments (MOS >= 4) and when the delay gets so annoying that they would decide to switch to another server (MOS <= 2). These values were requested for ten games (see Table 2).

The survey was distributed through the Internet by using game forums and mailing lists to reach a large number of online game players. While this does not create a representative sample of multiplayer gamers, it allowed us to reach a large group of experienced players. In the total we received 309 replies, of these the majority uses broadband DSL access networks to the Internet as seen in Figure 1a. The player's experience, not shown here, is roughly Gaussian distributed with an average of 2.9 years of general gaming experience. About two thirds visit LAN parties regularly while also two thirds do not participate in online leagues. Figure 1b shows these game related activities. More than one third plays Quake 3 while Counter Strike is the most-played game in this survey with more than half of all participants (see Figure 1c). The players were asked to give their experience regarding latency with a game if they are or have been playing that particular game. So Figure 1c also shows the number of samples regarding latency and MOS we collected for each game.

The minimum, maximum and average delay for the players' perceived latency limits for an unimpaired game are presented in Figure 2 and the limit before the game becomes unplayable in Figure 3. In both figures we eliminated the highest and lowest delays and used only the middle 95 percentile to minimise the effect of too high or too low results. As one can see, the results for different games are nearly identical for the unimpaired case. The latency given by the users is between 70.5 and 93.2 ms with an average of 80.7 ms. The lower limit is the same for all but one game. For Counter Strike ten people specified that they notice delays as low as 20 ms. These people were quite different in their experience, Internet connection and LAN parties or game leagues participation. As the survey was done anonymously we were not able to determine whether this group of people are related or not. Besides this, the player's assessments seem to confirm the results of [1], [7], and [8].

In the impaired case, for most of the games the players

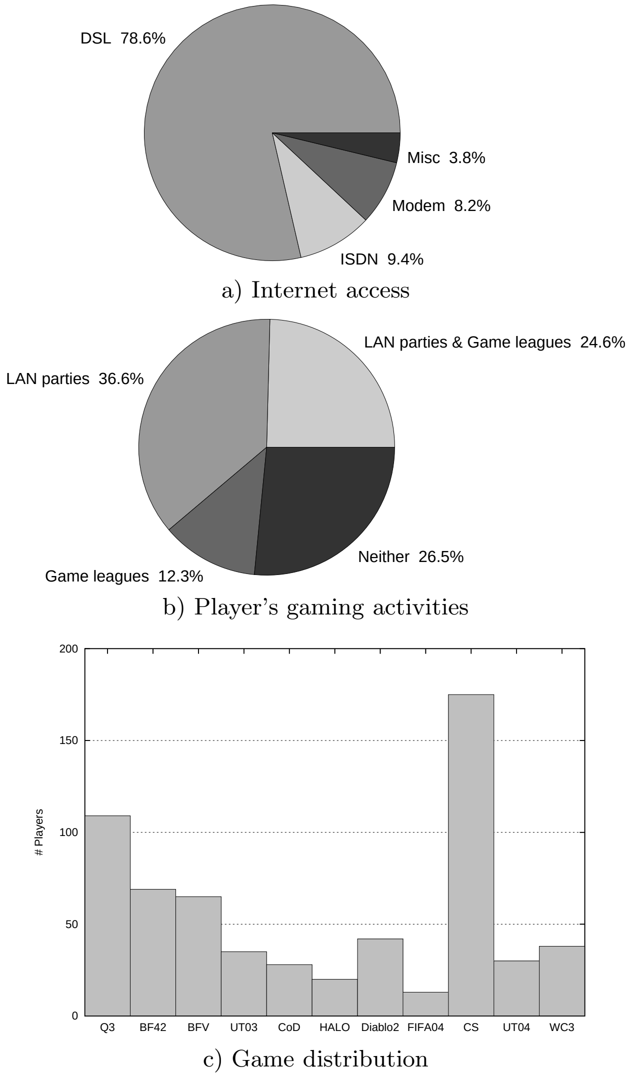

Figure 1: Survey statistics

state that they can tolerate network latency up to 150 ms, which also concurs with measurements from the literature. The average is 118.1 ms. While one can see increased variance for the average delay, there is no clear differentiation between first person shooters like Counter Strike or Quake and the real time strategy game or sports games. For FIFA 2004, all 12 replies within the middle 95 percent indicated the same maximum latency of 100 ms for this game.

As expected, there is a high correlation between latency and MOS for all games, the mean correlation coefficient for latency and MOS is -0.783. But the standard deviation of all correlation coefficients is only 0.053. We therefore presume that from the player's point of view, all multiplayer games have similar delay requirements on the network and that they do not differentiate between different classes of games such as first person shooters or realtime strategy games. However, the players have a good understanding at which point they can detect latency and when a game becomes unplayable. These numbers are in line with previously measured results as stated in the literature.

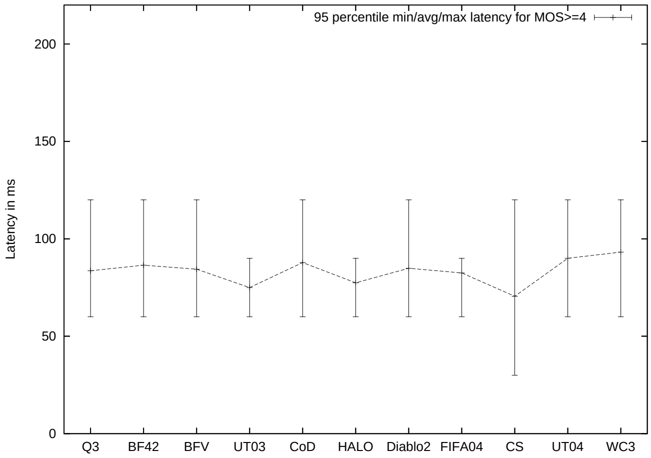

Figure 2: Player's idea on latency for an unimpaired game

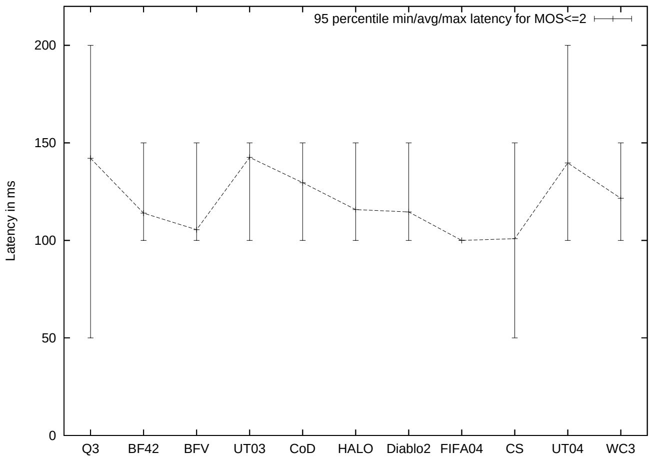

Figure 3: Player's idea on maximum tolerable latency

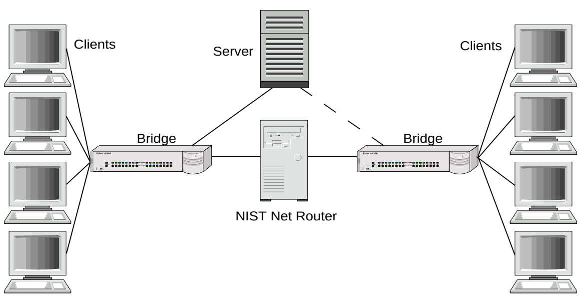

Figure 4: Setup of the practical evaluation

6. EXPERIMENTS

For our experiments we created a test lab with adequate gaming equipment in which we were able to control latency and jitter between two teams of players. Each team was put on a separate network with a 100 Mbit/s bridge as a collapsed backbone. Both networks were connected through a NIST Net [5] router as shown in Figure 4.

We used a dual Intel Xeon 3 GHz machine with 2 GB RAM as a dedicated game server, which could be connected to either of the team's networks. The NIST Net router, an Intel Pentium 4 with 2 GHz, was used to connect both networks and to create latency and jitter between the server and the team on the other side of the router.

We had a total of eight players at the same time which were separated into two teams of four players each. Since the traffic between the server and the clients on the same network was not affected by the NIST Net router, one team was unimpaired while the other suffered from additional delay and/or jitter. We did not tell the teams in advance if and how they were impaired to not influence their responses. Each test was done twice with the server connected to the other network as well. Hence, we examined the results of all players under every network condition. For our network settings, we used all combinations of 0, 50, 150, and 500 ms for delay and 0, 50, and 150 ms for jitter in ascending order. This was done for Counter Strike, Unreal Tournament 2004 and Need For Speed Underground 2. As a full game for Warcraft 3 took more than one hour, it was not possible to test every combination of delay and jitter. Therefore, we examined only the unimpaired case and a delay / jitter setting of 500 ms / 500 ms. We decided not to include a full quantitative analysis of Warcraft 3. With Need for Speed, we decided to use the ranking as GOS rather than the lap or total time.

In order to determine the effects on the impaired team only, the average GOS and MOS values were calculated for all members of this team. GOS and MOS of the players from the unimpaired team were collected but not evaluated besides a plausibility check. We asked each player for his skill as mentioned in section 4.

6.1 Experimental Results

The results of the experimental evaluation are split into two parts. The first part discusses the results of the unimpaired case, the effects of latency, and the problems with jitter. In the second part, we examine the results of a regression analysis of the measured data.

6.1.1 Results for an Unimpaired Network

During the first tests, we did not introduce any additional latency or jitter between the teams and the server. This was done to familiarise the players with the testing environment.

From Table 3 one can see the relation between the normalised game score and the skill given by each player. All games show a high correlation coefficient of approx. 0.6 or above. As expected, the player's skill is the most important factor in an environment with perfect network conditions.

For Warcraft 3 we recognise a very strong correlation between skill and game score. This value is also the case for an impaired environment and confirms [9].

6.1.2 Results Regarding Network Delay

To analyse the impact of latency, we introduced an ad-

Game	Correl. coeff.
CS	0.592
UT2004	0.766
NFS	0.591
WC3	0.974

Table 3: GOS vs. Skill

ditional fixed delay of 50, 150 and 500 ms for every game and analysed how the players perceived this increased latency as well as its effect on the game score. Figure 5 shows the results for the average MOS of all games. As one can see, network latency has a clear impact on all three games. The average player perceives delay in Counter Strike as least disturbing compared to the other games and the game stays playable even at delays of 500 ms. For Unreal Tournament 2004 the results are different. While as for Counter Strike the player notices only small differences for no or low delays, UT2004 becomes quickly unplayable at higher delays. For a delay of 500 ms, Unreal Tournament 2004 achieves the lowest MOS value (2.00) of all games. With Need For Speed even a small delay of 50ms is noticed by the player and at 150ms the game is clearly impaired. This also confirms [7].

For the objective game results, Figure 7 shows the relation between average GOS and latency. For Counter Strike the results are indetermined as there is no clear relation between the score and the latency. Surprisingly, the highest game score on average is achieved under the worst network conditions. For Unreal Tournament 2004 the score roughly follows the players' perception of the network condition. Need For Speed shows only a slight decrease with higher delays and the impact of delays up to 150 ms is low. In contrast to [7], we did not measure lap times but only the final position of the players after the races.

Therefore, we conclude that while latency can have great impact on the subjective users' perception, the objective game outcome is not always affected by latency. For Counter Strike and Unreal Tournament 2004 the differences in Figure 7 are clearly visible even though both games are first person shooters. We also observe that we get MOS values above 3 for latencies up to 150 ms, which means that the game play is still acceptable. This result is contradictory to the survey in which the average tolerable latency was 118.1 ms. Hence, we believe that users tend to overestimate their latency requirements.

6.1.3 *Results Regarding Jitter*

To examine the effect of jitter we created random delays between 0 and 50 or 150 ms for each packet with no correlation between consecutive packets. As for delay discussed in the previous section we can see a correlation between average MOS and jitter (Figure 6). However, even for jitter values of up to 150 ms the average MOS remains well above 4 indicating a very good environment. For Counter Strike the drop in MOS is larger than for the other games when jitter is introduced. This could mean that for Counter Strike the user's perceived network quality is more susceptible to jitter than for the other games, however, at 150 ms Counter Strike is at the same level as the other games.

For the average GOS in Figure 8 we can see no significant difference in the outcome of the game for Unreal Tournament 2004 and Need for Speed for different jitter settings.

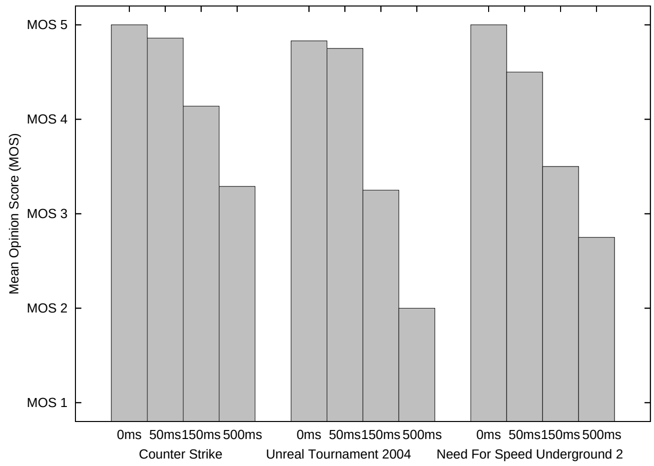

Figure 5: MOS vs. Latency in ms

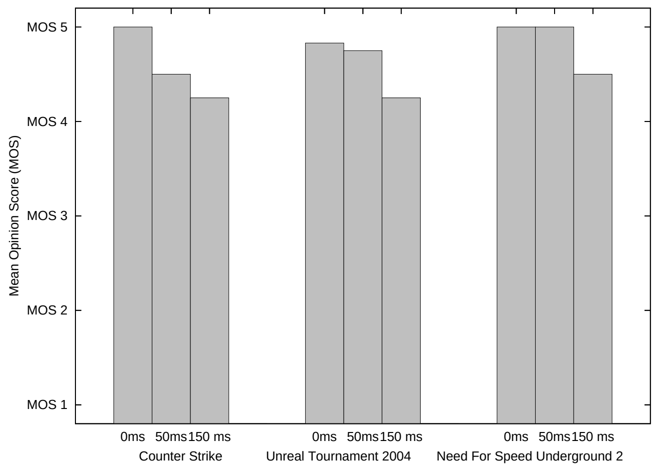

Figure 6: MOS vs. Jitter in ms

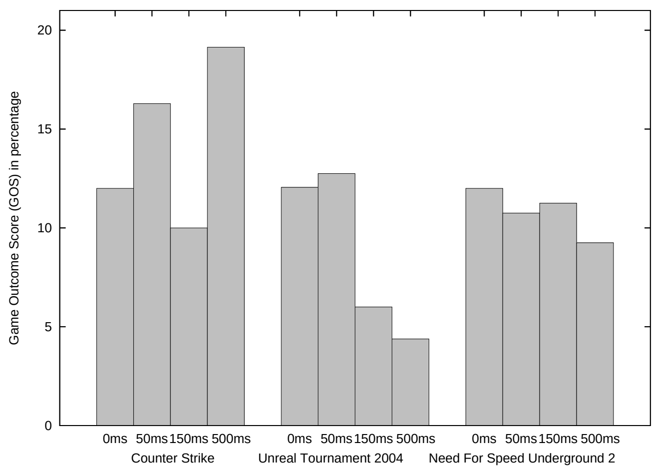

Figure 7: GOS vs. Latency in ms

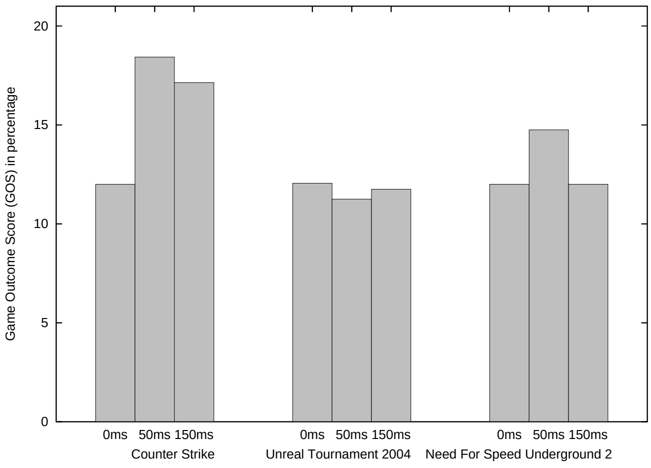

Figure 8: GOS vs. Jitter in ms

As for latency, the results for Counter Strike remains inconclusive.

6.1.4 Regression Analysis

In the previous sections we analysed the single impact of latency or jitter on multiplayer games. In order to evaluate their combined effect we use a multi-dimensional linear regression analysis to find correlations between the measured values. Together with the skill of the players we analysed their relations with MOS and GOS. This analysis does not only include the results from the previous two sections (delay only and jitter only) but all combinations of delay and jitter for every game. Table 4 shows the results of this analysis for all four investigated games. We analyzed the quality of the regression model and show only the values for which we could determine a statistical significance.

In the case of MOS the correlation coefficients for all four games are greater than 0.77 which indicates a high impact of the selected variables on the subjective users' perception. In particular, for FPS games the influence of network specific variables such as latency or jitter is fairly high while the skill seems to have no considerable impact on MOS. Latency, jitter, and skill have about the same impact on MOS in Counter Strike and Unreal Tournament 2004. However, latency seems to be the most significant factor for MOS in Unreal Tournament 2004 while in Counter Strike jitter plays the prominent role. One could guess that this behaviour is the result of a latency compensation scheme in Counter Strike which is more vulnerable to jitter than other FPS games. For Need for Speed Underground 2 latency is the primary factor for users' perception.

Regarding the game outcome score we observed different behaviours in all four games. In Counter Strike the only factor which affected the GOS was the player's skill, while the influence of network specific factors can be neglected. This also explains why the graphs for Counter Strike in Figures 7 and 8 were inconclusive. The correlation coefficient value of 0.411 is not very high which leads us to the assumption that there exist other factors which influence the objective game outcome but have not been considered for this analysis. With Unreal Tournament 2004 we observe that skill

Game	Function	Variables	regression coefficient	correlation coefficient
CS	MOS	skill	inconclusive	0.886
		latency	-0.381	
		jitter	-0.599	
	GOS	skill	0.366	0.411
		latency	inconclusive	
		jitter	inconclusive	
	GOS	MOS	0.364	0.364
UT 2004	MOS	skill	inconclusive	0.838
		latency	-0.861	
		jitter	-0.282	
	GOS	skill	0.599	0.742
		latency	-0.455	
		jitter	inconclusive	
	GOS	MOS	0.608	0.608
NFS	MOS	skill	inconclusive	0.771
		latency	-0.731	
		jitter	inconclusive	
	GOS	skill	inconclusive	0.189
		latency	inconclusive	
		jitter	inconclusive	
	GOS	MOS	inconclusive	0.225
WC3	GOS	skill	0.918	0.934
		latency	inconclusive	
		jitter	inconclusive	

Table 4: Regression analysis results

and low delay are of equal importance. The overall correlation value of 0.742 shows a considerable impact of these two variables. Furthermore, a correlation of 0.608 between GOS and MOS can be observed for this game as well. This is different from Counter Strike where the impact of MOS on the game score is decisively lower. In Need for Speed Underground 2 no considerable impact of the selected variables on GOS could be observed. However, this issue corresponds with the feedback from players who mentioned that in an impaired environment the game outcome is rather randomised. This is because even players who had unimpaired access to the game server felt suddenly distorted because cars from impaired players disappeared and reappeared in the game. This resulted in excessive steering movements and/or collisions. In Warcraft 3 we observed a very high influence of the player's skill on the game outcome under all network conditions while the impact of the network specific parameters seems to be non-existing. The concept of indirectly controlled characters is therefore more resistant against bad network conditions. As we have only investigated delays up to 500 ms and jitter values up to 150 ms, this result concurs with [9].

7. CONCLUSIONS & FUTURE WORK

In this paper we analysed the effect of latency, jitter and skill on users' perception in multiplayer games as well as the game score for four different games. We also introduced the subjective mean opinion score and the objective game outcome score as a normalised metric for multiplayer games.

With a survey we investigated the user's impression on delay for a variety of games. While game players have good knowledge about the threshold at which they can detect network-induced impairments, they tend to underestimate their maximum tolerance of network latency.

As our experiments covered only a group of eight players for each games, the actual statistical significance is quite small. However, we believe that the results shown here indicate the difference the between the games which were analyzed in this paper. We found out that various multiplayer games behave differently under the same network conditions. While this is common knowledge today, our results suggest that this statement is even true for games of the same category such as first person shooters. Skill, latency and jitter generally can have a large but also game-dependant influence on a good game quality. One reason for this may lie in the different algorithms or implementations of prediction and latency compensation techniques used in today's games. Only for Unreal Tournament 2004 we could find a strong correlation between the subjective perception and the objective outcome of the game. Also, depending on the game even unimpaired players can experience degraded game quality when the interaction with other players is high, e.g. with car racing simulations.

In future work we will look at other games to verify our results and to determine network categories for multiplayer games. As our experience with other games indicate that older games may have more problems with latency we plan to review that assumption. With different game scenarios we can improve our test scenarios with delay and jitter values that reflect existing networks. Furthermore, the possibility of asymmetric delays can be considered.

8. ACKNOWLEDGEMENT

The authors thank Michael Koch for the preparation of the survey as well as all other participants of the experimental studies and the survey.

9. REFERENCES

[1] T. Beigbeder, R. Coughlan, C. Lusher, J. Plunkett, E. Agu, and M. Claypool. The Effects of Loss and Latency on User Performance in Unreal Tournament 2003. In *Proc. Workshop on Network and System Support for Games*, pages 144–151, Aug. 2004.

[2] Entertainment Software Association. 2005 Sales, Demographics and Usage Data – Essential Facts about the Computer and Video Game Industry, 2005.

[3] International Telecommunication Union. Methods for the Subjective Assessment of small Impairments in Audio Systems including Multichannel Sound Systems. Rec. BS.1116-1, ITU-R, 1997.

[4] Johannes Faerber. Network Game Traffic Modelling. In *Proc. Workshop on Network and System Support for Games*, pages 53–57, Apr. 2002.

[5] National Institute of Standards and Technology, USA. NIST Net home page, http://www-x.antd.nist.gov/nistnet/.

[6] J. Nichols and M. Claypool. The Effects of Latency on Online Madden NFL Football. In *Proc. of the 14th ACM International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV)*, June 2004.

[7] L. Pantel and L. Wolf. On the impact of delay on real-time multiplayer games. In *Proc. of the 12th international Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV)*, pages 23–29, New York, USA, 2002.

[8] 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. In *Proc. Workshop on Network and System Support for Games*, pages 152–156, 2004.

[9] N. Sheldon, E. Girard, S. Borg, M. Claypool, and E. Agu. The Effect of Latency on User Performance in Warcraft III. In *Proc. Workshop on Network and System Support for Games*, May 2003.