2021 IEEE International Conference on Recent Advances in Systems Science and Engineering (RASSE) | 978-1-6654-3441-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/RASSE53195.2021.9686853

Modelling Network Latency and Online Video Gamers' Satisfaction with Machine Learning

Albert Wong
Mathematics and Statistics
Langara College
Vancouver, BC Canada
alwong@langara.ca

Chunyin Chiu
Mathematics and Statistics
Langara College
Vancouver, BC Canada
edwardchiu.in.ca@gmail.com

Gaétan Hains
LACL
Université Paris-Est Créteil
Paris, 94000 France
gaetan.hains@gmail.com

James Behnke
Computer Science
Okanagan College
Kelowna, BC Canada
jamesbehnke94@gmail.com

Youry Khmelevsky
Computer Science
Okanagan College
Kelowna, BC Canada
ykhmelevsky@okanagan.bc.ca

Chris Mazur
WTFast
Kelowna, BC Canada
chris.mazur@wtfast.com

***Abstract*—The Gamer's Private Network (GPN) technology improves and stabilizes latency of communication between players and servers in online video games, especially when players are distributed worldwide. Latency is known to be the most critical factor in gaming quality of experience. We investigate GPN latency improvement over normal internet and its relationship to player satisfaction using complex, massive data sets, machine learning techniques, and game *genres* or *types*. The conclusions confirm the added value of GPN technology for players but also quantify how it meets the exact needs of specific game types.**

***Index Terms*—video games, big data analytics, big data search and discovery, deep learning applications, industrial applications**

I. Introduction and previous work

The Gamer's Private Network (GPN) technology of WTFast improves normal internet connections between video game servers and players that are distributed worldwide and require stable, low latencies.

Latency reduction is both the heart of WTFast's business and the key quality feature of games networks. This dimension has been studied for many years [1] and in their recent paper [2], Saldana and Suznjevic confirm the necessity of low latency, even above that of bandwidth throughput, for player engagement in almost every kind of online game. The main genres/types for online video games according to [3] are: First Person shooters (FPS), Massively Multiplayer Online Role Playing Games (MMORPG), Real Time Strategy (RTS), Multiplayer Online Battle Arena (MOBA) and Sports games that simulate team sports such as racing.

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), WTFast and Okanagan College.

Authors in [2] quote existing surveys that for FPS games, a one-way delay of 80ms can be acceptable for most game users [4]. For MMORPG games, players started rating the game quality from "excellent" to "good" when one-way latency raised above 120ms. Geographical location of servers is correlated with latency due to transmission delays [4].

WTFast has accumulated massive and detailed datasets around round-trip latency for *ping* messages (whose volume is negligible and has no effect on gaming quality) that repeatedly cover long geographic distances between player clients and game servers. Real-time measurements provide details of every gamer-server connection like the game being played, IP and geolocation of player and server, routing hops, latency, time of day, etc. Equivalent data is also available for normal internet routing of the same messages, allowing multi-dimensional comparisons with GPN routing.

Our research team has studied many aspects of the technology, both in laboratory setups and now using massive datasets of real-time networking data [4]–[12]. Previous data analysis confirmed the superiority of GPN over normal internet. In this paper we report on more advanced analysis and modelling that brings GPN data analysis closer to a real-time closed loop inside of GPN routing techniques.

Descriptive statistics give a reliable and precise picture of the improvement in client-server latency that GPN provides over normal internet routing [4]. Confident in that conclusion, we then asked the question of how well can the other measured variables be used to predict the latency of messages in a given gaming session. Our experiments have identified the best type of machine learning (ML) models for such prediction.

Authorized licensed use limited to: The University of British Columbia Library. Downloaded on March 03,2024 at 05:30:43 UTC from IEEE Xplore. Restrictions apply.

Then we asked how our data analysis can be connected with the genres of games being played, so as to relate networking performance i.e. quality of service, with actual gaming quality of experience. That work provides a hint of the vast and hidden[1] set of factors that make low and stable latency (our measure of quality of service (QoS)), not only a necessary but also a *sufficient* factor for gamer satisfaction or *quality of experience (QoE)*.

II. Dataset and latency: quality of service analysis

The data used for the analysis described here was the record of the WTFast game sessions in July 2020. The original raw data has 64,000 rows with sixty-eight columns of which fifteen were selected. The total geographical distance between the gamer, nodes (intermediate routers), and the game server was derived, as well as a column to indicate whether a game session was started on either Friday, Saturday, or Sunday or not. The percentage difference of ping values was calculated between the sessions using the WTFast service and those without. This is defined as the measure of change in QoS.

III. Game genres and types: quality of experience analysis

Like music and films, video games are given informal categories called *genres* that are not unanimous but relatively standard. Some genres like "casual" (meaning easy to learn, suitable for irregular players) are not relevant for our study but most of them define so-called gameplay factors and are thus good overall descriptors of how intense, frequent, and critical are the small delays incurred by client-server messages. For example, strategy games are less demanding on the network (and therefore less sensitive to latency) than first-person shooter games. We have classified game sessions according to nine main genres of the online database `rawg.io`: action games, RPG games, strategy games, massively multiplayer games, shooting games, simulation games, adventure games, sports games, and casual games. To relate a WTFast session to the set of genres we normalized the session game name (for example, "Final fantasy IV Russia" to "Final fantasy") and then queried `rawg.io` about that game. This kind of primary key calculation was done by hand for lack of time and of a reliable ISBN-like table. We had the pleasant surprise to find out that our 1200 unique (long) game names could be classified into only 14 of the 512 possible subsets of the 9 genres. Figure 1 shows the resulting lattice of game genre subsets.

[1]For privacy reasons

We decided to call them game *types* so that an elementary type is a genre. We then associated a binary (1/0) *sensitivity* value for each of the game genres. Its value is 1 if the intuitive understanding of that game genre is that its player quality of experience is dependent on low and stable latency. Otherwise, its value is 0 if the player quality of experience is independent from the latency. For example sensitivity (RPG) is 0 and sensitivity (ACTION) is 1. We then propagated sensitivity to all game types by taking the fraction of a game type's genres that are latency-sensitive. For example ACTION-STRATEGY-SIMULATION games have sensitivity 1/3 because sensitivity (ACTION) is 1 and the other two are 0.

Finally, we evaluated the change in the quality of gamers' experience (QoE) by multiplying the percentage change in the quality of networking service (QoS) by game sensitivity. For example, since RPG-CASUAL games have sensitivity ((0+0)/2)=0, the improvement in QoE that GPN provides for them is 0, regardless of the improvement in QoS (latency). At the other extreme, ACTION-SHOOTER games have sensitivity ((1+1)/2)=1, so the improvement in QoS is completely effective in providing improved QoE. In summary, QoS is measured by reliable and low network latency, while QoE is QoS weighed by our estimate of the game's sensitivity to low and stable client-server latency.

IV. Results on Quality of Experience Analysis

We analyzed by categorization the percentage of change in QoE (GPN over internet) and its histogram i.e. count of categorized ΔQoE. We had the following groups (from 0 to 5) in % Range: <0; = 0; (0,0.1); [0.1, 0.25); [0.25, 0.5); >0.5.

Most of the game sessions (35%) in July 2020 had 10% to 25% improvement in gamers' experience. On the other hand, around 49% of the game sessions in July had lower than 10% or even negative improvement in the quality of gamers' experience. Those sessions benefit from no more than 10% improvement in the quality of gamers' experience. This implies the GPN service was wasted on nearly half of the gaming sessions in July. In the pure QoS analysis without game sensitivities this critical information was completely invisible.

By comparing the two histograms we see that they have similar, but not equal, right-skewed shapes, that is, a long upper tail in the distribution. But sensitivity does affect distribution of QoE vs QoS: the distribution of percentage change in QoE is shifted a little bit leftward compared with the distribution of percentage change in

Authorized licensed use limited to: The University of British Columbia Library. Downloaded on March 03,2024 at 05:30:43 UTC from IEEE Xplore. Restrictions apply.

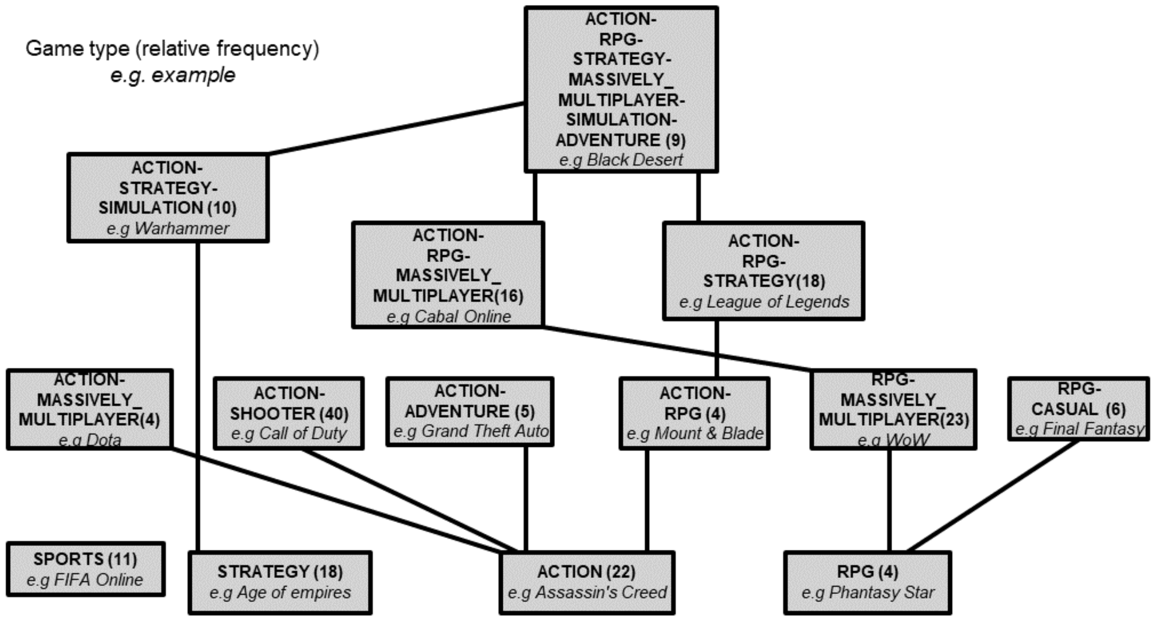

Fig. 1. Game types, or nodes in the lattice of game genres.

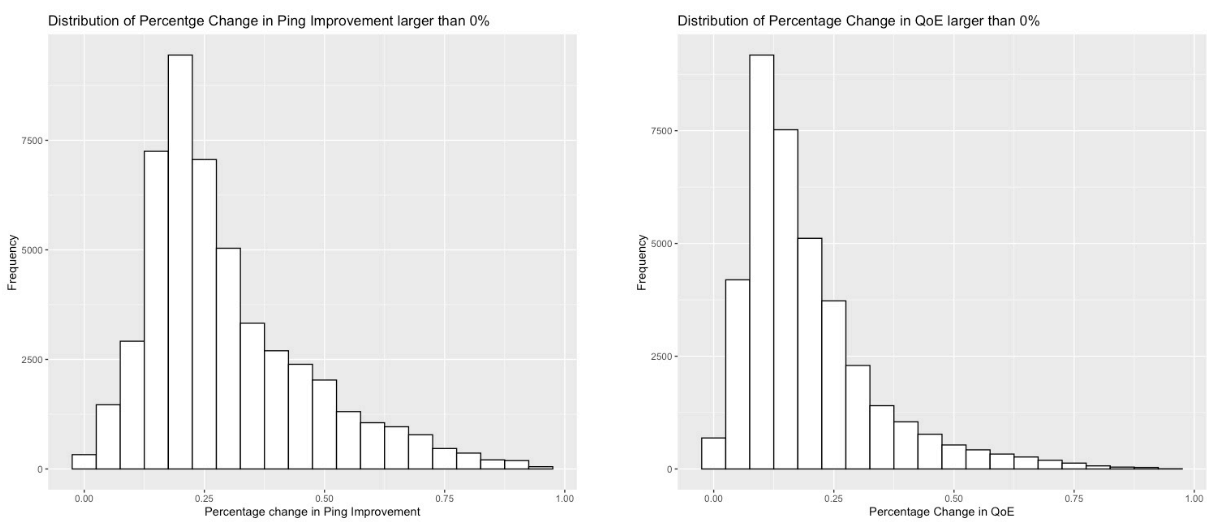

Fig. 2. Histogram of Ping Improvement (provided by GPN over internet) and Delta QoE where it is larger than 0%.

ping improvement. We can conclude that this leftward shift is the effect from game sensitivity to latency.

V. Machine learning Models for Classifying Network Latency

In addition to quality of experience analyses, the variables associated with the game sessions could be used to classify them in terms of network latency using a machine learning model. We found that the numerical variables are highly skewed to the right (high values). This is due to the fact that some game sessions lasted a very long time (the maximum was 662624 seconds or 7.6 days). A square-root transformation was therefore applied to all numerical variables so as to reduce the impact of skewness on the performance of the models developed. Moreover, since the machine learning algorithms are scaling sensitive, min-max scaling was applied to convert all numerical variables into the

Authorized licensed use limited to: The University of British Columbia Library. Downloaded on March 03,2024 at 05:30:43 UTC from IEEE Xplore. Restrictions apply.

same unit. We also used a binning technique on some of these transformed variables to convert them into categorical variables. This categorizing process results in information loss so only the following four variables are categorized: BYTES_UP_TCP, BYTES_UP_UDP, BYTES_DOWN_TCP, and BYTES_DOWN_UDP. To model the impact of these variables on network latency of game sessions, we use these variables as features to classify network latency categories (Very Fast, Fast, Medium, Slow, Ineffective) through a number of ML models based on ping value of the game sessions under the GPN.

Sixty percent of the dataset, which has over forty thousand records, was randomly selected and used as training data for the development of the models. The remaining forty percent were used for testing the models' performance. This sixty/forty percent split is considered typical in machine learning model development.

We developed many ML models for classifying network latency under GPN. These models were built around SVM, random-forest and 4-6 layer neural net algorithms. Many variants of these models were compared for accuracy of classification and the best average accuracy rate is around 91%, thus demonstrating the feasibility of such classification models, as seen in Table I [13]. In addition, the experiments also show that in general random forest models perform better than neural nets and SVM for this data set.

TABLE I
THE BEST MACHINE LEARNING MODELS FOR PING VALUE CLASSIFICATION BASED ON AVERAGE ACCURACY RATE

ML Algorithm	**Av. Acc.**
Random Forest (Trees: 175, Max. depth: 30)	90.94%
Random Forest (Trees: 200, Max. depth: 30)	90.94%
Multiple Layer Perceptron (Two Inputs Model)	90.12%
Support Vector Machine (Gamma: Scale, C:5)	87.93%

Future work on machine learning will aim at applying the lessons learned in the above experiments for the development of optimal ML models and integrating them into the real-time routing algorithms of GPN, as a kind of latency prevention system in the spirit of [14].

VI. CONCLUSIONS

Video gaming is a niche specialty for data analysis so our data analysis and machine learning experiments appear to be the only ones of their kind, at least in the public domain and with relatively large and detailed datasets.

From the point of view of the routing technology, we confirm the value of GPN for players but quantify its current imperfections in meeting the exact needs of specific game types. We also confirm that high-quality predictors for latency are possible, and that random forest algorithm seems to perform better than neural nets or support vector machines in this setting.

Finally, we have made our first steps in the direction of a rational study of game genres as they relate to latency, with the goal of defining a useful and objective quality of experience of GPN users without infringing on the privacy of their gaming experience.

REFERENCES

[1] T. Jehaes, D. De Vleeschauwer, T. Coppens, B. Van Doorselaer, E. Deckers, W. Naudts, K. Spruyt, and R. Smets, "Access network delay in networked games," in *Proceedings of the 2nd workshop on Network and system support for games*. ACM, 2003, pp. 63–71.

[2] J. Saldana and M. Suznjevic, "QoE and Latency Issues in Networked Games," *Handbook of Digital Games and Entertainment Technologies*, pp. 1–36, 2015.

[3] S. W. K. Lee and R. K. C. Chang, "Enhancing the Experience of Multiplayer Shooter Games via Advanced Lag Compensation," in *Proceedings of the 9th ACM Multimedia Systems Conference*, ser. MMSys '18. New York, NY, USA: ACM, 2018, pp. 284–293. [Online]. Available: http://doi.acm.org/10.1145/3204949.3204971

[4] B. Ward, Y. Khmelevsky, G. Hains, R. Bartlett, A. Needham, and T. Sutherland, "Gaming network delays investigation and collection of very large-scale data sets," in *11th Annual IEEE International Systems Conference, SysCon 2017 - Proceedings*, 2017.

[5] T. Alstad, J. R. Dunkin, R. Bartlett, A. Needham, G. Hains, and Y. Khmelevsky, "Minecraft computer game simulation and network performance analysis," *Second International Conferences on Computer Graphics, Visualization, Computer Vision, and Game Technology {(VisioGame 2014)}*, 11 2014.

[6] T. Alstad, J. R. Dunkin, S. Detlor, B. French, H. Caswell, Z. Ouimet, Y. Khmelevsky, and G. Hains, "Game Network Traffic Emulation by a Custom Bot." *2015 IEEE International Systems Conference (SysCon 2015) Proceedings*, pp. 675–680, 4 2015.

[7] Z. Ouimet, H. Caswell, Y. Khmelevsky, R. Bartlett, and A. Needham, "Game servers deployment automation case study," in *2016 Annual IEEE Systems Conference (SysCon)*, 2016, pp. 1–7.

[8] G. Hains, Y. Khmelevsky, R. Bartlett, and A. Needham, "Game private networks performance: Analytical models for very-large scale simulation," in *2016 IEEE International Conference on Cybercrime and Computer Forensic, ICCCF 2016*, 2016.

[9] ——, "Game private networks performance: From geolocation to latency to user experience," in *11th Annual IEEE International Systems Conference, SysCon 2017 - Proceedings*, 2017.

[10] M. Cocar, R. Harris, and Y. Khmelevsky, "Utilizing Minecraft bots to optimize game server performance and deployment," in *Canadian Conference on Electrical and Computer Engineering*, 2017.

[11] G. Hains, C. Mazur, J. Ayers, J. Humphrey, Y. Khmelevsky, and T. Sutherland, "The WTFast's Gamers Private Network (GPN®) Performance Evaluation Results," in *2020 IEEE International Systems Conference (SysCon)*. IEEE, 2020, pp. 1–6.

[12] C. Mazur, J. Ayers, J. Humphrey, G. Hains, and Y. Khmelevsky, "Machine Learning Prediction of Gamer's Private Networks (GPN®S)," in *Proceedings of the Future Technologies Conference*. Springer, 2020, pp. 107–123.

[13] A. Wong, C. Chiu, G. Hains, J. Behnke, Y. Khmelevsky, and C. Mazur, "Network Latency Classification for ComputerGames," in *The IEEE International Conference on Recent Advances in Systems Science and Engineering (submitted)*, 2021.

Authorized licensed use limited to: The University of British Columbia Library. Downloaded on March 03,2024 at 05:30:43 UTC from IEEE Xplore. Restrictions apply.

[14] S. Mallick, G. Hains, and C. S. Deme, "A resource prediction model for virtualization servers," in *High Performance Computing and Simulation (HPCS), 2012 International Conference on*, 2012, pp. 667–671.

Authorized licensed use limited to: The University of British Columbia Library. Downloaded on March 03,2024 at 05:30:43 UTC from IEEE Xplore. Restrictions apply.