

RESEARCH ARTICLE

Predicting Virtual World User Population Fluctuations with Deep Learning

Young Bin Kim¹, Nuri Park², Qimeng Zhang¹, Jun Gi Kim³, Shin Jin Kang³, Chang Hun Kim^{2*}

1 Interdisciplinary Program in Visual Information Processing, Korea University, Seoul, Korea, **2** Department of Computer and Radio Communications Engineering, Korea University, Seoul, Korea, **3** School of Games, Hongik University, Seoul, Korea

* chkim@korea.ac.kr



OPEN ACCESS

Citation: Kim YB, Park N, Zhang Q, Kim JG, Kang SJ, Kim CH (2016) Predicting Virtual World User Population Fluctuations with Deep Learning. PLoS ONE 11(12): e0167153. doi:10.1371/journal.pone.0167153

Editor: Wei-Xing Zhou, East China University of Science and Technology, CHINA

Received: August 17, 2016

Accepted: November 9, 2016

Published: December 9, 2016

Copyright: © 2016 Kim et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education, Science, ICT and future Planning (NRF-2014R1A2A2A01007143, NRF-2015R1A2A1A16074940, and NRF-2015R1A1A1A05001196) and the ICT R&D program of MSIP/IITP (R-20160404-003511, High performance computing (HPC) based rendering solution development). The funders had no role in

Abstract

This paper proposes a system for predicting increases in virtual world user actions. The virtual world user population is a very important aspect of these worlds; however, methods for predicting fluctuations in these populations have not been well documented. Therefore, we attempt to predict changes in virtual world user populations with deep learning, using easily accessible online data, including formal datasets from Google Trends, Wikipedia, and online communities, as well as informal datasets collected from online forums. We use the proposed system to analyze the user population of EVE Online, one of the largest virtual worlds.

Introduction

Evolution in network technology and computing power has enabled people to interact with one another over the Internet. In the same vein, a growing number of users interact with one another and engage in economic, educational, and artistic activities in large virtual worlds (e.g., Second Life). Additionally, massively multiplayer online games (MMOGs)—e.g., World of Warcraft and EVE Online—have attracted an increasing number of active users who build communities and participate in a range of interactions [1–6].

Research on virtual worlds or online societies has been extensive. Easily accessible data from Pardus, an MMOG, enables researchers to investigate social theories in large virtual populations [7–12]. In particular, research on the structure and dynamic evolution of socio-economic networks in virtual worlds has demonstrated significant results using diverse approaches [13–18]. Moreover, data from prior research has been analyzed from different perspectives [19–24]. In addition, research on forecasting the value of virtual currencies used in economic activity among virtual world users is currently underway [25, 26].

Previous studies focused on analyzing virtual world users based on social theories; research into user dynamics in virtual worlds is difficult to find. By contrast, extensive research on predicting real-world population fluctuations has been conducted [27–33]. Still, most researchers have adopted long-term perspectives regarding predictions of real-world population dynamics. These methods could be applied to virtual environments and combined with a range of

study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

internal components to find meaning in long-term population trends. In this study, we focus on the characteristics of virtual worlds such as MMOGs, whose rapid changes correlate with increasing or decreasing user populations [34–36].

The present paper focuses on EVE Online, one of the largest MMOG virtual worlds, which has attracted approximately 0.5 million subscribers since it was released in May 2003 [37]. We predict the daily fluctuations in the number of users of EVE Online through deep learning [38, 39], based on Google Trends [40], Wikipedia usage [41, 42], online forum usage [43] and sentiment data from online forum postings. The proposed method achieved a certain level of success in predicting daily user population fluctuations by drawing on easily available data closely associated with virtual worlds.

Methods

System overview

For the proposed system, we collected data associated with MMOG EVE Online and tagged each posting and reply in its online forum with a positive or negative sentiment value. Based on this data and deep learning [38], we created a model for predicting fluctuations in the number of users (Fig 1).

Data collection

To generate prediction models, we selected data considered to be associated with an increase or decrease in the number of EVE Online users. The data selected was easily accessible online and was gathered from three sources: Google Trends, Wikipedia, and EVE Online forums [40–43].

Google Trends measures search interest in keywords searched on Google over a given period on a scale from 1 to 100. Google Trends data are widely used to analyze relevant phenomena in a variety of fields [44–49]. We gathered Google Trends data about the keyword “EVE Online” for this paper.

Second, Wikipedia usage data [41, 42] was gathered, showing the number of page views for certain keywords on a given day. These data are also widely used to analyze phenomena on the Internet [46, 50, 51]. Again, we gathered the Wikipedia page data relevant to “EVE Online”.

We also gathered data from EVE Online forums [43], which are used by community members to upload postings and exchange opinions regarding topics of common interest [25, 52–54]. Therefore, online forums provide good sources for monitoring the daily responses of many users to certain MMOGs. Communities or forums are widely used in MMOGs for information exchange [52]. According to one study, EVE Online linked its forums with economic activity among users [25], which we found to be relevant when predicting fluctuations in the current number of virtual world users. Comments and relevant replies posted by users on general discussion boards in EVE Online forums were crawled, along with post time, comment and reply post times, the number of replies to each comment, and the number of views. Replies quoting previous comments and replies were crawled to exclude overlapping sentences.

Each HTML page was crawled using Python regex to parse HTML tags and extract the number of topics, the number of replies, the dates on which the topics and replies were posted, and the URL of each topic from the general discussion boards. Based on the URLs of extracted topics, content and replies to them were also extracted. These data were saved in.json format, which was in turn converted to other formats (e.g. csv and xlsx) for different purposes. The.json files from the EVE Online forums crawled can be viewed in the supporting information. One researcher carried out this data collection on a single PC for approximately 72 hours.

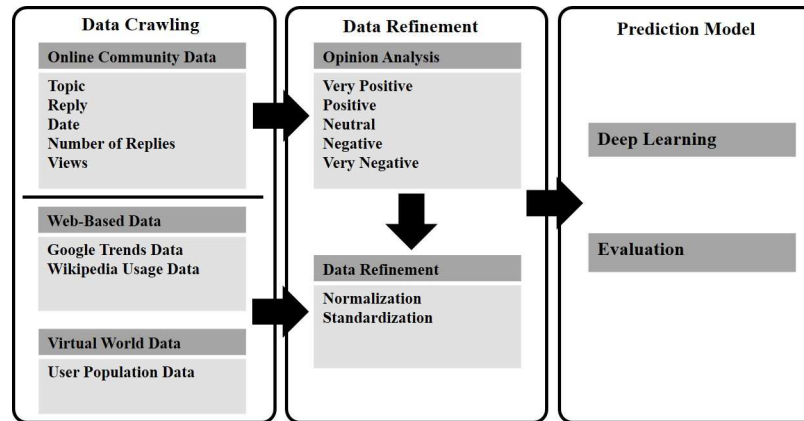


Fig 1. System overview.

doi:10.1371/journal.pone.0167153.g001

We collected data over a period extending from September 12, 2011 to April 29, 2016 (see Table 1) in a manner that complies with the terms and conditions stipulated by each service. Moreover, the collected data do not include any personal information.

Tagging user comment data and correlation analysis

We tagged the collected user comment data with positive or negative sentiment values. Previous studies have mostly focused on classifying user comments in certain fields. Emoticons, neologisms, and ungrammatical expressions are frequently seen on Internet forums. C.J. Hutto and Eric Gilbert [55] proposed an algorithm (VADER) to complement such informal expressions and suggested a method of analyzing social media text using a rule-based model. The online forums of interest here are comparable to social media text. Thus, we utilized the VADER algorithm to tag user comment data from the forums.

VADER normalizes negative and positive sentiment on a scale of -1 to 1. A comment was tagged as very negative, negative, positive or very positive, for $-1 \leq x < -0.6$, $-0.6 \leq x < -0.2$, $0.2 \leq x < 0.6$, and $0.6 \leq x \leq 1.0$, respectively (where x is a number). Each posting and reply was tagged (see the opinion analysis example in Table 2).

We verified the correlation between tagged sentiment values and the increase or decrease in user population. Here, the Pearson correlation coefficient [56] was used to determine the correlation between data sources and the increase or decrease in user population.

As shown in Eq (1), the results of opinion analysis based on the topics and replies (VADER-based tagged values) and population fluctuations were transformed each into z-scores for standardization against those of the previous 20 days. On a given date t ($t = 20$ in the

Table 1. Summary of crawled data

Source	Boundary	Data Volume
EVE Online Community	Sep. 12, 2011–Apr. 29, 2016	1,156,608 User Replies
		37,418 User Threads
Google Trends (EVE Online)	Sep. 12, 2011–Apr. 29, 2016	1,692 Google Trends Values (1 value per day)
Wikipedia Usage (EVE Online)	Sep. 12, 2011–Apr. 29, 2016	1,692 Wikipedia Usage Values (1 value per day)

doi:10.1371/journal.pone.0167153.t001

Table 2. EVE Online Forum Opinion Analysis Example

Opinion Criteria	Example topic sentences
Very Positive	"Pretty happy with Crucible" / "Happy New Year from CCP Games!" / "The days of "Erotica 1" are gone and I AM GLAD" / "Great job CCP!"
Positive	"Nice launcher" / "Better Missile graphics are good" / "CCP Thank you for new Capital Rats"
Neutral	"Guide to Fleet Commanding" / "Drone ships completely useless especially for new players." / "A New Player Guide" / "Worth coming back to eve?"
Negative	"Why does Something Awful forums looks so old and terrible?" / "Back to play this terrible game terribly" / "Server crashed?" / "Awful download speed since launcher is released."
Very Negative	"Those WTF moments" / "AFK Cloaking in System is a Terrible Mechanic" / "Verification Failure—Still Happening" / "Eve Installer not working"

doi:10.1371/journal.pone.0167153.t002

paper), the z-score of expected value \mathbb{E} , denoted by $Z_{\mathbb{E}}$, was defined as:

$$Z_{\mathbb{E}_t} = \frac{\mathbb{E} - \bar{x}(\mathbb{E})}{\sigma(\mathbb{E})} \tag{1}$$

where $\bar{x}(\mathbb{E})$ and $\sigma(\mathbb{E})$ represent the mean and standard deviation of each item for every date with a time granularity of 1 day. Fig 2 shows an example of test results comparing the fluctuations in population and opinion analysis z-scores.

Table 3 shows the Pearson correlation coefficients between the z-scores for the opinion analysis results and the z-scores for population fluctuation.

Overall, a positive linear relationship was found. Notably, there was a clear correlation between user replies and very positive topics.

Prediction Modeling

We used the data collected and the tagged opinion data for the model with the intent to predict fluctuations in user population based on deep learning. Deep learning is widely used for solving a range of problems [38, 39, 57–62]. Data sources have increased quantitatively and qualitatively in proportion to the history of virtual worlds. Still, little research on applying deep learning to virtual worlds for problem solving has been conducted. We developed a setting for applying deep learning based on collected data from a 4.5-year period.

First, we refined the data for the learning models. Specifically, the data gathered (as explained above) was standardized against its applicability to learning. We used long-term data and feature vectors by standardizing the data against the previous 20 days to lessen the impact of the significant changes in the range of data values over longer time periods. An example of applicable input data is shown in Table 4.

We built deep learning models to perform predictions based on the input data. We accumulated multiple hidden layers to learn the deep structure of the data. Here, we configured 1, 2, 3, 5, and 7 hidden layers; of these, we selected the one that returned the best prediction results. We allocated 1,500 neurons to the single hidden layer; when using 2 hidden layers, we allocated 1,024 neurons to each; when using 3 hidden layers, we allocated 1,024, 1,024, and 512 neurons; when using the 5 hidden layers, we allocated 2,048, 1,024, 1,024, 512 and 512 neurons to them; and finally, when using 7 hidden layers, we allocated 2,048, 1,024, 1,024, 1,024, 512, 512 and 512 neurons to the layers.

For the input layer, based on the input data in Table 4, we represented 15 input data as continuing vectors and allocated a different number of neurons to a different number of cumulative days used for learning (i.e., 45, 75, 105, 135, and 180 neurons were allocated to 3, 5, 7, 9,

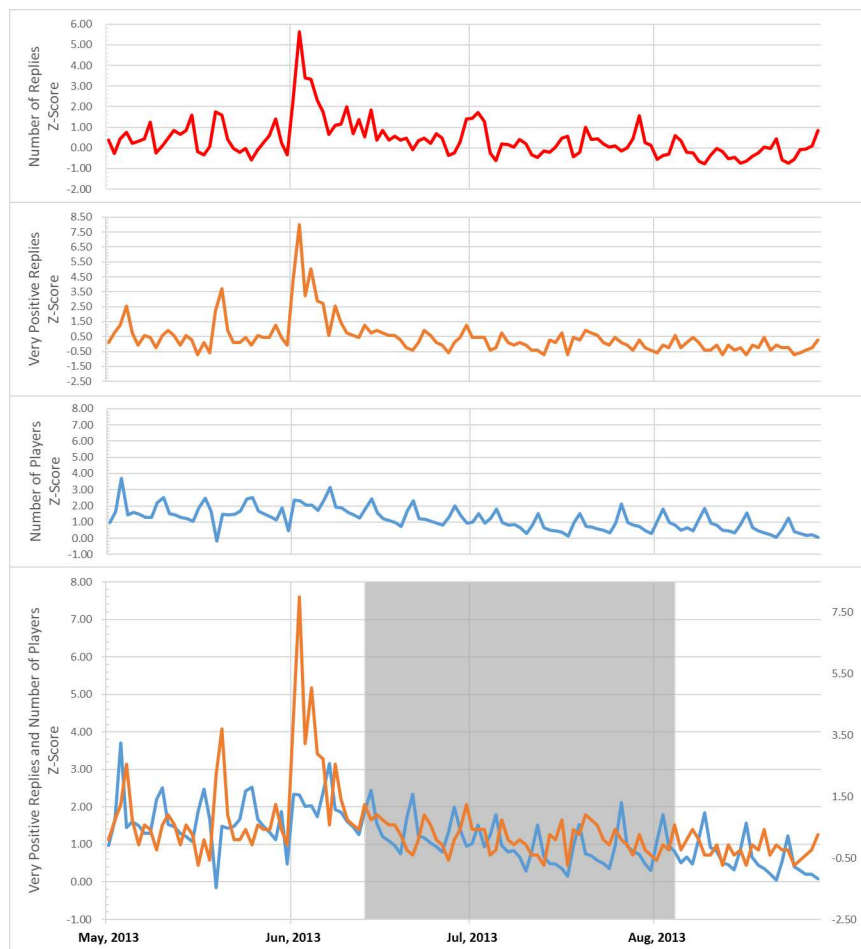


Fig 2. Z-scores of fluctuations in population and results of opinion analysis. Some opinions show a trend similar to that of fluctuations in the population.

doi:10.1371/journal.pone.0167153.g002

Table 3. Pearson Correlation Coefficient Result

Opinion	Pearson Correlation Coefficient between results of opinion analysis and the population
Very Negative Topic	0.2547
Negative Topic	0.2422
Neutral Topic	0.2136
Positive Topic	0.2784
Very Positive Topic	0.3037
Very Negative Reply	0.3178
Negative Reply	0.3667
Neutral Reply	0.3047
Positive Reply	0.3380
Very Positive Reply	0.3298

doi:10.1371/journal.pone.0167153.t003

Table 4. Example of a deep learning dataset. The z-score $Z_{e,t}$ for data from the previous 20 days was used as the values A–J, which indicate the value of the sum of forum opinion on a given date. V–Z denote formal data values (number of topics, sum of replies, sum of views, Google Trends value, and Wikipedia page views) on a given date.

Data Class	Opinion Data										Formal Data					
	Date	Very positiveTopic	PositiveTopic	NeutralTopic	NegativeTopic	Very negativeTopic	Very positiveReply	PositiveReply	NeutralReply	NegativeReply	Very negativeReply	Number of Topics	Sum of Replies	Sum of Views	Google Trends Value	Wikipedia Page Views
Crawled Raw Data	Apr 02, 2016	A	B	C	D	E	F	G	H	I	J	V	W	X	Y	Z
Input Learning Data	Apr 02, 2016	Z_{A_t}	Z_{B_t}	Z_{C_t}	Z_{D_t}	Z_{E_t}	Z_{F_t}	Z_{G_t}	Z_{H_t}	Z_{I_t}	Z_{J_t}	Z_{V_t}	Z_{W_t}	Z_{X_t}	Z_{Y_t}	Z_{Z_t}

doi:10.1371/journal.pone.0167153.t004

and 12 cumulative days, respectively). For the output layer, 2 neurons were allocated to represent the probabilities of population increase or decrease with the softmax function.

We implemented the model using the Google Tensorflow library [38] and accelerated the deep learning model with GPU operation (nVIDIA CUDA). The gap in the prediction values between the optimal model and other model configurations is discussed in the following section.

Experimental Results

We performed prediction modeling by means of deep learning based on the collected and refined data, and predicted fluctuations in the EVE Online user population. We used 90% of the data from the period between September 12, 2011 and April 29, 2016 for learning and 10% for validation. The accuracy rate, F-measure, and Matthews correlation coefficient (MCC) were used to evaluate the performance of the proposed models.

Table 5 and Fig 3 show the prediction results. The highest accuracy (87.57%) resulted from the 2-layer neural network model that used data from the previous 7 days for learning. Table 5 outlines the prediction results by layer configuration and learning data. If the number of hidden layers and days used were less than 2 and 7, respectively, learning was insufficient and prediction accuracy decreased slightly. Conversely, overfitting could occur (with the prediction accuracy failing to significantly improve) if these numbers exceeded 2 and 7, respectively.

Table 5. Experimental results of predicted fluctuations in the EVE Online user population

Data Set		Accuracy (%)	F1-Score	MCC
Hidden Layers	Learning Days			
1 Hidden Layer	3 Days	65.68%	0.6577	0.3098
	5 Days	72.04%	0.7199	0.4342
	7 Days	75.15%	0.753	0.5145
	9 Days	79.88%	0.8006	0.6042
	12 Days	80.06%	0.8027	0.6027
2 Hidden Layers	3 Days	69.23%	0.6933	0.3815
	5 Days	72.78%	0.7271	0.4495
	7 Days	87.57%	0.8763	0.7517
	9 Days	78.70%	0.788	0.5802
	12 Days	81.55%	0.8147	0.6127
3 Hidden Layers	3 Days	68.94%	0.6895	0.3756
	5 Days	62.13%	0.6213	0.2283
	7 Days	73.96%	0.74	0.4822
	9 Days	77.52%	0.7763	0.5438
	12 Days	79.76%	0.7988	0.5981
5 Hidden Layers	3 Days	68.34%	0.6844	0.3572
	5 Days	70.41%	0.7040	0.4078
	7 Days	75.15%	0.7521	0.4987
	9 Days	76.63%	0.768	0.5344
	12 Days	77.23%	0.7732	0.536
7 Hidden Layers	3 Days	69.23%	0.6935	0.408
	5 Days	73.37%	0.7337	0.4675
	7 Days	84.62%	0.8468	0.6897
	9 Days	75.74%	0.7578	0.5079
	12 Days	77.38%	0.7751	0.5386

doi:10.1371/journal.pone.0167153.t005

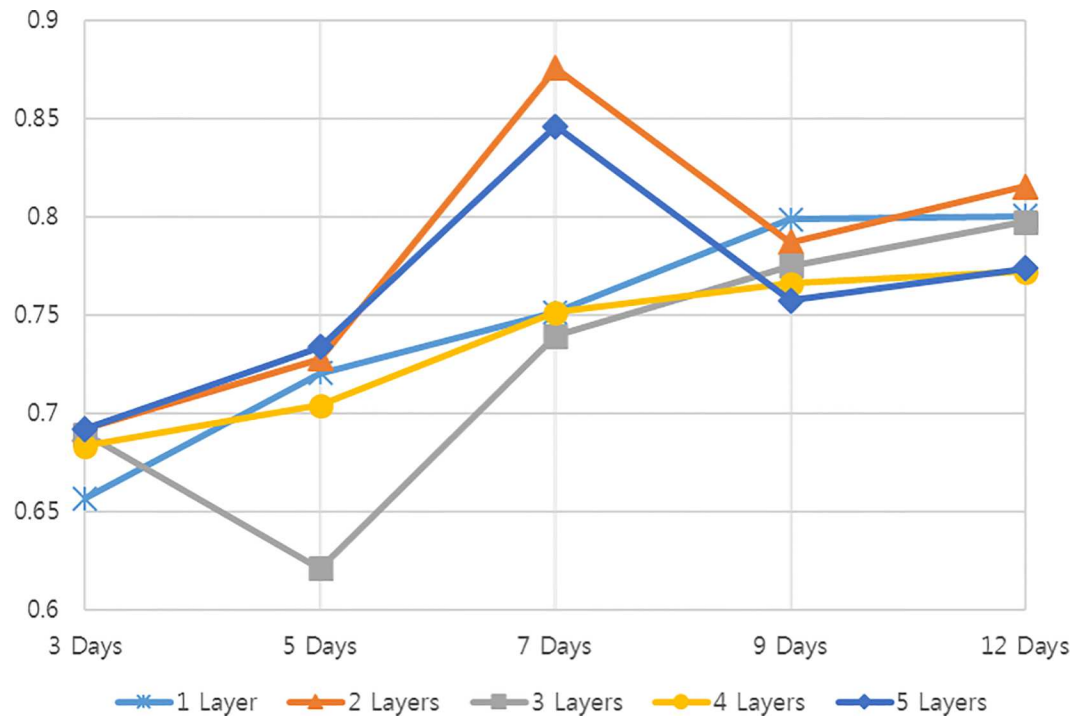


Fig 3. Experimental results. Values by data and layers used.

doi:10.1371/journal.pone.0167153.g003

Discussion and Conclusion

This paper proposed a new method for predicting fluctuations in the number of users in a massive virtual world, EVE Online, with a deep-learning prediction model based on data from a variety of sources. The proposed method successfully predicted fluctuations in the number of EVE Online users based on the easily accessible data relevant to the virtual world. User comments in online forums were found to affect user actions in the virtual world.

The proposed method could be applicable to diverse fields, e.g., verifying newly added content in the creation and management of virtual worlds and solving network problems by forecasting the number of users. The proposed method could also be applied to previous findings, e.g., space and NPC management in virtual worlds [63, 64] and to the optimization of virtual currency systems [25]. In addition, it could be used to apply social science theories to virtual worlds so as to understand the large, diverse user base.

Due to the paucity of previous findings in this field, our proposal has some limitations that need to be rectified in future work. First, the data need to be enriched further for better results. Diverse types of data would increase prediction accuracy. This research was limited to using data gathered for the prediction model only by means of deep learning; more diversified data would be applicable to feature selection as well. The VADER algorithm, optimized for social media analysis, performed well in this study. However, additional analysis of sarcastic or ironic language would improve our results. Moreover, Word2Vec [65] could also potentially improve these results. Sentiment analysis using the Word2Vec-based Doc2Vec [66] focuses on user reviews, making it difficult to apply directly to the analysis of online community users. Still, it might be applicable to and improve the analysis of user comments in online communities. As all data used for learning were formal data, data refinement methods should also be improved. Furthermore, other sizable virtual worlds, e.g., World of Warcraft and Second Life, could be used for prediction.

Virtual worlds have been growing in size and diversity. Considerable research on them has already been conducted. Still, the size of the virtual world user population has not been elucidated. With ongoing improvement for wider application, the proposed method of applying data from different sources to virtual world user population dynamics will contribute to enhancing the understanding of virtual worlds and their user bases.

Supporting Information

S1 File. Results of crawling EVE Online forum (in.json format)

(7Z)

S2 File. Python-based crawler source code for forum data collection

(ZIP)

S3 File. Community usage, Google Trends values, activity on Wikipedia, sentiment analysis results, and EVE Online population data used in the experiment (in.csv format)

(CSV)

Author Contributions

Conceptualization: YBK SJK.

Data curation: YBK JGK.

Formal analysis: YBK NP QZ.

Investigation: YBK NP JGK.

Methodology: YBK NP.

Project administration: YBK CHK.

Resources: YBK NP QZ.

Software: YBK NP QZ JGK.

Supervision: CHK.

Validation: YBK NP QZ.

Visualization: YBK.

Writing – original draft: YBK NP.

Writing – review & editing: YBK CHK.

References

1. De Lucia A, Francese R, Passero I, Tortora G. Development and evaluation of a virtual campus on Second Life: The case of SecondDMI. *Comput Educ.* 2009; 52(1):220–33.
2. Jarmon L, Traphagan T, Mayrath M, Trivedi A. Virtual world teaching, experiential learning, and assessment: An interdisciplinary communication course in Second Life. *Comput Educ.* 2009; 53(1):169–82.
3. Ratan RA, Chung JE, Shen CH, Williams D, Poole MS. Schmoozing and Smiting: Trust, Social Institutions, and Communication Patterns in an MMOG. *J Comput-Mediat Comm.* 2010; 16(1).
4. Kong JSL, Kwok RCW, Fang YL. The effects of peer intrinsic and extrinsic motivation on MMOG game-based collaborative learning. *Inform Manage-Amster.* 2012; 49(1):1–9.
5. Waddell JC, Peng W. Does it matter with whom you slay? The effects of competition, cooperation and relationship type among video game players. *Comput Hum Behav.* 2014; 38:331–8.

6. Pena J, Blackburn K. The Priming Effects of Virtual Environments on Interpersonal Perceptions and Behaviors. *J Commun.* 2013; 63(4):703–20.
7. Mryglod O, Fuchs B, Szell M, Holovatch Y, Thurner S. Interevent time distributions of human multi-level activity in a virtual world. *Physica A.* 2015; 419:681–90.
8. Szell M, Lambiotte R, Thurner S. Multirelational organization of large-scale social networks in an online world. *P Natl Acad Sci USA.* 2010; 107(31):13636–41.
9. Szell M, Sinatra R, Petri G, Thurner S, Latora V. Understanding mobility in a social petri dish. *Sci Rep-Uk.* 2012; 2.
10. Szell M, Thurner S. Measuring social dynamics in a massive multiplayer online game. *Soc Networks.* 2010; 32(4):313–29.
11. Szell M, Thurner S. Social Dynamics in a Large-Scale Online Game. *Adv Complex Syst.* 2012; 15(6).
12. Thurner S, Szell M, Sinatra R. Emergence of Good Conduct, Scaling and Zipf Laws in Human Behavioral Sequences in an Online World. *Plos One.* 2012; 7(1).
13. Bainbridge WS. The scientific research potential of virtual worlds. *Science.* 2007; 317(5837):472–6. doi: [10.1126/science.1146930](https://doi.org/10.1126/science.1146930) PMID: [17656715](https://pubmed.ncbi.nlm.nih.gov/17656715/)
14. Messinger PR, Strolulia E, Lyons K, Bone M, Niu RH, Smirnov K, et al. Virtual worlds—past, present, and future: New directions in social computing. *Decis Support Syst.* 2009; 47(3):204–28.
15. Xie W-J, Li M-X, Jiang Z-Q, Tan Q-Z, Podobnik B, Zhou W-X, et al. Division of labor, skill complementarity, and heterophily in socioeconomic networks. *arXiv preprint arXiv:150303746.* 2015.
16. Fuchs B, Sornette D, Thurner S. Fractal multi-level organisation of human groups in a virtual world. *arXiv preprint arXiv:14033228.* 2014.
17. Jiang Z-Q, Zhou W-X, Tan Q-Z. Online-offline activities and game-playing behaviors of avatars in a massive multiplayer online role-playing game. *EPL (Europhysics Letters).* 2009; 88(4):48007.
18. Xie W-J, Li M-X, Jiang Z-Q, Tan Q-Z, Podobnik B, Zhou W-X, et al. Skill complementarity enhances heterophily in collaboration networks. *Sci Rep-Uk.* 2016; 6.
19. Corominas-Murtra B, Fuchs B, Thurner S. Detection of the Elite Structure in a Virtual Multiplex Social System by Means of a Generalised K-Core. *Plos One.* 2014; 9(12).
20. Fuchs B, Thurner S. Behavioral and Network Origins of Wealth Inequality: Insights from a Virtual World. *Plos One.* 2014; 9(8).
21. Kang SJ, Kim YB, Park T, Kim CH. Automatic player behavior analysis system using trajectory data in a massive multiplayer online game. *Multimed Tools Appl.* 2013; 66(3):383–404.
22. Klimek P, Thurner S. Triadic closure dynamics drives scaling laws in social multiplex networks. *New J Phys.* 2013; 15.
23. Szell M, Thurner S. How women organize social networks different from men. *Sci Rep-Uk.* 2013; 3.
24. Xie WJ, Li MX, Jiang ZQ, Zhou WX. Triadic motifs in the dependence networks of virtual societies. *Sci Rep-Uk.* 2014; 4.
25. Kim YB, Lee SH, Kang SJ, Choi MJ, Lee J, Kim CH. Virtual World Currency Value Fluctuation Prediction System Based on User Sentiment Analysis. *Plos One.* 2015; 10(8):e0132944. PubMed Central PMCID: [PMCPMC4524693](https://pubmed.ncbi.nlm.nih.gov/PMCPMC4524693/). doi: [10.1371/journal.pone.0132944](https://doi.org/10.1371/journal.pone.0132944) PMID: [26241496](https://pubmed.ncbi.nlm.nih.gov/26241496/)
26. Kim YB, Kim JG, Kim W, Im JH, Kim TH, Kang SJ, et al. Predicting Fluctuations in Cryptocurrency Transactions Based on User Comments and Replies. *Plos One.* 2016; 11(8):e0161197. doi: [10.1371/journal.pone.0161197](https://doi.org/10.1371/journal.pone.0161197) PMID: [27533113](https://pubmed.ncbi.nlm.nih.gov/27533113/)
27. Brass W. Perspectives in Population Prediction—Illustrated by Statistics of England and Wales. *J Roy Stat Soc a Sta.* 1974; 137:532–83.
28. Wisniewski A, Smith PWF, Bijak J, Raymer J, Forster JJ. Bayesian Population Forecasting: Extending the Lee-Carter Method. *Demography.* 2015; 52(3):1035–59. doi: [10.1007/s13524-015-0389-y](https://doi.org/10.1007/s13524-015-0389-y) PMID: [25962866](https://pubmed.ncbi.nlm.nih.gov/25962866/)
29. Shang HL, Smith PWF, Bijak J, Wisniewski A. A multilevel functional data method for forecasting population, with an application to the United Kingdom. *Int J Forecasting.* 2016; 32(3):629–49.
30. Ward EJ, Holmes EE, Thorson JT, Collen B. Complexity is costly: a meta-analysis of parametric and non-parametric methods for short-term population forecasting. *Oikos.* 2014; 123(6):652–61.
31. Tayman J, Smith SK, Lin J. Precision, bias, and uncertainty for state population forecasts: an exploratory analysis of time series models. *Popul Res Policy Rev.* 2007; 26(3):347–69.
32. Chi GQ. Can Knowledge Improve Population Forecasts at Subcounty Levels? *Demography.* 2009; 46(2):405–27. doi: [10.1353/dem.0.0059](https://doi.org/10.1353/dem.0.0059) PMID: [21305400](https://pubmed.ncbi.nlm.nih.gov/21305400/)

33. Chi GQ, Voss PR. Small-area Population Forecasting: Borrowing Strength across Space and Time. *Popul Space Place*. 2011; 17(5):505–20.
34. O'Donnell C. On the Back of a Flying Gryphon: Soaring Over/Through the Global Game Industry Nick Dyer-Witheford and Greig de Peuter, *Games of Empire* Bonnie A. Nardi, *My Life as a Night Elf Priest* William Sims Bainbridge, *The Warcraft Civilization*. *Technol Cult*. 2012; 53(1):196–9.
35. Gursimsek RA. Being there together: Social interaction in virtual environments. *Convergence-U.S.* 2012; 18(1):112–4.
36. Daniel P, Chris G, editors. A measurement study of virtual populations in massively multiplayer online games. 6th ACM SIGCOMM workshop on Network and system support for games; 2007.
37. Bergstrom K, Carter M, Woodford D, Paul C. Constructing the ideal EVE online player. *Proceedings of DiGRA 2013: DeFragging Game Studies*. 2013.
38. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. *arXiv:1603.04467*; 2016.
39. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015; 521(7553):436–44. doi: [10.1038/nature14539](https://doi.org/10.1038/nature14539) PMID: [26017442](https://pubmed.ncbi.nlm.nih.gov/26017442/)
40. Google Trends: Google; [cited 2016 Apr 29]. Available from: <https://www.google.com/trends/>.
41. Wikipedia article traffic statistics [cited 2016 Apr 29]. Available from: <http://stats.grok.se/>.
42. Wikipedia Traffic Grapher [cited 2016 Apr 29]. Available from: <http://tools.wmflabs.org/traffic-grapher/Wikipedia-Traffic-Grapher/index.php>.
43. EVE Online Forums: CCP; [cited 2016 Apr 29]. Available from: <http://forums.eveonline.com>.
44. Choi HY, Varian H. Predicting the Present with Google Trends. *Econ Rec*. 2012; 88:2–9.
45. Preis T, Moat HS, Stanley HE. Quantifying Trading Behavior in Financial Markets Using Google Trends. *Sci Rep-Uk*. 2013; 3.
46. Kristoufek L. BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. *Sci Rep-Uk*. 2013; 3.
47. Kristoufek L. Can Google Trends search queries contribute to risk diversification? *Sci Rep-Uk*. 2013; 3.
48. Kang M, Zhong HJ, He JF, Rutherford S, Yang F. Using Google Trends for Influenza Surveillance in South China. *Plos One*. 2013; 8(1).
49. Choi H, Varian H. Predicting the present with Google Trends. *Econ Rec*. 2012; 88(s1):2–9.
50. Moat HS, Curme C, Avakian A, Kenett DY, Stanley HE, Preis T. Quantifying Wikipedia Usage Patterns Before Stock Market Moves. *Sci Rep-Uk*. 2013; 3.
51. Mestyan M, Yasserli T, Kertesz J. Early Prediction of Movie Box Office Success Based on Wikipedia Activity Big Data. *Plos One*. 2013; 8(8).
52. Hau YS, Kim YG. Why would online gamers share their innovation-conducive knowledge in the online game user community? Integrating individual motivations and social capital perspectives. *Comput Hum Behav*. 2011; 27(2):956–70.
53. Panzarasa P, Opsahl T, Carley KM. Patterns and Dynamics of Users' Behavior and Interaction: Network Analysis of an Online Community. *J Am Soc Inf Sci Tec*. 2009; 60(5):911–32.
54. Sing CC, Khine MS. An analysis of interaction and participation patterns in Online community. *Educ Technol Soc*. 2006; 9(1):250–61.
55. Hutto CJ, Gilbert E. Vader: A parsimonious rule-based model for sentiment analysis of social media text. Eighth International AAAI Conference on Weblogs and Social Media 2014.
56. Benesty J, Chen J, Huang Y, Cohen I. Pearson correlation coefficient. *Noise reduction in speech processing*: Springer; 2009. p. 1–4.
57. Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, et al. Human-level control through deep reinforcement learning. *Nature*. 2015; 518(7540):529–33. doi: [10.1038/nature14236](https://doi.org/10.1038/nature14236) PMID: [25719670](https://pubmed.ncbi.nlm.nih.gov/25719670/)
58. Silver D, Huang A, Maddison CJ, Guez A, Sifre L, Van Den Driessche G, et al. Mastering the game of Go with deep neural networks and tree search. *Nature*. 2016; 529(7587):484–9. doi: [10.1038/nature16961](https://doi.org/10.1038/nature16961) PMID: [26819042](https://pubmed.ncbi.nlm.nih.gov/26819042/)
59. CireşAn D, Meier U, Masci J, Schmidhuber J. Multi-column deep neural network for traffic sign classification. *Neural Networks*. 2012; 32:333–8. doi: [10.1016/j.neunet.2012.02.023](https://doi.org/10.1016/j.neunet.2012.02.023) PMID: [22386783](https://pubmed.ncbi.nlm.nih.gov/22386783/)
60. Deng L, Hinton G, Kingsbury B, editors. New types of deep neural network learning for speech recognition and related applications: An overview. 2013 IEEE International Conference on Acoustics, Speech and Signal Processing; 2013: IEEE.

61. Yu D, Yao K, Su H, Li G, Seide F, editors. KL-divergence regularized deep neural network adaptation for improved large vocabulary speech recognition. 2013 IEEE International Conference on Acoustics, Speech and Signal Processing; 2013: IEEE.
62. Krizhevsky A, Sutskever I, Hinton GE, editors. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*; 2012.
63. Kang S-J, Kim Y, Kim C-H. Live path: adaptive agent navigation in the interactive virtual world. *The Visual Computer*. 2010; 26(6–8):467–76.
64. Kang S-J, Kim YB, Park T, Kim C-H. Automatic player behavior analysis system using trajectory data in a massive multiplayer online game. *Multimed Tools Appl*. 2013; 66(3):383–404.
65. Goldberg Y, Levy O. word2vec Explained: deriving Mikolov et al.'s negative-sampling word-embedding method. arXiv preprint arXiv:14023722. 2014.
66. Le QV, Mikolov T, editors. Distributed Representations of Sentences and Documents. *ICML*; 2014.