

Research Article

Predicting the Currency Market in Online Gaming via Lexicon-Based Analysis on Its Online Forum

Young Bin Kim,¹ Kyeongpil Kang,² Jaegul Choo,² Shin Jin Kang,³ TaeHyeong Kim,¹ JaeHo Im,² Jong-Hyun Kim,⁴ and Chang Hun Kim²

¹*Interdisciplinary Program in Visual Information Processing, Korea University, Seoul, Republic of Korea*

²*Department of Computer Science and Engineering, Korea University, Seoul, Republic of Korea*

³*School of Games, Hongik University, Sejong, Republic of Korea*

⁴*Department of Software Application, Kangnam University, Yongin, Republic of Korea*

Correspondence should be addressed to Chang Hun Kim; chkim@korea.ac.kr

Received 6 May 2017; Revised 15 September 2017; Accepted 19 October 2017; Published 24 December 2017

Academic Editor: Katarzyna Musial

Copyright © 2017 Young Bin Kim et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Transactions involving virtual currencies are becoming increasingly common, including those in online games. In response, predicting the market price of a virtual currency is an important task for all involved, but it has not yet attracted much attention from researchers. This paper presents user opinions from online forums in a massive multiplayer online game (MMOG) setting widely used around the world. We propose a method for predicting the next-day rise and fall of the currency used in an MMOG environment. Based on analysis of online forum users' opinions, we predict daily fluctuations in the price of a currency used in an MMOG setting. Focusing specifically on the World of Warcraft game, one of the most widely used MMOGs, we demonstrate the feasibility of predicting the fluctuation in value of virtual currencies used in this game community.

1. Introduction

Individuals can engage in countless interactions over the Internet owing to advancements in network technologies and unprecedented computing power. The massive multiplayer online game (MMOG) environment is becoming popular with an increasing number of users. In this MMOG environment, many users carry out economic activities with a variety of purposes [1, 2]. In the World of Warcraft (WoW) MMOG, users engage in transactions using a particular virtual currency called Gold. In Second Life—another well-known virtual world—a virtual currency called Linden dollars is used in transactions related to user-created items, such as houses and clothes [1].

The virtual economies based on these virtual currencies enable users to engage more fully in MMOGs, sometimes even achieving real economic value [3, 4]. Transactions between virtual currencies and real currencies have been gradually increasing [1, 3, 4]. Most MMOG-related virtual currencies are exchanged with real currencies on eBay or

with crypto-currencies such as Bitcoin [1, 5]. As the number of users engaging in such transactions increases, the market size is expected to grow [5]. Many researchers have investigated the MMOG environment and related situations [1–7]. However, most of these researchers have focused on the transaction environment rather than on the currency value [1].

Real financial transactions have, of course, been extensively studied. Research on techniques for predicting the stock price, for example, dates to the commencement of stock trading [8]. More recently, stock market trends have been predicted using machine learning techniques, such as neural networks and support vector machines [9–13]. Data used for training some prediction models have been based on financial news [11, 12]. Several previous studies focused on the analysis of stock markets using web-based data [9, 10]. However, few researchers have attempted to value virtual currencies and predict their future values.

Numerous transactions involving virtual currencies occur between online game users, forming a large market

[1, 3, 4]. Thus, the ability to predict the next-day rise and fall of virtual currencies in conjunction with cash transactions has been considered important. The ability to predict the market prices of virtual currencies would help transaction agents engage in reasonable transactions and assist game developers in managing virtual environments as well as identifying and solving issues in virtual economies [1].

Virtual currencies for MMOGs exist in economic systems where an enormous amount of bidding competition is allowed, which is not the case for real currencies [14]. All economic agents are fully informed, to the extent that there is no information asymmetry [1, 14]. Suppliers and consumers of virtual currencies act primarily as rational profit maximizers. Fewer variables exist for moderating the value fluctuations of virtual currencies than is the case with real currencies; hence, it is possible to observe value fluctuations easily.

MMOG environments are sources of “Big Data” and thus are suitable for the study of many subjects. In the game *Pardus* [15], data acquisition makes it feasible to study social theories in large-scale virtual-world populations [16–22]. Various approaches to analyzing the structure and dynamic evolution of social networks in virtual worlds have been developed and have yielded significant findings [23, 24]. Related studies based on these findings have been conducted from a variety of perspectives [25–30].

Based on the above-mentioned studies, we have developed a method for predicting daily fluctuations in the value of a currency used in an MMOG environment via user opinion analysis. In this study, we demonstrated the application of this method for predicting daily fluctuations of the virtual currency used in the *WoW* game, the MMOG with the largest active user base.

The work by Kim et al. [1] parallels our research in terms of the objective; however, their predictions were based on trading data from only a small number of users who purchased virtual currencies in cash, which limits the quantity of available data. This makes it difficult to check the data on all daily transactions; therefore, the set price is not reliable. Tokens sold in *WoW* can earn in-game currency if sold by the developer via a qualified path; thus, experiments can be conducted based on more obvious data and using all transactions.

Furthermore, in the technical aspects, existing studies conducted sentiment analysis in a simple form, but the contribution of our study lies mainly in the application of our novel text mining approach based on a custom-built lexicon. That is, to extract a meaningful feature of a document or a time point (containing documents posted in it), we first build a set of keywords, or a lexicon, defining a particular concept or notion, which we call the process of *concept building*. Afterwards, we measure how strongly such a concept is manifested in a given document or a time point, by counting those keywords (and possibly other relevant keywords) occurring in it. Building a custom concept and its lexicon plays a crucial role in analyzing a domain-specific document corpus, and, in our paper, we aim at revealing the signals from textual data that can lead to price prediction. For example, if we use an off-the-shelf sentiment analysis method

that uses standard lexicon for positive and negative keywords, then there would be many domain-specific keywords not captured by the method, that is, false negatives.

Instead, we build our own lexicons closely related to the price of a virtual currency in the game of *WoW*. In this game, users buy tokens with cash at a fixed price (for example, 20 US dollars in the North America (NA) region or 20 Euro in the Europe (EU) region) and sell them to other users in exchange for the virtual currency referred to as Gold. The token/Gold exchange rate changes according to the demand and supply, as shown in Figure 1.

In general, acquiring Gold in the game environment requires an extended period of time and effort, but such time and effort can be reduced if tokens are bought in cash and then sold to another user for Gold. Tokens that are bought with Gold are generally used to buy time available for using the virtual world. Such tokens have two advantages. First, tokens prevent the virtual currency from being illegally transacted in cash. Second, the value of tokens can be traced continuously. We aim to predict the next-day rise and fall of the token/Gold exchange rate (also referred to as the token price) based on user opinion data analyzed on consecutive days. The proposed method can predict fluctuations in the value of virtual currency in the MMOG environment and can be applied to the selling/buying of virtual currency, allowing developers to determine the numerical values and probabilities of multiple effects that would otherwise be difficult to identify.

2. Proposed Approach

As shown in Figure 2, the proposed approach is characterized as follows. The data extracted include users’ opinions concerning the MMOG environment and the prices of the virtual currency used in the transactions. User opinions are sorted with respect to their scores of our custom-built concept and then the causation between the user opinion and the virtual currency price is determined by conducting a causality analysis. Based on the user opinions scored afterward and the machine learning model, we predict the rise and fall of the virtual currency price in the next day.

3. Data Crawling

Initially, data required to generate the prediction model is crawled. We gathered data from official *WoW* forums. People use these Internet forums to upload postings and exchange opinions regarding particular topics of common interest [1, 31–35]. Therefore, such online forums are good sources of information to gauge the daily reactions of many users to certain MMOGs. Communities or forums are widely used in MMOGs for information exchange [31]. According to a previous study [1], *WoW* links its forums to economic activities among users, which we found to be relevant to predicting the fluctuations in the current number of virtual-world users. Topics and relevant replies posted by users on general discussion boards in official *WoW* forums were crawled. We also crawled the time when each comment

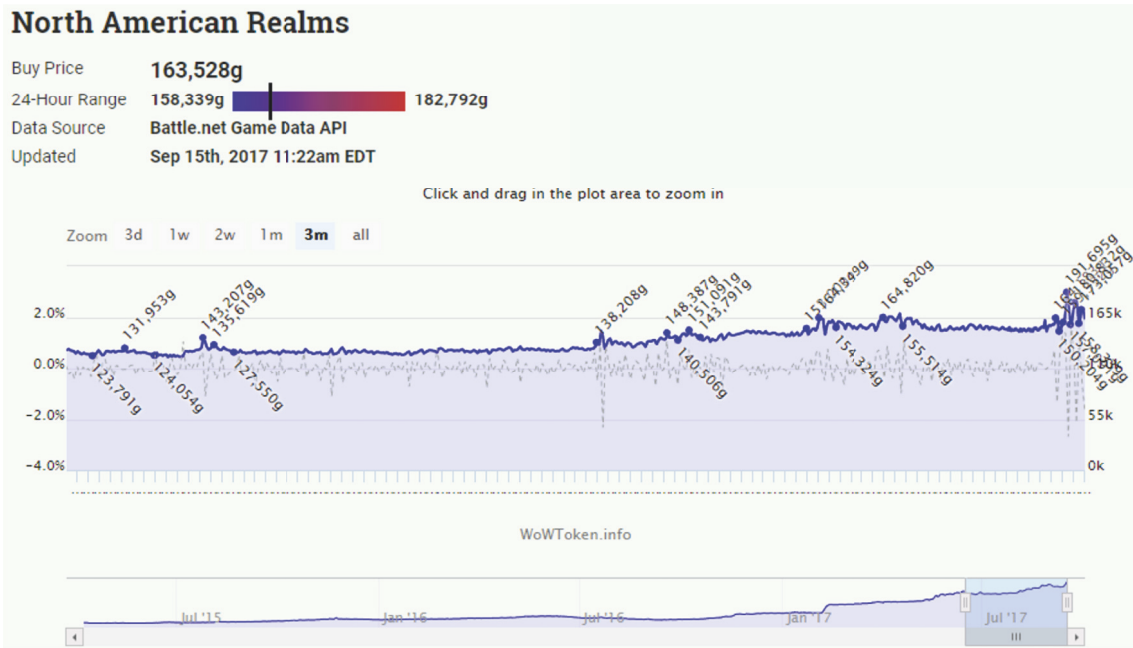


FIGURE 1: Token/Gold exchange rate example (from <http://wowtoken.info>).

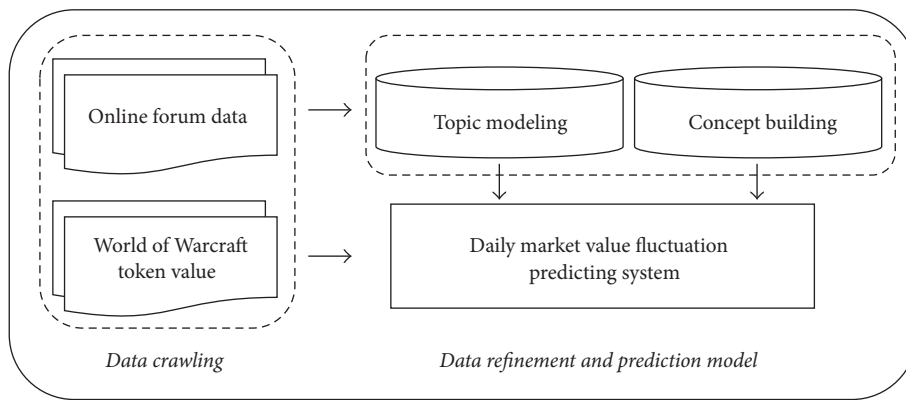


FIGURE 2: Overview of the proposed approach.

and reply to it were posted, the number of replies to each comment, and the number of views. Replies quoting previous comments and replies were crawled, excluding 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 and content, the replies to them were also extracted. The data were saved in json format, which was in turn converted to other formats (e.g., csv and xls) for different purposes. We collected data over a period of 460 days (April 23, 2015 to July 25, 2016). Within this period, the total number of topics collected was 166,651 (140,831 in the NA region and 25,820 in the EU region) and that of user replies was 2,931,748 (2,587,001 in the NA region and 344,747 in the EU region). We collected data in a manner that complied with the stipulated

terms and conditions of use of the forums. The collected data did not include any personal information.

WoW players use real currencies (e.g., USD and EUR) to purchase tokens and sell them for the virtual currency—Gold—in the auction house within the gaming environment. Tokens have only been in use for a short time; moreover, websites and tools continue to trace and show the value of the tokens. We employed one of these sites to crawl the daily closing price of the token. The use of such data does not violate WoW’s Terms of Use agreement or include users’ personal or identifying information.

4. Analysis of User Opinion Data

Our goal is to build lexicons of concepts, that is, meaningful keywords for predicting the price of the currency used in WoW from the crawled data. To this end, we initially ran

topic modeling using the entire user comments to extract the representative keywords, the subset of which will then be used for building our initial lexicons. Afterwards, we retrieved relevant keywords to these initial keywords that we selected based on the similarity measure via a kernel density estimation technique. The main idea behind the kernel density estimation is to compute the similarity score of a given word to each keyword in our lexicon using a Gaussian kernel function in a word embedding space and take the average value of these similarity scores. Those highly relevant keywords would have high similarity values computed in this manner. For more details, refer to the subsequent section.

4.1. Data Preprocessing. To remove noise or unnecessary information, we applied several preprocessing strategies for all user comments. First, we removed the URL string, stop words such as auxiliary verbs, prepositions, and special characters. Then, we tokenized the strings into words and lemmatized each word. In addition, we used only words with a frequency higher than f_{\min} , to exclude words used very sparsely. In this study, we set f_{\min} as 3.

4.2. Concept Building. Next, we built a lexicon representing a concept for our own purpose. The lexicon-based document analysis plays an important role in document analysis in various fields, such as economics, politics, and social sciences.

Our main goal is to find significantly relevant concepts existing in the entire document corpus by analyzing user comments. For example, a document in the WoW forum can be composed of concepts such as an item, a raid, and game contents. These concepts can be explained by those words corresponding to their lexicons, respectively. Hereby, we can quantitatively score how relevant a document is to each concept by analyzing how strongly the signal of the concept appears in it.

The lexicon building of a concept comprises two steps: (1) generation of a candidate word set that is potentially relevant to the concept and (2) manual refinement of it to finalize the lexicon. To extract candidate words for each concept, we constructed topic modeling with user comments. Topic modeling will be discussed later. Once representative keywords are generated for each topic in this manner, we used a union set of keywords from each topic as our candidate word set for concepts. Then, based on the prior knowledge of the features of virtual world [36–38] targeted by the authors who had a major related to game and the developers in the game companies making games in the MMOG genre, we selected words from candidate words and categorized them as some suitable lexicon for our desired concepts; each word is allowed to be assigned to multiple concepts. Thus, we can collect words that can explain the corresponding concept. Table 1 shows examples of these lexicons generated by this process.

4.3. Topic Modeling for Initial Lexicon Building. The topic modeling approach we used to extract representative keywords from a document corpus is nonnegative matrix factorization (NMF) [39], where the nonnegative constraint

TABLE 1: Concept building example created and used in this study.

Concepts	Keywords
General	Game, like, people, time, player, blizzard, play, think, server, ...
Patch	Legion, new, game, change, going, spec, patch, expansion, ...
Raid	Get, gear, raid, dungeon, time, heroic, group, guild, ...
Movie	Movie, Warcraft, legion, time, see, story, good, well, ...
Leveling	Player, level, quest, character, new, ability, reward, ...
PVP (player versus player)	Click, gladiator, game, time, gear, pvp, play, damage, ...
Gear	Get, cloak, glove, gear, belt, robe, damage, gauntlet, ...

gives the interpretability on the resulting weights from factor matrices as the relevance score of a word or a document to each topic.

In detail, given a term-document matrix $A \in \mathfrak{R}^{m \times n}$, where m and n represent the size of the vocabulary and the number of documents, respectively, we normalized each column of this matrix to have a unit L2-norm. Given this matrix, NMF approximately factorizes it into two matrices $W \in \mathfrak{R}^{m \times d}$ and $H \in \mathfrak{R}^{n \times d}$, where d represents the number of topics, that is,

$$\min_{W, H} \|A - WH^T\|_F, \quad \text{where } W \geq 0, H \geq 0. \quad (1)$$

In this equation, the subscript F indicates the Frobenius norm, for example, $\|A\|_F = (\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2)^{1/2}$ where $A \in \mathfrak{R}^{m \times n}$. NMF has nonnegative constraints, $W \geq 0, H \geq 0$, as shown in (1), which makes the element in W and H nonnegative, thus maintaining the interpretability. The columns in the resulting matrix, W , correspond to different topics, and the keywords correspond to the indices of the k largest value in each column function as the representative keywords of the topic.

In our analysis, we constructed a term-document matrix, A , from 140,831 threads in the US and 25,820 threads in the EU collected from the WoW forum. Each article has content and date features, such that we can calculate the scores per day by summing up the frequency of each word in sentences or articles generated in the corresponding day. In topic modeling, we set the number of topics, d , as 10. We also set the number of representative keywords for each topic, k , as 30.

4.4. Lexicon Expansion of Concepts and Relevance Score Calculation. In this subsection, we will describe details of how to expand concept keywords using kernel density estimation and to calculate the relevance scores with these concept keywords. Owing to the lack of expression resulting from the limited number of keywords a person could manage and the difficulty in determining the relevance to a concept in a user's

mind, we utilized kernel density estimation (KDE), a statistical metric to estimate the probability density function, which can be smoothed using multiple kernels, to infer what other words' concepts are. In other words, we selected some keywords for each concept, but the problem is that the number of keywords for each concept is relatively small compared to the number of the total vocabulary. To overcome this problem, we trained word embedding vector representations for all words using Word2Vec [40], which can provide semantically and syntactically meaningful vector representations. Afterwards, we calculated all distances between concept keywords and other words. Then, we computed the conditional probabilities or distributions for all words given each concept by utilizing KDE.

In particular, we adopted a Gaussian kernel adopts as follows [41]. For concept c , the conditional probability can be calculated by the distance function, d , which represents the distance between embedding vectors of words and vectors of word set in each concept, and the kernel, K , which ensures proper balance between the given word and others. The conditional probability of the embedding vector of the keyword, z , for a concept, c_i , that contains the embedding vector of concept keywords, $x_j \in c_i$, can be computed as follows:

$$p(z | c = c_i) = \frac{1}{\#(x \in c_i)} \sum_{x_j \in c_i} K(d(x_j, z)). \quad (2)$$

Here, we used Euclidean distance and Gaussian kernel and these equations are as follows.

$$d(x, y) = \|x - y\|_2 = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3)$$

$$K(d) = \exp\left(-\frac{d^2}{\gamma}\right), \quad \text{where } \gamma = 2\sigma^2$$

and we set the parameter, γ , as 1.

The conditional probability can also be considered as the relevance score to each class. Therefore, the score of the given sentence, s , for concept, c_i , is defined as

$$\text{score}_i(s) = \sum_{w_t \in s} p(w_t | c = c_i). \quad (4)$$

The range of the score is $[0, \infty)$ because the score of the sentence sums up all the scores of the words in the sentence, and all the probabilities are nonnegative. In practice, the sentence is not very long, and the score of each word is less than 1; therefore, the score of the sentence is not very large.

After scoring all the comments, we calculated scores for each day by summing up the articles generated in that day. Thus, we can obtain the scores for each day. Through this analysis, we found that the scores were rising in August 2015 and July 2016 when a new expansion pack and patch to change virtual world greatly was announced. Figure 3 shows the line graph of each score in each concept, as presented in Table 1 in some periods.

5. Causality Analysis

A Granger causality test [42] was conducted to assess the relation between standardized token prices and scores of each concept. The Granger causality test is based on the assumption that if variable X causes Y , then changes in X will consistently occur before changes in Y [1, 9]. We did not seek to test for actual causation but rather whether the time series of scores of each concept contained some predictive information regarding token prices.

Our time series for the selling price of tokens, denoted S_t , reflects daily changes in the price of tokens. We tested whether the time series of the collected data can predict changes in the token price by comparing the variance explained by two linear models. The first model uses only n lagged values of S_t (i.e., S_{t-1}, \dots, S_{t-n}) for prediction, while the second model uses n lagged values of both S_t and the time series of scores of each concept, denoted by X_{t-1}, \dots, X_{t-n} . We conducted the Granger causality test according to the models described in

$$S_t = \alpha + \sum_{i=1}^n \beta_i S_{t-i} + \epsilon_t \quad (5)$$

$$S_t = \alpha + \sum_{i=1}^n \beta_i S_{t-i} + \sum_{i=1}^n \gamma_i X_{t-i} + \epsilon_t.$$

Based on the results of the Granger causality tests, we rejected the null hypothesis that the time series of scores of each concept does not predict token prices—that is, $\beta \neq 0$ —with a high level of confidence. By analyzing their relevance, we extracted the concepts with the highest Granger causality relations (p value < 0.05).

6. Configuration of the Prediction Model

Using the collected data and the analyzed and rated comment data, we built machine learning models for predicting the fluctuation in the WoW token price using gradient boosting trees, random forest, and support vector machines (SVM), which are widely used in binary classification problems, used to investigate sensitivity to different machine learning algorithms. We used R packages called *xgboost* and *cforest* for the implementation in gradient boosting and random forest. The implementation of SVM was based on libSVM [43] with the radial basis function (RBF) kernel. Using LibSVM, the learning data was cross-validated to search for optimum parameters for the RBF kernel. We created a setup to apply machine learning to data spanning over a period of 460 days.

As the first step, we standardized the data to improve its applicability to the learning model. The z -score, $Z_{E_t} = (E - \bar{x}(E))/\sigma(E)$, where $\bar{x}(E)$ and $\sigma(E)$ represent the mean and the standard deviation for every date, respectively, of data for the previous 12 days ($t = 12$) was used. An example of applicable input data is shown in Table 2. As for the input nodes, based on the input data provided in Table 2, 10 input data points were represented as serial vectors to allocate neurons based on the cumulative number of days spent on learning; that is,



FIGURE 3: Scores of each concept for days.

TABLE 2: Example of a machine learning data set.

Data class	Date	KDE-based concept scoring data							Formal data		
		concept							Number of topics	Sum of replies	Sum of views
		General	PVP	Raid	Patch	Movie	Leveling	Gear			
Crawled and analyzed data	May 1, 2016	A	B	C	D	E	F	G	V	W	X
Input learning data	May 1, 2016	Z_{A_t}	Z_{B_t}	Z_{C_t}	Z_{D_t}	Z_{E_t}	Z_{F_t}	Z_{G_t}	Z_{V_t}	Z_{W_t}	Z_{X_t}

TABLE 3: Statistical significance (p values) of bivariate Granger causality test between WoW token price (NA Region) and concepts of forum opinions.

Time lag	WoW token price								
	General	PVP	Raid	Patch	Movie	Leveling	Gear	Number of topics	Number of replies
1 day	0.0116	0.0032	0.0089	0.0074	0.0052	0.0073	0.0034	0.0016	0.0018
2 days	0.0286	0.0059	0.0138	0.0243	0.0144	0.0131	0.0049	0.0073	0.0169
3 days	0.0181	0.0035	0.0081	0.0091	0.0177	0.0075	0.0156	0.0163	0.0252
4 days	0.0494	0.0088	0.0163	0.0315	0.0174	0.0161	0.0332	0.0246	0.0391
5 days	0.0886	0.0138	0.0603	0.0363	0.0657	0.0346	0.0221	0.0286	0.0683
6 days	0.1245	0.0476	0.0813	0.121	0.1109	0.0676	0.0061	0.0341	0.1244
7 days	0.2668	0.1325	0.2188	0.3087	0.2317	0.2106	0.0322	0.0665	0.4006

TABLE 4: Statistical significance (p values) of bivariate Granger causality test between WoW token price (EU Region) and concepts of forum opinions.

Time Lag	WoW token price								
	General	PVP	Raid	Patch	Movie	Leveling	Gear	Number of topics	Number of replies
1 day	0.4323	0.1714	0.4791	0.1193	0.1806	0.3695	0.1352	0.0026	0.2684
2 days	0.5641	0.2162	0.5267	0.1496	0.2411	0.4839	0.1117	0.0031	0.1178
3 days	0.1518	0.0499	0.1643	0.0647	0.0812	0.1301	0.0515	0.0107	0.1542
4 days	0.0924	0.011	0.1257	0.0148	0.0214	0.0722	0.0132	0.0255	0.2404
5 days	0.1384	0.0176	0.2236	0.0303	0.0524	0.1185	0.0067	0.0221	0.2501
6 days	0.2431	0.0427	0.3235	0.061	0.0896	0.2303	0.0114	0.0143	0.2077
7 days	0.2766	0.0391	0.2266	0.0337	0.114	0.2278	0.0062	0.0076	0.1791

20, 30, 50, and 700 neurons were allocated to cumulative 2, 3, 5, and 7 days.

7. Experimental Results

7.1. Results of Granger Causality Test. The Granger causality test was conducted on the Bitcoin transaction count and the price for time lag of 1 to 7 days. The time lag was omitted because it produced less significant results after 8 days. Tables 3 and 4 list the test results.

From the results, it was observed that in the NA region, most user opinions had an influence when the time lag was small and user opinions related to gear and PVP had causation when the time lag became larger. In the EU region, it was observed that user opinions related to gear, PVP, and patches had causation. This process was only used for verification. The entire data set was used to build the actual learning model for prediction.

7.2. Prediction Results. We built and applied the machine learning model based on the gathered and scoring data to

predict the daily fluctuation of the WoW token price. From April 23, 2015, to July 25, 2016, tests are conducted with 10-fold cross validation. The accuracy rate, Matthews correlation coefficient (MCC), and F -measure were used to evaluate the performance of the proposed model.

Table 5 presents the prediction results. The most accurate prediction model for the WoW token price in the NA region (accuracy rate = 82.55%) is based on the gradient boosting and previous seven days' learning data. The most accurate prediction model for the WoW token price in the EU region (accuracy rate = 81.52%) is based on the gradient boosting and the previous twelve days' learning data. Table 5 presents the results relative to the different machine learning models and learning data structures. Cumulative learning data for seven days or longer resulted in negligible difference, and cumulative learning data for less than five days proved insufficient for learning and compromised the prediction accuracy.

8. Discussion and Conclusion

In this paper, we presented a method for predicting the value fluctuation of the virtual currency in an MMOG, a

TABLE 5: Experimental results of predicted WoW token prices fluctuation.

Region	Learning method	Gradient boosting			Random forest			Support vector machine		
	Learning days	Accuracy (%)	F1-score	MCC	Accuracy (%)	F1-score	MCC	Accuracy (%)	F1-score	MCC
North America region	3 days	75.65%	0.7576	0.5731	69.13%	0.6984	0.4799	64.13%	0.6432	0.3445
	5 days	79.56%	0.7955	0.5921	76.08%	0.7784	0.5704	74.48%	0.7539	0.318
	7 days	82.55%	0.8255	0.6372	78.26%	0.793	0.5768	76.95%	0.7765	0.3106
	12 days	80.65%	0.8045	0.6009	78.69%	0.7858	0.579	78.69%	0.7996	0.3773
Europe region	3 days	69.56%	0.6957	0.3917	71.52%	0.7138	0.4375	75.43%	0.7966	-0.0219
	5 days	71.73%	0.7174	0.4346	73.91%	0.7367	0.4907	70.65%	0.6225	0.2742
	7 days	77.17%	0.7718	0.5429	77.39%	0.77	0.5723	71.3%	0.6346	0.2989
	12 days	81.52%	0.815	0.6322	77.6%	0.7791	0.5779	73.91%	0.6656	0.3416

subject that had previously been minimally investigated. Our results demonstrated that the value fluctuation of the virtual currency in the WoW MMOG environment can be predicted. This paper demonstrated that the proposed method could be applied to virtual currency selling/buying.

The results of refined user data were used in the predictions, which revealed that users' opinions could be effectively used in currency value fluctuation predictions. Granger causality test results showed that users' opinions affect the token value, regardless of the region of the world.

The proposed method provides developers with market price trends and enables them to determine the numerical values and probabilities of multiple effects that would otherwise be difficult to identify. These include users' economic activities in MMOG environments and the relevance of currency value fluctuations, which are conducive to adjusting the overall balance in the MMOG and thus improving the environment. Validated currency value fluctuation predictions would enable users to pursue profits, in a narrower sense, and to perceive the overall flows of a given virtual currency in a broader sense.

The proposed prediction system could be improved by taking the following aspects into consideration. First, a more sophisticated user characterization could yield more revealing findings. For example, a considerable number of reasons exist for users playing games [44], and these reasons are likely to be relevant to their transaction activities. An analysis of user motives and their relation to transaction activities could improve the understanding of economic activities in MMOG settings. In addition, data on factors, such as changes in the gaming environment, virtual currency system updates, and user comments, including chats, are worth analyzing to thoroughly understand the MMOG setting and increase prediction accuracy. Furthermore, if a sufficiently long time period of data could be gathered, multivariate time series analysis will ensure competitive results. Our plan is to enhance our prediction system in future research to incorporate these considerations with improved reliability and efficiency of the proposed method.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2015R1A1A1A05001196, NRF-2016R1E1A2A02946052, and NRF-2017R1A2B2005380), by an Institute for Information & Communications Technology Promotion (IITP) grant funded by the Korean government (MSIP; 2016-0-00285, High performance computing [HPC] based rendering solution development), and by Linewalks Inc.

References

- [1] Y. B. Kim, S. H. Lee, S. J. Kang, M. J. Choi, J. Lee, and C. H. Kim, "Virtual world currency value fluctuation prediction system based on user sentiment analysis," *PLoS ONE*, vol. 10, no. 8, Article ID e0132944, 2015.
- [2] D. H. Shin, "Understanding purchasing behaviors in a virtual economy: Consumer behavior involving virtual currency in Web 2.0 communities," *Interacting with Computers*, vol. 20, no. 4-5, pp. 433-446, 2008.
- [3] Y. Guo and S. Barnes, "Why people buy virtual items in virtual worlds with real money," *ACM SIGMIS Database*, vol. 38, no. 4, pp. 69-76, 2007.
- [4] S. Papagiannidis, M. Bourlakis, and F. Li, "Making real money in virtual worlds: MMORPGs and emerging business opportunities, challenges and ethical implications in metaverses," *Technological Forecasting & Social Change*, vol. 75, no. 5, pp. 610-622, 2008.
- [5] S. Scarle, S. Arnab, I. Dunwell, P. Petridis, A. Protopsaltis, and S. de Freitas, "E-commerce transactions in a virtual environment: Virtual transactions," *Electronic Commerce Research*, vol. 12, no. 3, pp. 379-407, 2012.
- [6] V. Lehdonvirta, "Virtual item sales as a revenue model: Identifying attributes that drive purchase decisions," *Electronic Commerce Research*, vol. 9, no. 1-2, pp. 97-113, 2009.
- [7] D. Y. Wohn, "Spending real money: Purchasing patterns of virtual goods in an online social game," in *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems, CHI 2014*, pp. 3359-3368, May 2014.
- [8] T. Kimoto, K. Asakawa, M. Yoda, and M. Takeoka, "Stock market prediction system with modular neural networks," in *Proceedings of the 1990 International Joint Conference on Neural*

- Networks (IJCNN '90)*, vol. 1, pp. 1–6, Washington, DC, USA, June 1990.
- [9] J. Bollen, H. Mao, and X. Zeng, “Twitter mood predicts the stock market,” *Journal of Computational Science*, vol. 2, no. 1, pp. 1–8, 2011.
 - [10] I. Bordino, S. Battiston, G. Caldarelli, M. Cristelli, A. Ukkonen, and I. Weber, “Web search queries can predict stock market volumes,” *PLoS ONE*, vol. 7, no. 7, Article ID e40014, 2012.
 - [11] Y. Cohen-Charash, C. A. Scherbaum, J. D. Kammeyer-Mueller, and B. M. Staw, “Mood and the market: can press reports of investors’ mood predict stock prices?” *PLoS ONE*, vol. 8, no. 8, Article ID e72031, 2013.
 - [12] R. P. Schumaker and H. Chen, “Textual analysis of stock market prediction using breaking financial news: The AZFin text system,” *ACM Transactions on Information and System Security*, vol. 27, no. 2, article no. a12, 2009.
 - [13] Z. Yudong and W. Lenan, “Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network,” *Expert Systems with Applications*, vol. 36, no. 5, pp. 8849–8854, 2009.
 - [14] E. Kosminsky, “World of Warcraft: the viability of massively multiplayer online role-playing games as platforms for modeling and evaluating perfect competition,” *Journal For Virtual Worlds Research*, vol. 2, 2009.
 - [15] Pardus Massive Multiplayer Online Browser Game, <https://www.pardus.at/>.
 - [16] Y. B. Kim, N. Park, Q. Zhang, J. G. Kim, S. J. Kang, and C. H. Kim, “Predicting virtual world user population fluctuations with deep learning,” *PLoS ONE*, vol. 11, no. 12, Article ID e0167153, 2016.
 - [17] O. Mryglod, B. Fuchs, M. Szell, Y. Holovatch, and S. Thurner, “Interevent time distributions of human multi-level activity in a virtual world,” *Physica A: Statistical Mechanics and its Applications*, vol. 419, pp. 681–690, 2015.
 - [18] M. Szell, R. Lambiotte, and S. Thurner, “Multirelational organization of large-scale social networks in an online world,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 31, pp. 13636–13641, 2010.
 - [19] M. Szell, R. Sinatra, G. Petri, S. Thurner, and V. Latora, “Understanding mobility in a social petri dish,” *Scientific Reports*, vol. 2, article no. 457, 2012.
 - [20] M. Szell and S. Thurner, “Measuring social dynamics in a massive multiplayer online game,” *Social Networks*, vol. 32, no. 4, pp. 313–329, 2010.
 - [21] M. Szell and S. Thurner, “Social dynamics in a large-scale online game,” *Advances in Complex Systems (ACS)*, vol. 15, no. 6, Article ID 1250064, 2012.
 - [22] S. Thurner, M. Szell, and R. Sinatra, “Emergence of good conduct, scaling and zipf laws in human behavioral sequences in an online world,” *PLoS ONE*, vol. 7, no. 1, Article ID e29796, 2012.
 - [23] W. S. Bainbridge, “The scientific research potential of virtual worlds,” *Science*, vol. 317, no. 5837, pp. 472–476, 2007.
 - [24] P. R. Messinger, E. Stroulia, and K. Lyons, “Virtual worlds—past, present, and future: new directions in social computing,” *Decision Support Systems*, vol. 47, no. 3, pp. 204–228, 2009.
 - [25] B. Corominas-Murtra, B. Fuchs, and S. Thurner, “Detection of the elite structure in a virtual multiplex social system by means of a generalised K-core,” *PLoS ONE*, vol. 9, no. 12, Article ID e112606, 2014.
 - [26] B. Fuchs and S. Thurner, “Behavioral and network origins of wealth inequality: Insights from a virtual world,” *PLoS ONE*, vol. 9, no. 8, Article ID e103503, 2014.
 - [27] S.-J. Kang, Y. B. Kim, T. Park, and C.-H. Kim, “Automatic player behavior analysis system using trajectory data in a massive multiplayer online game,” *Multimedia Tools and Applications*, vol. 66, no. 3, pp. 383–404, 2013.
 - [28] P. Klimek and S. Thurner, “Triadic closure dynamics drives scaling laws in social multiplex networks,” *New Journal of Physics*, vol. 15, article 063008, 2013.
 - [29] M. Szell and S. Thurner, “How women organize social networks different from men,” *Scientific Reports*, vol. 3, article no. 1214, 2013.
 - [30] W.-J. Xie, M.-X. Li, Z.-Q. Jiang, and W.-X. Zhou, “Triadic motifs in the dependence networks of virtual societies,” *Scientific Reports*, vol. 4, article no. 5244, 2014.
 - [31] Y. S. Hau and Y.-G. Kim, “Why would online gamers share their innovation-conducive knowledge in the online game user community? Integrating individual motivations and social capital perspectives,” *Computers in Human Behavior*, vol. 27, no. 2, pp. 956–970, 2011.
 - [32] P. Panzarasa, T. Opsahl, and K. M. Carley, “Patterns and dynamics of users’ behavior and interaction: Network analysis of an online community,” *Journal of the Association for Information Science and Technology*, vol. 60, no. 5, pp. 911–932, 2009.
 - [33] C. C. Sing and M. S. Khine, “An analysis of interaction and participation patterns in online community,” *Journal of Educational Technology and Society*, vol. 9, article 250, 2006.
 - [34] Y. B. Kim, J. G. Kim, W. Kim et al., “Predicting fluctuations in cryptocurrency transactions based on user comments and replies,” *PLoS ONE*, vol. 11, no. 8, Article ID e0161197, 2016.
 - [35] Y. B. Kim, J. Lee, N. Park et al., “When Bitcoin encounters information in an online forum: Using text mining to analyse user opinions and predict value fluctuation,” *PLoS ONE*, vol. 12, no. 5, p. e0177630, 2017.
 - [36] V. H.-H. Chen, H. B.-L. Duh, and H. Renyi, “The Changing Dynamic of Social Interaction in World of Warcraft: The Impacts of Game Feature Change,” in *Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology, ACE 2008*, pp. 356–359, December 2008.
 - [37] A. Golub, “Being in the World (of Warcraft): Raiding, realism, and knowledge production in a massively multiplayer online game,” *Anthropological Quarterly*, vol. 83, no. 1, pp. 17–46, 2010.
 - [38] C. A. Paul, “Welfare epics? The rhetoric of rewards in world of warcraft,” *Games and Culture*, vol. 5, no. 2, pp. 158–176, 2010.
 - [39] J. Choo, C. Lee, C. K. Reddy, and H. Park, “UTOPIAN: User-driven topic modeling based on interactive nonnegative matrix factorization,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 1992–2001, 2013.
 - [40] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient estimation of word representations in vector space,” arXiv preprint <https://arxiv.org/abs/1301.3781>.
 - [41] D. Park, S. Kim, J. Lee, J. Choo, N. Diakopoulos, and N. Elmquist, “ConceptVector: text visual analytics via interactive lexicon building using word embedding,” *IEEE Transactions on Visualization and Computer Graphics*, 2017.
 - [42] C. Hiemstra and J. D. Jones, “Testing for linear and nonlinear Granger causality in the stock price-volume relation,” *The Journal of Finance*, vol. 49, no. 5, pp. 1639–1664, 1994.

- [43] C. Chang and C. Lin, "LIBSVM: a Library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, article 27, 2011.
- [44] J. Billieux, M. Van der Linden, S. Achab et al., "Why do you play World of Warcraft? An in-depth exploration of self-reported motivations to play online and in-game behaviours in the virtual world of Azeroth," *Computers in Human Behavior*, vol. 29, no. 1, pp. 103–109, 2013.




Hindawi

Submit your manuscripts at
<https://www.hindawi.com>

