

Research Article

Visualization of Online-Game Players Based on Their Action Behaviors

Ruck Thawonmas¹ and Keita Iizuka^{1,2}

¹ Intelligent Computer Entertainment Laboratory, Graduate School of Science and Engineering, Ritsumeikan University, Kusatsu, Shiga 525-8577, Japan

² Solution Development Team, Solution Development Department, Bandai Networks Co., Ltd., Tokyo 111-8081, Japan

Correspondence should be addressed to Ruck Thawonmas, ruck@ci.ritsumei.ac.jp

Received 1 February 2008; Revised 8 May 2008; Accepted 11 June 2008

Recommended by Jouni Smed

We propose a visualization approach for analyzing players' action behaviors. The proposed approach consists of two visualization techniques: classical multidimensional scaling (CMDS) and KeyGraph. CMDS is for discovering clusters of players who behave similarly. KeyGraph is for interpreting action behaviors of players in a cluster of interest. In order to reduce the dimension of matrices used in computation of the CMDS input, we exploit a time-series reduction technique recently proposed by us. Our visualization approach is evaluated using log of an online game where three-player types according to Bartle's taxonomy are found, that is, achievers, explorers, and socializers.

Copyright © 2008 R. Thawonmas and K. Iizuka. 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.

1. INTRODUCTION

The market size of online games continues to experience surging growth [1]. At the same time, competitions among them are also becoming very high. The quality of player service plays an important role in winning such competitions. It is therefore inevitable to online-game developers and publishers to know their player behaviors so that they can develop game contents that fulfill player demands. Visualization techniques have been recently applied to discover in-game player behaviors.

Most work in the literature focuses on visualization of player trails or time series of visited locations for examining the distance over time among the members of a social group [2], discovering playing strategies in a combat game [3], and analyzing movement patterns [4]. Other work focuses on extracting pathways [5] and on locating clusters of similar players based on their movement patterns [6].

This research, however, focuses on visualizing player behaviors based on their actions. According to Bartle's taxonomy [7], online-game players can be typically identified based on their action behaviors into achievers, explorers,

killers, and socializers. Player-type information should, therefore, be exploited to provide game contents that players favor, for example, a wider variety of collectable items for achievers, longer missions for explorers, more hunting opportunities for killers, and a higher frequency of social events for socializers. It has been recently reported in [8] that Bartle's taxonomy is also applicable to social data in a web-based application.

In this paper, we propose an approach for visualizing players' action behaviors using classical multidimensional scaling (CMDS) [9] and KeyGraph [10], both described in Section 2. First, CMDS is used for locating clusters of similarly behaving players. KeyGraph is then used for interpreting playing behaviors of players in a cluster of interest. The input to CMDS is derived based on time-series matrices of players' action sequences which are needlessly long due to noise and redundancy, leading to high computational cost. We, therefore, compute the CMDS input based on reduced time-series matrices obtained by our recently proposed time-series reduction technique in [11]. To make this paper self-contained, this technique is described in Section 3. Evaluation of our visualization approach is

given in Section 4, where achievers, explorers, and socializers are found in play log from an online-game used in the evaluation.

2. PLAYER VISUALIZATION

In this section, we describe CMDS, KeyGraph, log format, and visualization metrics. As with most other tools for information visualization [12], subjective interpretation is required for KeyGraph. The described visualization metrics are used for facilitating this task in Section 4.2.

2.1. Multidimensional scaling

CMDS is a prevailing technique for mapping pair-wise relationships to coordinates and has been applied to several areas such as statistics, psychology, sociology, political sciences, and marketing [9]. Recently, this technique has been successfully applied to clustering of online-game players based on their movement patterns [6]. CMDS takes as its input matrix \mathbf{D} , indicating dissimilarities between player pairs, and outputs a coordinate matrix whose configuration minimizes a loss function in preserving all interpoint distances. Two time series of interest are considered similar if they have similar rise and fall patterns, although they might have different scales on the time axis. A good measurement for deriving the distance or dissimilarities between such series is the dynamic time warping (DTW) distance [13].

In our research, the ij th element in \mathbf{D} is the DTW distance between the reduced time-series matrices of action sequences of players i and j . In addition, we use the function `cmdscale` in the Statistical Toolbox of Matlab for performing CMDS and select only the first two dimensions of the constructed coordinates for plotting players.

2.1.1. Action coding

Here, we describe how action sequences are numerically coded into time-series matrices for computation of DTW distances. Let O denote the set of action symbols of interest and $|O|$ its cardinality. Action sequence $s = s(1), s(2), \dots, s(L)$ is numerically coded into $|O| \times L$ time-series matrix $\mathbf{X} = [X(1), X(2), \dots, X(L)]$, where $X(i)$ is a column vector with the element indexing the action symbol of $s(i)$ being 1 and other elements 0.

Consider, for example, the set of action symbols $O = \{A, B, C\}$, and thus $|O| = 3$, where symbols A, B, and C are represented by column vectors $[100]^t$, $[010]^t$, and $[001]^t$. In an action sequence such as $s = C, A, B, C$, it is coded into $\mathbf{X} = [[001]^t, [100]^t, [010]^t, [001]^t,]$.

2.1.2. Dynamic time warping

The DTW distance between time-series matrices \mathbf{X} and \mathbf{Y} , $\text{dtw}(\mathbf{X}, \mathbf{Y})$, having lengths L_X and L_Y , is defined as follows:

$$\text{dtw}(\mathbf{X}, \mathbf{Y}) = g(L_X, L_Y), \quad (1)$$

| | A | C | B | E | F | | A | B | C | E | D | F | B |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |

FIGURE 1: Time-series matrices \mathbf{X} and \mathbf{Y} .

| | | | | | | | |
|---|-----|-----|-----|-----|-----|-----|-----|
| F | 5.6 | 4.2 | 4.2 | 4.2 | 4.2 | 4.2 | 5.6 |
| E | 4.2 | 2.8 | 2.8 | 2.8 | 4.2 | 5.6 | 7 |
| B | 2.8 | 1.4 | 2.8 | 2.8 | 4.2 | 5.6 | 5.6 |
| C | 1.4 | 1.4 | 1.4 | 2.8 | 4.2 | 5.6 | 7 |
| A | 0 | 1.4 | 2.8 | 4.2 | 5.6 | 7 | 8.4 |
| | A | B | C | E | D | F | B |

FIGURE 2: Derivation of dynamic time warping distance between \mathbf{X} and \mathbf{Y} .

where

$$g(i, j) = \min \begin{cases} g(i, j-1) + d(i, j), \\ g(i-1, j-1) + d(i, j), \\ g(i-1, j) + d(i, j), \end{cases} \quad (2)$$

$$g(i, 0) = \begin{cases} 0 & i = 0, \\ \infty & i > 0, \end{cases}$$

$$g(0, j) = \begin{cases} 0 & j = 0, \\ \infty & j > 0, \end{cases}$$

and $d(i, j)$ is the Euclidean distance between $X(i)$ and $Y(j)$.

Consider, for example, the set of symbols $O = \{A, B, C, D, E, F\}$ and two action sequences $x = A, C, B, E, F$ and $y = A, B, C, E, D, F, B$. The DTW distance between corresponding time-series matrices \mathbf{X} and \mathbf{Y} (c.f., Figure 1), $\text{dtw}(\mathbf{X}, \mathbf{Y})$, is 5.6, derived as shown in Figure 2.

2.2. KeyGraph

KeyGraph is a visualization tool for discovery of relations among text-based data. Its underlying concept is based on a building construction metaphor. As shown in [10], the precision-recall curve of KeyGraph is superior to TFIDF and NGRAM [14], well-known techniques for information retrieval, in extraction of correct keywords from a set of documents. KeyGraph has been later applied to visualize the relations among Web pages, among products in markets, among earthquake faults, and so forth [15]. It has also been successfully applied to identification of player types in an online-game simulator [16].

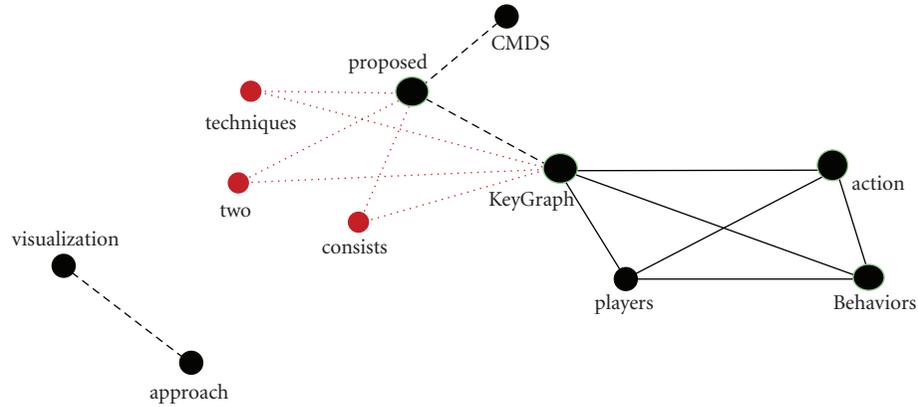


FIGURE 3: KeyGraph applied to the abstract of this paper.

Three major components of KeyGraph are as follows.

- (1) Foundations: subgraphs of highly associated and frequent terms representing basic concepts in the data.
- (2) Roofs: terms highly associated with foundations.
- (3) Columns: associations between foundations and roofs used for extracting keywords or main concepts in the data.

Associations between terms are defined as the co-occurrence among them in same sentences, and keywords are the terms in either foundations or roofs that are connected to strong columns. Under KeyGraph representation, solid lines and their touching black nodes depict foundations, dotted lines depict columns, red nodes depict roofs excluding those in the foundations, and double circles depict keywords. We use a tool called Polaris [17], publicly available, for generating KeyGraphs (<http://www.chokkan.org/software/dist/polaris-0.19alpha.zip>).

Figure 3 shows an example of KeyGraph when it is applied to the text data taken from the abstract of this paper, where common preprocessing for text data such as removing of conjunctions, determiners, and prepositions is performed. From this figure, it can be seen that there is one foundation consisting of four terms, that is, “KeyGraph,” “action,” “behaviors,” and “players.” The first three terms are also keywords. Another keyword is “proposed.” Three roof terms are “consists,” “two,” and “techniques.” These terms well represent the messages in the abstract.

2.3. Log format and visualization metrics

Player action sequences in our work are sequences of action symbols extracted from game log, of an online game discussed in Section 4, that has the following format:

$$\text{time stamp} \parallel \text{player ID} \parallel \text{event} \parallel \text{start position} \parallel \text{stop position} \parallel, \quad (3)$$

where an event consists of an action and its object, if any. This format is based on those adopted in an MMOG simulator

in [18] and 3D virtual worlds in [2]. In commercial online games, this kind of game log is stored in the monitoring database [19].

According to a recent work in [20], there are three categories of play motivations in online games as follows.

- (i) Achievement consisting of three subcomponents, that is, advancement, mechanics, and competition.
- (ii) Social consisting of three subcomponents, that is, socializing, relationship, and teamwork.
- (iii) Immersion consisting of four subcomponents, that is, discovery, role-playing, customization, and escapism.

The achievement, social, and immersion categories correspond to Bartle’s achievers, socializers, and explorers, respectively, although the above ten motivations overlap among player types.

In our work, we focus in particular on

- (i) advancement described in [20] as the desire to gain power progresses rapidly and accumulates in-game symbols of wealth or status,
- (ii) socializing described in [20] as having an interest in helping and chatting with other players, and
- (iii) discovery described in [20] as finding and knowing things that most other players do not know about.

This is because we anticipate that they should be identifiable using our action sequences and KeyGraphs. Below, we verify this anticipation with simplified data sets and their KeyGraphs, which serve as our visualization metrics for facilitating interpretation of KeyGraph results in Section 4.2.

Let us consider a set of action symbols $\{c, w, m, n, r\}$, standing for chat, walk, interaction with a mission master, interaction with a nearby object (item, NPC, or monster), and interaction with a remote object, respectively. The symbol w is a fundamental action and thus should be a frequent symbol in all action sequences. It is therefore removed from our consideration. For achievers motivated by advancement, interactions with mission masters should

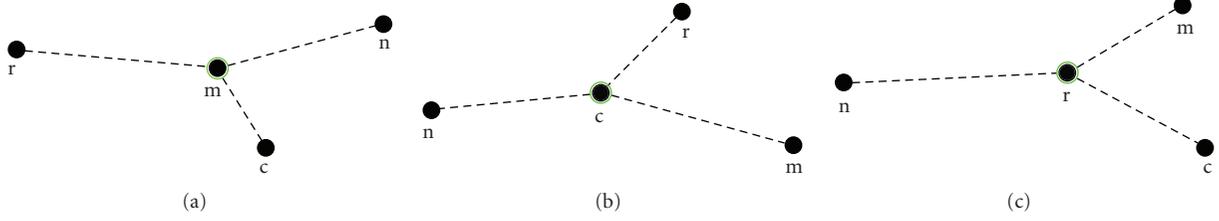


FIGURE 4: (a) KeyGraphs for achievers, (b) socializers, and (c) explorers generated from the given sample data sets, where m (interaction to a mission master), c (chat), and r (interaction to a remote object) are the keywords in (a), (b), and (c), respectively.

be frequently seen in their action sequences, and thus all possible sets of frequent action symbols for them are $\{c, m\}$, $\{m, n\}$, $\{m, r\}$, $\{c, m, n\}$, $\{c, m, r\}$, $\{m, n, r\}$, and $\{c, m, n, r\}$. For socializers motivated by socializing, chats should be commonly seen among them, leading to sets of frequent action symbols $\{c, m\}$, $\{c, n\}$, $\{c, r\}$, $\{c, m, n\}$, $\{c, m, r\}$, $\{c, n, r\}$, and $\{c, m, n, r\}$. For explorers motivated by discovery, interactions with remote objects should be commonly seen among them and thus their sets of frequent action symbols should be $\{c, r\}$, $\{m, r\}$, $\{n, r\}$, $\{c, m, r\}$, $\{c, n, r\}$, $\{m, n, r\}$, and $\{c, m, n, r\}$.

Figure 4 shows the resulting KeyGraphs for the three player types, where each KeyGraph was generated from the data sets of the corresponding player type. Note that m, c, and r are the keyword nodes in the KeyGraphs of achiever, socializer, and explorer, respectively, and henceforth these findings are used as visualization metrics.

3. TIME-SERIES REDUCTION

Our technique for obtaining compact sequences representing major player behaviors is based on Haar wavelet transform [21]. The Haar wavelet transform technique has a wide range of time-series applications including classification of DNA sequences [22], which motivated us to apply the technique to action sequences. Below, we first give an outline of the Haar wavelet transform and then describe the time-series reduction technique.

In the wavelet transform concept, decomposition involves obtaining wavelet coefficients from a sequence of interest. Reconstruction involves recovering the original sequence from obtained coefficients. Henceforth, it is assumed that the length L of a sequence is adjusted so that L is a power of 2 and $q = \log_2(L)$. The i th Haar wavelet coefficient at resolution order k , $d_{(k,i)}$, is derived as

$$d_{(k,i)} = \frac{x_{(k+1,2i-1)} - x_{(k+1,2i)}}{2}, \quad (4)$$

where $x_{(k,i)} = (x_{(k+1,2i-1)} + x_{(k+1,2i)})/2$ is the i th average at order k between two corresponding adjacent values in order $k+1$. With this representation, $k_{\max} = q$, the original sequence is represented by $x = x_{(q,1)}, x_{(q,2)}, \dots, x_{(q,L)}$. An example of Haar wavelet decomposition of the sequence 6, 8, 2, 7, 6, 5, 4, 3 is shown in Table 1.

TABLE 1: Example of Haar wavelet transform.

| Resolution | Averages $x_{(k,i)}$ | Coefficients $d_{(k,i)}$ |
|------------|--------------------------|--------------------------|
| $k = 4$ | (6, 8, 2, 7, 6, 5, 4, 3) | — |
| $k = 3$ | (7.0, 4.5, 5.5, 3.5) | (-1.0, -2.5, 0.5, 0.5) |
| $k = 2$ | (5.75, 4.5) | (1.25, 1) |
| $k = 1$ | (5.125) | (0.625) |

| | A | B | C | E | D | F | B | A |
|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |

FIGURE 5: Time-series matrix X .

Reconstruction of a given sequence from its Haar wavelet coefficients and averages is done as follows:

$$\begin{aligned} x_{(k,2i-1)} &= x_{(k-1,i)} + d_{(k-1,i)}, \\ x_{(k,2i)} &= x_{(k-1,i)} - d_{(k-1,i)}. \end{aligned} \quad (5)$$

Now, we describe our procedure for reducing the length of the time-series matrix of an action sequence of interest. For explanation, we use action sequence $x = A, B, C, E, D, F, B, A$ as an example, where $O = \{A, B, C, D, E, F\}$ and thus $|O| = 6$.

(i) Derive time-series matrix X for action sequence x of interest having length L . Figure 5 shows resulting time-series matrix X in our example.

(ii) Decompose each row in X to obtain Haar wavelet coefficients. Figure 6 shows resulting averages and coefficients for X in our example.

(iii) Reconstruct each row in X with selected Haar wavelet coefficients as follows.

Reconstruction of each row in X starts from the coefficient at the lowest resolution order, that is, $d(1,1)$, to those at the next higher order, and so forth. At a given resolution order, when the number of remaining coefficients is less than the number of coefficients in that order, they are selected

| | | | | | | | |
|-------|--------|-------|-------|------|------|------|------|
| 0.25 | 0 | 0.25 | -0.25 | 0.5 | 0 | 0 | -0.5 |
| 0.25 | 0 | 0.25 | -0.25 | -0.5 | 0 | 0 | 0.5 |
| 0.125 | 0.125 | -0.25 | 0 | 0 | 0.5 | 0 | 0 |
| 0.125 | -0.125 | 0 | 0.25 | 0 | 0 | 0.5 | 0 |
| 0.125 | 0.125 | -0.25 | 0 | 0 | -0.5 | 0 | 0 |
| 0.125 | -0.125 | 0 | 0.25 | 0 | 0 | -0.5 | 0 |

Average ← Coefficients →

$k = 1$ $k = 2$ $k = 3$

FIGURE 6: \mathbf{X} after decomposition.

| | | | | | | | |
|-------|--------|-------|-------|---|---|------|---|
| 0.25 | 0 | 0.25 | -0.25 | 0 | 0 | 0 | 0 |
| 0.25 | 0 | 0.25 | -0.25 | 0 | 0 | 0 | 0 |
| 0.125 | 0.125 | -0.25 | 0 | 0 | 0 | 0 | 0 |
| 0.125 | -0.125 | 0 | 0.25 | 0 | 0 | 0.5 | 0 |
| 0.125 | 0.125 | -0.25 | 0 | 0 | 0 | 0 | 0 |
| 0.125 | -0.125 | 0 | 0.25 | 0 | 0 | -0.5 | 0 |

Average ← Coefficients →

$k = 1$ $k = 2$ $k = 3$

FIGURE 7: \mathbf{X} after coefficient selection.

based on their total energy value in decreasing order, where total energy of $d_{(k,i)}$, $E_{(k,i)}$, is defined as

$$E_{(k,i)} = \sum_{n=1}^{|O|} d_{(n,k,i)}^2, \quad (6)$$

where $d_{(n,k,i)}$ is $d_{(k,i)}$ decomposed at row n of \mathbf{X} . All other unselected coefficients are then reset to zero. Figure 7 shows resulting \mathbf{X} after selection of four coefficients in our example. Following the recipe in [21], the number of Haar wavelet coefficients used in our performance evaluation is heuristically set to $\min(L - 1, \lfloor \log_2 L \times 4 \rfloor)$.

(iv) Reconstruct \mathbf{X} with the above coefficients (c.f., Figure 8 for our example).

(v) Reduce the size of \mathbf{X} by sampling down a group of repetitive and consecutive elements at each reconstructed row to one element. Figure 9 shows the reduced \mathbf{X} in our example.

Note that the DTW distance between the reduced time-series matrices \mathbf{X} of players i and j is assigned to the ij th element of the CMDS input matrix \mathbf{D} discussed in Section 2.1.

4. EVALUATION

4.1. Settings

We obtained player log from the online game The ICE [23], under development at our laboratory. A screen shot of The ICE and the game map in use are shown in Figures 10 and 11, respectively. The main game objects were nonplayer characters (NPCs), statically positioned at different locations, with whom player characters (PCs) must interact—(chat,

| | | | | | | | |
|-----|-----|-----|-----|---|---|-----|-----|
| 0.5 | 0.5 | 0 | 0 | 0 | 0 | 0.5 | 0.5 |
| 0.5 | 0.5 | 0 | 0 | 0 | 0 | 0.5 | 0.5 |
| 0 | 0 | 0.5 | 0.5 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 0 | 0 | 0.5 | 0.5 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |

FIGURE 8: Reconstructed \mathbf{X} .

| | | | | |
|-----|-----|---|---|-----|
| 0.5 | 0 | 0 | 0 | 0.5 |
| 0.5 | 0 | 0 | 0 | 0.5 |
| 0 | 0.5 | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 0 |
| 0 | 0.5 | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 0 |

FIGURE 9: Reduced \mathbf{X} .

TABLE 2: Action list of The ICE.

| Action | Symbol |
|-------------------------|----------------------------------|
| Attack with a snow ball | (c.f., top half part of Table 3) |
| Chat | c |
| Walk | w |
| Trade | t |
| Talk | (c.f., bottom half of Table 3) |
| Pick up potion | p |
| Use potion | u |
| Dead | d |
| Warp | x |

help, trade)—to receive and complete missions; the item-shop, from which PCs bought items and monsters, randomly positioned throughout the game world for PCs to attack with snowballs. Major missions in the game are as follows.

- (i) Item delivery where the PC must deliver an item from the mission issuing NPC to a specified NPC.
- (ii) Item trade where the PC must trade with NPCs to increase the amount of money initially provided by the mission issuing NPC.
- (iii) Monster extermination where the PC must help the mission issuing NPC by exterminating monsters.

Actions available in The ICE are summarized in Tables 2 and 3. All NPCs are involved in missions, except NPCs 1, 13, 14, and 16. In the resulting KeyGraphs given in Section 4.2, the symbols for these four nonmission NPCs and monsters are preceded by “n” for those residing in Town 1, that is, nH, and “r” for those in Town 2 or the eastern border of the map, that is, rT, rU, rW, rA, and rD. This is done in order to utilize the visualization metrics in Section 2.3.

A group of 20 players, on average 20 years of age, participated in this evaluation. These players consisted of third-year and fourth-year computer science undergraduate students who were familiar with online-games but had no experience in playing The ICE. After a brief introduction

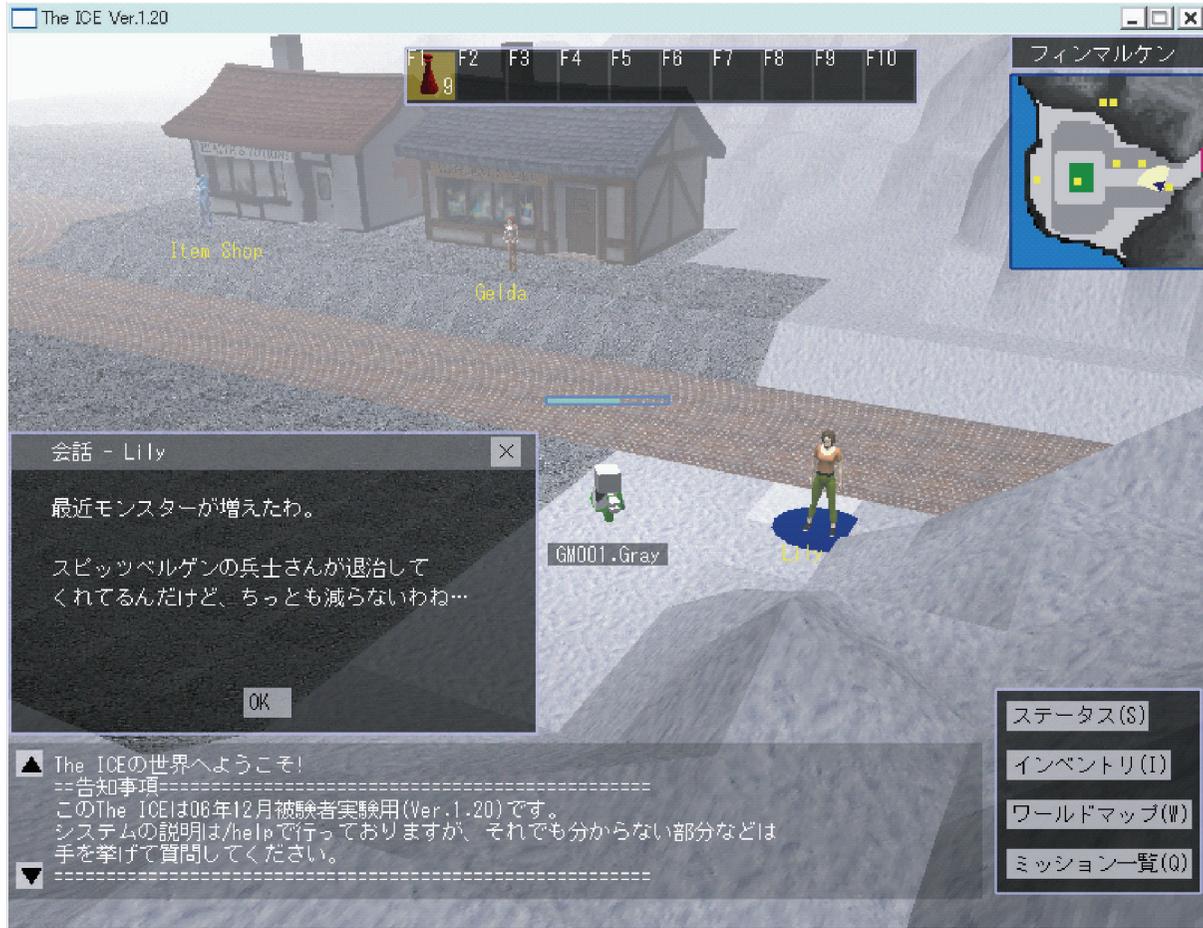


FIGURE 10: A screen shot of The ICE.

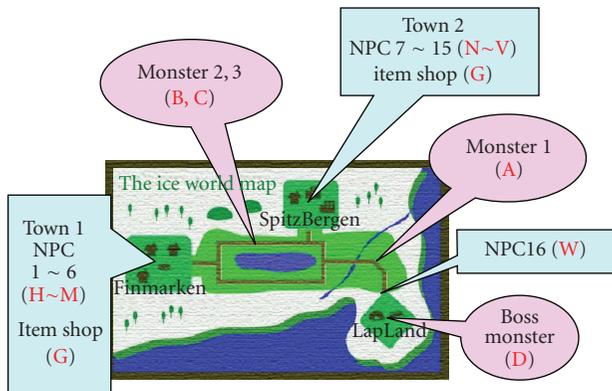


FIGURE 11: Map and the positions of NPCs and monsters.

to the game, they were asked to arbitrarily play it, starting from Town 1. In addition to these 20 players, labeled p1–p20, three game masters, JOJO, Justice, KURO, also participated in the event. In the rest of our evaluation, the symbol w was removed from the log because it was frequently present in all players' action sequences and thus bared no information.

TABLE 3: List of additional symbols related to actions Attack and Talk.

| Symbol | Description |
|--------|------------------------------|
| A | Attack to monster 1 |
| B | Attack to monster 2 |
| C | Attack to monster 3 |
| D | Attack to boss monster |
| E | Attack to other game objects |
| G | Talk to the item shop |
| H–M | Talk to NPCs 1–6 |
| N–V | Talk to NPCs 7–15 |
| W | Talk to NPC 16 |

4.2. Results and discussions

Table 4 shows the mean and variance of time-series matrices of action sequences before and after the time-series reduction technique is applied.

Figure 12 plots all players on two-dimensional space obtained by CMDS. Most players form a cluster on the right half of the figure. The rests can be considered as outliers, that is, p1, p5, p8, p9, p17. To remove the effect of these outliers,

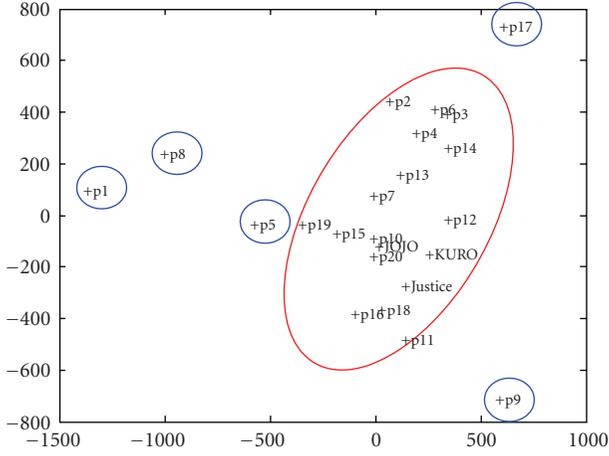


FIGURE 12: MDS result for all data.

TABLE 4: Mean and variance of the data lengths before and after applying the time-series reduction technique.

| | MEAV | VAR |
|--------|-------|--------|
| Before | 262.7 | 7688.6 |
| After | 25.5 | 13.1 |

we excluded them from the log and obtained a new result in Figure 13. From Figure 13, most players can be divided into three clusters: cluster 1 of Justice, JOJO, KURO, p10, p15, p20; cluster 2 of p2, p3, p4, p6; cluster 3 of p7, p16, p18, p19. Each cluster has different player behaviors as discussed below through interpretation of KeyGraph visualization results.

4.2.1. Cluster 1: explorers

Figure 14 shows the KeyGraph of cluster 1 from which salient features are summarized in the following.

- (i) They moved away from town 1 and fought monsters 2 and 3.
- (ii) They also went to the end of the map and fought the boss monster.
- (iii) They were not active in pursuing missions.

The above summary is based on our interpretation of this KeyGraph as follows. First, it can be seen that the foundation of this KeyGraph is mainly composed of warp and attack (monsters 2 and 3) nodes. Next, the symbol rD is a keyword indicating that these players went far away to the end of the map and fought the boss monster there. In addition, because there is only one NPC symbol J in the keywords, these players were not active in receiving missions, from NPCs, and in pursuing them.

Consequently, it can be stated that the players in cluster 1 like to explore the world map and that these players have no interest in pursuing missions and only fight monsters when they find them. This type of players fits Bartle’s explorer.

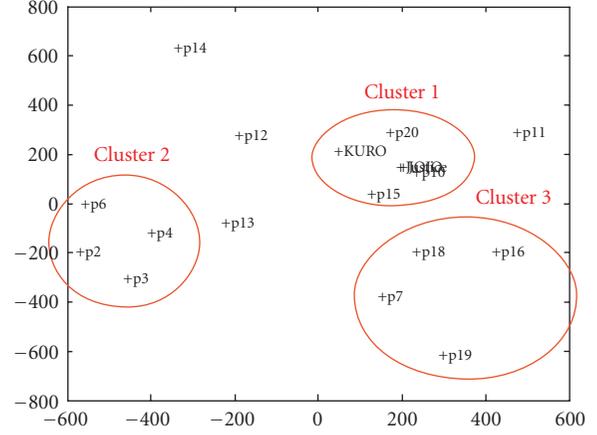


FIGURE 13: MDS result for data after exclusion of the outliers.

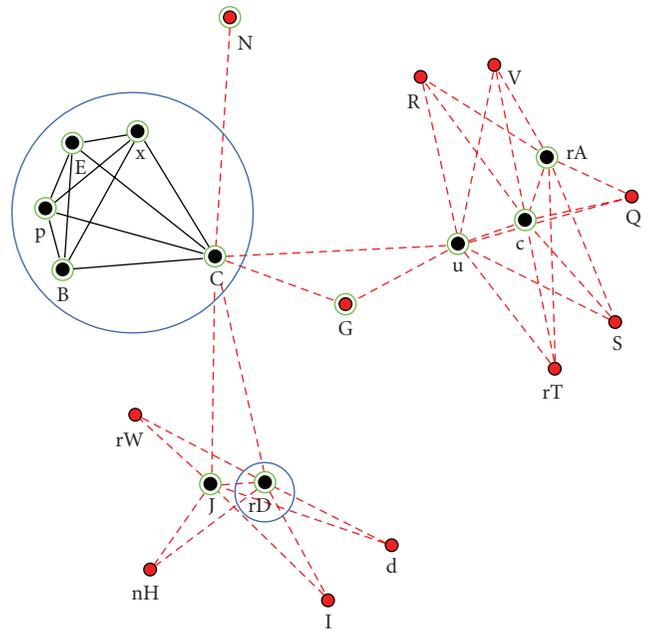


FIGURE 14: KeyGraph of cluster 1, where rD (attacks the boss monster residing in the most remote map from the initial map) is one of the keywords.

4.2.2. Cluster 2: achievers

Figure 15 shows the KeyGraph of cluster 2 from which salient features are summarized in the following.

- (i) They mainly moved within town 1.
- (ii) They also fought monsters 2 and 3.
- (iii) They received a lot of missions.

The above summary is based on our interpretation of this KeyGraph as follows. First, it can be seen that besides nodes related to fighting (B, C, E, u, p), nodes of NPCs residing in town 1 (I, J, K) are included in the foundation of this KeyGraph. This indicates that these players were mainly in town 1. In addition, the keywords include symbols L and R which denote NPCs who are involved in several missions.

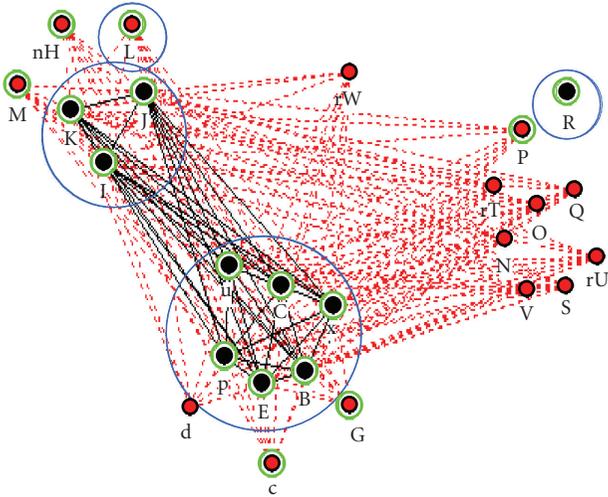


FIGURE 15: KeyGraph of cluster 2, where L and R (talk to mission NPCs) are among the keywords.

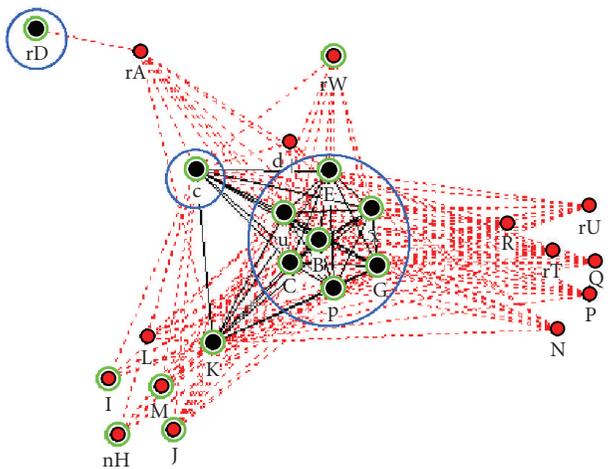


FIGURE 16: KeyGraph of cluster 3, where c (chat) is one of the keywords on the foundation.

As a result, it can be stated that the players in cluster 2 are aggressive in pursuing missions, especially those completable within or not far away from town 1. This type of players fits Bartle's achiever.

4.2.3. Cluster 3: socializers

Figure 16 shows the KeyGraph of cluster 3 from which salient features are summarized in the following.

- (i) They chatted a lot.
- (ii) They mainly fought monsters 2 and 3.
- (iii) They also fought the boss monster.

The above summary is based on our interpretation of this KeyGraph as follows. First, the foundation of this KeyGraph includes the symbol c, not seen in the foundation

or keywords of the previous two clusters. This indicates that these players chatted a lot among each other. Next, the symbol rD is a keyword showing that the players also fought the boss monster. We have confirmed through directly investigating the log that a group of three players (p5, p7, p8) and another group of four players (p16, p17, p18, p19) frequently chatted among their group members and that each group went together to the end of the map to fight the boss monster.

From the above interpretation, it can be stated that the players in this cluster like to communicate with others via chats. This type of players fits Bartle's socializers.

4.3. Computational complexity

We give here the computational complexity of the techniques used in our approach.

- (i) *CMDS*: for m players, the time complexity of the original CMDS is $O(m^3)$. To cope with very large m , a recently proposed approximation [24] taking an $O(m)$ time can be used.
- (ii) *DTW*: the time complexity of DTW for computing the distance between two time series of length l_x and l_y is $O(l_x l_y)$. We coped with this issue with the time-series reduction technique in Section 3. This technique can also be used together with an approximation technique in [25] that introduces lower bounding based on warping constraints.
- (iii) *KeyGraph*: the time complexity of KeyGraph is $O(n \log n)$, where n is the number of action symbol types.
- (iv) *Wavelet*: for a time-series of length l , the Haar wavelet transform has an $O(l)$ time.

5. CONCLUSIONS AND FUTURE WORK

Understanding the player behaviors is an important issue in improving the service quality of online games. We have proposed a visualization approach that first locates clusters of players who have similar action behaviors using CMDS and then interprets such behaviors of a cluster of interest using KeyGraph. To increase the efficiency in computation of the CMDS input, we have described the use of the time-series reduction technique proposed recently by us in [11]. Evaluation of the proposed approach has been done using log from The ICE, where three clusters have been found to fit three of the four Bartle's player types, that is, achievers, explorers, and socializers.

Our future work is to apply the proposed approach to log from commercial online games and to examine if Bartle's player types can be found. It might also be interesting to investigate log formats whose information can be used for automatically identifying other types of Nick Yee's play motivations.

ACKNOWLEDGMENTS

This work was supported in part by Grant-in-Aid for Scientific Research (C) no. 20500146 from the Japan Society for Promotion of Science. The authors would like to thank the anonymous reviewers for their invaluable comments.

REFERENCES

- [1] P. Harding-Rolls, "Western World MMOG Market: 2006 review and forecasts to 2011," *Screen Digest Management Report*, Screen Digest, London, UK, March 2007.
- [2] K. B"orner and S. Penumarthy, "Social diffusion patterns in three-dimensional virtual worlds," *Information Visualization*, vol. 2, no. 3, pp. 182–198, 2003.
- [3] N. Hoobler, G. Humphreys, and M. Agrawala, "Visualizing competitive behaviors in multi-user virtual environments," in *Proceedings of the 15th IEEE Visualization Conference (VIS '04)*, pp. 163–170, Austin, Tex, USA, October 2004.
- [4] L. Chittaro, R. Ranon, and L. Leronutti, "VU-Flow: a visualization tool for analyzing navigation in virtual environments," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 6, pp. 1475–1485, 2006.
- [5] R. Thawonmas, M. Hirano, and M. Kurashige, "Cellular automata and Hilditch thinning for extraction of user paths in online games," in *Proceedings of 5th ACM SIGCOMM Workshop on Network & System Support for Games (NetGames '06)*, Singapore, October 2006.
- [6] R. Thawonmas, M. Kurashige, and K. T. Chen, "Detection of landmarks for clustering of online-game players," *The International Journal of Virtual Reality*, vol. 6, no. 3, pp. 11–16, 2007.
- [7] R. Bartle, "Hearts, clubs, diamonds, spades: players who suit MUDs," *The Journal of Virtual Environments*, vol. 1, no. 1, 1996.
- [8] M. Wattenberg and J. Kriss, "Designing for social data analysis," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 4, pp. 549–557, 2006.
- [9] I. Borg and P. Groenen, *Modern Multidimensional Scaling: Theory and Applications*, Springer Series in Statistics, Springer, New York, NY, USA, 2nd edition, 2005.
- [10] Y. Ohsawa, N. E. Benson, and M. Yachida, "KeyGraph: automatic indexing by co-occurrence graph based on building construction metaphor," in *Proceedings of the IEEE International Forum on Research and Technology Advances in Digital Libraries (ADL '98)*, pp. 12–18, Santa Barbara, Calif, USA, April 1998.
- [11] R. Thawonmas and K. Iizuka, "Haar wavelets for online-game player classification with dynamic time warping," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 12, no. 2, pp. 150–155, 2008.
- [12] F. D. Fracchia, "Is visualization struggling under the myth of objectivity?" in *Proceedings of the 6th IEEE Visualization Conference (VIS '95)*, pp. 412–415, IEEE Computer Society Press, Atlanta, Ga, USA, October–November 1995.
- [13] P. J. Somervuo, "Online algorithm for the self-organizing map of symbol strings," *Neural Networks*, vol. 17, no. 8–9, pp. 1231–1239, 2004.
- [14] J. D. Cohen, "Highlights: language- and domain-independent automatic indexing terms for abstracting," *Journal of the American Society for Information Science*, vol. 46, no. 3, pp. 162–174, 1995.
- [15] Y. Ohsawa and P. McBurney, Eds., *Chance Discovery—Foundation and Its Applications*, Springer, New York, NY, USA, 2003.
- [16] R. Thawonmas and K. Hata, "Aggregation of action symbol sub-sequences for discovery of online-game player characteristics using KeyGraph," in *Proceedings of the 4th International Conference on Entertainment Computing (ICEC '05)*, F. Kishino, Y. Kitamura, H. Kato, and N. Nagata, Eds., vol. 3711 of *Lecture Notes in Computer Science*, pp. 126–135, Sanda, Japan, September 2005.
- [17] N. Okazaki and Y. Ohsawa, "Polaris: an integrated data miner for chance discovery," in *Proceedings of the 3rd International Workshop of Chance Discovery and Its Management in conjunction with International Human Computer Interaction Conference (HCI '03)*, pp. 27–30, Crete, Greece, June 2003.
- [18] A. Tveit, Ø. Rein, J. V. Iversen, and M. Matskin, "Scalable agent-based simulation of players in massively multiplayer online games," in *Proceedings of the 8th Scandinavian Conference on Artificial Intelligence (SCAI '03)*, pp. 80–89, Bergen, Norway, November 2003.
- [19] W. White, C. Koch, N. Gupta, J. Gehrke, and A. Demers, "Database research opportunities in computer games," *SIGMOD Record*, vol. 36, no. 3, pp. 7–13, 2007.
- [20] N. Yee, "Motivations for play in online games," *Cyberpsychology & Behavior*, vol. 9, no. 6, pp. 772–775, 2006.
- [21] F. K.-P. Chan, A. W.-C. Fu, and C. Yu, "Haar wavelets for efficient similarity search of time-series: with and without time warping," *IEEE Transactions on Knowledge and Data Engineering*, vol. 15, no. 3, pp. 686–705, 2003.
- [22] C. C. Aggarwal and P. S. Bradley, "On the use of wavelet decomposition for string classification," *Data Mining and Knowledge Discovery*, vol. 10, no. 2, pp. 117–139, 2005.
- [23] The ICE Online Game, <http://www.ice.ci.ritsumei.ac.jp/mmog.html>.
- [24] J. Tzeng, H. Lu, and W.-H. Li, "Multidimensional scaling for large genomic data sets," *BMC Bioinformatics*, vol. 9, article 179, 2008.
- [25] C. A. Ratanamahatana and E. Keogh, "Three myths about dynamic time warping," in *Proceedings of SIAM International Conference on Data Mining (SDM '05)*, pp. 506–510, Newport Beach, Calif, USA, April 2005.



Hindawi

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

