Detection of Landmarks for Clustering of Online-Game Players

Ruck Thawonmas, Masayoshi Kurashige and Kuan-Ta Chen

***Abstract*—Understanding of player behaviors is an important issue to keep online games interesting to their players. Focusing on player movement, in our previous work, we proposed a method for clustering online-game players based on the transition probabilities between manually determined landmarks in a game map of interest. In this paper, we first propose a method for automatically detecting landmarks from player trails based on weighted entropy of the distribution of visiting players. Next, we describe how to visualize player clusters using multidimensional scaling and prefuse, respectively. Their inputs are the Euclidean distances between players based on their weighted transition probabilities between the derived landmarks. The effectiveness of our approach is confirmed when it is evaluated using trail logs from both an online game developed at the first two authors' laboratory and a commercial massively multiplayer online role-playing game (MMORPG).**

***Index Terms*—Online-game player clustering, landmark detection, multidimensional scaling, prefuse.**

I. INTRODUCTION

The quality of player service plays an important role in online games. To maintain high player satisfaction, online-game developers and publishers need to know their player behaviors so that they can develop game contents accordingly based on such information.

Player or user movement represents important player behaviors and has been studied recently in the literature. In [3], player trails, or time series of visited locations, were used for examining the distance over time among the members of a social group. A visualization tool was developed in [13] for visualizing player flows in virtual environments. Another tool visualizing player trails in a combat game was developed in [10]. Two research groups independently developed a method for automatically extracting important parts from trails in [12] and [17]. The former focused on extraction of the most visited parts from repeatedly visited ones such as bus driving routes, while the latter focused on extraction of the most visited parts from more diverse data as commonly seen in online-game player trails. Recently, Chen and Hong proposed to use player movement behavior as a biometric to detect account hijacking and sharing [5].

Analysis of user movement is also crucial in other application domains such as virtual museum [9] and virtual heritages [16]. Other related researches in [1, 8] aim at prediction of user movement using GPS data.

According to the study in [19, 6], there exist typical movement patterns of users in museums and virtual museums, respectively. We conjectured that movement patterns should exist in online games, and this kind of information should be useful in (re)designing a game map of interest and the content therein. In [18], we therefore proposed how to cluster online-game players based on their trails using Self Organization Map (SOM) [14], where the SOM inputs are the transition probabilities between landmarks in a given game map. However, landmarks had to be selected manually.

II. PLAYER CLUSTERING

In this section, we describe improvements to our previous work [18]. We first propose a method for detecting of landmarks, and then describe how to compute the weighted transition probabilities between landmarks (WTPL) for being used in visualization of player clusters with multidimensional scaling (MDS) [2] and a visualization tool called prefuse [7], respectively. According to the work in [11], MDS is superior to SOM in clustering of multi-dimensional data. In prefuse, we utilize a force-directed model that in general has equivalence to MDS [15].

2.1 Landmark Detection

A map of interest is first divided into $m \times m$ grids. Let us assume that N players have visited the map. Let $v_{i,j}(k)$ denote the number of visits to grid (*i, j*) of user k. In our research, we consider a grid as a landmark if it has a relatively high number of visits by most players. For grid (*i, j*), this criterion can be represented by a weighted entropy *H* of the distribution of the players visiting this grid as follows:

$$H_{i,j} = -\frac{V_{i,j}}{V^*} \sum_{k=1}^{N} \frac{v_{i,j}(k)}{V_{i,j}} \log\left(\frac{v_{i,j}(k)}{V_{i,j}}\right), \qquad (1)$$

where $V_{i,j} = \sum_{k=1}^{N} v_{i,j}(k)$ and $V^* = \arg\max_{i,j} V_{i,j}$.

The proposed algorithm for detecting *L* landmarks is given below as follows:

Step 1 Unmark all grids and empty the landmark set.

Manuscript Received on May 12, 2007.

Ruck Thawonmas is a full professor in the Department of Human and Computer Intelligence at Ritsumeikan University, Japan. E-Mail: ruck@ci.ritsumei.ac.jp.

Masayoshi Kurashige is a graduate student in the Graduate School of Science and Engineering, Ritsumeikan University, Japan. He received a B. Eng. in Computer Science from the same university in 2006. E-Mail: rs016025@se.ritsumei.ac.jp.

Kuan-Ta Chen received his Ph.D. in Electrical Engineering from National Taiwan University in 2006. His research interests include multimedia networking, Internet measurement, network security and multimedia networking. E-Mail: cychen@iis.sinica.edu.tw.

Step 2 Mark the grid with the highest H among the unmarked grids and add this grid to the landmark set. Then mark its eight neighbor grids.

Step 3 Repeat Step 2 until the landmark set is filled by L members.

Note that marking the neighbor grids of a given landmark in Step 2 is to prevent obtaining consecutive landmarks.

2.2 Weighted Transition Probabilities between Landmarks

Given L landmarks detected by the above algorithm, the matrix of the WTPL of user k is denoted as $X(k)$ and

$$X(k)=\begin{bmatrix} p_{1,1}(k) & p_{1,2}(k) & \cdots & p_{1,L}(k) \\ p_{2,1}(k) & p_{2,2}(k) & \cdots & p_{2,L}(k) \\ \vdots & \ddots & \ddots & \vdots \\ p_{L,1}(k) & p_{L,2}(k) & \cdots & p_{L,L}(k) \end{bmatrix} \tag{2}$$

In the above formula, $p_{a,b}(k)$ is the weighted transition probability that player k moves from landmark a to landmark b and is calculated as follows:

$$p_{a,b}(k)=\frac{c_{a,b}(k)}{c^*_{a,b}}\frac{c_{a,b}(k)}{C(k)}, \tag{3}$$

where $c_{a,b}(k)$ is the number of times that player k moved from landmark a to landmark b in his/her trail,

$c^*_{a,b}=\arg\max_k c_{a,b}(k)$, and $C(k)=\sum_{i=1}^{L}\sum_{j=1}^{L}c_{i,j}(k)$.

The elements of the WTPL matrix differ from those of the TPL matrix in [18] by presence of the weight term $\frac{c_{a,b}(k)}{c^*_{a,b}}$ in (3).

In the current version, for computation of the elements of $X(k)$, a player coordinate residing in his/her trail is represented by its nearest landmark. Let us suppose that five landmarks (A, B, C, D, E) have been detected, then a player trail will be like BBBAAAACEEEE, leading to the transitions B $\rightarrow$ A $\rightarrow$ C $\rightarrow$ E. At present, we do not consider loop transitions such as A $\rightarrow$ A or B $\rightarrow$ B.

2.3 Multidimensional Scaling

Here, to visualize player clusters based on their movement patterns, we use simple and computationally efficient Classical Multi Dimensional Scaling (CMDS) [1]. The CMDS method takes an input matrix D, giving dissimilarities between pairs of players, and outputs a coordinate matrix whose configuration minimizes a loss function in preserving all inter-point distances. In our research, the ijth element in D is the Euclidean distance between $X(i)$ and $X(j)$. We select only the first two dimensions of the constructed coordinates and use the function *cmdscale* in the Statistical Toolbox of Matlab for performing CMDS.

2.4 Prefuse

We also use a forced-based model in prefuse for producing a graph layout where a node represents a player, a link between any two nodes indicates a connection between them, and the link length is proportionate to the distance between the corresponding players, i.e., the ijth element in D. In order to reduce computational complexity, we remove in advance all links between nodes whose distances are beyond a given threshold τ, in the paper set to $0.5\times$ the median of the distance. With prefuse, the user can interactively manipulate a resulting graph layout while he/she is attempting to locate clusters. An example of such interactive data analysis systems can be found in [4].

Fig. 1. A screen shot of the ICE.

III. RESULTS AND DISCUSSIONS

We obtained player-movement logs from an online game called The ICE[2], under development at the first two authors' laboratory, as well as those from a commercial MMORPG called Angle's Love[3].

3.1 The ICE

A screen shot of The ICE is shown in Fig. 1. The targeted map in use is shown in Fig. 2, where the fixed locations of seven non-player characters (NPCs) and the warping point (Warp) are also displayed. The map's size is 600×600 pixels, where a player can move from the current pixel to another by clicking on the destination pixel, and in this work the map was partitioned into 50×50 grids. Twenty players, 17 male and 3 female students in the first two authors' university, participated in the test play for 45 minutes. In this test play, typical missions such as trading items with NPCs, delivering items to NPCs, and collecting items from monsters, were available so as to have a variety of movement patterns.

Fig. 3 shows the locations of the resulting landmarks when L was set to 5. Landmark 1 is the location near Marlen and Lagi, and landmark 2 is the location near the place where all players entered this map. Landmark 3 is near the location of Gelda, landmark 4 is between Amanda and Village Master, and landmark 5 is near both Shop and Gelda. These landmarks are important locations in the game because most players if not all

[2] http://www.ice.ci.ritsumei.ac.jp/mmog.html
[3] http://al.gameflier.com/

must visit in order to enter the game, receive missions, and purchase items.

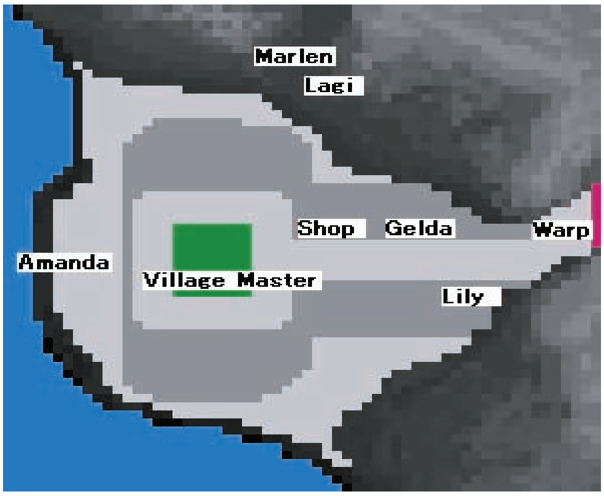


Fig. 2. The map in use of the ICE. See Color Plate 1.

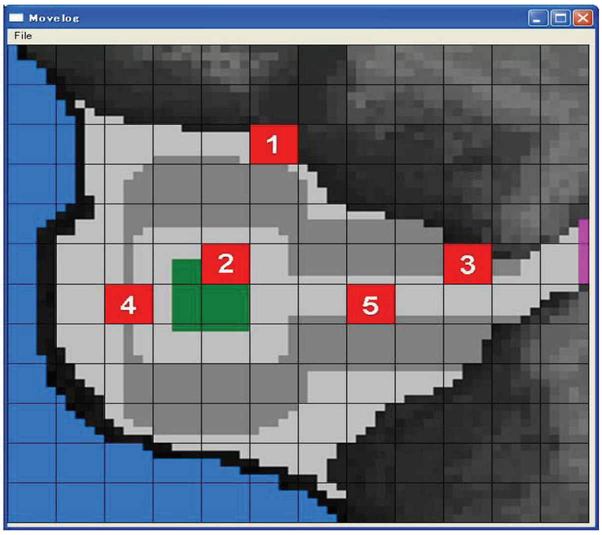


Fig. 3. Detected landmarks. See Color Plate 2.

Fig. 4 plots all players, player i represented by pi, on two dimensional space obtained by CMDS. Note that there is one distinct outlier, i.e., p11, causing all other players positioned on the right side. To remove the effect of this outlier, we did not consider this player in CMDS and obtained a new result in Fig. 5.

From Fig. 5, most players can be divided into four clusters: A, B, C, and D. Each cluster has a different movement pattern as discussed below, where the movement pattern of each cluster is visualized by representing each element of the averaged of X in the cluster by an arrow whose thickness is proportionate to its value.

Fig. 6 shows the visualized movement patterns for each cluster. Players in cluster A mainly moved between landmarks 2 and 5 and between landmarks 3 and 5. Players in cluster B most frequently moved between landmarks 2 and 4 and between landmarks 1 and 2. Players in cluster C most frequently moved between landmarks 3 and 5, and also between 2 and 5. Players in cluster D most frequently moved between landmarks 1 and 2, as well as between 2 and 5.

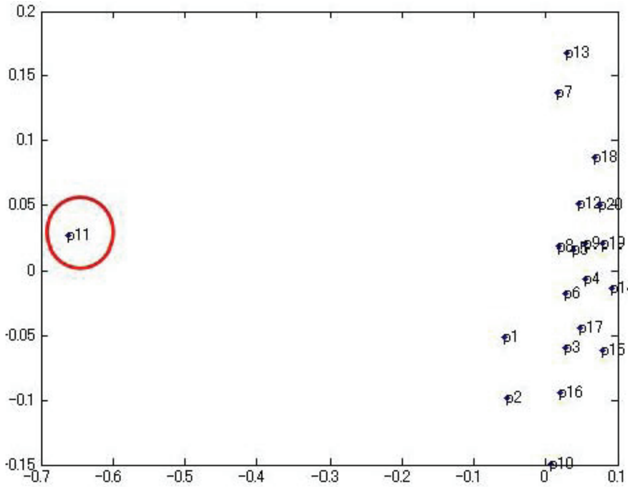


Fig. 4. MDS result prior to removing of the outlier p11.

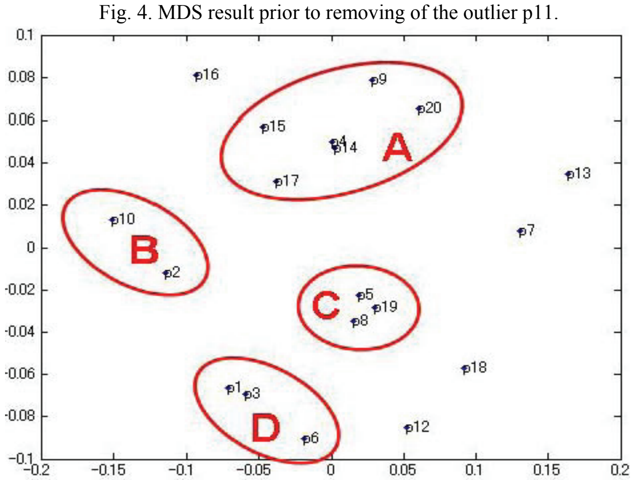


Fig. 5. MDS result after removing the outlier p11.

3.2 Angle's Love

Fig. 7 shows the targeted map of Angle's Love and the locations of the detected five landmarks. The map's size is 300×190 pixels, and this map was partitioned into 10×10 grids. There were 397 players who played in this map. Most of them traveled over the map to get into another map.

Among these five landmarks, three of them, i.e., landmarks 1, 2, and 3, are barriers that players must turn around and go back from the same trail. This may explain why these locations were detected as landmarks, as players moved around them at least twice (back and forth). Fig. 8 shows a screen shot of the landmark 3's area.

Fig. 9 shows the resulting graph layout of prefuse. A large cluster has been formed in the middle of the figure, indicating that most players have similar movement patterns. However, it is interesting to see that there are two clusters of a much smaller size on the top and bottom consisting of three players and four players, respectively.

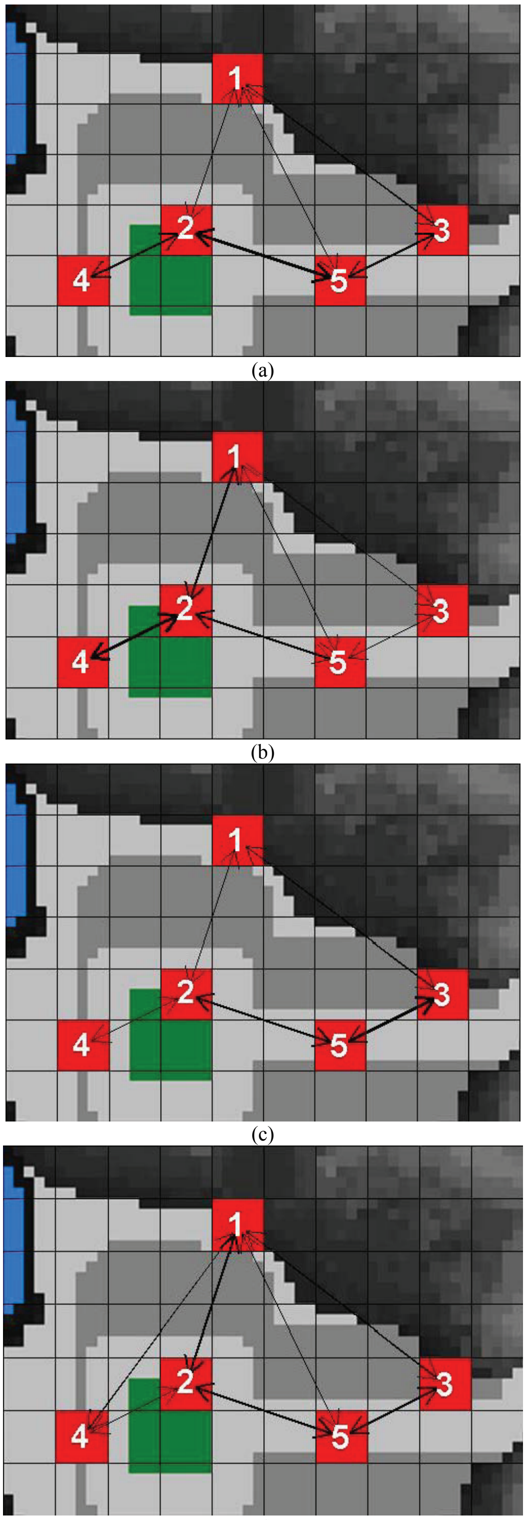


Fig. 6. Visualized movement patterns of clusters A (a), B (b), C (c) and D (d).

Fig. 10 and Fig. 11 visualize both the actual trail and the transitions between landmarks of each player in the top cluster and the bottom cluster, respectively. It can be seen that the players in each cluster moved similarly, especially with respect to the transitions between landmarks.

Fig. 7. The map in use of Angle's Love. See Color Plate 3.

Fig. 8. A screen shot of landmark 3 in Fig. 7. See Color Plate 4.

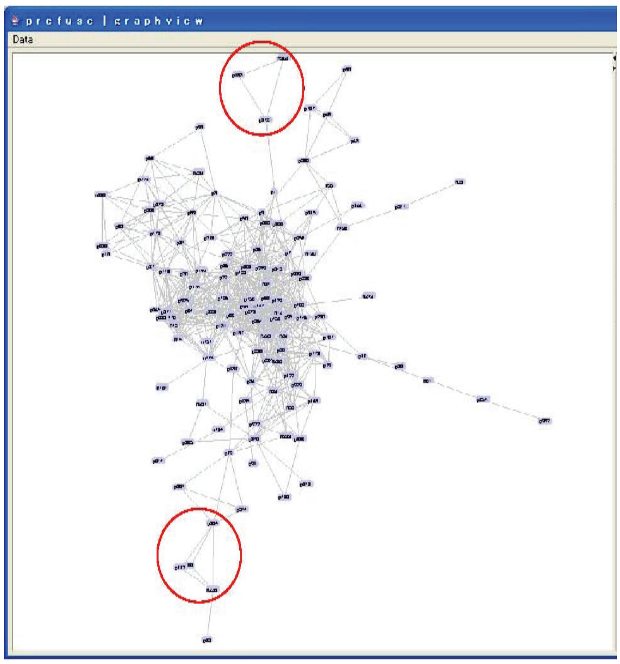


Fig. 9. prefuse's graph layout.

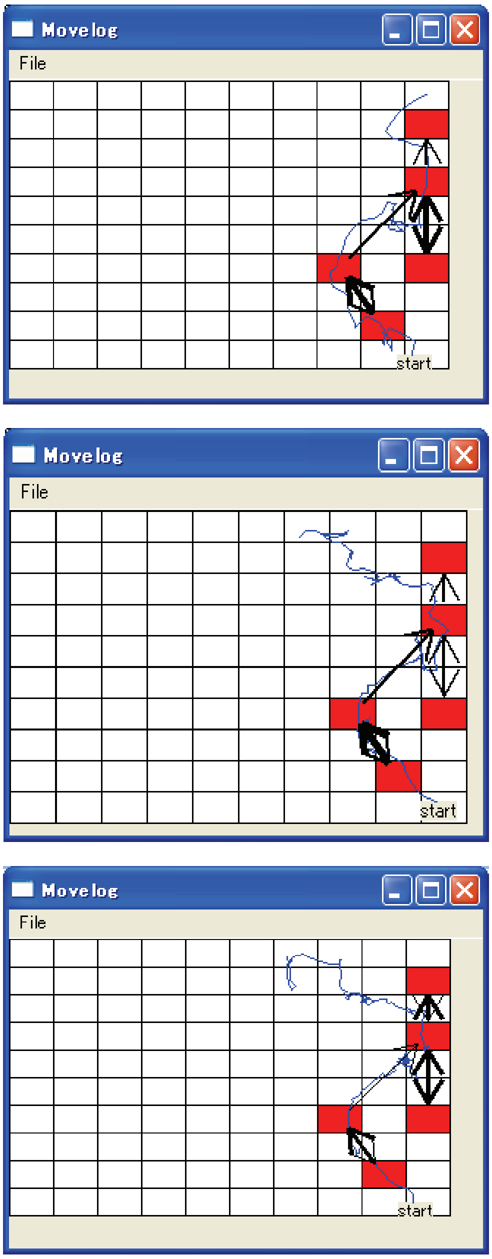


Fig. 10. Visualized movement patterns of each of the three players in the top cluster of Fig. 9.

IV. CONCLUSIONS AND FUTURE WORK

In order to understand movement patterns of players in online games, we proposed a method for automatically detecting landmarks from player trails based on a weighted entropy of the probability that each candidate location is visited. We then proposed to use multidimensional scaling and prefuse, respectively, for visualizing player clusters based on the newly proposed weighted transition probabilities between the derived landmarks. Our approach was validated using trail logs from both The ICE and Angle's Love. Evaluation results confirmed that our approach successfully identified player clusters having different movement patterns. These moving patterns imply players' playing styles and thus can be exploited for maintaining or increasing players' satisfaction.

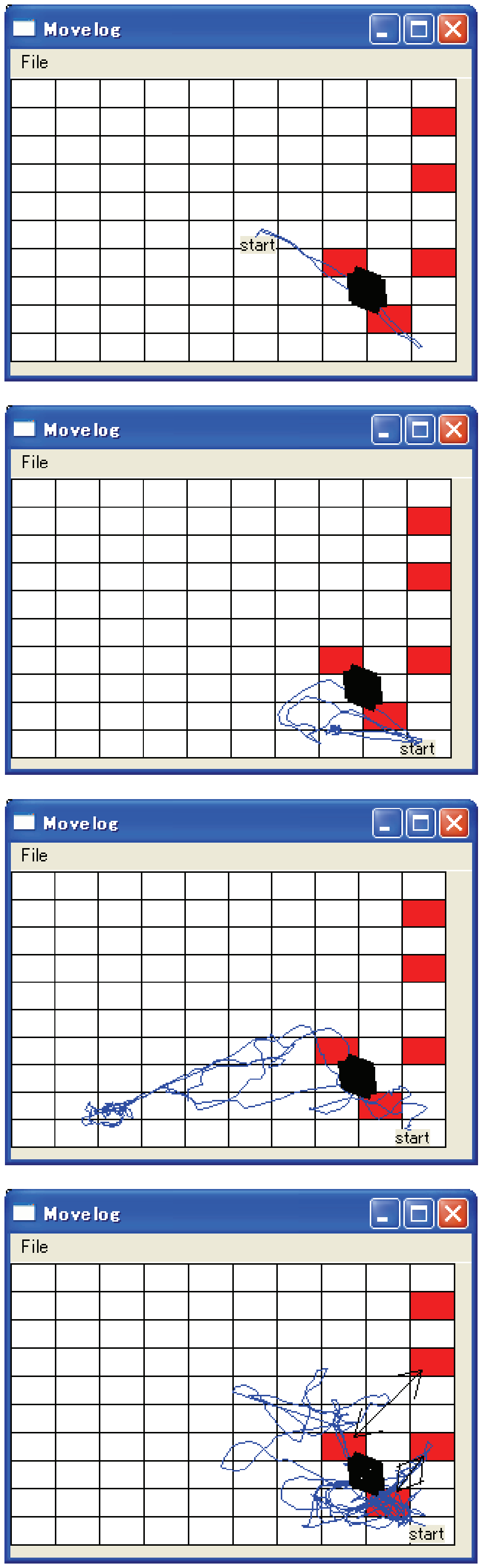


Fig.11. Visualized movement patterns of each of the four players in the bottom cluster of Fig. 9.

As our future work, we plan to derive methods for determining the optimal grid size as well as the optimal number of landmarks. In addition, we will examine how to incorporate information on actions into this approach.

ACKNOWLEDGEMENT

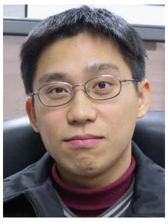
This work would not have been possible without the extensive traces of Angel's Love. The authors are much indebted to Tsing-San Cheng at UserJoy Tech. Co. Ltd, who helped us gather the trail logs.

REFERENCES

[1] D. Ashbrook and T. Starner. Using GPS to learn significant locations and predict movement across multiple users, *Personal and Ubiquitous Computing*, vol. 7, no.5, pp. 275–286, 2003.
[2] I. Borg and P. Groenen. Modern Multidimensional Scaling: Theory and Applications, Second Edition, *Springer Series in Statistics*, Springer, 2005.
[3] K. Borner and S. Penumarthy. Social Diffusion Patterns in Three-Dimensional Virtual Worlds, *Information Visualization Journal*, vol. 2, no. 3, pp. 182-198, 2003.
[4] J. X. Chen, D. B. Carr, H. Wechsler and Z. Pan. Interactive Visualization of Multivariate Statistical Data, *The International Journal of Virtual Reality*, vol. 5, no. 3, pp. 67-73, 2006.
[5] K. T. Chen (C. Y. Chen) and L. W. Hong. User Identification based on Game-Play Activity Patterns, *Proc. of ACM NetGames 2007*, 2007.
[6] L. Chittaro and L. Ieronutti. A Visual Tool for Tracing Behaviors of Users in Virtual Environments, *Proc. of AVI 2004: 7th International Conference on Advanced Visual Interfaces*, pp. 40-47, 2004.
[7] J. Heer, S. K. Card and J. Landay. prefuse: a toolkit for interactive information visualization, *Procs. of the ACM Conference on Human Factors in Computing Systems (CHI 2005)*, pp. 421-430, 2005.
[8] J. Hightower, S. Consolvo, A. LaMarca, I. Smith and J. Hughes. Learning and Recognizing the Places We Go, *Proc. UbiComp 2005*, 2005.
[9] M. Hirose. Title Virtual Reality Technology and Museum Exhibit, *The International Journal of Virtual Reality*, vol. 5, no. 2, pp. 31-36, 2006.
[10] N. Hoobler, G. Humphreys and M. Agrawala. Visualizing Competitive Behaviors in Multi-User Virtual Environments, *Proc. IEEE Visualization 2004*, pp. 163-170, 2004.
[11] S. Huang, M. Ward and E. Rundensteiner. Exploration of Dimensionality Reduction for Text Visualization, *Proc. The Third International Conference on Coordinated & Multiple Views in Exploratory Visualization (CMV 2005)*, pp. 63-74, 2005.
[12] C. Huiping, N. Mamoulis and D. W. Cheung. Mining Frequent Spatio-temporal Sequential Patterns, *Proc. The Fifth IEEE International Conference on Data Mining (ICDM 2005)*, 2005.
[13] L. Ieronutti, R. Ranon and L. Chittaro. 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.
[14] T. Kohonen. Self-Organizing Maps, Second Extended Edition, Springer Series in Information Sciences, *Springer*, 1997.
[15] C. W. Pak, H. Foote, G. Chin, P. Mackey and K. Perrine. Graph Signatures for Visual Analytics, *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 6, pp. 1399-1413, 2006.
[16] M. S. Sunar, A. M. Zin and T. M. T. Sembok. Effective Range Detection Approach for Ancient Malacca Virtual Walkthrough, *The International Journal of Virtual Reality*, vol. 5, no. 4, pp. 31-38, 2006.
[17] R. Thawonmas, M. Hirano and M. Kurashige. Cellular Automata and Hilditch Thinning for Extraction of User Paths in Online Games, *CD-ROM Proc. NETGAMES 2006 & 5th Workshop on Network & System Support for Games*, 2006.
[18] R. Thawonmas, M. Kurashige, K. Iizuka and M. Kantardzic. Clustering Online Game Users Based on Their Trails Using Self-Organizing Map, *Lecture Notes in Computer Science (ICEC 2006)*, Richard Harper et al. Eds., vol. 4161, pp. 366-369, 2006.
[19] J. Umiker-Sebeok. Behavior in a Museum: A Semio-Cognitive Approach to Museum Consumption Experiences, *Signifying Behavior*, vol. 1, no. 1, pp. 52-100, 1994.

Ruck Thawonmas received B.Eng. in Electrical Engineering from Chulalongkorn University, Thailand, in 1987, M.Eng. in Information Science from Ibaraki University, Japan, in 1990, and D.Eng. in Information Engineering from Tohoku University, Japan, in 1994. During 1987-1993, he was a recipient of the Japanese Government (Monbusho) Scholarship. Since April 2002, Dr. Thawonmas has been at Ritsumeikan University in Japan and has become a full professor in the Department of Human and Computer Intelligence since April 2004. Prior to that, he was with Hitachi, Ltd., RIKEN, University of Aizu, and Kochi University of Technology. His research interests include online games, story generation, and entertainment computing. He is a senior member of IEEE.

Masayoshi Kurashige received B.Eng. in Computer Science from Ritsumeikan University, Japan, in 2006. Since April 2006, he has been a graduate student in the Graduate School of Science and Engineering at the same university. His research interests include online games and data mining.



Kuan-Ta Chen (also known as Chun-Yang Chen) received his B.S. and M.S. in Computer Science from National Tsing Hua University in 1998 and 2000, respectively. He received his Ph.D. in Electrical Engineering from National Taiwan University in 2006. Since then he joined the Institute of Information Science, Academia Sinica as an assistant research fellow. His research interests include multimedia networking, Internet measurement, network security, and multimedia networking. Much of his recent work focuses on the analysis and design of networked multimedia systems, including traffic characterization, transport protocols, quality of service, human factors and security issues. He is a member of ACM and IEEE.