

C-24 オンラインゲームにおける再訪問パターンを用いた ゲーム休止・継続の判別

Prediction of Game Termination and Continuation Using Revisitation Patterns in Online Games

吉田 圭佑^{*1} ラック ターウォンマツト^{*1} Kuan-Ta Chen^{*2}

Yoshida Keisuke^{*1} Ruck Thawonmas^{*1} Kuan-Ta Chen^{*2}

1. はじめに

本研究の目的はオンラインゲームを利用しているユーザーがゲームを休止・継続するかを予測することである。オンラインゲームを運営する上でユーザーがゲームを休止・継続するかどうかを予測することは、収益の予測やサーバーの管理などに役立ち、ゲームを休止するユーザーに対し何らかの対策を立てられるようになると考えられる。本稿では、ユーザーがゲームに訪問する間隔を解析し、再訪問間隔を元に休止・継続の判別を行い認識率 75.1%、適合率 79.0%、再現率 73.2%の精度で休止するユーザーを判別することが出来た。

我々は Web ページの再訪問に関する研究[1]を参考にユーザーのオンラインゲームに対する再訪問についても何らかの特徴があることを確認した[2]。それを元に休止・継続を判別することが出来るのではないかと考え解析を行った。例えば、ゲームを継続するユーザーは短い間隔で定期的に再訪問し、休止するユーザーは訪問する間隔が長くなるなどが考えられる。関連研究として、オンラインゲームのキャラクターの行動から、ゲームを休止・継続するキャラクターの特徴を解析した[3]がある。

2. 対象オンラインゲームのログ

オンラインゲームユーザーの休止・継続の判別に用いたデータは、ShenZhou Online[4]の 5 万ユーザーが対象となっている。判別を行う際は、ユーザー毎のログを 30 日毎に分割し 15 日間隔でオーバーラップさせた。またゲームを休止したと判断する条件は分割したデータの最終ログイン日からゲームに 30 日以上ログインしていない場合とし、それ以外のデータは継続データとした。

十分な再訪問パターンを得るために、ゲームを初めて開始してからゲームを休止するまでの期間が 30 日に満たないユーザーを除外した。ゲームの課金期間が 1 カ月単位である為、削除対象を 1 カ月とした。削除対象のユーザーはログイン自体が少なく、ゲームの利用時間も少ないためゲームのお試し期間だけ利用したユーザーとみなす。除外した後のユーザー数が 3596 ユーザーとなった。

3. 再訪問曲線

異なるユーザーの再訪問パターンを比較するために X 軸を訪問間隔、Y 軸を訪問数とした再訪問曲線の概念を用いた。再訪問数を日常的な時間間隔である 10, 30, 60 (1 時間), 100, 140, 210, 420, 850, 1700 (1 日+), 3400, 6780, 13570 (1 週+), 27140, 27140 以上 (単位: 分) の 14 の区間で集計し再訪問曲線を求めた。ユーザーがゲームを休止・継続する際に特定のパターンがあるのならば、その特徴的なパターンが再訪問曲線に表れる。各ユーザーの再訪問曲線を求める際には、各区間に対して同区間の全ユーザーの平均値で正規化した各ユーザーの再訪問曲線は (1) より得られる。この正規化により、各ユーザーの再訪問数が全ユーザーの平均と比較した比が求められ、全ユーザーに対しての相対的な値がわかる。

$$\frac{\text{ユーザー } j \text{ の再訪問数}[i]}{\text{区間 } i \text{ の全ユーザー平均}} \quad (1)$$

ただし、 j はユーザー、 i は該当区間を表す。

4. 代表データ判別

実際に全データを用いて判別を行う前に予備実験として、休止・継続のデータが得られたユーザー毎に代表的な休止・継続のデータを 1 組選択し判別を行った。代表データの決定方法はユーザー毎に休止・継続データの cosine 距離 (2) を計算し、距離の総和が最少となるデータを休止・継続データの代表とした (図 1)。図 1 から、休止する場合と継続する場合とで、再訪問曲線の形に大きな違いが表れてくる。休止する場合の再訪問曲線は短い間隔での再訪問が少なく、長い間隔での再訪問が多くなっている。継続する場合の再訪問曲線は短い間隔での再訪問が平均以上あり、長い間隔での再訪問が少ない。

代表データを元に休止・継続の判別を行った。判別データは全データを 10 分割し 9 のデータを学習データ、1 のデータをテストデータとした 10 交差検証法を用いて作成した。再訪問曲線の各区間を入力ベクトルとした SVM (Support Vector Machine) [5]によって休止・継続を判別 (カーネル関数: rbfdot) し、継続期間の長さによって比較を行った (図 2)。図 2 から判別精度は、継続している期間が長ければながいほど高くなっている、ユーザーがゲームに訪問する間隔によって休止・継続を判別出来ている。

^{*1}:立命館大学大学院 理工学研究科
ruck@ci.ritsumeai.ac.jp

^{*2}:Multimedia Networking and Systems Lab, IIS, Academia Sinica
swc@iis.sinica.edu.tw

$$\text{コサイン距離} = \frac{\text{再訪問曲線 A} \cdot \text{再訪問曲線 B}}{|A||B|} \quad (2)$$

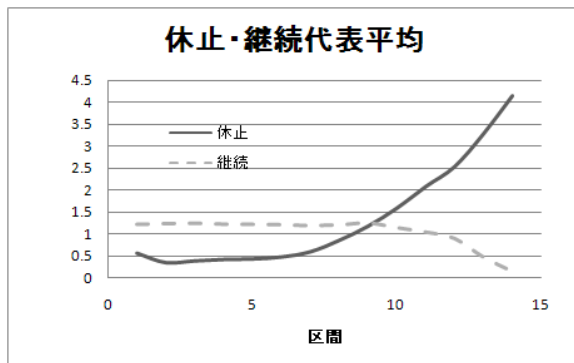


図1 休止・継続代表の平均再訪問曲線

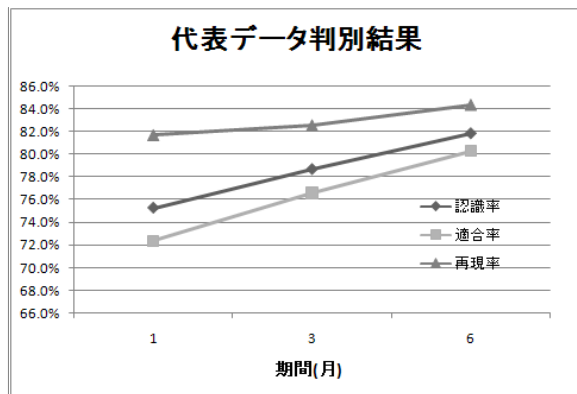


図2 代表データ判別結果

5. 全データ判別

4章で代表データに対して再訪問間隔による判別が有効であることが示されたので、休止・継続の全データに対して判別を行う。2章で行ったフィルタリング後のユーザー3596人の分割された休止・継続の全データが判別対象となっている。しかし、休止データ67111、継続データ70953とデータ数に大きな偏りがある(約11.6倍)。そのためデータの偏りを解消する為に、継続データを休止データと同じ数だけランダムに選択し休止・継続データが1対1の割合になるデータセットを11組作成した。作成した11組のデータ全てに対して10交差検証法を適用して、SVMを用いて判別を行った結果を表1、表2に示す。

	休止と判断	継続と判断
休止	53007	14093
継続	19373	47727

表1 判別結果

認識率	再現率	適合率
75.1%	79.0%	73.2%

表2 判別結果：認識率，適合率，再現率

6. 考察と応用

今回の判別の結果、オンラインゲームの再訪問パターンによって、認識率75.1%、適合率79.0%、再現率73.2%の精度で休止するユーザーを判別することが出来た。再現率が79.0%と8割近い精度で休止ユーザーを発見することが出来ている。この手法の応用例として休止するユーザーを予測することによって、ユーザー数を推測し収益の予測やサーバーの管理などに役立てられると考えられる。休止すると判断されたユーザーに対しては、ゲームに興味を持ってもらえるようなメールを送る、または別のゲームを勧めるなどの対処が可能になり、運営しているゲームのユーザー離れを防ぐことが可能になると考えられる。

7. 終りに

本稿では再訪問パターンを元にしたオンラインゲームにおけるユーザーの休止・継続を判別する手法を提案し、SVMによって判別を行った。判別実験の結果、再訪問間隔によって休止・継続するユーザーを判別することが出来た。今後の課題として、認識率、再現率、及び適合率のさらなる向上を図る。

参考文献

- [1] Eytan Adar, Jaime Teevan, and Susan Dumais, Large Scale Analysis of Web Revisitation Patterns <http://www.cond.org/revisit1.html>
- [2] Ruck Thawonmas, Keisuke Yoshida, Jing-Kai Lou and Kuan-Ta Chen, "Analysis of Area Revisitation Patterns in World of Warcraft," Proc. IFIP 8th International Conference on Entertainment Computing (ICEC 2009), Paris, France, Sep. 3-5 2009.
- [3] Pin-Yun Tarng, Kuan-Ta Chen, and Polly Huang, "An Analysis of WoW Players' Game Hours," Proc. of the Seventh Annual Workshop on Network and Systems Support for Games (NetGames 2008), 2008.
- [4] UserJoy Technology Co., Ltd., "ShenZhou Online," 2004. <http://www.ewsoft.com.tw/>
- [5] Alexandros Karatzoglou, Alex Smola, Kurt Hornik, Achim Zeileis, "kernlab-An S4 Package for Kernel Methods in R," Journal of statistical Software, 11(9), pp.1-20, 2004.