Predicting Player Churn in the Wild

Fabian Hadiji¹, Rafet Sifa², Anders Drachen³, Christian Thurau⁴, Kristian Kersting¹, Christian Bauckhage²,⁵

¹TU Dortmund University
²Fraunhofer IAIS
³Aalborg University
⁴Game Analytics
⁵University of Bonn

CIG, Dortmund, August 27th, 2014
Take Away Message

Churn prediction for mobile games can be realized with decision trees and game independent features.
What is Churn Analysis?

- Generally: detect and define subscribers of a system that will leave in the future
  - e.g. players of a game
- Churn starts with players losing interest in the game
For the final session it has been observed that session time decreases while intersession time increases [Feng et al., 2007, Workshop at ACM SIGCOMM]
Playtime, Intersession Time & Churn

- For the final session it has been observed that session time decreases while intersession time increases [Feng et al., 2007, Workshop at ACM SIGCOMM]
- Weibull distribution fits total playtimes well: implies an underlying power law process on the player’s interest [Bauckhage et al., 2012, CIG]
Playtime, Intersession Time & Churn

- For the final session it has been observed that session time decreases while intersession time increases [Feng et al., 2007, Workshop at ACM SIGCOMM]
- Weibull distribution fits total playtimes well: implies an underlying power law process on the player’s interest [Bauckhage et al., 2012, CIG]
- Here, we seek for a classifier for individual players, possibly exploiting these regularities
Why is Churn Analysis important?

www.newzoo.com/infographics/infographic-the-global-mobile-landscape/
Why is Churn Analysis important?

- New revenue model for mobile games is Free-to-Play (F2P)
  - The mobile game market is worth billions of dollars across the entire world
  - Monetization is based on advertisement or in-game purchases
  - Churn prediction can improve re-targeting campaigns
  - Decreasing the churn rate increases revenue
Why is Churn Analysis important?

- New revenue model for mobile games is Free-to-Play (F2P)
  - The mobile game market is worth billions of dollars across the entire world
  - Monetization is based on advertisement or in-game purchases
  - Churn prediction can improve re-targeting campaigns
  - Decreasing the churn rate increases revenue

- Obtain data-driven insights to inform design and development
Why is Churn Analysis important?

- New revenue model for mobile games is Free-to-Play (F2P)
  - The mobile game market is worth billions of dollars across the entire world
  - Monetization is based on advertisement or in-game purchases
  - Churn prediction can improve re-targeting campaigns
  - **Decreasing the churn rate increases revenue**

- Obtain data-driven insights to inform design and development

- See also [Runge et al., 2014, CIG]
Player Churn More Precisely

- F2P does most often not require a subscription
- Therefore, churn is not explicitly visible
- Makes a precise definition of churn necessary
- Learn a binary classifier that labels players as churners or non-churners
Churn Definition (P1)

- Churner: a player without a single session after the cutoff date
- No chance to reactivate these players again in mobile games
Churn Definition (P2)

- Churner: a player with a low number of sessions to play after the cutoff date
- Introduces sliding window $d_{rlx}$
- Transition to separate engaged and no longer engaged players
Dataset

- Data was provided by GameAnalytics (www.gameanalytics.com)
- Game developer can transmit various kinds of messages
- Observations of player behavior in 5 different games
- Data contains about twenty million sessions
- Data also contains in-game purchase information
- We normalize these raw messages to make them game-independent and to generate training/test instances
We chose $d_{chrn} = 7$ to cover every weekday

- $< 5\%$ of the players in our dataset return after a 7-day absence
- Negative instances are randomly sampled from positive ones
- Classical Machine Learning point of view

Data Generation (M1)
Data Generation (M2)

- Choose training date and split user in churners and non-churners
- We can easily incorporate “absence time” as feature
- Respects overall game life-cycle
- More suitable in real-world scenario
Building a Classifier

- Binary classification task
- Compared different classifiers (Logistic Regression, Neural Networks, Naive Bayes & Decision Trees)
- We found that Decision Trees to perform best
- For M1, we used a cross-validation with 50,000 users per game
- For M2, training- and test-datasets were chosen according to a training date for each game
- For each user, we generated a feature vector
Features

- Focus on game-design independent features to maximize applicability
- Temporal features
  - Number of Sessions, Number of Days
- Session-wise temporal information
  - Average Playtime per Session, Average Time Between Sessions
- Playtime models
  - Playtime Model Parameters, Retention Value
- Virtual economy features
  - Premium User Flag, Predefined Spending Category, Number of Purchases & Average Spending per Session
- With M2: absence time
Experimental Results for P1

- Generally, the classes are not balanced (typical for F2P)
- **M1**
  - About 71% churners
  - Decision Trees with M1-based training achieve an average F1-score of 0.64
  - However, fails to capture the negative class well (F1-score 0.14)
- **M2**
  - About 75% churners
  - Decision Trees with M2-based training achieve an average F1-score of 0.91
  - Shows to be balanced in both classes (0.92/0.81)
- **M2** with “absence time” feature is more effective
Feature Importance for P1 (I)

The chart shows the relative importance of various features for P1 (I) using two data generation methods: Method 1 and Method 2.

The features are categorized into different groups, including:
- Number of Sessions
- Number of Days
- Avg. Playtime per Session
- Avg. Time Between Sessions
- Playtime Model Parameters
- Retention Value
- Premium User Flag
- Predefined Spending Cat.
- Number of Purchases
- Avg. Spending per Session
- Current Absence Time

Each feature is represented by a bar, with the height indicating its relative importance. The bars are color-coded to distinguish between Method 1 (red) and Method 2 (teal). The chart highlights the importance of each feature across the two methods, aiding in the analysis and understanding of user behavior and preferences.
Feature Importance for P1 (II)

- Besides the number of occurrences, the position in the tree matters as well
  - for M1 “Number of Days” was always the root node
  - for M2 “Average Time Between Sessions” was most often the root node
Experimental Results for P2

- With increasing size of the window $d_{rlx}$, the class imbalance becomes more extreme
Experimental Results for P2

- With increasing size of the window $d_{rlx}$, the class imbalance becomes more extreme
- $d_{rlx} = 1$
  - 86% of players in the test-set churn in the window
  - the averaged F1-score over 5 games is 0.91, however, F1-score for the non-churners drops to 0.41
Experimental Results for P2

- With increasing size of the window $d_{rlx}$, the class imbalance becomes more extreme
- $d_{rlx} = 1$
  - 86% of players in the test-set churn in the window
  - the averaged F1-score over 5 games is 0.91, however, F1-score for the non-churners drops to 0.41
- $d_{rlx} = 2$
  - 90% of players in the test-set churn in the window
  - the averaged F1-score over 5 games is 0.93, however, F1-score for the non-churners is at 0.39
Experimental Results for P2

- With increasing size of the window $d_{rlx}$, the class imbalance becomes more extreme

- $d_{rlx} = 1$
  - 86% of players in the test-set churn in the window
  - the averaged F1-score over 5 games is 0.91, however, F1-score for the non-churners drops to 0.41

- $d_{rlx} = 2$
  - 90% of players in the test-set churn in the window
  - the averaged F1-score over 5 games is 0.93, however, F1-score for the non-churners is at 0.39

- Except for one game, the games suffer from high churn rates and finding characteristics of the few non-churners is challenging
Conclusion

- Two definitions of churn
  - P1: hard churn
  - P2: soft churn with $d_{rlx}$
Conclusion

- Two definitions of churn
  - P1: hard churn
  - P2: soft churn with $d_{rlx}$

- Two data generation methods
  - M1: sampling based
  - M2: date centric

- Usage of Decision Trees with universal behavioral features
  - Not only descriptive, but also predictive
  - Evaluation on 5 different mobile games achieves high quality results (F1-scores between 86% and 91%)
Conclusion

- Two definitions of churn
  - P1: hard churn
  - P2: soft churn with $d_{rlx}$

- Two data generation methods
  - M1: sampling based
  - M2: date centric

- Usage of Decision Trees with universal behavioral features
Conclusion

- Two definitions of churn
  - P1: hard churn
  - P2: soft churn with $d_{rlx}$
- Two data generation methods
  - M1: sampling based
  - M2: date centric
- Usage of Decision Trees with universal behavioral features
- Not only descriptive, but also predictive
Conclusion

- Two definitions of churn
  - P1: hard churn
  - P2: soft churn with $d_{rlx}$

- Two data generation methods
  - M1: sampling based
  - M2: date centric

- Usage of Decision Trees with universal behavioral features

- Not only descriptive, but also predictive

- Evaluation on 5 different mobile games achieves high quality results (F1-scores between 86% and 91%)
Take Away Message — Revised

Players who have been playing with a low average time between sessions for a few days are likely to be non-churners if they have been playing in the last hours.
“Players who have been playing with a low average time between sessions for a few days are likely to be non-churners if they have been playing in the last hours.”
Thank You For Your Attention!

Questions?
References I

How players lose interest in playing a game: An empirical study based on distributions of total playing times.
In Proceedings of the IEEE Conference on Computational Intelligence and Games (CIG).

A long-term study of a popular mmorpg.

Churn prediction for high-value players in casual social games.
In Proceedings of the IEEE Conference on Computational Intelligence and Games (CIG).