Inferring Personality of Online Gamers by Fusing Multiple-View Predictions

Jianqiang Shen, Oliver Brdiczka, Nicolas Ducheneaut, Nicholas Yee, and Bo Begole

Palo Alto Research Center, 3333 Coyote Hill Road, Palo Alto, CA 94304, USA
{jianqiang.shen, oliver.brdiczka, nicolas, nyee, bo}@parc.com

Abstract. Reliable personality prediction can have direct impact on many adaptive systems, such as targeted advertising, interface personalization and content customization. We propose an algorithm to infer a user's personality profile more reliably by fusing analytical predictions from multiple sources including behavioral traces, textual data, and social networking information. We applied and validated our approach using a real data set obtained from 1,040 *World of Warcraft* players. Besides behavioral and social networking information, we found that text analysis of character names yields the strongest personality cues.

Keywords: personality, behavior analysis, social networks, sentiment analysis, virtual worlds

1 Introduction

Computer systems and devices become "smarter" every day thanks to enhancements in usability and adaptivity. In order for computer systems to further adapt to different users, there is a growing need for fine-grained modeling of preferences and, in particular, the personality of users. Modeling personality based on online behavior could enable better personalization, collaboration and targeted advertising, among others. For instance, personalizing user interfaces and content could improve work efficiency by steering users to the right information. In the workplace, employers could form efficient teams based on compatible personalities. Matching personality types also has commercial applications. For example, if online dating service providers had a better knowledge of a user's personality, they could match him/her with other users with a higher chance of success.

The popularity of online games offers a great opportunity to examine personality inference using significant amounts of data. Indeed, recent online games offer a wealth of behavioral indicators ranging from combat statistics to sociability that could reflect a player's personality. In this paper, we leverage this behavioral richness by attempting to infer the personality of the individual behind a game character. We use data from the popular massively multiplayer online game (MMOG), *World of Warcraft* (WoW), one of the most successful in its genre with close to 11 million subscribers worldwide (http://mmodata.net) who collaboratively accomplish a wide range of activities, from group combat (against tough computer-controlled "bosses" or other players) to crafting virtual items. Participation in the game world requires considerable time from players [6] and social bonds formed in these virtual worlds often translate to lasting relationships in and out of the game [20].

Recent research has shown that when meeting a stranger face-to-face for the first time, it is possible to quickly assess their personality with some accuracy thanks to verbal and non-verbal cues [14]. In this paper, we attempt a similar personality assessment based on a different set of cues, namely, digital traces generated by a player's activities in the virtual world. To maximize our chances of predicting personality accurately, we consider three data sources: behavioral metrics (e.g. achievements, number of kills), textual data (e.g. names chosen for a character) and social network information (e.g. a player's position in guilds). We build on the intuition that, much like during face-to-face encounters, the activities of the players are cues to their personality. For instance, shy and quiet players might prefer solo activities such as cooking and fishing. Outgoing players might prefer large-scale "raids" involving up to 40 players. Predictions from each data source provides a partial and complementary view of personality, which we then fuse with the others to get the most accurate personality profile possible.

2 Related Work

Personality Profiling. In personality psychology, the Big-5 model is the gold standard. The model measures five traits: Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness to Experience. Personality profiles are usually constructed through surveys based on proven inventories of questions, e.g. International Personality Item Pool [8]. Studies have shown that judgments of personality at zero acquaintance are moderately accurate and based on consensual indicators [14]. For instance [14], observers generally agree that Extraverted individuals speak louder, with more enthusiasm and energy, and that they are more expressive with gestures. Interestingly, personality cues can also be found in the physical world: observers can predict the personality of strangers by looking at their offices and bedrooms [9], or even by examining their top ten favorite songs [19]. There has been some success at predicting personality in a meeting scenario using visual and acoustic indicators [17].

Communication & Personalities. Researchers have started to explore the possibility of predicting personality from electronic cues. It turns out that websites can be used to predict their owner's personality with high levels of consensus and accuracy from the observers [22]. Facebook profiles can be similarly revealing [2] and personality also influences the way one writes electronic text [7] - for instance, the nature and structure of email messages reveals personality traits, down to the simplest indicators: observers were able to infer personality solely on the basis of an email address [1]. We note however that most of these studies rely on human coders to categorize personality from text data. In this paper, we will use automated text analysis techniques instead.

Online Gaming Studies. The depth and breadth of activities available in contemporary MMOGs, coupled with their widespread adoption, has led researchers to use them as "virtual laboratories" for social science research [5]. Research has explored a wide array of issues emerging in these online gaming communities: their unique culture [21], players' motivations and psychology [26], their economic importance [3], their social life [23] - and, most relevant to this paper, the link between a player's online and offline personalities [27]. The latter, however, only considered behavioral indicators available on Blizzard's Armory, a public database of character statistics. In this paper, we extend Yee et al.'s [27] approach by also considering textual and social networking data. We also attempt to predict a player's personality based on information from a single character, whereas Yee et al. grouped information about a player across all their characters. We adopt this more challenging single-character approach since, in general, there is no way to determine if two characters belong to the same player. Only single-character predictions enable practical applications like personalization or targeted advertising.

Our Contribution and Its Implications. Our contribution is threefold. First, instead of aggregating multiple characters, we show it is possible to reliably infer a player's personality based on the activities of a single character. Second, in addition to behavioral features, we explore social networking and textual features. To the best of our knowledge, this is the first attempt at utilizing such rich, multi-pronged information for personality predictions in virtual worlds. Third, we present the constructed features and their predictive power in detail, in order to inform future work on personality. Besides improving personalization and recommendations, incorporating personality estimations into adaptive systems can benefit a wide range of functionalities: for instance, our personality predictor is part of a larger system being developed to detect anomalies and prevent malicious behaviors in corporate networks. Personality profiling enables us to focus on individuals having the motivation and capability to carry out attacks.

3 Personality in Virtual Worlds

To provide some important context for our findings, we give below a brief introduction to the Big 5 personality traits, provide a brief overview of WoW, and discuss our approach to collecting data from the game.

3.1 The Big 5 model

Personality traits are consistent patterns of thoughts, feelings, or actions that distinguish people from one another [13]. The Big 5 model was developed using factor analytic techniques on adjectives and descriptive phrases of people culled from an English corpus. The corresponding five factors have been shown to account for most individual differences in personality [10, 13]:

- **Extraversion** implies an energetic approach to the social world and includes traits such as sociability and positive emotionality. High scorers tend to be sociable, friendly, talkative; low scorers tend to be reserved, shy, quiet.

- **Agreeableness** is a tendency to be compassionate and cooperative rather than suspicious and antagonistic towards others. High scorers tend to be friendly, caring, cooperative; low scorers tend to be critical, rude, suspicious.
- **Conscientiousness** describes socially prescribed impulse control that facilitates task and goal-directed behavior. High scorers tend to be reliable, self-disciplined; low scorers tend to be disorganized, negligent.
- **Neuroticism** relates to emotional stability. High scorers tend to be nervous, sensitive, vulnerable; low scorers tend to be calm, relaxed, secure, confident.
- **Openness** describes the breadth, depth, originality, and complexity of an individual's mental life. High scorers tend to be original, curious, complex; low scorers tend to be conventional, narrow interests, uncreative.

3.2 World of Warcraft

WoW is one of the most popular commercial online games (http://mmodata.net). It is based on a typical leveling up formula seen in many role-playing games. Players start at level 1 and kill monsters to become higher level and acquire better equipment in order to kill bigger monsters. The game encourages players to collaborate - for instance, an important game mechanic is that players must create characters with different skill sets that complement each other: heavily-armored tank classes shield the group from enemy attacks while lightly-armored damage classes deal damage to enemies and healing classes restore health lost in combat. Players must choose to belong to one of two factions: Alliance or Horde. Each faction has five distinct races, e.g., Gnomes or Orcs. A variety of rules dictate where and when players may attack and kill each other. Thus, a distinction is made between PvP (player-vs-player) and PvE (player-vs-environment) activities. PvP activites can range from one-to-one duels to large 40 vs. 40 battlegrounds (BGs). And in general, it is a player's choice as to how much PvP activity they want to engage in. Players in WoW communicate via typed chat and might also use VoIP tools to communicate via speech. The game also provides a modest set of "emotes" (e.g., hug). Players are able to specialize in crafting professions and convert collected raw ingredients into finished goods, such as in tailoring or cooking. There is also a system of Achievements that tracks a variety of combat and non-combat objectives, including Achievements for zones explored, number of hugs given, and cooking proficiency. These Achievement scores provide a sense of how a player chooses to spend their time in WoW. Thus, overall, WoW offers a wide and varied set of rich behavioral cues to draw from. From class choice to amount of PvP activity, from number of emotes used to amount of world exploration, the game context offers a range of measurable behaviors.

3.3 Data Collection

1,040 game players were recruited from WoW forums, mailing lists, publicity on popular gaming sites and social media like Twitter. The age range was 18-65 and the average was 27.03 (SD = 8.21). 26% of participants were women. Participants were asked to list up to 6 WoW characters they were actively playing. This resulted in a total of 3,050 active characters. A 20-item survey measuring the Big-Five was drawn from the International Personality Item Pool [8]. Participants completed a web-based survey that gathered their demographic and personality information. Participants rated themselves on the items using a scale ranging from 1 (Very Inaccurate) to 5 (Very Accurate). For all personality traits, the distributions of participants' scores are roughly Gaussian, with means around 3 and standard deviations around 0.8.

Blizzard, the developer of WoW, provides public access to much of their game data at a website known as the Armory. By searching for a character's name, anyone can view details about their past activities, including how many hugs they have given, their equipment, etc. Using custom data collection software, we collected information about each of our participants' characters on a daily basis from 11/22/2010 to 05/29/2011 (the Armory is updated once a day). Armory profiles consist of hundreds of variables, often in a hierarchy. To avoid being inundated by low-level variables or including overlapping variables, we adopted an analytic strategy of looking at or generating high level variables where possible. This in turn produces more stable variables that map to psychologically meaningful concepts. For example, a notion of geographical exploration would seem to be better tracked by the overall count of zones explored rather than looking at any one particular zone. The resulting variables represent our first data source: behavioral metrics.

Blizzard designed WoW to be extensible through the use of addons: small programs written by players to extend or refine the game's user interface. One addon function can be used to gather limited but valuable data: the *who* command. For a character, typing *who* lists characters in the same game zone who are roughly the same level (± 5 levels), with an upper limit of 49 results returned. The intent is to facilitate the formation of groups. The command can be expanded to include additional parameters such as specifying a different game zone or a level range. It therefore becomes possible to conduct a census of the entire population at a given time by progressively cycling through small segments of the population, aggregating batches of 49 players or less to cover all players. Based on rate limits for each *who* query and server load, our designed addon captures a list of all active players on a server every 5 to 15 minutes. We then use this data to create our second data source: social networks, constructed on the basis of which characters are seen playing with each other.

Finally, we build on the intuition that character names are not chosen at random: they are a key marker of identity and players often come up with inventive or humorous uses of names [4]. The same logic applies to guild names. As such, character and guild names constitute our third data source: text data.

4 Personality Inference

We use the three data sources above to build personality predictors and fuse the classifiers. This leverages the different representations of the patterns for classification, which has been shown to increase efficiency and accuracy [15], especially when classifiers are diverse.

4.1 Behavioral Information

To reduce the number of variables to a manageable and meaningful set, we follow the aggregation strategy described by Yee et al. [27]. This yields 68 high-level behavioral metrics that can be extracted from any character's profile on the Armory. The corresponding nine broad categories are as follows:

- Achievements: total achievements, profession achievements (cooking, first aid), achievements from group tasks (10-man dungeons), and their ratios.
- Death: we count the number of deaths in different game areas (such as raids, 10-man dungeon, 25-man dungeon, falling, fatigue, drowning, fire).
- Respecs: a player can switch their character's skill set by paying a fee. We count the corresponding number of these "respecs".
- Travel: we count the uses of each game transportation system (such as summon, flight, portal, hearthed)
- Emote: a character can communicate with other characters through emotes (such as hug, LOL, wave). We count the number of each different emote.
- Equipment and pets: a character can collect or buy equipment and pets in the game. We also differentiate purchased equipments from looted equipment.
- Need/greed rolls: valuable equipment drops from monsters are given to players according to dice rolls. Players select to roll based on "Need" or "Greed", of which the former is given higher priority.
- PvP scenarios: we count participation in PvP events of each type (such as arenas, duels, battlegrounds).
- Damage and healing: we count the sum and ratio of damage/healing points.

We chose to use regression trees to learn personality predictors from those features because of their simplicity, efficiency and accuracy.

4.2 Text Analysis

We use *sentiment analysis*, *keyword lists* and *n-grams* to generate features from character and guild names. Sentiment analysis is the task of identifying positive and negative opinions, emotions, and evaluations [24]. To do so, we use two sentiment polarity dictionaries. First, we use a dictionary from Wilson et al. [24] containing human-annotated polarity information on 8,221 distinct words. Second, we also use an in-house sentiment dictionary in which the overall polarity of a word was determined by a statistical approach. This dictionary has higher coverage (76,400 words) but lower precision, since the polarities of many words are context dependent. Each dictionary is used to produce separate features. Using these dictionaries, we scan each guild and character name and count how many positive/negative/neutral words appear. A sentiment word can be an adjective, adverb, noun, verb or any part of speech. We count the frequency of each case. We also count the frequency of strongly/weakly subjective words. A word that

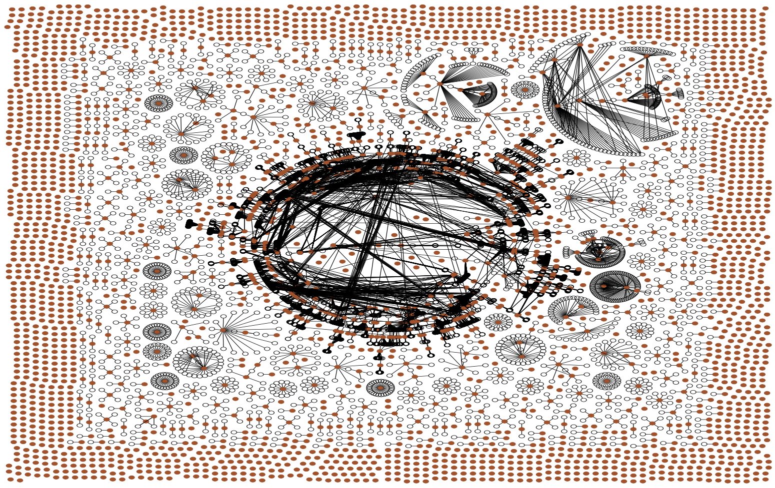


Fig. 1. Part of the constructed social network, including the 3,050 targeted characters and their direct neighbors. Filled nodes are our targeted characters. The entire graph contains more than 135,000 nodes and the frequency threshold to set an edge is 4.

is subjective in most contexts is considered strongly subjective (e.g., "abusive", "naive") ; otherwise it is weakly subjective (e.g., "accept", "neat").

We further created a game keyword list and check if a name contains those keywords. They include race names (e.g., elf, gnome), role names (e.g., priest, warrior), actions (e.g., kill, wave), failures (e.g., drown, fatigue), scenarios (e.g., arena, dungeon) and other frequent words. We currently collect 80 keywords.

In the textual analysis domain, n-gram analysis is a popular technique that uses sliding window character sequences in order to aid classification. To capture other hidden patterns in character and guild names, we also construct n-grams from names. An n-gram is a subsequence of n letters from a given sequence. For example, if the character name consists of 4 letters – ABCD, then we will have bigram AB, BC, CD. We limit n to 4, i.e., we only consider bigrams, trigrams and 4-grams. A larger n adds too much computation complexity and does not improve accuracy much. In many cases, a character's name is related to the player's other choices in virtual worlds, such as race and gender. Thus we include the character's region, virtual gender, race, role and faction as additional features. We train regression trees to profile personality from the text information.

4.3 Social Network Analysis

We hypothesize that personality traits can be detected through the nature and structure of a character's social activities (for instance, [27] shows that Extraverted characters have higher social connectivity on average than Introverts). We therefore attempt to analyze a character's social network to generate predictive features. We use a simple heuristic to build social networks from activity logs: in the networks, each node represents a distinct character; if the frequency that two characters were observed playing for the same guild, at the same location, at the same time is more than a specified threshold θ, we add an undirected

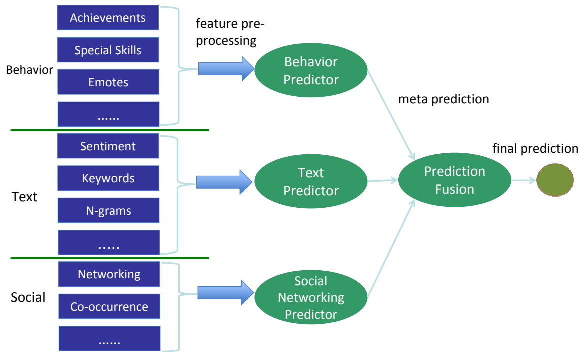


Fig. 2. We fuse several predictions together to get the final prediction of personality.

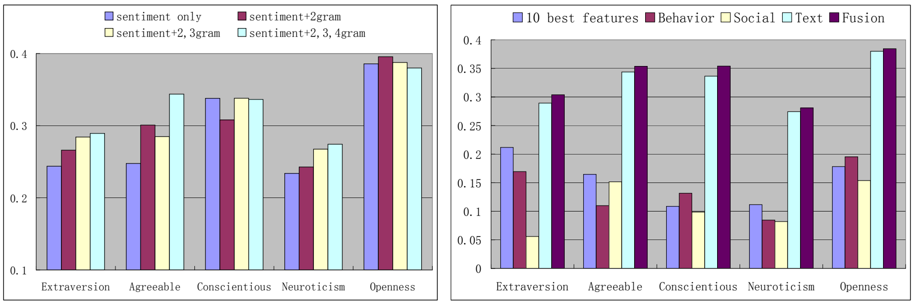
edge between those two characters. By specifying different θ values (we used 9, 6, 4 here), we get different networks. A partial graph by specifying θ=4 is shown in Figure 1. We then analyze the network to compute the following graph characteristics [18] for each node:

- **Degree centrality**: number of edges attached to a node, i.e. a measure of network activity for a node.
- **Betweenness centrality**: nodes that occur on many shortest paths between other nodes have higher betweenness than those that do not. Nodes with high betweenness have greater influence over what flows in the network.
- **Closeness centrality**: nodes that tend to have short geodesic distances to other nodes in the graph have higher closeness. They are in an excellent position to monitor the information flow in the network.

We enhance these social networking features by calculating co-occurrence heuristics. We hypothesize that a socially active person is likely to visit crowded places. For each character, we count the total number of characters playing in the same zone for a given play session. We normalize these values by taking into account the size of zones. We calculate their maximum values, minimum values and histograms as our features and input into regression trees for prediction.

4.4 Fusing Predictors

Classifier fusion has received considerable attention for pattern recognition in the past decade [15, 16]. By combining individual outputs, classifier fusion aims for a higher accuracy than that of the best classifier. It has been observed that, although one of the classifiers could yield the best performance, the sets of patterns misclassified by the different classifiers would not necessarily overlap [15]. Thus different classifier designs potentially offer complementary information about the patterns to be classified, which could be harnessed to improve accuracy. Therefore, fusing classifiers is particularly useful if they are inherently different - hence our decision to fuse predictors from behavioral information, text analysis and social network analysis. The corresponding diagram is shown in Figure 2. We train a separate predictor on each information source, and predictions from each predictor are then fused into one final prediction through linear regression.



(a) Correlation from text analysis. (b) Correlation from different methods.

Fig. 3. Correlation of personality predictions with the real values.

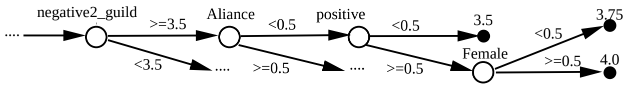


Fig. 4. Part of the regression trees for predicting Openness.

5 Experimental Results

We evaluate our approach with the aforementioned data from 3,050 WoW active characters. Personalities are coded as real numbers. It is important for the prediction ranking to be close to the real ranking, i.e., if a person has a high "real" value on a personality trait, it is good if the prediction is correspondingly high. We evaluate our approach with the *Pearson correlation* defined as the covariance of the two variables divided by the product of their standard deviations [11]. Results are based on 10-fold cross validation [12].

We found that character and guild names contain rich information on personality. For example, "sin" in names seems to correspond to low Agreeableness, and "hall" and "warrior" seem to correspond to low Extraversion. Sentiment analysis is especially powerful here. The results are plotted in Figure 3(a). We include features from the keyword list in all results, since names are related to the player's other choices in the virtual world. Sentiment analysis plus keyword features generates very good results. Adding n-grams can generally improve accuracy. As we increase n, we get better results in most cases except for Openness. But when n is larger than 4, it introduces too many tokens and decreases the efficiency. We also did not see too much accuracy benefit from large n. Thus we limit n to 4. We reviewed the generated trees and found that parts of them look meaningful. For example, part of the trees for Openness are shown in Figure 4. This tree suggests that if a guild's name contains many negative words from the in-house dictionary, a character with a name containing positive words from Wilson et al.'s dictionary [24] would have higher Openness than other characters.

The complete generated social network when θ=4 contains 135,547 nodes and 30,922 edges. Part of the constructed graph is shown in Figure 1. The graph is relatively sparse and many nodes are singletons. Co-occurrence heuristics are an important supplement and are used to enhance social networking features.

Table 1. Rankings and information gain values of some top informative features

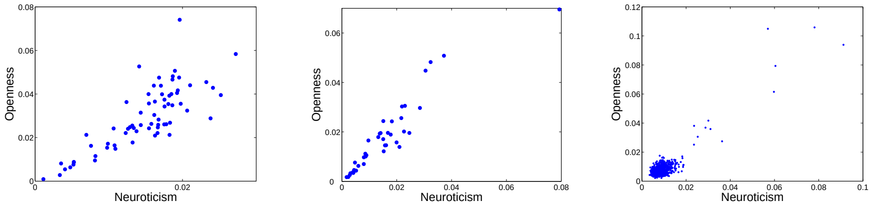
Rank	IG	Feature
Behavior		
1	0.0423	Dungeon-based achievements divided by the sum of all achievements
2	0.0421	Ratio of the count of "need" rolls to the count of all rolls
4	0.0289	Ratio of won duels divided by the count of all duels
5	0.0287	Count of reaching the highest status with a specific faction
8	0.0279	Sum of damage this character created
Text		
1	0.091	Count of negative words in the guild name (in-house dictionary)
2	0.090	Count of positive words in the guild name (in-house dictionary)
3	0.066	Count of negative words in the character name (in-house dictionary)
4	0.057	Count of positive words in the character name (in-house dictionary)
6	0.037	Count of strong subjective words in the guild name (UPitt dictionary)
Social		
1	0.073	Degree centrality when the frequency threshold to set an edge is 1
2	0.041	Frequncy that this character played with fewer than 5 characters
5	0.027	Frequncy that this character played with more than 20 characters
6	0.026	Frequncy that this character played with fewer than 5 guild members
8	0.017	Closeness centrality when the frequency threshold to set an edge is 4

Performance from different methods is shown in Figure 3(b). Yee et al. [27] suggested 10 behavioral features that have highest correlation with each personality trait among all game players. We show results from training a regression tree with those features. Our fusion methods always significantly outperform the "10 best features" method (p <0.05).

Text analysis gives better performance than behavioral and social networking information (p <0.05). We hypothesize that there could be two reasons. First, behavioral and social networking information were collected from a 6-month period and might not capture the whole picture. Second, behavioral and social networking information can be noisy whereas the character's name, guild, virtual gender, race, role and faction are usually chosen by the player after some careful thinking. Analyzing such relatively clean, stationary data can therefore let us gain some insight about the player's personality. Behavioral features work best for personality traits like Extraversion and Openness, while social network features work best for personality traits like Agreeableness and Openness. Extraversion is reflected in group and solo activities, such as dungeon, raid, questing, and cooking. Agreeableness is related to emotes (e.g. hugs) and player-vs.-player activities such as arenas and duels. Openness is reflected in exploration and also through activities like professions and dungeons. Prediction fusion takes advantage of the diversity of different predictors and gives the best performance in all Big 5 personality traits. Considering this is a difficult problem, it is exciting that the fusion method can achieve correlations higher than 0.35 in general.

Our algorithms can capture common characteristics between characters belonging to the same player and the predictions are well correlated. There are 813 players having at least 2 characters and we randomly sampled two groups. Each group sampled one character from each player and we checked the predictions. The two groups had correlation of 0.538, 0.485, 0.666, 0.480 and 0.623 for Extraversion, Agreeableness, Conscientiousness, Neuroticism and Openness.

Information Gain (IG) [25] is a well-known criterion to measure a feature's power. It calculates the reduction of entropy in the predicted class distribution provided by knowing the value of a feature. We discretize our features by simple



(a) Behavior features (b) Social features (c) Text features

Fig. 5. Information gains of different features. Each point corresponds to a feature.

binning [12] and measure the information gain for each personality trait. We sort features based on their average IGs of 5 traits. Some top features are shown in Table 1, with some similar features skipped for brevity. We note that different personality traits usually have similar impact on cues, i.e., for most features, their IG values usually have the same magnitude across different traits. We calculated the correlation between features' IGs for different traits and found it is high. For example, for Neuroticism and Openness, the correlation for behavioral features is 0.803, for textual features it is 0.941, and for social features it is 0.934. Figure 5 shows scatter plots with axis values as IGs. It is clear that, in general, good features for Openness are also good for Neuroticism while bad features for Openness are also bad for Neuroticism, although their strength is different, which might contribute to the accuracy difference of predictions. It is also worth noting that though combining *n*-gram text features together provides good personality cues, many of them alone have weak IGs.

6 Conclusion

Reliable personality inference has important personal and commercial applications. The depth and breadth of activities in online games, coupled with their widespread adoption, make them a good platform to examine personality inference approaches. In this paper, we attempted to infer the personality of the player behind a game character based on data from the popular MMOG, *World of Warcraft*. We profile a person's personality by fusing analytic predictions from multiple sources, including behavioral metrics, textual analysis and social networking information. Each source provides a partial and complementary view about the player's personality. In addition to behavioral and social networking information, we found that names contain strong personality cues.

Acknowledgments

We gratefully acknowledge support for this work from DARPA through the ADAMS (Anomaly Detection At Multiple Scales) program Contract W911NF-11-C-0216. This research was also sponsored by the Air Force Research Laboratory. Any opinions, findings, and conclusions or recommendations in this material are those of the authors and do not necessarily reflect the views of the government funding agencies. We thank Kurt Partridge and anonymous reviewers for the constructive comments.

References

1. M. D. Back, S. C. Schmukle, and B. Egloff. How extraverted is honey.bunny77@hotmail.de? Inferring personality from e-mail addresses. *Journal of Research in Personality*, 42(4):1116–1122, 2008.
2. M. D. Back, J. M. Stopfer, S. Vazire, S. Gaddis, S. C. Schmukle, B. Egloff, and S. D. Gosling. Facebook profiles reflect actual personality, not self-idealization. *Psychological Science*, 21:372–374, 2010.
3. E. Castronova. On virtual economies. *Game Studies*, 3(2), July 2002.
4. L. Cherny. *Conversation and community: chat in a virtual world.* CSLI Pub, 1999.
5. N. Ducheneaut. Massively multiplayer online games as living laboratories: opportunities and pitfalls. *Online Worlds: Convergence of the Real and the Virtual*, Springer, 2010.
6. N. Ducheneaut, N. Yee, E. Nickell and R. Moore. Building a MMO with mass appeal: a look at gameplay in World of Warcraft. *Games and Culture*, 1:4:1–38, 2006.
7. S. D. Gladis. *Writetype: Personality Types and Writing Styles.* HRD Press, 1994.
8. L. R. Goldberg, J. A. Johnson, H. W. Eber, R. Hogan, M. C. Ashton, C. R. Cloninger, and H. C. Gough. The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40:84–96, 2006.
9. S. D. Gosling, S. J. Ko, T. Mannarelli, and M. E. Morris. A room with a cue: Personality judgments based on offices and bedrooms. *Journal of Personality and Social Psychology*, 82(3):379–398, 2002.
10. R. Hogan, J. Johnson, and S. Briggs. *Handbook of Personality Psychology.* Academic Press, third edition, 1997.
11. R. V. Hogg. *Introduction to Mathematical Statistics.* Prentice Hall, 5th edition, 1994.
12. T. Joachims. *Learning to Classify Text Using Support Vector Machines.* Kluwer, 2001.
13. O. P. John, R. W. Robins, and L. A. Pervin. *Handbook of Personality: Theory and Research.* The Guilford Press, third edition, 2010.
14. D. A. Kenny, C. Horner, D. A. Kashy, and L. chuan Chu. Consensus at zero acquaintance: Replication, behavioral cues, and stability. *Journal of Personality and Social Psychology*, pages 88–97, 1992.
15. J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas. On combining classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20:226–239, 1998.
16. L. I. Kuncheva. A theoretical study on six classifier fusion strategies. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24:281–286, 2002.
17. B. Lepri, N. Mana, A. Cappelletti, F. Pianesi, and M. Zancanaro. Modeling the personality of participants during group interactions. In *UMAP-09*, pages 114–125, 2009.
18. M. Newman. *Networks: An Introduction.* Oxford University Press, 2010.
19. P. J. Rentfrow and S. D. Gosling. Message in a ballad: The role of music preferences in interpersonal perception. *Psychological Science*, 17(3):236–242, 2006.
20. A. F. Seay, W. J. Jerome, K. S. Lee, and R. E. Kraut. Project massive: A study of online gaming communities. In *Proc. of CHI-04*, pages 1421–1424, 2004.
21. T.L. Taylor. *Play between worlds.* Cambridge, MA: The MIT Press.
22. S. Vazire and S. D. Gosling. E-perceptions: Personality impressions based on personal websites. *Journal of Personality and Social Psychology*, 87:1:123–132, 2004.
23. D. Williams, N. Ducheneaut, L. Xiong, Y. Zhang, N. Yee, and E. Nickell. From treehouse to barracks: The social life of guilds in WoW. *Games and Culture*, 1(4), 338–361, 2006.
24. T. Wilson, J. Wiebe, and P. Hoffmann. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of HLT-EMNLP*, pages 347–354, 2005.
25. Y. Yang and J. O. Pedersen. A comparative study on feature selection in text categorization. In *Proc. of ICML-97*, pages 412–420, 1997.
26. N. Yee. The Norrathian Scrolls: A Study of EverQuest. *www.nickyee.com/eqt/report.html*
27. N. Yee, N. Ducheneaut, L. Nelson, and P. Likarish. Introverted elves & conscientious gnomes: The expression of personality in world of warcraft. In *CHI-2011*, pages 753–762.