Proceedings, The Eleventh AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE-15)

A Data-Driven Approach for Computationally Modeling Players' Avatar Customization Behaviors

Chong-U Lim and **D. Fox Harrell**
Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology
77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA
{culim, fox.harrell}@mit.edu

Abstract

Avatar customization systems enable players to represent themselves virtually in many ways. Research has shown that players exhibit different preferences and motivations in how they customize their avatars. In this paper, we present a data-driven analytical approach to modeling player behavioral patterns exhibited during the avatar customization process. We used our data mining tool *AIRvatar* to analyze telemetry data obtained from 190 players using an avatar creator of our own design. Using non-negative matrix factorization (NMF) and N-gram models, we demonstrate how our approach computationally models behavioral patterns exhibited by players such as "regular shopping," "engaged shopping," or "bored browsing". Our models obtained significant effect sizes ($0.12 <= R^2 <= 0.54$) when validated with multiple linear regressions for players' time spent engaging in activities within the avatar creator. The NMF model had comparably high performance and ease of interpretation compared to control models.

Introduction

In videogames, players often construct virtual identities in the form of characters and avatars. Despite being virtual, these identities can reveal aspects of a player's real-world identity (e.g., preferences, control, appearance, understanding of social categories, etc.,) by being projected onto the actual implemented avatars in what is termed a "blended identity" (Harrell 2013). Research has demonstrated that the way players behave in both the real and virtual worlds can be influenced by these virtual avatars (Harrell and Harrell 2012; Yee and Bailenson 2007; Yee et al. 2012). Given the importance of these virtual representations to players, we thus seek to model and better understand the behavioral patterns of players during the avatar customization process.

In this paper, we used our data mining tool *AIRvatar* to analyze telemetry data obtained from players using an avatar creator of our own design. Our approach computationally models interpretable behavioral patterns exhibited by users using non-negative matrix factorization and N-gram models. To the best of our knowledge, using data-driven approaches to model behavioral patterns during avatar customization to gain insight into players' values and preferences for their virtual representations has not been previously undertaken.

Previous Work In our previous work, we studied "infrastructural values" that were built into systems by analyzing the end-product of constructed avatars, including how statistical attributes and visual characteristics reflected social phenomena potentially symptomatic of developer bias or implicitly shared worldviews (Lim and Harrell 2015a; 2015b; 2015c). Building on those findings, here we focus on modeling "user values" through behaviors enacted out within systems, which particularly include the temporal properties of performing sequences of actions within a given system.

Related Work

We discuss related work useful for better understanding and distinguishing our aims and approaches used in this paper.

Avatars and Identity

A study of player behaviors in avatar customization for three virtual worlds conducted in (Ducheneaut et al. 2009) used self-reported surveys and timing-related data. It showed that players exhibited preferences for different features of their avatars (e.g., more importance on hair versus body.) Other studies on players' behavioral characteristics in avatar customization have shown various motivations for customization such as virtual exploration, social navigation, contextual adaptation, and identity representation (Lin and Wang 2014). Related studies in (Yee and Bailenson 2007; Yee et al. 2011) covered the behavioral effects that virtual identities had on the players and their real-world identities, such as conforming to expectations of their avatars appearance and correlations with personalities (e.g., conscientious players explored more of the game world). We hypothesize that, before immersion within the game world, behaviors exhibited during the customization can be computationally modeled to provide insight on players' identities and preferences.

Player Modeling and Game Telemetry Data Mining Game telemetry data mining is used to provide information to help designers gain insight into player behaviors exhibted within the game. These data are used to construct player models and we use the taxonomy proposed in (Smith et al. 2011), which are: *domain* (game actions or human reactions), *purpose* (generative or descriptive), *scope* (individual, class, universal, or hypothetical), and *source* (induced, interpreted, analytic, or synthetic). As detailed later, the models we develop are universal, descriptive models of game actions from both induced, interpreted, and synthetic sources. We use data clustering, which categorizes large data of players into smaller discrete categories, enabling designers to model patterns of players and use them for tasks like dynamically adapting to different styles, quantitatively evaluating user performance, and improving player experience and satisfaction (Drachen et al. 2012; Yannakakis and Hallam 2009; Yannakakis and Togelius 2011).

A study of common clustering methods of *World of Warcraft* characters in (Drachen et al. 2014) found that clusters differed in (1) whether they were easily interpretable, (2) how distinct they were from each other, (3) whether they depicted legal/valid states in the game, and (4) how representative of the original data set they were. We used non-negative matrix factorization (NMF) for its interpretable and distinct clusters (Hoyer 2004). NMF has been successfully applied to images (e.g., parts of faces), text (e.g., topic modeling), procedural content generation (e.g., level generation), and modeling players' identities and values (Lee and Seung 1999; Shaker and Abou-Zleikha 2014; Lim and Harrell 2015a). For analysis of sequences of actions performed by players, we used N-gram models, which have been effective in modeling sequences of player actions, solution features, and platformer level styles (Harrison and Roberts 2014; Butler et al. 2015; Dahlskog, Togelius, and Nelson 2014).

AIRvatar

The *AIRvatar* system collects analytical data as players interact with a given customization system. The two main types of data are (1) timing-related data (e.g., time duration spent interacting with aspects of the system) and (2) interaction-related data (e.g., item selection mouse clicks.)

Case-study: Heroes of Elibca

We developed an avatar customization system set in the context and style of a traditional computer role-playing game (RPG) called *Heroes of Elibca*. Resources and assets were obtained from publicly available sources (Mack Looseleaf Creator 2015; Liberated Pixel Cup 2015). A text-based introductory story provided players with the motivation of creating a character for a fantasy-style setting and game genre.

Customization Interface A screen shot of the customization interface is shown in Figure 1. Players first choose a **gender** for their character out of two[1] available choices: male and female. Customization options then appear for players to select appearance options across five categories: hair, head (face), body, arms, and legs. For each **category**, several sub categories provide more fine-grain selection options, for example, the "head" category may provide "eye color", "eye shape" and "facial hair" sub-categories. The image of each customization **item** is shown, which the players can click to select. For some items, a color palette appears, whereby clicking on them refreshes the avatar with different **color** variations of the currently selected item. An animated preview of the character in a walk-cycle is shown against a backdrop. Players can use the **animation-control** buttons to cycle through the four directional views of the character: right, left, back, and front, and to start or stop the animation.

[1] We recognize the distinction between gender and sex, but follow RPG conventions here. Other models of gender go far beyond male/female gender binary and is an important area for future work.

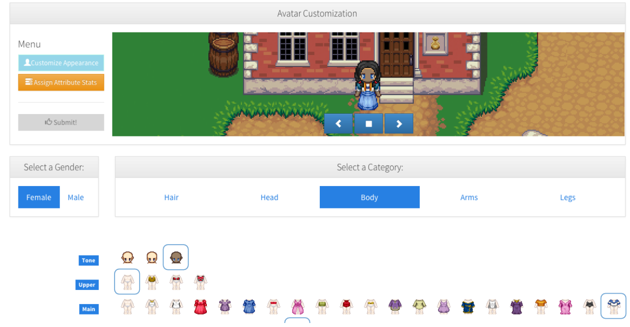

Figure 1: The customization interface of *Heroes of Elibca*.

User Study We conducted a user study, approved by the human subjects research committee at our institution, with participants from the social news and discussion site *Reddit*. Participants were informed its research nature and that analytical data would be anonymously collected. 104 participants (54%) identified as "Male", 81 (43%) identified as "Female", and 6 (3%) listed "Other." 154 participants (80%) were between "18-21" years old, 32 (17%) were between "25-34" years old, and the other age groups were $< 1\%$.

Methods

We describe how the data collected on behaviors exhibited during customization were computationally modeled.

Constructing Interaction Sequence Vectors

Interaction sequences were modeled using only the top-level click event (e.g., "color-selection") without the specific choice being made (e.g., "blue"). With five types of click events, a player who starts of by (1) selecting a "male" gender, (2) choosing to customize the "hair" category, (3) picking a choice of hair fringe style, (4) selecting a red variant of the hair fringe, (5) then picking a choice of the top of the hair, and (6) selecting a blue variant of the top of the hair would have a interaction sequence of: `<gender, category, item, color, item, color>`. The interaction sequence is converted into numeric vector using the 1-of-k encoding scheme. Thus, the vector of the previous interaction sequence would be: `<00010 01000 00001 00100 00001 00100>`.

Developing the NMF Model

Non-negative matrix factorization (NMF) is an algorithmic process for representing data as a combination of derived factors. Given data points $V = \{x_1, x_2, ..., x_n\}$, $V \in \mathbb{R}_{n \times m}$, NMF results in an approximation $\hat{V}$ that is the product of two matrices $W \in \mathbb{R}_{n \times k}$ and $H \in \mathbb{R}_{k \times m}$, with elements $v_{ij} \in V$, $w_{ij} \in W$, and $h_{ij} \in H$ all ≥ 0. The value k is the number of sub components desired (k≪n,m). NMF minimizes the difference $||V - \hat{V}||_F^2$, where $||X||_F^2 = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{m} |x_{ij}|^2}$ is the Frobenius norm of the matrix X. Each row in matrix H is an m-dimension *basis vector* and each column in matrix W relates each sample in v_i with each basis vector with *coefficients* $w_{ij} \in W$, representing the weight of basis vector j in sample x_i.

The data set of all interaction sequences is represented by the matrix $V \in \mathbb{R}_{n \times m}$, where $n = 190$ participants and m is the maximum interaction sequence length of the data set. In model construction, we specify the number k of desired basis vectors, each represented by rows in matrix $H \in \mathbb{R}_{m \times k}$. The matrix $W \in \mathbb{R}_{n \times k}$ represents the weights of each interaction sequence in terms of the k basis vectors. Because the result of the matrix decomposition may yield basis vectors $h_j \in \mathbb{R}^m$ with multiple non-zero values, we used the digit with the maximum value to perform the inverse transformation from a vector to an event sequence, enabling each basis vector to be meaningful represented as an interaction sequence. We minimized the Frobenius norm by increasing k, but also considered the performance of each basis vector when modeled as N-grams, which we describe next.

Pattern Discovery using N-grams

Let a sequence of n events from i to j be s_i^n and a single event at i be s_i. The probability of observing a given sequence is given as $P(s_i^n) = P(s_n|s_1^{n-1})$. An N-gram model uses the Markov assumption that the probability of observing an event at a given point of sequence depends only previous $N - 1$ events. Thus, $P(s_n|s_1^{n-1}) \approx P(s_n|w_{n-N+1}^{n-1})$. Given an interaction sequence, we used an N-grams to identify patterns in behaviors for interpretation and to assess its effectiveness. We varied $2 \leq N \leq 6$ during model construction and used the following steps decide find optimal values:

1. Created k N-gram models M_j using each basis vector.
2. Selected the top three individuals with the highest weight coefficient w_{ij} from NMF for each basis vector.
3. Identified the top five N-gram patterns to appear in them.
4. Evaluated each model M_j by calculating the probabilities of generating each of the top five N-gram patterns.

Optimal values of N and k were determined based on the best average performance of the N-gram models. N-gram model interpretation for each base vector h_j used the top five N-gram patterns (1) occurring in its sequence and (2) in the sequences of its top three individuals.

Correlations between NMF Model and Variables

We calculated Pearson's correlations between the weights in matrix W and each of the following data variables:

- t_{Total} (total session duration)
- $t_{Category_i}$ (time spent customizing a item category) $Category_i \in \{Hair, Head, Body, Arm, Leg\}$
- t_{Gender_j} (time spent customizing with avatar gender) $Gender_j \in \{Female, Male\}$
- $t_{Orientation_j}$ (time spent in rotation orientation) $Orientation_k \in \{Front, Right, Left, Right\}$

Model Validation with Multiple Linear Regressions

Multiple Linear Regression (MLR) was performed for each of the above dependent data variables. Backward-selection and t-tests were used to identify the basis vectors that contributed most to each model. The adjusted goodness-of-fit R^2 was used to evaluate the performance of each NMF model and demonstrating its explanatory power with the values 0.1 being low, 0.3 being medium, and 0.5 being high.

Control Models and Comparisons

In order to provide a comparison of our approach using NMF models, we conducted two sets of control experiments:

- Control #1 - Archetypal Analysis[2] (AA) (Cutler and Breiman 1994): Performs a similar matrix decomposition of a data set to NMF, but imposes convexity constraints on the resultant weight matrix W and basis vectors H, such that archetype $h \in H$ lies on the convex hull.
- Control #2 - Principal Component Analysis (PCA) (Jolliffe 2005) performs a matrix decomposition such that H is set to the eigenvectors of the data covariance.

The criteria for comparing the models were the (1) ease of interpretation of basis vectors and (2) explained variance when validated with MLR on dependent variables.

Results & Analysis

We present the results from the analyzing the data collected and assessing the constructed NMF and N-gram models.

Descriptive Statistics

Table 2 shows the frequency distributions of the collected events and the histogram in Figure 2 shows the distribution of interaction sequence lengths. We omitted one user with an unusually long sequence length and had N=190 remaining.

Click Event	min	max	mean	sd
e_{Gender}	1	15	1.5	1.6
$e_{Category}$	4	53	8.7	6.2
e_{Icon}	4	342	77	56
e_{Color}	3	280	57	48
$t_{Animation}$	0	240	28	40

Table 2: The frequency distributions of different event data.

NMF and N-gram Model Selection

From a scree plot of the Frobenius norm values, the elbow of the plot occurred between $7 \leq k \leq 9$. While using such a k can be viewed as more accurate (i.e., lower Frobenius norm), we aimed for a NMF model with a smaller k to (a) regularize our model against over-fitting and (b) favor a simpler model for interpretation. Thus, for $2 \leq k \leq 6$, we constructed N-gram models for $3 \leq N \leq 6$, as shown in Figure 3. We see that the 6-gram model performs best for all k, the 3-gram model worst for $k \geq 3$, and both 4-gram and 5-gram models fluctuate in performance. The relative performances of

[2]A variant called Convex Hull Non-negative Matrix Factorization (CHNMF) (Thurau, Kersting, and Bauckhage 2009) was used.

	Base Vector 1		
	Sequence (First 80 *events)*	*w*	*l*
	SCIAINCCICICCCICCCCICCICICCCCCACICIICAICIIIIIIIIIICIIIIAIICIIAIIIIIIIIIIIIIIICII	-	-
#1	SAAAAAAAACIIICICGIIICCICICIICICCIIIICICICGGIICIIICCCICICICCIICIIIIGICICCIICICIII	1.82	265
#2	SCICAAAIIIICCIIAAAAAIIACAAAAICAAAAAACACCICAACIIIIIIIIICCCCCICCICICICICIIIIIIIIII	1.80	245
#3	SCICICIIICCICCIICIICICICICICICICICICICICICICICICICICICICICICICIIIIIIIIIIIIIICICI	1.75	292
	Base Vector 2		
	Sequence (First 80 *events)*	*w*	*l*
	SCICCCNNNICIICNINICICICICICIICIIINNICIIICICICACACNNNCNNICICNCCCNNICACNCNANANNCAC	-	-
#1	SCIICCAAACIIAAAAICCCICICCICIIIIICICIIIICCICCIICICIICICICICIICIIICIICICCCICICICIC	3.76	792
#2	SGGCICGCCCCICCIICCCICICICCCICCICIIIIIGIIIGGGGGGGGGGICCICIIICICICICICCICIIICCCCCC	3.75	742
#3	SCICICICIICIICCCCCCICICCICCCICICICCIIICICCICCICCCCCCCICICICICCICIICAAAAAAAAIICAA	3.39	660
	Base Vector 3		
	Sequence (First 80 *events)*	*w*	*l*
	SCICICIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIICIIIIIIIIIIIICCIIIIIIIIICIIIICI	-	-
#1	SCICICIICICICICICICCCICIIIIIIIIICIICIIIIIICIIICCICIICICIIICICIICICICIIIICIICCICI	1.43	112
#2	SCICICICCCCCICCICICICICICIIIIIIIIIIIIIIIIICIICCCIIIICIIICCICICCCICICICICICCICCCC	1.42	112
#3	SCICICCICICICIIIIIICICCICIIICICICCICIICICIICCICCCCIICCCCICCCCCCICICCICCICIIICICI	1.36	114
	Key: G Gender C Category I Icon C Color A Animation Control		

Table 1: The table shows the first 80 events for each obtained basis vector and its top three weighted individuals.

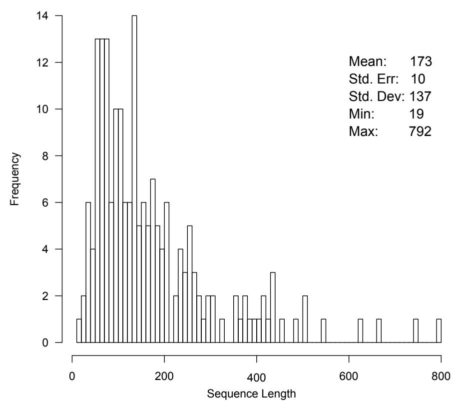

Figure 2: Variation of lengths of interaction sequences.

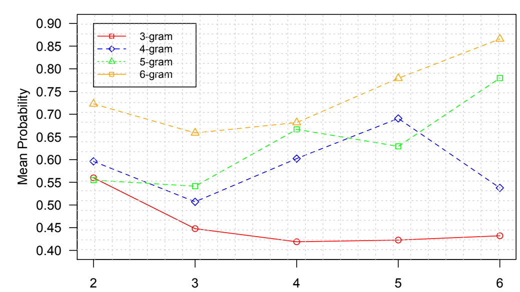

Figure 3: Performance comparisons of different N-gram models with varying number of NMF basis vectors k.

the models are most consistently in-order within the range $k = [3, 4]$. To balance the trade-off between increasing k for more representative basis vectors (lower Frobenius norm), increasing N for N-gram performance, and favoring a simpler model (lower k), we settled on values $k = 3$ and $N = 6$.

Interpreting Basis Vectors

Table 1 shows the first 80 events for each resultant basis vector and the top-3 weighted sequences. Top sequences of basis vector 2 had the longest sequence length followed by basis vector 1 and basis vector 3. For interpretation, we use the N-gram frequency tables for each basis vector (Table 3) and the top 3 sequences (Table 4). We discuss them as follows:

1. Basis vector 1 has N-grams that feature Icon events with $\leq$ 1 Color event within the sequence (e.g., IIIIII or IIIIIC). The N-grams from its top three individuals share these properties. The distinctive N-gram is the sequence of Color events (i.e., CCCCCC). It depicts the behavior of trying different color variations regularly. We thus interpret this basis vector as "regular shopping."
2. Basis vector 2 has N-grams that have $\geq$ 2 Color events. The N-grams from its top three individuals show particularly interesting characteristics. There is a high occurrence of a sequence of AnimationControl events (e.g., AAAAAA and AAAAIC) and the sequence of consecutive Color events. It depicts trying many different options and rotating the avatar. We thus interpret this basis vector to reflect the behavior of "engaged shopping."
3. Basis vector 3 has similar N-grams to basis vector 1. The N-grams from its top three individuals have patterns that are common to both basis vector 1 and 3's top individuals (e.g., ICICIC and CICICI). The main differences are the low frequency counts overall and a sequence with a Category event (i.e., CICICI). This seems to depict simply cycling through categories without trying out options. We thus interpret this basis vector as "bored browsing."

Base Vector 1		Base Vector 2		Base Vector 3	
N-gram	f	N-gram	f	N-gram	f
N N N N N N	510	N N N N N N	267	N N N N N N	664
I I I I I I	59	I C I C I C	9	I I I I I I	49
I I I I I C	12	C I C I C I	9	I I I C I I	5
I I I C I I	12	I C I C I I	8	I C I I I I	4
I I C I I I	12	C I I I C I	7	I I C I I I	4

Table 3: The N-gram frequency tables for each of the basis vectors. Only the top five N-grams are shown.

Base Vector 1		Base Vector 2		Base Vector 3	
N-gram	f	N-gram	f	N-gram	f
I I I I I I	59	A A A A A A	62	I I I I I I	18
C C C C C C	46	I C I C I C	57	I C I C I C	12
I C I C I C	37	C I C I C I	55	C I C I C I	9
C I C I C I	37	A A A A I C	51	I C I C I I	8
I I I I I C	16	C C C C C C	51	C I C I C I	7

Table 4: The N-gram sequences obtained from the top 3 individual sequences with the highest weights (i.e., most representative) for each basis sequence vector.

	Basis 1	Basis 2	Basis 3
Basis 1	1.0	0.15	-0.53
Basis 2	0.15	1.00	-0.57
Basis 3	-0.47	-0.57	1.00

Table 5: Correlation coefficients between basis vectors β weights of the NMF model. (All results at $p < .05$).

	Hair	Face	Body	Arm	Leg
Basis 1	0.52	0.23	0.30	-0.06	0.20
Basis 2	0.50	0.22	0.43	-0.02	0.06
Basis 3	-0.52	-0.24	-0.28	0.08	-0.13

Table 6: Correlation coefficients between basis vectors and durations customizing each category. (Non-gray: $p < .05$).

	Front	Back	Left	Right
Basis 1	0.02	0.40	0.18	0.23
Basis 2	0.05	0.67	0.23	0.29
Basis 3	0.01	-0.55	-0.19	-0.36

Table 7: Correlation coefficients between basis vectors and durations in different rotation views (Non-gray: $p < .05$).

NMF Model – Correlations

We found no significantly large correlations between the basis vectors and (1) the total time spent in the system or (2) time spent customizing avatars of particular genders. We omit them and focus on the remaining data variables.

Basis behavioral patterns appear to coherently depict the different interpretations of customization behaviors. The calculated correlations between the weights w_j for each basis vector in our data set are shown in Table 5. We observe that basis vector 1 and 2 are weakly positively correlated (α= .15, p<.05) with each other, while basis vector 3 is highly negatively correlated with basis vector 1 (α= -.53, p< .05) and basis vector 3 (α= .57, p $<$ .05), coinciding with our earlier interpretations. A player exhibiting "bored browsing" (Basis 3) is less likely to demonstrate "regular shopping" (Basis 1) behavior, and even less likely to have "engaged shopping" (Basis 2) behaviors. It also shows that players may exhibit different base patterns throughout a given session, but are less likely to transit from one extreme ("bored browsing") to the other ("engaged shopping").

Different basis behavioral patterns correlate differently based on the part of the avatar being customized. The calculated correlations between the weights w_j for each basis vector and the time durations spent customizing different categories of the avatar are shown in Table 6. We observe that basis vector 1 ("regular shopping") is correlated across more categories (4) than the other two (3) and basis vector 3 ("bored browsing") is negatively correlated across all significant categories. While mostly comparable, Basis vector 1 has higher correlations with the "leg" category while basis vector 2 has higher correlations with the "body" category. It suggests that regular customization behaviors try out different customization categories in general, but more engaged behaviors occur during the customization of the "body".

Rotating characters to different views correlates with behavioral patterns. The calculated correlations between the weights w_j for each basis vector and the time durations in different view rotations are shown in Table 7. We observe that basis vector 3 is highly negatively correlated with "back"-view rotations and moderately negatively correlated with "right"-view rotations. This makes sense as a player that is just "bored browsing" would not likely spend time rotating their characters. We observe that basis vector 2 ("engaged shopping") correlates positively with viewing the character from the three available alternate rotations and is particularly high (α= .67) for the "back"-view rotation. This seems to be in line with our basis vector interpretations, since while being engaged, a player would likely be carefully considering how their avatar looks from multiple different rotational views. Basis vector 1 ("regular shopping") has moderate positive correlations with the "back" and "right"-rotation views, indicating some levels of interest in viewing the avatar from at least the side and back profiles.

NMF Model – Multiple Linear Regressions

The results of the MLR models constructed for each of the dependent data variables are shown in Table 8. We only included models that were significant ($p < .05$) and listed them in decreasing order of effect size (adjusted R^2).

Engaged behaviors can particularly predict time spent customizing avatar bodies and rotating avatars. While both coefficients for basis vector 1 and basis vector 2 are significant across most of the depedent variables with the exception of t_{right} and t_{leg}, the coefficients for basis vector 2 (β_2) are higher than basis vector 1 for time spent on the avatar "body" and time spent in the "back" and "left"-view rotations. This indicates that engaged behaviors contribute more towards time spent immersed with particular aspects of

	$\beta_1(\times 10^3)$	$\beta_2(\times 10^3)$	$\beta_3(\times 10^3)$	*adj.* R^2
t_{back}	19.6***	31.2***	-	.542
t_{hair}	62.6***	47.9***	-	.450
t_{body}	113***	133***	-	.242
t_{right}	-	-	-31.8*	.129
t_{face}	81.5**	60.1**	-	.080
t_{left}	4.57*	5.16**	-	.064
t_{leg}	98.7**	-	-	.033

Table 8: MLR results of the NMF model.

	adj. R^2		
	NMF	AA	PCA
Valid/Interpretable?	Yes	Yes	No
t_{back}	.542	.320	.550
t_{hair}	.450	.147	.466
t_{body}	.242	.153	.251
t_{right}	.129	.033	.099

Table 9: Comparisons of NMF, AA, and PCA models based on MLR effect sizes and validity of the basis vectors.

the system such carefully considering customization options for the avatars' body by constantly rotating the character.

Regular shopping behaviors can predict time spent on customizing the avatar's head region and also general interest in customization avatar . Coefficients for basis vector 1 (β_1) are higher for time spent customizing avatar "hair" and "face". This indicates that regular behaviors of trying out different options contribute towards the head region of the avatar. We note that β_1 is less significant on both the "body" of the avatar and time spent in the "left"-view rotation, while being the only significant coefficient for t_{leg}.

Behavior patterns exhibiting boredom or disinterest can predict engagement with the customization system. The only variable where the coefficient for basis vector (β_3) was significant was for time spent in the "right"-view rotation. The negative sign indicates that "bored browsing" behavioral patterns contribute towards less time spent on observing avatars from multiple different views, which depicts a disengaged, disinterested, or dis-satisfied player.

Performance Comparisons with Control Models

Table 9 shows a comparison of the NMF approach with the two control models. We only show significant MLR results for dependent data variables that showed at least small effect sizes (*adj.*$R^2 \geq .1$). We observed that NMF and PCA performed significantly better than AA, suggesting that interaction sequences are better modeled as sub-patterns occurring within a sequence, as opposed to convex combinations of existing data points. While PCA slightly outperforms NMF, its basis vectors are not interpretable due to the nature of the decomposition in PCA resulting in principal component axes rather than points in the same representation space.

Discussion

We have found that computational models constructed from avatar customization interaction data can help to us gain insight about players' behavioral patterns in a quantifiable manner compared to relying on self-reported surveys. Since customization often occurs early, it can be useful for constructed models to be used in tailoring subsequent experiences for the player without requiring explicit input from them. For example, a player exhibiting high "engaged shopping" behavioral patterns could be presented with a detailed tutorial on subsequent gameplay, while one with "bored browsing" patterns could be given a less detailed, more action-focused tutorial. If models are constructed in real-time, they could be used to dynamically adapt to players, maximizing their engagement, retention, or prevent churn.

Limitations & Future Work

The N-gram models were trained on the NMF basis vectors and evaluated based on the their performance in predicting top N-grams from the top weighted individuals. Future work could evaluate each model against entire interaction sequences using metrics like perplexity of an N-gram model as a performance measure. With N-grams, we assumed that interaction behaviors would exhibit partial sequences with a Markovian assumption. This distinguished between patterns, such as cycling through categories versus approaching customization hierarchically (e.g., first hair, then head, etc.) However, non-Markovian models like the Sudden Relaxation Model (SRM) (Shushin 2005) or CMRules (Fournier-Viger et al. 2012) could provide alternative insights.

The avatar customization system for *Heroes of Elibca* featured retro-style, 16-bit 2D RPG sprites set in a fantasy setting. Behaviors would likely differ in other settings (e.g., a space themed game might show less focus on an avatar's head and body, but more on arms and legs.) We plan to apply *AIRvatar* to customization systems from other genres and of greater fidelity (e.g., 3D assets.) Beyond data variables from player behaviors (e.g., time durations), we plan to study other aspects of players. Preliminary results of NMF model validation against real-world aspects of identity (e.g., personality using the BIG-5 personality tests or demographic information) did not show significant results. A more tailored survey would improve results and enable us to collect information on more specific aspects of players.

Conclusion

We have presented an approach to computationally reveal and model player behavioral patterns during avatar customization for characters in a fantasy-themed 2D RPG context. Three basic patterns – regular shopping, engaged shopping, and bored browsing, were modeled using non-negative matrix factorization on the data set of interaction sequences collected using our data-mining system *AIRvatar*. N-gram models were used to interpret and evaluate each of these patterns. We validated these models with timing-related data measuring players' engagements with different parts of the system. We found that these patterns could effectively model the different levels of engagement players had and their preferences. It shows that data-driven AI approaches can be effective in quantifying the implicit behaviors of players.

Acknowledgment

We thank the anonymous reviewers for their valuable feedback. This material is based upon work supported by the National Science Foundation under Grant No. 1064495. We also acknowledge the support of a QCRI-CSAIL Collaboration Grant.

References

Butler, E.; Andersen, E.; Smith, A. M.; Gulwani, S.; and Popović, Z. 2015. Automatic game progression design through analysis of solution features. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, 2407–2416.

Cutler, A., and Breiman, L. 1994. Archetypal analysis. *Technometrics* 36(4):338–347.

Dahlskog, S.; Togelius, J.; and Nelson, M. J. 2014. Linear levels through n-grams. *Proceedings of the 18th International Academic MindTrek*.

Drachen, A.; Sifa, R.; Bauckhage, C.; and Thurau, C. 2012. Guns, swords and data: Clustering of player behavior in computer games in the wild. In *Computational Intelligence and Games, 2012 IEEE Conference on*, 163–170. IEEE.

Drachen, A.; Thurau, C.; Sifa, R.; and Bauckhage, C. 2014. A comparison of methods for player clustering via behavioral telemetry. *arXiv preprint arXiv:1407.3950*.

Ducheneaut, N.; Wen, M.-H.; Yee, N.; and Wadley, G. 2009. Body and mind: A study of avatar personalization in three virtual worlds. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1151–1160.

Fournier-Viger, P.; Faghihi, U.; Nkambou, R.; and Nguifo, E. M. 2012. Cmrules: Mining sequential rules common to several sequences. *Knowledge-Based Systems* 25(1):63–76.

Harrell, D. F., and Harrell, S. V. 2012. Imagination, computation, and self-expression: Situated character and avatar mediated identity. *Leonardo Electronic Almanac* 17(2).

Harrell, D. F. 2013. *Phantasmal Media: An Approach to Imagination, Computation, and Expression*. MIT Press.

Harrison, B. E., and Roberts, D. 2014. Analytics-driven dynamic game adaption for player retention in a 2-dimensional adventure game. In *Tenth Artificial Intelligence and Interactive Digital Entertainment Conference*.

Hoyer, P. O. 2004. Non-negative matrix factorization with sparseness constraints. *The Journal of Machine Learning Research* 5:1457–1469.

Jolliffe, I. 2005. *Principal component analysis*. Wiley Online Library.

Lee, D. D., and Seung, H. S. 1999. Learning the parts of objects by non-negative matrix factorization. *Nature* 401(6755):788–791.

Liberated Pixel Cup. 2015. http://lpc.opengameart.org.

Lim, C.-U., and Harrell, D. F. 2015a. Developing computational models of players' identities and values from videogame avatars. In *Proceedings of the 10th International Conference on the Foundations of Digital Games*.

Lim, C.-U., and Harrell, D. F. 2015b. Toward telemetry-driven analytics for understanding players and their avatars in videogames. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '15, 1175–1180.

Lim, C.-U., and Harrell, D. F. 2015c. Understanding players' identities and behavioral archetypes from avatar customization data. In *Proceedings of the IEEE Conference on Computational Intelligence and Games*.

Lin, H., and Wang, H. 2014. Avatar creation in virtual worlds: Behaviors and motivations. *Computers in Human Behavior* 34:213–218.

Mack Looseleaf Creator. 2015. http://www.geocities.jp/kurororo4/looseleaf/.

Shaker, N., and Abou-Zleikha, M. 2014. Alone we can do so little, together we can do so much: A combinatorial approach for generating game content. In *Tenth Artificial Intelligence and Interactive Digital Entertainment Conference*.

Shushin, A. 2005. The non-markovian stochastic liouville equation and anomalous quantum relaxation kinetics. *Journal of Experimental & Theoretical Physics* 101(4):577–583.

Smith, A. M.; Lewis, C.; Hullet, K.; and Sullivan, A. 2011. An inclusive view of player modeling. *Proceedings of the 6th International Conference on Foundations of Digital Games* 301–303.

Thurau, C.; Kersting, K.; and Bauckhage, C. 2009. Convex non-negative matrix factorization in the wild. In *Data Mining, 2009. ICDM'09. Ninth IEEE International Conference on*, 523–532. IEEE.

Yannakakis, G. N., and Hallam, J. 2009. Real-time game adaptation for optimizing player satisfaction. *Computational Intelligence and AI in Games, IEEE Transactions on* 1(2):121–133.

Yannakakis, G. N., and Togelius, J. 2011. Experience-driven procedural content generation. *Affective Computing, IEEE Transactions on* 2(3):147–161.

Yee, N., and Bailenson, J. 2007. The proteus effect: The effect of transformed self-representation on behavior. *Human communication research* 33(3):271–290.

Yee, N.; Ducheneaut, N.; Nelson, L.; and Likarish, P. 2011. Introverted Elves & Conscientious Gnomes : The Expression of Personality in World of Warcraft. *In Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems* 753–762.

Yee, N.; Ducheneaut, N.; Shiao, H.-T.; and Nelson, L. 2012. Through the azerothian looking glass: Mapping in-game preferences to real world demographics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, 2811–2814.