Visual Analytics of Dynamic Interplay Between Behaviors in MMORPGs

Junhua Lu* [1] Xiao Xie† [1] Ji Lan‡ [1] Tai-Quan Peng§ [2] Wei Chen¶ [1]
Yingcai Wu‖ [1]

[1] State Key Lab of CAD&CG, Zhejiang University
[2] Department of Communication, Michigan State University

Abstract

The rapid development of massively multiplayer online role-playing games (MMORPGs) has led operators to record huge amounts of fine-grained data from the in-game activities of players. These data provide considerable opportunities with which to study the dynamic interplay among player behaviors and investigate the roles of various social structures that underlie such interplay. However, modeling and visualizing these behavioral data remain a challenge. In this study, we propose a novel influence-susceptible model to measure the dynamic interplay among multiple behaviors. Based on this model, we introduce a new visual analytics system called BeXplorer. BeXplorer enables analysts to interactively explore the dynamic interplay between player purchase and communication behaviors and to examine the manner in which this interplay is bound by social structures where players are embedded.

Index Terms: Human-centered computing—Visualization—Visualization application domains—Visual analytics

1 Introduction

Massively multiplayer online role-playing games (MMORPGs) create a virtual society where players interact by assuming various characters. Millions of players invest monetary, cognitive, and social resources in MMORPGs as interaction and self-participation are embedded in these games. Consequently, MMORPGs have become an important component of the electronic business with an overall revenue of $19.8 billion USD worldwide in 2016 [34]. The profitability and sustainability of MMORPGs are contingent on the experiences and commitment of the players. Hence, understanding the patterns and dynamics of player behaviors is economically significant. Moreover, the self-initiated behaviors function as a "social telescope" that allows social scientists to examine how others influence player behaviors and how such influence is bound or facilitated by the virtual social fabric in MMORPGs [9].

Prior studies on behavioral patterns merely examined a snapshot of the evolving behavior of individuals. These studies decontextualized players from virtual social contexts or neglected the complex nature of player behaviors [5,21]. Therefore, they do not support the analysis of dynamic interplay among multiple types of behaviors.

Two challenges should be overcome to address these issues. First, the correlations among different behaviors are highly dynamic and often hidden in an enormous amount of complex data. Detecting correlations in a vast search space without computational support is difficult. Thus, a mining model is necessary to leverage the required computational power to reduce search space. However, deriving such a model that can quantitatively measure the dynamic interplay among multiple types of behaviors and model the roles of various social structures in the interplay is a challenge. Second, the detected patterns can be complicated and difficult to understand. For instance, the dynamic interplay between purchase and communication behaviors in MMORPGs occurs with multiple types of time-varying relations, such as the influence relations among different groups of players and the influence of different groups of players on various products. Creating a discernible visual summary of these relations is difficult. Moreover, the patterns may not be clearly defined or accurately captured. Human knowledge and expertise are necessary to assist in the analysis and interpretation of these patterns.

To address the first challenge, this study explores the social structures that can influence the behavior of MMORPG players. It specifically focuses on the dynamic interplay between communication and purchase behaviors, as discussed in Section 2. An influence-susceptible model is proposed to investigate the effect of social structures on the dynamic interplay among player behaviors. To address the second challenge, we introduce a visual analytics system called BeXplorer with two major views: (1) a flow view that provides an overview of the dynamic interplay between the communication behavior of influential players and their peers' purchase behavior on varying product categories and (2) a correlation view that supports inter-group and intra-group correlation analyses of player behaviors.

The main contributions of our study are as follows:

- a model that can measure and disentangle the effects of various social structures on behavior dynamics,
- a visual analytics system that can assist in uncovering the behavioral and social dynamics of players, and
- interesting insights and lessons on the visual analysis of dynamic user behaviors gained in this study.

2 Related Works

Visualization for Games. There have been several works utilize visualization to analyze online game data. Most of these visualizations and analysis are based on logs from the game. Pao et al. [26] visualized the trajectory of bots and human and discover different characteristics. They derived a bot detection scheme based on this insight. Thawonmas and Iizuka [36] used classical multidimensional scaling to visualize players behaviors and find similar players. They further investigated the detailed behaviors of those players in a cluster with KeyGraph. MMOSeer [15] is a visual analytics system for exploring the evolution of egocentric player intimacy network. It depicts the evolution of MMORPG players' behaviors and social interaction with different views. User interactions are supported to help users browse and analyze the data. Li et al. [16] proposed a system to show the evolution of players' status and positions in a MOBA (Multiplayer Online Battle Arena) game. Different events and game trend are also displayed. Their information are visualized in multiple coordinated views are used to reveal the players' strategies and performance, and enable users to find patterns interactively. Kuan et al. [14] also propose a similar system which shows summary

*e-mail: akiori@zju.edu.cn
†e-mail: xxie@zju.edu.cn
‡e-mail: lanjisteven@zju.edu.cn
§e-mail: pengtaiq@msu.edu
¶e-mail: chenwei@zju.edu.cn
‖e-mail: ycwu@zju.edu.cn.
Wei Chen and Yingcai Wu are the co-corresponding authors.

information of players' status in a RTS (Real-Time Strategy) game with different views. Previous studies display different behaviors or status with different views, which fail to explicitly show the dynamic interplay between them. Our system manages to quantitatively measure and visualize the dynamic interplay among multiple types of behaviors.

Behavior Analysis in Social Networks. Out of multiple possible behaviors in MMORPGs, two behavioral types are of particular importance: purchase and communication. The importance of players' purchase behavior is straightforward, as it is directly linked to the revenue of developers and their continuous investment in MMORPGs. However, players' purchase behavior is heavily susceptible to the influence from peer players, especially from influential ones [2, 22]. Influence from peer players can be realized through three types of social interaction mechanisms, namely, *direct communication* [2], *social influence* [25], and *triadic closure* [28], which are explained in Section 4.1. Players in MMORPGs can exert their impacts on other players via direct communication. Communication ties between players serve as the primary pathway along which ideas or products can become contagious in the virtual social system created in MMORPGs [2]. Moreover, the impact of direct communication can be confounded by two structural mechanisms of communication networks in which the players are situated [29], namely, social influence mechanism [3] and triadic closure mechanism [28]. A more detailed introduction and thorough discussions on these mechanisms can be found in Section 4.1. The impacts of direct communication, social influence, and triadic closure mechanisms are interdependent. They work jointly to influence players' purchase behavior. There have been many works analyze behavior in networks using visualization. For example, Prieto et al. [27] proposed a visual design to facilitate the analysis of location-based network data from spatial and temporal aspects. This data contains information about inhabitants' behavior, which helps improve urban planning. Our study aims to employ longitudinal data from a large-scale MMORPG to unravel the net and joint influences of these three mechanisms.

Flow-based Temporal Data Visualization. Flow is a frequently used visual metaphor for temporal data to help visualize temporal trends (e.g. [13, 23, 30, 37]). Recently, the use of flow-based designs to display evolving text/topics (e.g., CiteRivers [11] and TopicStream [17]) and events (e.g., EventRiver [18] and Outflow [39]) has become popular. CiteRivers [11] employs the flow metaphor to visually represent the contents and citations of scientific documents. Outflow [39] aggregates multiple event sequences into a graph, and the edges which represent the transitions are displayed simultaneously. Wu et al. [40] utilized the flow metaphor to encode the temporal patterns of detected risks from modules in factories. Compared with the data used in the prior studies, the game data used in the present study include the dynamics of avatar attributes, game activities, and numerous forms of behaviors. These characteristics prohibit this study from applying the existing methods directly because the data are not organized in a flat or hierarchical structure. Therefore, extending the existing methods is necessary to allow multiple types of time-varying relations.

3 Background and System Overview

This section presents the problem characterization, user requirements, and system overview.

3.1 Problem Characterization and User Requirements

The data were provided by the operator of a famous MMORPG called "A Chinese Ghost Story 2". The game runs on dozens of servers and has attracted millions of players. The operator selected one server for this study to enable a thorough analysis of players' diverse behaviors and interactions over a long period. After data cleaning, the remaining data consisted of 60,000 players' information over a time span of 49 weeks. Various avatar attributes and activities with timestamps were stored in isolated CSV files. Each server of the game is treated as a virtual world. A player's avatar is identified by its *AvatarID*. Avatar attributes can be categorized into **static attributes** (unchanged during their lifetime) and **dynamic attributes** (changing over time) as follows. **Static attributes**:

- *avatar_gender*: The gender of an avatar is determined by the player when the role is created.
- *avatar_class* and *avatar_clique*: Eight classes of avatars exist, and each class belongs to one *clique*. These cliques are ***B**arrack*, ***M**ountain*, ***T**emple*, and ***G**oong*, which are denoted by $\bar{B}$, $\bar{M}$, $\bar{T}$, and $\bar{G}$, respectively. Avatars in different classes have various task settings, fighting skills, means of upgrading their levels, and needs for different equipments. Avatars in a class would also prefer chatting with specific classes of players to complete specific tasks. As classes of the same clique share similarities in functions and evolution patterns, player behaviors are studied at the clique level.

Dynamic attributes:

- *online_time*: This attribute denotes the cumulative time a player spends on the game.
- *grade*: An avatar's *grade* starts at 1 and ends at 150.
- *chat_frequency*: This attribute represents the frequency of the players' chats with others.
- *count_pvp*: Players are allowed to fight with others in activities called "Player versus player" (*pvp*). *Count_pvp* is the frequency of a player's *pvp* activity.
- *purchase_records*: Players buy products in a virtual mall using their virtual wealth. The mall houses 23 types of products classified into five categories according to the operators of the MMORPG under study: C1 *Miscellaneous* (3 types), C2 *Facilities Enhancement* (6 types), C3 *Capacity Enhancement* (3 types), C4 *Advanced Capacity Enhancement* (2 types), and C5 *Context Enhancement* (5 types).
- *VIP_level*: The more the player pays cumulatively, the higher *VIP_level* that the player achieves.
- *Practice*: Learning and improving skills and finishing tasks increase the *Practice* level.
- *Mastery*: A player's *Mastery* is zero before joining a guild. *Mastery* increases as the avatar masters additional skills and contributes further to the guild.

In the past year, we closely worked with a scholar of communication and media studies and two senior analysts from the game operator. Typically, the analysts utilize basic statistical methods. For example, once a new strategy is launched by the game designers, they analyze the participation rate and completion rate. If a new fashion is introduced, then the analysts compute the percentage of players who buy the fashion and that of consumers who buy repeatedly. The data warehouse tool *Hadoop Hive* is used to accomplish these tasks, though it can only provide basic statistics.

A user-friendly interactive tool that can help analysts gain high-level insights remains unavailable. This study aims to build a visual analytics system to facilitate their analysis. Biweekly meetings were held with the analysts to discuss their analysis requirements and present several preliminary designs. Modifications were iteratively implemented on the system prototype on the basis of their feedback and comments. The collected requirements and feedback were condensed and transformed into the following design goals.

T **Time-oriented analysis of evolving behaviors.**

- T1 *How do the behaviors of players evolve?* Among different forms of in-game behaviors, analysts are particularly interested in purchase and social interaction behaviors. The system should help explore the temporal changes in purchases and social interaction behaviors.
- T2 *How do the communication and purchase behaviors of players co-evolve?* Social interactions between players exert an impact on their purchase dynamically through

multiple social interaction mechanisms. Analysts aim to examine the dynamic interplay.

T3 *What are the time-varying similarities and differences among the influences of communication exerted on the purchase of various categories of products?* The game offers various types of virtual products, which have a wide variety of prices, usages, and purposes. The system should allow analysts to visually compare social influences on the purchase of different products over time.

T4 *How can other groups play roles in a particular kind of communication influence on the purchase of a particular group of players? How do these roles change over time?* The communication influence of different groups of players can have a time-varying impact on the purchase behavior of a group of players. Analysts intend to track the changes of the influential power and identify which groups of players are more influential or susceptible.

G Inter-Group and intra-group analysis of behaviors.

G1 *What are the differences in communication and purchase behaviors between groups of players?* Different groups of players have distinct communication and purchase behaviors. Inter-group comparison and analysis are necessary to allow analysts to compare behaviors between groups.

G2 *What are the behavioral characteristics within a group?* Players may also exhibit various behavioral characteristics within their own group. Analysts aim to examine, analyze, and understand characteristics within a group of players.

P Multi-Perspective detailed exploration of behaviors.

P1 *What are the reasons for the discovered patterns? Can preliminary hypotheses be formed for possible causes?* Analysts aim to unfold the detected patterns and examine more details to form preliminary hypotheses or explanations. For example, can explanations be formed given the temporal trends in the rise and fall of the patterns?

3.2 System Overview

BeXplorer is a web application with three parts: data pre-processing, data analysis, and interactive visualization. AngularJS is employed as the front-end framework, while Node.js with MongoDB is used as the back-end and database. D3.js is used as the visualization library, while Python and R are used as data processing and modeling tools. The data pre-processing component is used to store data, extract basic information, and compute the derived attributes of players from massive CSV files. The data analysis component supports real-time retrieval of information and pre-computation of the model. The visualization component accepts results from the data pre-processing and analysis components as input to enable visual analysis.

4 Model

Fig. 1 presents the pipeline of our proposed model. The model first identifies the *sophisticated players* and computes three *social interaction variables* in the pre-processing stage (Fig. 1(B)). A group-specific and product-specific equation system is built by considering three *social interaction mechanisms*, namely, *direct communication*, *social influence*, and *triadic closure mechanisms* (Fig. 1(C)). These mechanisms are explained further in Section 4.2. The equation system is estimated using seemingly unrelated regression (SUR), which estimates the *influential power* of different groups of players (Fig. 1(D)). SUR [42] is a type of generalized least squares estimation which allows the residuals to be contemporaneously correlated and thus provides an efficient estimator for the coefficients in the equation system. To avoid confusion between the variables and influential power, their formal definitions are provided as below.

- ◇ **Social interaction variables** consist of direct communication, social influence, and triadic closure variables, which are denoted by $\boldsymbol{N}$s, $\boldsymbol{S}$s, and $\boldsymbol{SH}$s, respectively. They are independent variables in the system.
- ◇ **Influential power** of direct communication, social influence, and triadic closure represents the power of the three social interaction mechanisms. The effects represent the explanatory power of the social interaction variables on purchase behavior (Section 4.3).

4.1 Pre-processing

Sociological theories suggest that players' purchase behavior in an MMORPG is highly susceptible to the influence of their peers, especially from influential peers [2, 22, 35]. Therefore, classifying players into groups according to their degree of sophistication in the game is necessary as it allows researchers to capture the unique roles of influential and general players. Traditionally, player sophistication in MMORPG is simply operationalized as a behavioral attribute (e.g., experience, skills, etc.). However, this kind of attribute-based measure neglects the interactive nature of MMORPGs. In this study, a new measure called *Communication-Weighted Behavioral Sophistication* (*CWBS*) is developed. CWBS is a composite measure that weighs players' behavioral sophistication by their communicative importance in an MMORPG. Fig. 1(B)(upper left) shows the model elaborated as follows.

The behavioral sophistication of players is operationalized as a composite score of four behavioral attributes, namely, current grade, mastery degree, practice degree, and VIP level in an MMORPG. These four attributes of behavioral sophistication appear to have adequate internal consistency with Cronbach's alpha equal to 0.67 [7]. Cronbach's alpha [7] is a widely used measure of internal consistency in social science which conveys how closely related a set of measurement items are as a group. Principal component analysis (PCA) with varimax rotation is adopted to reduce the four attributes to a composite score of behavioral sophistication. This is a factor analytical technique which can extract a number of uncorrelated latent factors from observed variables based on the correlation among observed variables [1]. Players' communicative importance in the MMORPG is measured by their eigenvector centrality scores [4] in the communication network among players in a specific timeframe. The CWBS score of a player is then calculated by multiplying his/her behavioral sophistication by respective communicative importance score. Based on players' CWBS scores, the top 10% of players of each clique in a specific timeframe (i.e., two weeks in this study) are classified as *sophisticated players*, whereas the remaining 90% of the players are classified as *general players* (lower left of Fig. 1(B)). Hence, all players in the MMORPG are assigned to eight groups according to their CWBS scores and cliques. **B**, **M**, **T**, and **G** represent the sophisticated players from the four cliques, namely, $\bar{B}$, $\bar{M}$, $\bar{T}$, and $\bar{G}$ respectively, while **b**, **m**, **t**, and **g** represent general players of these cliques.

Before introducing the model, we first explain the three social interaction mechanisms which will be used, and how these social interaction variables are computed during pre-processing.

Direct communication (DC): Communication among players serves as the primary pathway along which ideas or products can become contagious [2]. Information regarding function and quality of products can be transmitted from one player to another via direct communication, thereby updating the expectations of players on those products and further impacting their purchase decision.

Social influence (SI): Individuals will change their behaviors to align themselves more closely with their partners [8, 19]. By observing the performance and skills of other players, a player is expected to purchase additional products to improve his/her performance and skills, thereby making him/her homophilous to his/her communication partners [3]. The greater the gap in the performance and skills between a player and others, the stronger the motivation this player will have to purchase added products to improve his/her performance and skills [25].

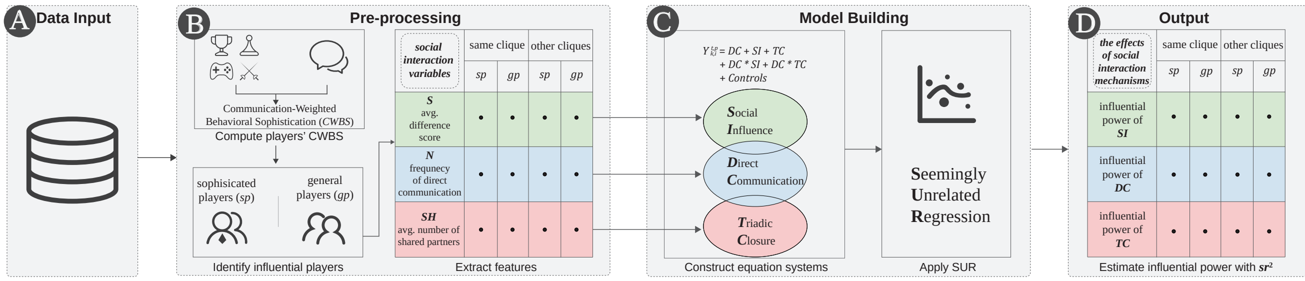

Figure 1: Pipeline of the model. (A) Data input; (B) data pre-processing that identifies sophisticated players, and then computes *social interaction variables* of each group on each category of products for the equation system; (C) model building that applies seemingly unrelated regression [42] to estimate the equation system; (D) output, i.e., the influential power of various groups of players are estimated via sr^2.

Triadic closure (TC): This is the property among three nodes, i.e., A, B, and C, such that if B and C are friends of A, then B and C tend to become friends as well. If one player has more shared partners with others in his/her network, this player is expected to have stronger motivation to purchase further to improve his/her skills and performances and maintain their ties with other players [28].

The frequency of communication between two players is calculated to capture direct communication mechanism. The difference in CWBS scores between two players who communicate with each another is calculated to capture social influence mechanism. The Jaccard coefficient, which measures the number of shared communication partners between two players, is calculated to capture triadic closure mechanism.

4.2 An Influence-Susceptible Model

We introduce a novel influence-susceptible model to examine the influence of players' communication behavior on their purchase behavior. This model considers communication both between players from different cliques and within the same clique. The model also empowers us to examine the unique roles of various social structures that may facilitate or hinder such influence. The reasons for modeling the influence between cliques are twofold. First, a *clique* creates a social division among players. Hence, the influence from a player's own clique should be different from those of other cliques. Second, a *clique* determines the cooperative or competitive relationship among players, which in turn would influence their purchase. The susceptible-influence relationship among players between cliques can be realized using direct communication, social influence, and triadic closure, as explained above.

The proposed model is defined as follows.

$$Y_{k^i}^{t,p} = DC + SI + TC + DC * SI + DC * TC + Ctrl \tag{1}$$

where $Y_{k^i}^{t,p}$ is product p purchased by player k^i in clique i at timeframe t which is explained by six blocks of factors. The first three blocks of factors (i.e., DC, SI, and TC) are included to examine the main effects of direct communication (DC), social influence (SI), and triadic closure (TC). The fourth and fifth blocks (i.e., $DC{*}SI$ and $DC{*}TC$) examine the joint effects between DC and SI and between DC and TC, respectively. Including these two blocks controls the confounding effects of social influence and triadic closure over the effect of direct communication, which has been well documented in empirical studies in various fields [12, 29]. The last block (i.e., $Ctrl$) is a list of variables whose effects on the purchase behavior of an individual player should be controlled. We abstract the factors included in *DC, SI,* and *TC* in the table in Fig. 1(B). Specifically, these factors are elaborated using the following equations. For simplicity, cliques $\bar{B}$, $\bar{M}$, $\bar{T}$, and $\bar{G}$ are denoted as clique 1, 2, 3, and 4 in this section, respectively.

$$DC = \overset{(\mathbf{i})}{\boxed{N^t_{\mathrm{sp}^i,k^i}\beta_1}} + \overset{(\mathbf{ii})}{\boxed{\sum_{j=1,j\neq i}^{4} N^t_{\mathrm{sp}^j,k^i}\beta_{2,j}}} + \overset{(\mathbf{iii})}{\boxed{N^t_{\mathrm{gp}^i,k^i}\beta_3}} + \overset{(\mathbf{iv})}{\boxed{\sum_{j=1,j\neq i}^{4} N^t_{\mathrm{gp}^j,k^i}\beta_{4,j}}} \tag{2}$$

$$SI = \boxed{S^t_{\mathrm{sp}^i,k^i}\beta_5} + \boxed{\sum_{j=1,j\neq i}^{4} S^t_{\mathrm{sp}^j,k^i}\beta_{6,j}} + \boxed{S^t_{\mathrm{gp}^i,k^i}\beta_7} + \boxed{\sum_{j=1,j\neq i}^{4} S^t_{\mathrm{gp}^j,k^i}\beta_{8,j}} \tag{3}$$

$$TC = \boxed{SH^t_{\mathrm{sp}^i,k^i}\beta_9} + \boxed{\sum_{j=1,j\neq i}^{4} SH^t_{\mathrm{sp}^j,k^i}\beta_{10,j}} + \boxed{SH^t_{\mathrm{gp}^i,k^i}\beta_{11}} + \boxed{\sum_{j=1,j\neq i}^{4} SH^t_{\mathrm{gp}^j,k^i}\beta_{12,j}} \tag{4}$$

where βs (as well as those in the following equations) represent the estimated coefficients in the equation system, and $N^t_{\mathrm{sp}^j,k^i}$ and $N^t_{\mathrm{gp}^j,k^i}$ refer to the frequency of direct communication between player k^i in clique i and sophisticated players (sp) and general players (gp) in clique j at timeframe t, respectively. Similarly, $S^t_{\mathrm{sp}^j,k^i}$ and $S^t_{\mathrm{gp}^j,k^i}$ refer to the average difference scores in CWBS between player k^i and his/her sp and gp contacts in clique j, respectively. $SH^t_{\mathrm{sp}^j,k^i}$ and $SH^t_{\mathrm{gp}^j,k^i}$ refer to the average number of shared partners between player k^i and his/her sp and gp contacts in clique j, respectively.

The (**i**) and (**iii**) terms in the three equations represent the main influential power of DC, SI, and TC mechanisms from sp and gp in the same clique i, while the (**ii**) and (**iv**) terms capture the main influential power of DC, SI, and TC mechanisms from sp and gp from a different clique j. This decomposition empowers us to examine and compare the unique roles of sophisticated players and peer players within a clique and across cliques.

The (**i**) and (**iii**) terms in the three equations represent the main influential power of DC, SI, and TC mechanisms from sp and gp in the same clique i, while the (**ii**) and (**iv**) terms capture the main influential power of DC, SI, and TC mechanisms from sp and gp from a different clique j. This decomposition empowers us to examine and compare the unique roles of sophisticated players and peer players within a clique and across cliques.

4.3 Estimation and Measurement of Influential Power

The proposed model is a group-specific and product-specific equation system, that allows the estimation of the role of different groups of players in the purchase of a specific category of products. With the game data provided by the MMORPG operator, eight equation systems are obtained for eight groups of players; each system includes five equations to account for the purchase of five categories of products by a group of players. As shown in Fig. 1(C), each equation system is estimated using SUR [42], which considers the covariance residual structure and produces efficient estimates.

To compare the influential power of different groups of players and reveal the influence that can be purely attributed to a specific mechanism, squared semi-partial correlation (sr^2) is adopted to measure the influential power of sophisticated players and general players. sr^2 is frequently used as a measure of net explanatory power of an independent variable on the dependent variable when the impacts of other independent variables are controlled [6]. This measure is additive, normalized, and comparable within an equation

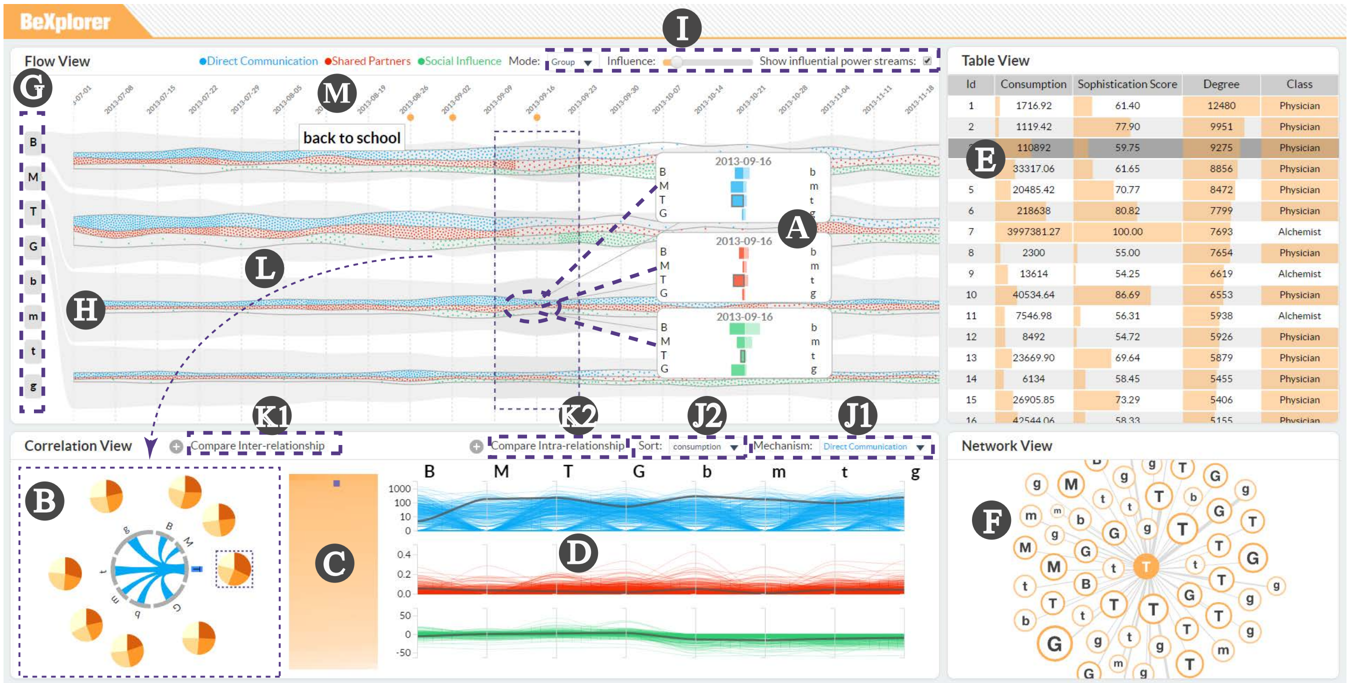

Figure 2: BeXplorer interface. The interface comprises a flow view (including purchase flows (H), purchase nodes (G), influential power streams in different colors, and influence bar charts (A)), a correlation view (including a radial visualization (B), a pixel bar chart (C), and parallel coordinate plots (D)), a table (E), and an ego-network (F). (I) and (J) show two control panels for the flow and correlation views, respectively.

or among the equations of a system, which has been employed to measure the coopetition power of social issues [33, 41].

Stepwise regression is adopted to identify the influential power of player groups. It is a method of fitting regression models by adding or subtracting a variable from the set of independent variables, which allows researchers to estimate the net explanatory power of each independent variable [6]. The influential power of sophisticated players in clique i (sp^i) is operationalized as the summation of sr^2 of relevant terms (i.e., $N^t_{\mathrm{sp}^i,k^i}$, $S^t_{\mathrm{sp}^i,k^i}$, and $SH^t_{\mathrm{sp}^i,k^i}$), while the influential power of general players in a clique i (sp^i) is operationalized as the summation of sr^2 of relevant terms (i.e., $N^t_{\mathrm{gp}^i,k^i}$, $S^t_{\mathrm{gp}^i,k^i}$, and $SH^t_{\mathrm{gp}^i,k^i}$). The influential power produced by the three social interaction mechanisms is captured by summing up the sr^2 of relevant terms in Equation (2), (3) and (4), respectively. These outputs are shown Fig. 1(D).

The resultant measures of influential power can help examine the evolution and compare the relative importance of the three social interaction mechanisms in affecting players' purchase of a specific category of products. Moreover, such measures of influential power warrant the pairwise comparisons of influential power from different groups of players over the purchase behavior of different groups of players and the purchase of various categories of products.

5 Visual Design

BeXplorer has four visualizations, namely, flow, correlation, network, and table views (Fig. 2) to support exploration of the resultant measures of influential power. The flow view uses a flow metaphor to enable the time-oriented analysis of evolving behaviors (T1–T4). When a timeframe is selected in the flow view, the correlation view will be updated to allow inter-group and intra-group correlation analyses of player behaviors, which are characterized by the social interaction variables within that period (G1 and G2). Visual cues based on purple rectangles and arrows are used in the layout of the tool to indicate the process of drilling down into the data (Fig. 2(L)). With these views, analysts can develop hypotheses and seek explanations for these patterns through in-depth analysis (P1).

5.1 Flow View

Results in Section 4.3 constitute complex time series data that characterize the dynamics of multiple player behaviors and the time-varying relation data that capture the interplay among multiple behaviors. However, creating a concise, informative, and discernible visual summary of the behavioral dynamics and multiple relation data to address the questions listed in T is difficult.

Therefore, we propose using a flow metaphor to visually analyze the data (T1–T3). The flow metaphor is commonly used for time series data in visualization [10, 18], and it can create an easy-to-understand and clutter-free layout for time series data (T1). In particular, we adopt a composite flow-based design that can display the co-evolutionary relationship regarding the purchase of products and the influences of different groups of players (T2 and T3).

We encode the purchase behavior as ***purchase flow***, and represent the influential power of social interaction mechanisms as ***influential power streams*** (see Fig. 2(H) and Fig. 3). The co-evolutionary relationship can be intuitively revealed using this composite design, which overlays different colored streams on flows. Streams in different colors that visually encode the types of social interaction mechanisms fluctuate with flows over time. Two analysis modes, *group* and *product* modes, are supported by the composite design to analyze the group-specific and product-specific measures estimated by our model. Moreover, ***influence bar charts*** (Fig. 2(A)) are used to explore the influential power of different social interaction mechanisms on accounting for player purchase of various categories of products and examine the roles of different groups of players (T4). The detailed encodings are described as follows.

Purchase Flow. Two analysis modes (group and product modes) are supported by purchase flow. The purchase behavior of each group is visualized as a gray flow (Fig. 3(A)), which is drawn in a vertically symmetrical manner. The width of a flow is proportional to the average purchase value. The flows can illustrate the temporal purchases of different groups of players (Fig. 2(H)) and that of

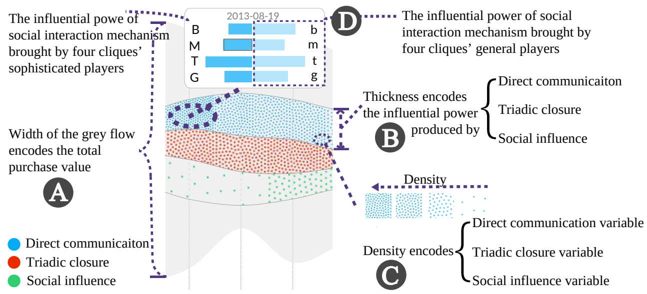

Figure 3: Visual encodings of a purchase flow (A), influential power streams (B and C), and an influence bar chart (D) in the flow view.

different categories of products (Fig. 7) in the group mode and the product mode, respectively. For space efficiency, users can enable or disable any flow for analysis by clicking the *purchase nodes* (Fig. 2(G)) located to the left of the flows. For example, only the flows of group **B**, **M**, **T**, and **G** are enabled and shown in Fig. 2(H).

Influential Power Streams. As shown in Fig. 3, three colored streams, which represent the three social interaction mechanisms, namely, direct communication, triadic closure, and social influence, are uniformly overlaid on each purchase flow in a consistent order. For presenting users with abstractions with meaning, we replaced triadic closure with *shared partners* used for the legend on the top of the flow view (to the left of Fig. 2(I)). The thickness of a stream at a certain timeframe is proportional to the degree of influential power of the associated mechanism at that time (Fig. 3(B)). The degree of influential power is estimated by the measure described in Section 4.3. The three streams are stacked to show the joint degree of influential power of the mechanisms.

In the group mode, the design allows analysts to compare the social influences of different groups of players on product purchase over time. We use stippling to encode the normalized values of the three social interaction variables (Fig. 3 (C)) which are computed by the measures described in the last paragraph of Section 4.1. Denser dots indicate higher values of the variables. Stippling is a specific type of texture and its density channel can be separated effectively from the size (thickness) channel [24, 38]. Thus, users can simultaneously perceive the degrees of influential power and values of social interaction variables.

In the product mode, the design can help compare social influence on the purchase of different categories of products over time (**T3**), as shown in Fig. 7. No social interaction variable is associated with virtual products. However, stippling with constant density on each stream is used to maintain the consistency of visual encoding.

Influence Bar Charts. The influential power of a social interaction mechanism produced by eight groups of players is visualized through a bar chart (**T4**), as shown in Fig. 3(D). The associated chart can be displayed by clicking a stream on the flow. Bar charts, instead of pie charts, are used because proportion and absolute value information are important for analysis. Analysts do not only wish to see the proportion at a particular timeframe but also intend to compare the absolute values at different timeframes. The bars placed at the left and right sides of a baseline represent the sophisticated players and general players of the four cliques (see Fig. 3(D)). The length of a bar indicates the degree of influential power of a specific social interaction mechanism contributed by a specific group of players.

5.2 Correlation View

We use a **radial visualization** for inter-group analysis (**G1**) and a **pixel bar chart** and three parallel coordinate plots (**PCPs**) for intra-group analysis (**G2**).

Inter-group Analysis. We adopt a **radial visualization** (Fig. 2(B)) that combines a chord diagram and eight pie charts for inter-group analysis (Fig. 4). A chord diagram and its inner

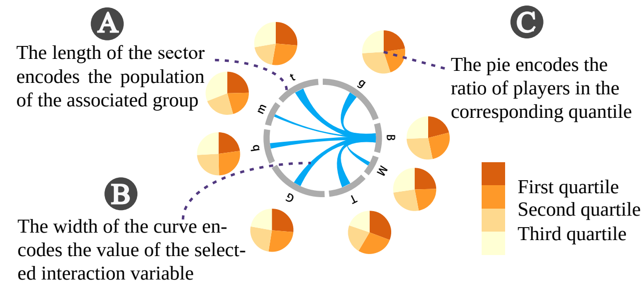

Figure 4: A chord diagram displays the population of eight groups of players. Each group is represented by a sector of the diagram.

curves display the population of eight groups of players and the social interaction between different groups of players. The pie charts surrounding the chord diagram represent the distribution of players on a selected behavioral attribute in each group of players.

Chord diagram. In the chord diagram, each group is represented by a sector, and the population size is encoded according to the sector length (Fig. 4(A)). Chord diagrams are adopted for two reasons. First, they have a reading accuracy that is no worse than that of pie charts [32], while pie charts are equal to bar charts for tasks such as estimating the proportion of a part to a whole [31]. Second, a chord diagram has a central area to display inter-group relationships (**G1**).

As shown in Fig. 4(B), the curves inside the chord diagram encode the social interaction among different groups, which helps analysts identify the inter-group communication, sophistication score, and shared partners, and thereby relating player social interactions to their attributes. The thickness of a curve visually represents the values of the selected social interaction variable between two groups.

Pie charts. The distribution of groups on selected behavioral attributes are analyzed and compared to understand the behaviors of the groups (**G1**). Two major tasks remain to be solved when analyzing the distribution. (1) *What is the distribution in a group? (2) What are the differences of distribution between groups?* To complete these tasks, we improve the chord diagram by adding a pie chart outside each sector to show the distribution of a selected attribute of the group (Fig. 4(C)). Pie charts are used for the tasks because of their high reading accuracy of proportion data. Given an attribute that analysts are interested in, we divide the attribute values of the players in the eight groups according to the quartiles. Hence, we divide players into four levels. The proportions of the player levels in a group are regarded as the distribution of the group.

Intra-group Analysis. We adopt a **pixel bar chart** (Fig. 2(C)) and three **PCPs** (Fig. 2(D)) for intra-group analysis (**G2**). The pixel bar chart, which is a scalable and compact visualization method, is used to visualize a large amount of attribute data of a player group selected by clicking the corresponding sector of the chord diagram. A **pixel bar chart** is placed to the right of the radial visualization to display the finest granularity of individual attributes. Each pixel represents a player of the group, with the level of brightness corresponding to the attribute. The pixels are sorted based on a selected player attribute. Each player has three social interactions with eight groups of players. Three **PCP**s, which are widely used to display multi-dimensional data, are adopted to visualize the three types of social interaction variables. Eight axes are placed in each PCP, where the lines match the color scheme used in the influential power streams for consistent encoding of the various social interaction mechanisms. Once an analyst selects a player group on the pixel bar chart, the lines of the players are highlighted in the PCPs. Brushing the lines highlights the associated pixels in the pixel bar chart.

5.3 Other Views and User Interactions.

A **table view** (Fig. 2(E)) and a **network view** (Fig. 2(F)) display detailed information regarding interesting players and outliers detected

Table 1: McElroy's R^2 and Adjusted R^2 of the Simultaneous Equation Systems. Notes: gp: general players; sp: sophisticated players.

Equation Systems	Mean of McElRoy's R^2	Mean of Adjusted R^2				
		C1	C2	C3	C4	C5
gp in $\bar{B}$	0.75	0.81	0.79	0.81	0.80	0.76
gp in $\bar{M}$	0.80	0.83	0.85	0.88	0.85	0.79
gp in $\bar{T}$	0.73	0.81	0.81	0.80	0.76	0.74
gp in $\bar{G}$	0.66	0.75	0.75	0.72	0.67	0.65
sp in $\bar{B}$	0.79	0.85	0.84	0.83	0.82	0.79
sp in $\bar{M}$	0.87	0.90	0.88	0.91	0.90	0.87
sp in $\bar{T}$	0.74	0.81	0.81	0.78	0.76	0.74
sp in $\bar{G}$	0.68	0.76	0.76	0.73	0.68	0.66

from other views. Each row of the table indicates a player, while each column represents a player attribute. A bar is placed in each cell to visually encode its value. A force-directed graph draws the ego-network of a selected player. The letter of a node represents its group identity, while its size represents its associated sophistication score. The width of a link encodes the communication frequency between players. A set of user interactions is supported as follows.

The **flow view** can be enlarged when a user double clicks on it. A user can switch between the group or product modes in the view and show or hide the influential power streams by clicking on a checkbox (Fig. 2(I)). Streams on the flows vary over time, and an influence filter (Fig. 2(I)) enables analysts to remove small streams. Important events (e.g., "back to school." in Fig. 2(M)) can be manually annotated as orange nodes in the timeline to visually remind analysts of the events.

The **correlation view** allows a user to specify which social interaction variable to analyze in the chord diagram through a drop-down menu (Fig. 2(J1)). If a sector of the diagram is clicked, then the corresponding pixel bar chart is displayed. The pixel chart is organized based on the selected attribute of a user (Fig. 2(J2)). Juxtaposed views are provided on demand (Fig. 2(K1) and (K2)) for side-by-side visual comparison of inter-group relationships (Fig. 8(B1) and (B2)) and intra-group relationships (Fig. 8(C1) and (C2)).

We render the visual encoding explanation part of the system as an on-demand help. When a user clicks on the name of a view, a visual explanation of its visual encoding will be displayed.

6 Evaluation and Discussion

The experiments were conducted in the Google Chrome browser on a desktop PC using the MMORPG data described in Section 3.1.

6.1 Model Evaluation

The simultaneous equation systems estimated in the study are evaluated by two measures: overall goodness of fit (McElroy R^2) of the system [20] and equation-specific goodness of fit (Adjusted R^2). Both measures are adequately good, which suggests that the influence-susceptible model performs well.

McElroy's R^2 provides an overall goodness of fit measure for SUR estimation models in which the residuals are auto-correlated or heteroskedastic. As shown in Table 1, the average McElroy's R^2 for the equation systems of the four groups of general players in the four cliques ranges from 0.66 to 0.80, whereas that of sophisticated players ranges from 0.68 to 0.87. The results indicate that the eight equation systems are effective and robust in accounting for the purchase behavior of eight groups of players.

The adjusted R^2 shows the explanatory power of a specific equation in an equation system as measured by the proportion of the variance in the dependent variable, which is explained by the vector of independent variables of the equation. The mean values of the adjusted R^2 of the five equations in the equation system of the general players in $\bar{G}$ and the sophisticated players in $\bar{M}$ range from 0.65 to 0.75 and from 0.87 to 0.91, respectively. In other words, at least 65% of the variance in the purchase of a specific type of products

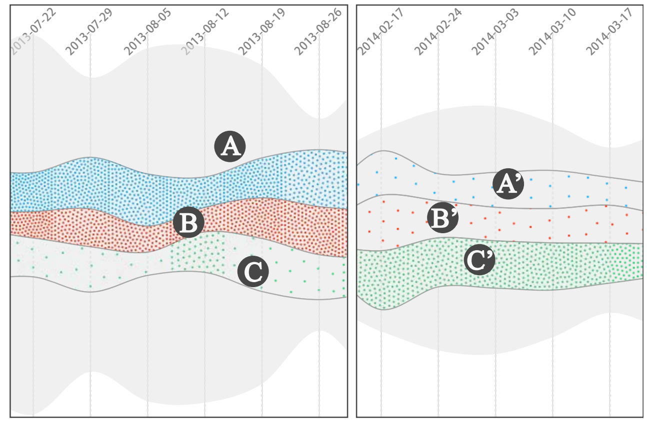

Figure 5: Purchase flow and influential power streams of clique $\bar{M}$ at early (left) and later (right) stages. The dot density from (A) to (A') and (B) to (B') decreases, whereas the density from (C) to (C') increases.

by the general players in $\bar{G}$ is explained by the model, while at least 87% of the variance in the purchase of a specific type of products by the sophisticated players in $\bar{M}$ is explained by the model.

6.2 Case Studies

We conducted two case studies with domain experts to demonstrate the efficiency and effectiveness of our system.

6.2.1 Visual tracking and comparison of dynamic patterns

We conducted the first case study to show the use of BeXplorer to visually track and compare dynamic communication and purchase behaviors between cliques and products. We started by looking at the evolution of behaviors of each player group. We found that the communication and purchase behaviors of both player groups show a downtrend over time. For example, Fig. 5 displays the communication and purchase behaviors of group **M**. Fig. 5(A) shows that the purchase flow of the sophisticated players in clique $\bar{M}$ demonstrates a clear downtrend (the width of flow decreased), suggesting that the players in the clique consume less and less over time. The dot density of the blue and red streams decreases over time (from Fig. 5(A) to Fig. 5(A') and from Fig. 5(B) to Fig. 5(B')), whereas that of green stream increases over time (from Fig. 5(C) to Fig. 5(C')). In other words, the direct communication becomes less frequent, and the average shared partners between players of the group **M** and others lessen (T1). On the contrary, the difference of the sophistication scores between the players of group **M** and other players becomes larger, indicating the decline of sophistication scores of the players of group **M**. These trends suggest that the overall activeness of sophisticated players in this clique is on a downward trend and those sophisticated players are likely to quit the game.

Game analysts commented that they noticed only the purchase downtrend, not the decline of communication and sophistication score of sophisticated players. The new observation is important for the MMORPG operator, because it suggests that they should not only take some timely measures to re-invigorate players' interests in investments, but also promote communication among players in the game. Influential power, which is indicated by the thickness of the three influential power streams, remains stable over time, as the thickness does not change significantly in Fig. 5. Despite the declining trends of the key variables, the explanatory power of our model is quite robust (T2).

We then analyzed behaviors across different player groups (G1). Fig. 6 shows four purchase flows of sophisticated players in different cliques between July 29 and October 7, 2013. Flows **B** and **M** demonstrate more fluctuations than flows **T** and **G**. Flow **M** fluctuates most dramatically whereas flow **G** is tranquil. Such fluctuations

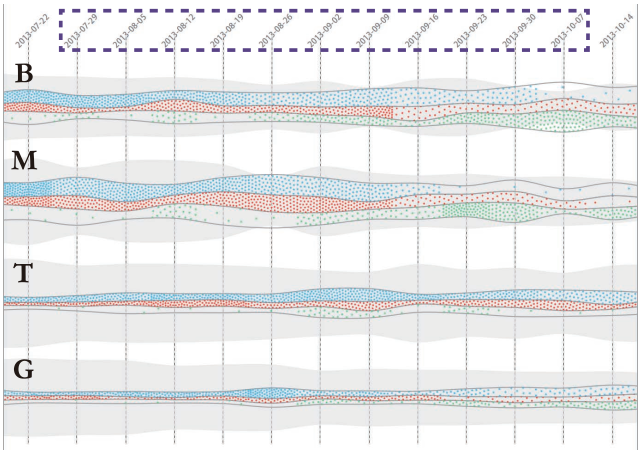

Figure 6: Purchase flows and influential power streams of player groups **B, M, T,** and **G**.

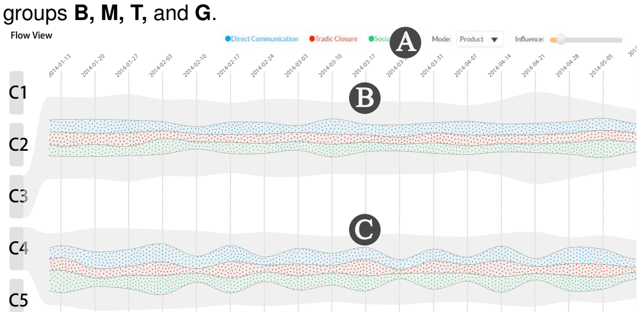

Figure 7: Players' susceptibility to the influence on different categories of products. (A) Group mode is selected; (B) and (C) show the purchase flows of product categories C3 and C4, respectively.

imply instability in the purchase of virtual products among players in cliques $\bar{B}$ and $\bar{M}$, especially $\bar{M}$. To understand which mechanisms could account for the instability of purchase flows (T2, P1), we analyzed the streams on the flows which encode the influential power associated with the three social interaction mechanisms included in our model (T4). The thickness of streams on the flows reveals that players in clique $\bar{B}$ and $\bar{M}$ are relatively much more susceptible to influence. The influential power stream of group **M** is thickest, indicating that the players of $\bar{M}$ are most susceptible. Hence, a positive correlation may exist between the changing rate of the purchases and the susceptibility to the influence of this group. The analysts were informed of these findings, which shed light on the further direction of the investigation. The analysts have planned to collect fine-grained information and explore the related behavioral data of the sophisticated players in clique $\bar{M}$ to study the correlation between the changing rate of purchase and the susceptibility to communication influence.

We also observed that the density of red and blue streams in clique $\bar{T}$ was larger than those in other cliques. Thus, clique $\bar{T}$ had relatively higher communication frequency with more shared partners compared to other cliques. Moreover, the large purchase flow of clique $\bar{T}$ indicates large purchases. Both large purchases and high-frequency communication behaviors of sophisticated players in the clique remain stable over time. We suggest that the MMORPG operator should dig deep into the game setting for this clique to better understand the reasons behind the phenomena and be able to properly adjust the game settings of other cliques to promote high purchase and high-frequency communication behaviors.

After switching the default group mode to the product mode (T3), player susceptibility to influence on different categories of products are displayed (Fig. 7). Upon selecting purchase nodes C3 (Capacity Enhancement) and C4 (Advanced Capacity Enhancement), we found that the susceptibility of influence of C4 fluctuates frequently (the width of the C4 influential power stream keeps changing). The finding suggests that influencing the purchase of C3 through promoting communication among players would be more consistently cost-effective for the game operator. This find is attributed to the fact that the susceptibility of this category to influence is more stable and thus is more reliable.

6.2.2 Multi-level visual exploration of multiple behaviors

We conducted the second case study to demonstrate the effectiveness of our system in supporting multi-level visual exploration of influence-susceptible relations between players. Fig. 8(A) shows four purchase flows and influential power streams of sophisticated players in different cliques. The direct communication frequency and shared partners of players in **B**, **M**, and **G** dropped sharply between September 9 and September 23, 2013. On the contrary, **T** was quite stable. The period was the end of the summer break, and youngsters who constituted the majority of players needed to return to their study. Players who did not go to school also had less communication and were less active as many of their partners or team members became less active. To understand why players in **T** withstood the shock during the period, we examined the influence bar charts by clicking the streams. We did not find meaningful explanations with the charts. Hence, we clicked on the timeframe to further explore the behaviors of sophisticated players of clique $\bar{T}$.

We used the juxtaposed view to visually compare inter-group communication relationships between different player groups. The blue curves in Fig. 8(B1) indicate that players in **T** had considerable communication with all the other groups and the frequency was similar among them (all the blue curves are thick). After comparing the pie charts (Fig. 8(B2)) that indicate the contribution each group of players to the purchase, we found that the darker slices of the charts of clique $\bar{T}$ are larger than those of other cliques. Thus, more players in clique $\bar{T}$ purchased more.

To learn more about clique $\bar{T}$ (G2, P1), we utilized the juxtaposed view of PCPs to visually compare the individual social interaction variables of high-purchase players (the blue rectangle in Fig. 8(C1)) with those of low-purchase players (the blue rectangle in Fig. 8(C2)). The social interaction variables of high-purchase players are located at the upper part of the axes (the lines are denser at the upper part of the axes in Fig. 8(C1)), whereas those of the low-purchase players are distributed near the bottom (the lines are denser around zero in Fig. 8(C2)). High-purchase players in clique $\bar{T}$ tended to communicate more frequently, whereas low-purchase players seldom communicated with others. This finding also suggests that the operator should promote communication among players in clique $\bar{T}$, which is likely to boost purchases.

By selecting these high-purchase and high-communication players, we noticed that most of them were physicians. This is the only class of players that could heal other players. Players of this class have good sense of teamwork and tend to work closely with others in the game. Thus, they have many friends and always join group activities as a team member. Fig. 8(E) shows an example in which the physician ego communicated frequently (thick links) with many partners from different cliques. Many of these partners are sophisticated players. We speculated that the physicians in clique $\bar{T}$ withstood the shock because they had a large number of friends and joined teams. When the students went to school, the physician players continued building connections with other partners and were much less susceptible to the influence. Based on this discovery, we suggest that the game operator should promote the class of physicians at the end of the summer break to have more physicians in the game. The campaign could build up more social ties, which would help avoid the loss of players.

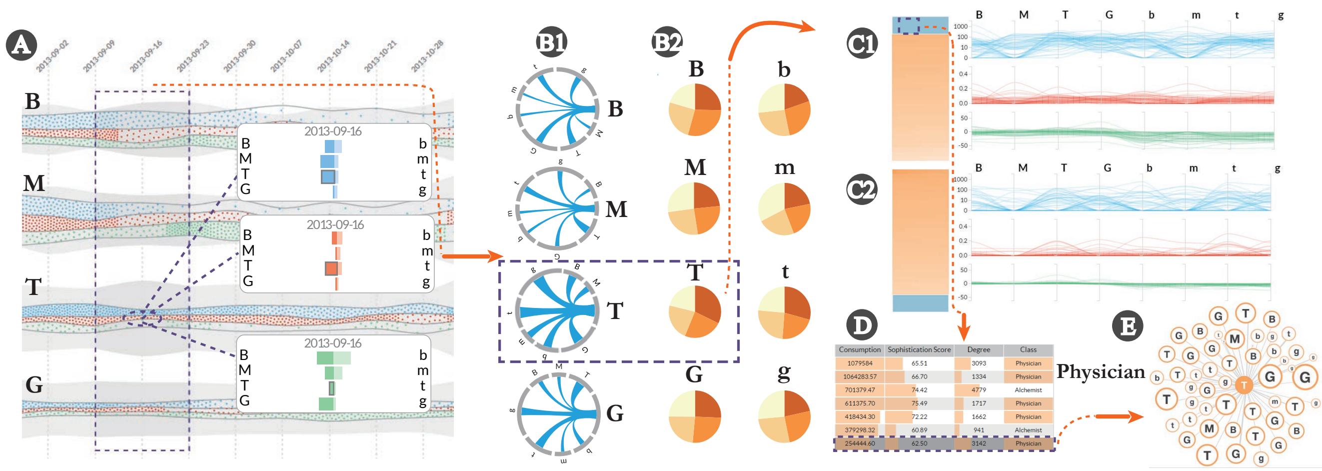

Figure 8: Analysis pipeline of the second case study. (A) shows the purchase flows of **B**, **M**, **T**, and **G** with the influence bar charts of **T** displayed. (B1) and (B2) are decomposed from the radial visualization. Individual attributes and social interaction variables of high-purchase and low-purchase players of **T** are depicted in (C1) and (C2), respectively. More detailed information of players of group **T** is provided in (D) and (E).

6.3 User Feedback

We invited three analysts denoted by E1, E2, and E3 from our collaborating MMORPG operator to evaluate our system. Their routine analysis methods and tools are described in Section 3.1. We conducted a semi-structured interview with each analyst. Each interview lasted one and a half hours. An interview started by introducing the visual encodings used as well as the interface and interactions of the system. We then walked through the case studies described above. The analysts also explored data with BeXplorer by themselves. Their feedback is summarized as follows.

Visual design. The analysts were impressed by the visual design of the system. They reacted positively to the flow-based visualization because it is aesthetically attractive and it provides an intuitive means to understand the dynamic changes of behaviors. E1 also appreciated the radial visualization for its explicit depiction of player communication and attribute distribution, which can help provide plausible explanations for interesting patterns. E3, who also has visualization background, found the PCPs and the network view became extremely dense under certain circumstances. “An appropriate filter may be useful to reduce the visual clutter in the network view,” he suggested.

System usability. The system was received well by the users. The analysts were excited by the capability of BeXplorer to visually explore different types of player behaviors. E1 commented, “The system looks extremely valuable to our work. It opens a new door for us to see through hidden behaviors that have never been noted before.” E2 pointed out that BeXplorer has considerable potential for evaluating the influences of different social mechanisms on purchase. She added, “Game designers in our companies will love to try BeXplorer for interactive analysis.” E3 suggested, “The user interface can be simplified by, for example, hiding unnecessary views.”

7 Discussion

To evaluate our work, we have quantitatively evaluated the model (Section 6.1) and collected feedback from analysts from the game operator (Section 6.3). Case studies are also done with the help of game analysts (Section 6.2). The number of volunteers is enough to do a statistical analysis for demonstrating the usability and intuitiveness.

Novelty. Our work focuses on analysis and visualization of dynamic interplay between behaviors in MMORPGs. Although the novelty mainly comes from its analysis model and the visualization system rather than the visual design, we also present a tailored flow metaphor which overlays stippling streams on a streamgraph to visually represent and explore the co-evolutionary relationship between different behaviors.

Generalizability. BeXplorer can be extended to address similar problems in other domains concerned with relationships among dynamic human behaviors. The generalizability of the work is deeply rooted in the universal presence of the social phenomena under study, the widespread applicability of theoretical models, and the methodological feasibility of our measurements. First, the influence-susceptible relationship examined in the study is common in human society [2, 22, 35]. Second, the three social interaction mechanisms, through which the influence-susceptible relationship can be realized, are fundamental mechanisms guiding human behavior, which have been widely examined in empirical studies on different human behaviors [2, 25, 28]. Third, the developed measurement devices can also be generalized to many other scenarios. The measurement of the three mechanisms is not limited to MMORPGs. With the enriching digital traces on social media, it is technically feasible to observe direct communication between individuals and observe structural variables (i.e., social influence and triadic closure) in a direct and valid manner. Therefore, our work has great potential to be applied to other settings beyond MMORPGs.

8 Conclusion and Future Work

This study introduces an influence-susceptible model that examines the roles of multiple social interaction mechanisms in the dynamic interplay of communication and purchase behaviors of players. BeXplorer was designed based on the model to facilitate the analysis and visualization of the behaviors in a large MMORPG. This study has two major theoretical implications. First, the model simultaneously examines the roles of three intertwined social interaction mechanisms in the purchase behavior of players in MMORPGs, which capture the unique individual and joint contribution of the mechanisms. Second, the study represents the first attempt, to our knowledge, to examine the mutual influence between sophisticated and general players through three social interaction mechanisms.

We plan to record the user logs of the system and analyze the logs for comprehensive evaluation of BeXplorer. We will also continue to work with the game operator to deploy the system such that domain experts can use our tool for their daily analysis. With the closer collaboration, we hope to gain more insight and report our new findings in the future work.

Acknowledgments

The work was supported by NSFC (61761136020, 61502416, 61772456), NSFC-Zhejiang Joint Fund for the Integration of Industrialization and Informatization (U1609217), Zhejiang Provincial Natural Science Foundation (LR18F020001) and the 100 Talents Program of Zhejiang University. This project was also partially funded by Microsoft Research Asia. The authors would also like to thank Zhipeng Hu, Bai Liu, and Luyi Xie from NetEase Inc. for their generous support and help with the user requirement collection and user evaluation.

References

[1] H. Abdi and L. J. Williams. Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, 2(4):433–459, 2010.

[2] S. Aral and D. Walker. Identifying influential and susceptible members of social networks. *Science*, 337(6092):337–341, 2012.

[3] A. V. Banerjee. A simple model of herd behavior. *The Quarterly Journal of Economics*, 107(3):797–817, 1992.

[4] R. S. Burt and J. Merluzzi. Embedded brokerage: Hubs versus locals. In *Research in the Sociology of Organizations: Contemporary Perspectives on Organizational Social Networks*, vol. 40, pp. 161–177. Emerald Group Publishing Limited, 2014.

[5] E. Castronova, D. Williams, C. Shen, R. Ratan, L. Xiong, Y. Huang, and B. Keegan. As real as real? macroeconomic behavior in a large-scale virtual world. *New Media & Society*, 11(5):685–707, 2009.

[6] J. Cohen, P. Cohen, S. West, and L. Aiken. *Applied multiple regression/correlation analysis for the behavioral sciences*. Psychology Press, 2014.

[7] L. J. Cronbach. Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3):297–334, 1951.

[8] D. Easley and J. Kleinberg. *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge University Press, 1 ed., 2010.

[9] S. A. Golder and M. W. Macy. Digital footprints: Opportunities and challenges for online social research. *Sociology*, 40(1):129–152, 2014.

[10] S. Havre, B. Hetzler, and L. Nowell. ThemeRiver: Visualizing theme changes over time. In *Proceedings of IEEE Symposium on Information Visualization*, pp. 115–123, 2000.

[11] F. Heimerl, Q. Han, S. Koch, and T. Ertl. CiteRivers: Visual analytics of citation patterns. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):190–199, 2016.

[12] G. Kossinets and D. J. Watts. Origins of homophily in an evolving social network. *American Journal of Sociology*, 115(2):405–450, 2009.

[13] R. Krueger, G. Sun, F. Beck, R. Liang, and T. Ertl. Traveldiff: Visual comparison analytics for massive movement patterns derived from twitter. In *Proceedings of the IEEE Pacific Visualization Symposium (PacificVis)*, pp. 176–183, 2016.

[14] Y.-T. Kuan, Y.-S. Wang, and J.-H. Chuang. Visualizing real-time strategy games: The example of starcraft ii. In *Proceedings of IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 1–10, 2017.

[15] Q. Li, Q. Shen, Y. Ming, P. Xu, Y. Wang, X. Ma, and H. Qu. A visual analytics approach for understanding egocentric intimacy network evolution and impact propagation in mmorpgs. In *Proceedings of the IEEE Pacific Visualization Symposium (PacificVis)*, pp. 31–40, 2017.

[16] Q. Li, P. Xu, Y. Y. Chan, Y. Wang, Z. Wang, H. Qu, and X. Ma. A visual analytics approach for understanding reasons behind snowballing and comeback in moba games. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):211–220, 2017.

[17] S. Liu, J. Yin, X. Wang, W. Cui, K. Cao, and J. Pei. Online visual analytics of text streams. *IEEE Transactions on Visualization and Computer Graphics*, 22(11):2451–2466, 2016.

[18] D. Luo, J. Yang, M. Krstajic, W. Ribarsky, and D. Keim. EventRiver: Visually exploring text collections with temporal references. *IEEE Transactions on Visualization and Computer Graphics*, 18(1):93–105, 2012.

[19] P. V. Marsden and N. E. Friedkin. Network studies of social influence. *Sociological Methods & Research*, 22(1):127–151, 1993.

[20] M. B. McElroy. Goodness of fit for seemingly unrelated regressions: Glahn's $r^2_{y.x}$ and Hooper's r^2. *Journal of Econometrics*, 6(3):381–387, 1977.

[21] C. M.Hughes, B. J.Griffin, and E. L.Worthington. A measure of social behavior in team-based, multiplayer online games: The sociality in multiplayer online games (SMOG) scale. *Computers in Human Behavior*, 69(C):386–395, 2017.

[22] E. Moretti. Social learning and peer effects in consumption: Evidence from movie sales. *The Review of Economic Studies*, 78(1):356–393, 2011.

[23] C. Muelder, B. Zhu, W. Chen, H. Zhang, and K.-L. Ma. Visual analysis of cloud computing performance using behavioral lines. *IEEE Transactions on Visualization and Computer Graphics*, 22(6):1694–1704, 2016.

[24] T. Munzner. Texture and stippling. In *Visualization Analysis & Design*, chap. 10.4.6, pp. 239–240. CRC Press, 2014.

[25] M. E. Newman. Assortative mixing in networks. *Physical review letters*, 89(20):208701, 2002.

[26] H.-K. Pao, K.-T. Chen, and H.-C. Chang. Game bot detection via avatar trajectory analysis. *IEEE Transactions on Computational Intelligence and AI in Games*, 2(3):162–175, 2010.

[27] D. F. Prieto, E. Hagen, D. Engel, D. Bayer, J. T. Hernández, C. Garth, and I. Scheler. Visual exploration of location-based social networks data in urban planning. In *Proceedings of the IEEE Pacific Visualization Symposium (PacificVis)*, pp. 123–127, 2015.

[28] A. Rapoport. Spread of information through a population with socio-structural bias: I. assumption of transitivity. *The bulletin of mathematical biophysics*, 15(4):523–533, 1953.

[29] C. R. Shalizi and A. C. Thomas. Homophily and contagion are generically confounded in observational social network studies. *Sociological methods & research*, 40(2):211–239, 2011.

[30] Y. Shi, C. Bryan, S. Bhamidipati, Y. Zhao, Y. Zhang, and K.-L. Ma. MeetingVis: Visual narratives to assist in recalling meeting context and content. *IEEE Transactions on Visualization and Computer Graphics*, 24(6):1918–1929, 2018.

[31] D. Simkin and R. Hastie. An information-processing analysis of graph perception. *Journal of the American Statistical Association*, 82(398):454–465, 1987.

[32] D. Skau and R. Kosara. Arcs, angles, or areas: Individual data encodings in pie and donut charts. *Computer Graphics Forum*, 35(3):121–130, 2016.

[33] G. Sun, Y. Wu, S. Liu, T.-Q. Peng, J. J. Zhu, and R. Liang. EvoRiver: Visual analysis of topic coopetition on social media. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1753–1762, 2014.

[34] SUPERDATA. MMO market report 2015. `https://www.superdataresearch.com/market-data/mmo-market`, 2015.

[35] M. Szell, R. Lambiotte, and S. Thurner. Multirelational organization of large-scale social networks in an online world. *Proceedings of the National Academy of Sciences*, 107(31):13636–13641, 2010.

[36] R. Thawonmas and K. Iizuka. Visualization of online-game players based on their action behaviors. *International Journal of Computer Games Technology*, 2008:5, 2008.

[37] D. Thom, R. Krüger, T. Ertl, U. Bechstedt, A. Platz, J. Zisgen, and B. Volland. Can twitter really save your life? A case study of visual social media analytics for situation awareness. In *Proceedings of the IEEE Pacific Visualization Symposium (PacificVis)*, pp. 183–190, 2015.

[38] C. Ware. Texture: Theory and data mapping. In *Information Visualization: Perception for Design*, chap. 6, pp. 199–210. Morgan Kaufmann, 2012.

[39] K. Wongsuphasawat and D. Gotz. Exploring flow, factors, and outcomes of temporal event sequences with the Outflow visualization. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2659–2668, 2012.

[40] W. Wu, Y. Zheng, K. Chen, X. Wang, and N. Cao. A visual analytics approach for equipment condition monitoring in smart factories of process industry. In *Proceedings of the IEEE Pacific Visualization Symposium (PacificVis)*, pp. 140–149, 2018.

[41] P. Xu, Y. Wu, E. Wei, T.-Q. Peng, S. Liu, J. J. Zhu, and H. Qu. Visual analysis of topic competition on social media. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2012–2021, 2013.

[42] A. Zellner. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association*, 57(298):348–368, 1962.