Delft University of Technology

Understanding and recommending play relationships in online social gaming

van de Bovenkamp, R; Shen Jie, S; Iosup, A; Kuipers, FA

DOI
https://doi.org/10.1109/COMSNETS.2013.6465556

Publication date
2013

Document Version
Accepted author manuscript

Published in
Proceedings 5th International Conference on Communication Systems and Networks (COMSNETS 2013)

Citation (APA)
van de Bovenkamp, R., Shen Jie, S., Iosup, A., & Kuipers, FA. (2013). Understanding and recommending play relationships in online social gaming. In D. Kotz, K. Shivkumar, & B. Raman (Eds.), *Proceedings 5th International Conference on Communication Systems and Networks (COMSNETS 2013)* (pp. 1-10). IEEE Society. https://doi.org/10.1109/COMSNETS.2013.6465556

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Understanding and Recommending Play Relationships in Online Social Gaming

Ruud van de Bovenkamp, Siqi Shen, Alexandru Iosup, and Fernando Kuipers
Delft University of Technology
Email: {R.vandeBovenkamp, S.Shen, A.Iosup, F.A.Kuipers}@tudelft.nl

***Abstract*—Online Social Networking (OSN) applications such as Facebook's communication and Zynga's gaming platforms service hundreds of millions of users. To understand and model such relationships, social network graphs are extracted from running OSN applications and subsequently processed using social and complex network analysis tools. In this paper, we focus on the application domain of Online Social Games (OSGs) and deploy a formalism for extracting graphs from large datasets. Our formalism covers notions such as game participation, adversarial relationships, match outcomes, and allows to filter out "weak" links based on one or more threshold values. Using two novel large-scale OSG datasets, we investigate a range of threshold values and their influence on the resulting OSG graph properties. We discuss how an analysis of multiple graphs—obtained through different extraction rules—could be used in an algorithm to improve matchmaking for players.**

I. Introduction

An increasing number of complex network studies and social network analyses use graphs to represent (the structure in) data. Often, two extreme approaches for building a graph from raw (observed) data are deployed. At one extreme, graphs are straightforwardly extracted when the raw data specifies links, such as in the case of friendship relationships for Facebook data. At the other extreme, graphs are extracted by applying a single, domain-specific and usually threshold-based, rule for mapping raw data to links. The latter approach has been used, for example, for building the graphs of mail and email exchanges of various communities and organizations, of messaging in Twitter and the Microsoft Messenger Network, of hyperlinking in the Web and in the blogosphere, etc. The impact of the choice of graph extraction rules and thresholds has received much less attention and constitutes the focus of this study.

We focus on the social networks formed by the actions of players involved in online social gaming (OSG), that is, in online gaming where the social element affects positively the gameplay experience. Online social games such as Defense of the Ancients (DotA) and League of Legends are each played by tens of millions of gamers. The game is fragmented into hundreds of thousands of non-communicating instances (matches [1]), and groups of only about ten players are involved in any instance of the game at any one time. Players can find partners for a game instance through the use of community web sites and online tools, which may include services that matchmake players to a game instance. An interesting feature of OSGs, as opposed to many online social networks, is that not only friendship relations are formed, but also adversarial relationships may manifest as useful social relationships.

In practice, it may be difficult to obtain a clear social structure from an online social game. Data that has proven economic value or private status, such as expressed friendships, are rarely shared by game operators; instead, third-parties may have access only to proxies, such as a list of players for each game instance. Furthermore, for OSG networks, traditional relationships such as expressed friendship relationships may not be a good indicator of joint activity. This may be a consequence of the classes of prosocial emotions involved in OSGs, for example vicarious pride and happy social embarrassment [2, Ch. 5], which are both derived from competition. These emotions complement other prosocial emotions that precede traditional friendship, such as admiration and devotion, and may require other tools for expression.

Without a clearly specified social structure, the analyst of OSG networks has to select and tune the *extraction rule*, that is, the rule for extracting graph links from *play relationships* recorded in the logs of completed and ongoing game instances. For example, it is common to extract OSG graphs through the simple rule of forming a link between gamers that have played at least once together. However, this simple rule overemphasises the importance of a single in-game encounter, which may have been casual or the result of an automated matchmaking mechanism. OSG networks derived using this rule have been recently investigated, for example for EverQuest [3], World of Warcraft [4], Fighters Club (a Facebook application) [5], etc. In contrast, in Section III, we generalise this formalism for graph extraction.

We further analyse, in Section IV, the interplay between various concepts and thresholds that are used when defining extraction rules in OSGs, and the characteristics of the resulting OSG graphs. Based on two large OSG datasets, we study various rules and thresholds for extracting graphs. Using this approach, we are able to show that even small changes in the thresholds used for various rules, especially for the small values that have been used by previous OSG graph studies, can lead to significant changes in the structure and use of the resulting graph. Thus, our work provides the tools to obtain a broader picture of OSG graph analysis than previous work.

The results of OSG studies are useful for tuning the distributed systems on which the games operate, for improving the game experience, and for understanding the individual and group psychology of players. As an application of the proposed analysis technique, we study the quality of matchmaking players in game instances.

Our main contributions are:

1) We introduce a broad formalism for graph extraction (Section III).
2) We apply our formalism to extract and analyse OSG graphs from two large datasets, using six different extraction strategies (Section IV).
3) We utilise our graph extraction analyses for improving player matchmaking (Section V).

II. Data Sets

We select for our study Defense of the Ancients (DotA), a free and popular game in which social relationships, such as same-guild membership and even friendship, can improve the gameplay experience. DotA is a 5-against-5-player game, that is, it consists of independent matches played by two contesting teams of five players each. DotA is a multiplayer online battle arena (MOBA) game, in which each player controls an in-game representation (here, the hero), and teams have as objective the conquest of the opposite side's main building. The game includes many strategic elements, from the team operation to the management of resources and the creation of helper troops.

DotA, which is the representative of the MOBA game subgenre, is played today by an estimated[1] number of players above 20 million, world-wide. DotA has featured in several tournaments with wide appeal to gamers and game-watchers, such as the World Cyber Games (WCG) and the Electronic Sports World Cup (ESWC). (To understand the phenomenon of watching DotA games, we recommend the seminal book on gaming culture of Rossignol [6].) Other successful MOBA games are Blizzard DotA, Valve's DotA2, League of Legends, and Heroes of Newerth. Each of these games is played by millions of players, which are loosely grouped into large communities. In turn, most communities operate their own game servers, maintain lists of tournaments and results, and publish information such as resulting player rankings via common websites.

We have collected data over multiple years for two DotA communities, Dota-League and DotAlicious. Both communities identify each of their matches with a unique number in increasing order, and for each match dedicated information (such as the names of the players participating in the match and the duration of the match) is available on a corresponding webpage. We have crawled all the unique matches played within these communities via their openly accessible websites, by gradually increasing the identifier number from 1 to the total number of matches played; we obtained the latter from the main page of each website. Some of the webpages with a match identifier in the crawling range appeared to be broken. We have crawled each web page at least twice, at different times, to reduce the effect of possible temporary unavailability and traffic shaping of the website.

[1] http://www.playdota.com/forums/blog.php?b=892

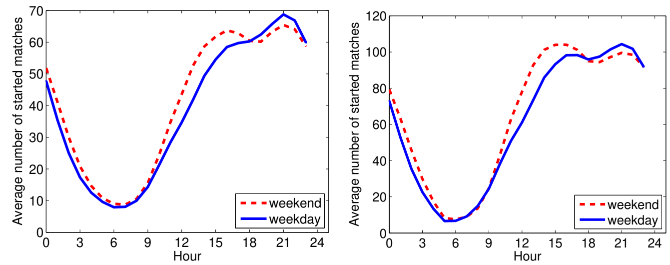

Fig. 1: Average number of matches per hour for DotAlicious (left) and Dota-League (right).

The collected Dota-League dataset consists of 3,744,753 matches that were played between July 2006 and July 2011. The dataset reflects for each match the names of the players (61,198 in total) on each team, the active time, the start and end times, various gameplay statistics per team, and the team that has won. The *active time* refers to the time at which the match opens and players can enter, while the *start time* refers to the time when sufficient players are present and the match commences. The *end time* refers to the time at which the match terminates. To sanitise the Dota-League dataset, we introduce the concept of *played* matches. Because we can be sure that matches have actually been played only for matches with correct start and end timestamps, we only consider these matches for our study. Although the match active time stamp is available for all matches, the match start and end time stamps were only available from November 2008 onwards, which corresponds to 1,470,786 played matches.

For DotAlicious, our dataset consists of 625,692 played matches, which represents the complete set of matches played from April 2010 to February 2012. Each match entry in the dataset records the names of the players (62,495 in total) of that match, the countries from which they are playing, the result of the match (winning team, draw, or abort), the start and end times, various gameplay statistics per player, and the number of points obtained for each player. To sanitise this dataset, we filtered out matches whose duration was zero, obtaining 617,069 played matches.

To understand the general characteristics of our datasets, we conducted an analysis of the game activity they represent. We find that both datasets exhibit similar time patterns, along the lines of other games and of other typical Internet applications. The number of matches started per hour of the day shows a clear diurnal pattern in weekdays and weekends, as can be seen from Figure 1. In weekdays, the number of matches played per hour rises from 6AM onwards and reaches a peak around 9PM. In weekends, the number of matches played rises from 6AM and reaches a peak around 4PM, where it stabilises until 9PM, with only a small drop around dinner time. The average number of matches played per day of the week is fairly stable (approaching 1000 matches/day), with slightly more matches played in the weekends.

We find that the two datasets are also similar regarding the inter-arrival time of matches. The time between the

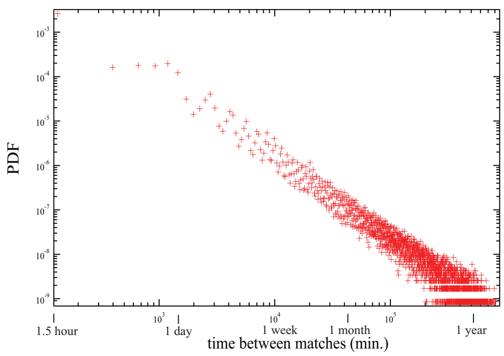

Fig. 2: Probability density of the time between consecutive matches for each player in the DotAlicious dataset.

consecutive matches of individual players is shown, for the DotAlicious dataset, in Figure 2. The time between matches shows a power-law-like behaviour, with many matches being separated by less than an hour, but also consecutive matches being separated by more than a year.

III. A Formalism for Graph Extraction

A common approach in online social network studies is to model a dataset as a graph. Formally, a dataset D is mapped onto graph G via a mapping function $M(D)$. A simple undirected and unweighted graph $G = (\mathcal{N}, \mathcal{L})$ consists of a set $\mathcal{N}$ of N nodes and a set $\mathcal{L}$ of L links. In a *weighted* graph a link weight w is associated to every link in $\mathcal{L}$. In a directed network, a link between two nodes also has a direction.

A *mapping* is a set of rules that define the nodes and links in a graph. Entities are often mapped to nodes, while relations between entities are mapped to links. Entities are usually readily identifiable as persons, events, or objects and are therefore intuitively mapped onto nodes. Mapping relations to links, however, is more challenging.

Entities can be related to each other in many different and often subtle ways and due care should be given to which relations produce insightful graphs. For example, some relations between entities may be the result of chance, while others have a clear origin. The difference between random and meaningful relations is often expressed by a notion of strength. Strength, represented by a link weight, adds another dimension of complexity to representing and understanding the characteristics of a network.

The datasets used to create many of the well-known "real-world" graphs often use a single explicit mapping. For instance, all graphs in the Stanford Network Analysis Project's (SNAP)[2] large graphs collection are based on explicit relations in the source datasets with one exception [7], where graphs are built using the co-occurrence of phrases on news sites and blogs. For this exception, nodes (phrases) and links (co-occurrences) have to meet a set of criteria to be added to the graph, although the authors merely state which criteria they used and do not elaborate on the systematic effects of these criteria.

In this paper we explore how the mapping influences the resulting graph. Since a dataset usually comprises different types of information, e.g. location, time, etc., mapping to only one graph might misrepresent the dataset, unless clearly put in perspective. We consider multiple graph representations. Although we use the gaming datasets as examples, the techniques developed here are generally applicable.

Presently, many proposed complex network metrics (see, for example, Section IV-A) only apply to unweighted graphs. As a result, relations are often only expressed as links if the strength is within a desired range by applying a threshold. *Thresholding*, therefore, is a powerful and popular tool to deal with weighted networks and has an important impact on the resulting graph. For our OSG datasets, the individual players in the dataset are always mapped to nodes. Nodes without adjacent links are removed from the extracted graph. We explore six different strategies to map play relationships to links. For each strategy, a link in the extracted graph corresponds to a different type of gaming relationship between two players, which is likely not expressed directly in the raw (input) data. Each mapping includes a threshold n; as a consequence, each mapping may produce a different graph per threshold value.

By applying different mappings and different thresholds to the dataset, we investigate how a particular mapping and/or threshold affects the resulting graph. We have used the following mappings:

- SM: The number of times two players are in the *Same Match* is greater than n.
- SS: The number of matches played on the *Same Side* is greater than n.
- OS: The number of matches played on *Opposing Sides* is greater than n.
- ML: The number of *Matches* played and *Lost* together is greater than n.
- MW: The number of *Matches* played and *Won* together is greater than n.
- PP: This mapping is a directed version of the other five mappings. In the PP mapping, a *directed* link exists from player A to B if player A has played at least $n\%$ of all his matches (either on the same team, or opposing team etc.) with player B.

The different mappings are related to each other. For example, applying mapping function SM to the Dota-League dataset with a (low) threshold $n = 10$ creates graphs such that for two players p_1 and p_2 a link between them is formed if p_1 and p_2 occur in at least 10 different matches, while applying the SS mapping adds the extra condition that the players played on the same side. Our formalism can support more complex mappings, e.g., players played against each other at least 10 times, during the winter, while located in the same country.

[2]http://snap.stanford.edu/

	Dota-League							DotAlicious					
	SM	SS	OS	ML	MW	PP (SM)		SM	SS	OS	ML	MW	PP (SM)
N	31,834	24,119	26,373	18,047	18,301	29,500		31,702	29,377	11,198	22,813	21,783	34,523
N_{lc}	27,720	16,256	19,814	6,976	8,078	33		26,810	20,971	10,262	10,795	13,382	3,239
L	202,576	62,292	85,581	30,680	33,289	53,514		327,464	108,176	92,010	43,240	54,009	125,340
L_{lc}	199,316	54,186	79,523	17,686	21,569	120		323,064	99,063	91,354	29,072	44,129	17,213
d ($\times 10^{-4}$)	4.00	2.14	2.46	1.88	1.99	0.62		6.52	0.49	14.7	1.66	2.28	1.05
d_{lc} ($\times 10^{-4}$)	5.19	4.10	4.05	7.27	6.61	1100.00		8.99	2.51	17.4	4.99	4.93	16.4
μ	0.0301	0.0060	0.0114	0.0040	0.0032	-		0.0385	0.0120	0.0403	0.0095	0.0194	-
$\bar{h}$	4.42	6.30	5.40	8.09	7.67	3.70		4.24	5.3	3.97	6.80	5.95	18.45
D	14	24	21	28	26	9		17	19	12	20	22	74
$\bar{C}$	0.37	0.41	0.40	0.41	0.41	-		0.43	0.47	0.27	0.47	0.49	-
ρ	0.13	0.25	0.26	0.27	0.28	-0.10		0.08	0.25	0.01	0.27	0.29	0.20
B_m	0.04	0.09	0.09	0.17	0.12	1.21		0.03	0.04	0.05	0.06	0.06	0.37
co_m	85	41	55	19	22	3		131	48	68	16	20	7

TABLE I: Metrics for the Dota-League and DotAlicious datasets for a threshold value of 10, including: number of nodes N, number of nodes in largest connected component N_{lc}, number of links L, number of links in largest connected component L_{lc}, link density d, link density of largest connected component d_{lc}, algebraic connectivity μ, average hop count $\bar{h}$, diameter D, average clustering coefficient $\bar{C}$, assortativity ρ, maximum betweenness B_m, and maximum coreness c_m.

IV. The Formalism in Practice: An Analysis of the Extracted Graphs

In this section, we study the structure of the graphs formed by using different mappings. As each mapping extracts a different type of relationship between players, it is interesting to know whether these relationships give rise to significantly different graph structures. The various mappings can help finding different properties of the two seemingly similar datasets (see Section III). We compare the obtained graphs using a broad selection of graph metrics, which are discussed in Section IV-A. We discuss selected findings from these comparisons in Section IV-B.

A. Graph Metrics

To compare the graphs extracted using different mappings, we calculate a number of network metrics for the largest component of each extracted graph. The selected metrics all reflect properties related to the degrees and paths between players, and allow us to study the social relations in the gaming community. We also include a spectral metric. The metrics used in this section are explained in the following. For a more in-depth explanation we refer to [8], [9].

Size(s) of the connected component(s) (N, L): The size of the largest and other connected components indicates how many fellow players a player can reach in the network.
Link density (d): The link density is obtained by dividing the number of links in the network by $\binom{N}{2}$ and indicates how densely connected the network is. For directed graphs the density is obtained by dividing the number of links by $2\binom{N}{2}$.
Degree distribution: The degree distribution characterises the number of direct neighbours a node has.
Algebraic connectivity: The algebraic connectivity is the second smallest eigenvalue of the Laplacian matrix[3]. This is a spectral graph metric that indicates how well connected a graph is.
Average hop count ($\bar{h}$): The average hop count indicates how many hops, on average, players are removed from each other.
Diameter (D): The diameter is the longest shortest path, in terms of hops, in the network.
Average clustering coefficient ($\bar{C}$): The average clustering coefficient is a measure for how many neighbours of a node are also neighbours of each other.
Betweenness centrality (B): The betweenness centrality score of a node indicates on what fraction of shortest paths it is present and is a measure of node importance.
Coreness (c): A node of coreness k has at least k neighbours with at least k neighbours.
Assortativity Coefficient (ρ): The assortativity coefficient measures to what extent nodes link to other nodes with similar degrees.

In addition to focussing on the largest component, we further analyse the effect of *thresholding* (both on the link weight and the play percentage for the PP mapping) on the size and number of network components. As thresholding removes random and weak relations between players; graphs extracted with a high threshold value will reveal the strongest social structures. Among the graph component representations of these social structures, many may be of similar size, without a clear largest component, while the collection of those components offers insight into strong social ties.

B. Graph Analysis

In this section, based on different mapping functions and their thresholds, we analyse the size and social structure of the graphs extracted from our two OSG datasets (see Section III). We compare the results for Dota-League and DotAlicious to indicate that, while not seemingly different for one general mapping, different mappings can reveal fundamentally different graph characteristics, and as such should be considered by the careful practitioner to fully

[3]The Laplacian matrix is obtained as $L = D - A$, where D is a diagonal matrix of the node degrees and A the adjacency matrix.

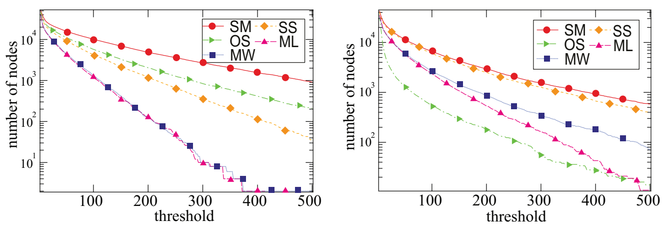

Fig. 3: Number of nodes in the network as a function of the threshold for Dota-League (left) and DotAlicious (right).

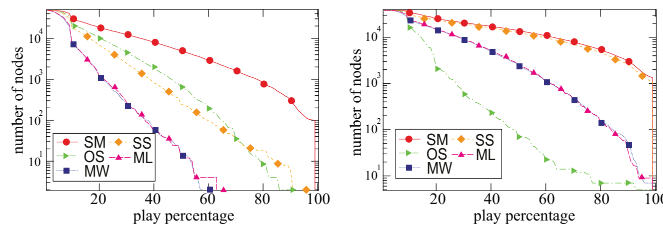

Fig. 4: Number of nodes in the network as a function of the play percentage for Dota-League (left) and DotAlicious (right).

capture the various facets of any data set.

1) Network Sizes: We first analyse the basic sizes in terms of nodes and links of the extracted graphs, and summarise the results in Table I. Under the most general mapping, the SM mapping, graphs extracted from the Dota-League and DotAlicious datasets have different sizes (in Table I, rows N and L, the respective columns "SM"). The network extracted from the DotAlicious dataset contains more links, while the two networks contain an equal number of nodes. Based on this general mapping (column "SM" in Table I, for each of the two datasets DotA-League and DotAlicious), the graph extracted from the DotAlicious dataset is denser ($d = 4.00 \times 10^{-4}$ and $d = 6.52 \times 10^{-4}$, respectively), although the number of matches in the DotAlicious dataset is half that in the Dota-League dataset. We conclude from this that players who play regularly together in the DotAlicious dataset do so in more diverse combinations, because they create more links with fewer matches, than the players in the Dota-League dataset do.

Not only do the graphs obtained for different datasets differ, the graphs obtained using different mappings also highlight differences between the datasets (compare, for each row in Table I, the respective values obtained for each dataset). In general, it seems that players from the Dota-League dataset, when they appear in the same game, play on opposing sides; in contrast, for the DotAlicious dataset they play mostly on the same side. This can be seen by contrasting the sizes of the networks extracted using the SS and OS mappings (rows N and L in Table I). The DotAlicious network contains almost 3 times more nodes and links for the SS mapping than for the OS mapping. In contrast, for the Dota-League dataset, the networks extracted using the OS mapping are larger than those extracted using the SS mapping.

The tendency of DotAlicious players to play on the same side and that of Dota-League players to play on opposing sides can be seen more clearly from Figure 3, where the number of nodes in the network as a function of the threshold is shown for all mappings. The Dota-League graphs are larger (have more nodes) for the OS mapping compared to the SS mapping throughout the range of thresholds. In the DotAlicious dataset, however, the OS mapping results in the smallest graph of all mappings. Moreover, forming links between players that played on the same side results in graphs that are almost as large as those produced with the SM mapping. This indicates that whenever DotAlicious players play many matches together, they almost surely play those matches on the same side. Arguably, playing together forms a stronger social bond than playing against each other; an observation that could not have been made based without studying different mappings and thresholds.

In addition to playing together, DotAlicious players also like to win together. Up from a threshold of 100 the lines for ML and MW markedly take different slopes in Figure 3 (right). The larger networks for MW compared to ML show that winning together forms true friendships—it leads to long-lasting relationships.

In contrast to the graphs we extract from DotAlicious, in the graphs we extract from the Dota-League dataset, the network sizes for the ML and MW mappings are the same, indicating a 50-50 win ratio for matches played together. There is no indication that play relations are strengthened by winning in the Dota-League dataset. This can be explained by the fact that players in Dota-League cannot choose on which side they play; instead, the game mechanism offered in this community only allows joining the waiting queue (see Section V). The 50-50 win ratio we observe for matches played together indicates that matches are generally well balanced.

By using different mappings we have revealed clear differences between the seemingly similar datasets used in this study. The most general SM mapping does not display those differences; for example, the curves for SM in both the left and right hand graph in Figure 3 are very similar. The PP mapping, however, can also reveal a difference between the two SM mappings, as can be observed in Figure 4. The number of nodes in the network for the higher percentages of games played together, is higher by an order of magnitude. This difference indicates that although both datasets contain pairs of players that play a high *number* of matches together, only the DotAlicious dataset contains players that play a high *percentage* of their matches with a select group of other players.

Although the PP mapping reveals differences between the DotAlicious and Dota-League datasets, it also clouds information that was visible using the regular mappings. In the DotAlicious dataset, the difference between the network sizes for the OS and SS mappings is no longer clear when using the PP mapping. The PP mapping highlights that players have, on

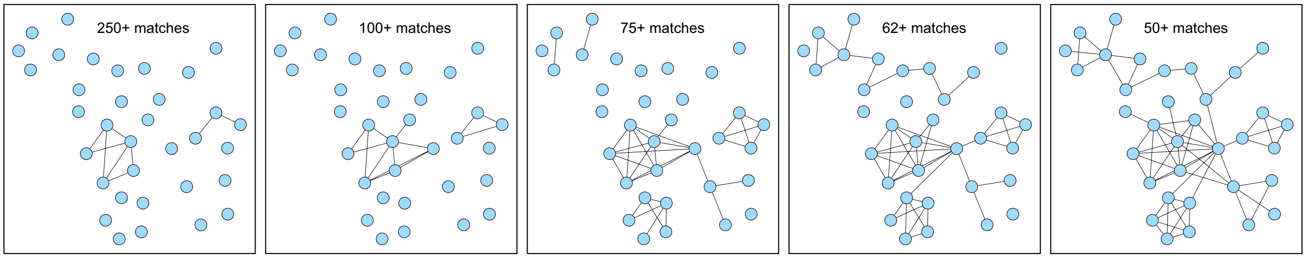

Fig. 5: Clusters for the SS DotAlicious graph, for various values of the mapping threshold.

average, a win ratio of 50%. Because the PP mappings produce directed graphs, the individual win ratio results in equally large networks for ML and MW.

2) Social Network Structure: The metrics in the lower half of Table I offer a better insight into the social structure of the network. We look at three path length metrics: the average hop count ($\bar{h}$), the diameter (D), and the maximum betweenness centrality (B_m). These three metrics relate to how easily nodes can reach each other in the network and whether some nodes are more important with respect to reachability. Social networks are characterised by a low average hop count as a function of the number of nodes, and a relatively high clustering coefficient; Watts and Strogatz [10] call this the small-world property. In Table I, the SM mapping of both datasets shows very similar values for the metrics related to path length. Based on this general mapping function, the small average hop count and high clustering coefficient suggest that the extracted graphs indeed show small-world properties rather than the properties expected of random graphs.

The relatively high clustering coefficient that is often found in social networks is often the result of the fact that a friend of your friend is likely to also be your friend, too. For both the Dota-League and DotAlicious graphs the clustering coefficient of players that played in the same match is around 0.4 (0.37 for Dota-League and 0.43 for DotAlicious), indicating that the players you play with are also likely to play among each other. The higher clustering among DotAlicious players might be the result of the greater control players have over with whom they play, or might be the result of player behaviour.

Contrary to many online social networks, the competitive element in online social games may result in foe relations, in addition to friendship relations, which can be studied via the OS and SS mappings. As can be seen in Table I for DotAlicious, the clustering coefficient for the SS mapping is higher than for the SM mapping. The increase in clustering coefficient indicates that, if we interpret being on the same side in a match as being friends, the friends of your friends are indeed likely to be friends of yours. The clustering of players increases a bit further if we use the MW mapping, showing the strengthening effect of winning matches together. Where winning together strengthens relations, losing matches does, from a clustering perspective, lead to slightly weaker relations as can be seen from the clustering coefficient for the ML mapping. The overlap in links between the ML and MW graphs is about 30%, i.e. 30% of the links in the ML graph also occur in the MW graph and vice versa.

Whereas the SS mapping results in graphs with a higher clustering coefficient for the DotAlicious dataset, the OS mapping creates graphs with a markedly lower clustering coefficient. Intuitively this ties in with the idea that although a friend of your friend is also a friend of yours, an enemy of your enemy is also your friend. Because players on opposing sides in a match instance will most likely not dislike each other outside the scope of that match, we nonetheless see a different graph structure for *foe* relations. The much lower clustering coefficient cannot be explained by other graph properties such as link density. Indeed, the link density in the OS graph is *higher* than in the SS graph.

As in Section IV-B1, we find again that the different mappings allow us to see differences between the two datasets. The clustering coefficients are the same for all graphs extracted from the Dota-League dataset. In particular, there is no evident difference between friend and foe relations in the Dota-League dataset. In contrast, the difference is prominent in the DotAlicious dataset. As observed before in this study, the PP extension of the SM mapping shows vastly different results for the two datasets (see Table I, the columns “PP” corresponding to each dataset). The difference in the number of nodes is not large, but the difference in the size of the largest strongly connected component is. The Dota-League dataset shows no large strongly connected component, whereas the DotAlicious dataset shows a largest connected component of 3,000 nodes. However, the largest strongly connected component extracted from the DotAlicious dataset is poorly connected: it has a diameter of 74 hops and an average hop count of 18.45.

3) Impact of the Mapping Threshold on the Resulting Network Connectedness: An important property of networks is whether the network is connected. A disconnected network indicates that the nodes in the network cannot come into contact, either directly or via intermediate nodes. The number of connected components and the size of the largest connected component in a network are measures of how connected or fragmented the network is.

We find that the threshold value has a large impact on the number of connected components, as depicted in Figure 5. For a threshold $n = 0$, the extracted graphs contain a single connected component. For increasing values of n, the extracted networks become smaller and more indicative of clusters of players who are actively involved in relationships with other cluster members. The reverse process—lowering the

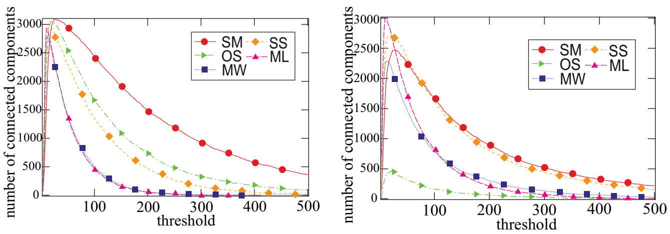

Fig. 6: Number of connected components for Dota-League (left) and DotAlicious (right).

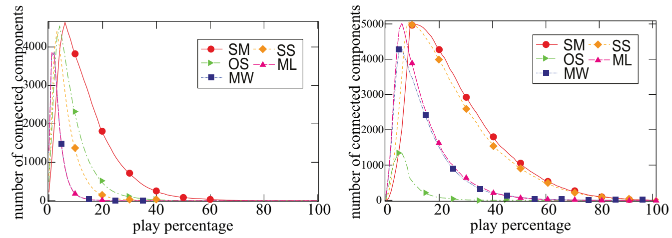

Fig. 7: Number of strongly connected components with PP mapping for Dota-League (left) and DotAlicious (right).

threshold—makes the isolated clusters grow. We illustrate the dependency between cluster growth and threshold reduction in Figure 5, where plots from left to right depict the results for decreasing values of the mapping threshold. As the threshold value decreases, the few connected nodes in the left-most image grow into a single, large connected component. In practical terms, players first organise in smaller clusters before these smaller clusters all connect. The player in-between the two initial clusters depicted in the left-most image of Figure 5 clearly functions as a hub.

Although it is possible to extract an almost fully connected network for all mappings, applying a threshold reduces the relative size of the largest component significantly. Figure 6 shows the number of connected components excluding isolated nodes. As the threshold value increases, the dominance of the largest component diminishes and the number of connected components in the graph increases rapidly. This indicates that the network falls apart quickly in many small components (players with a strong gaming relationship), rather than that nodes are stripped of the giant component or that it splits into a few large subcomponents. The small components are all between two and five nodes in size, which is consistent with the maximum size of five players per team that is typical of DotA (see Section II). At a threshold value of 28, half of the nodes are located in small clusters. At a threshold value of 100, 85% of the nodes are in small clusters, yet the distribution of cluster sizes does not change much. The peak value at a threshold of 28 is in the same region where the number of nodes in the graph stops decreasing dramatically and enters a regime of more steady decrease as was shown in Figure 3.

In Figure 7, the number of strongly connected components in the directed graphs extracted with the PP mapping is shown for the Dota-League dataset (left) and DotAlicious dataset (right). The number of strongly connected components has higher peaks in the networks extracted from the DotAlicious dataset than for the Dota-League and also stays much higher for higher play percentages. This indicates that although the network falls apart in many small clusters in both datasets, only in the DotAlicious dataset these clusters are larger than 1 node. We argue that using the largest component to characterise the network may provide valuable insight for low threshold values, but still is too crude. Scaling up the range of the threshold values could, in an easier manner than most clustering algorithms, bring forth new and significant (strong) gaming relations.

The betweenness centrality (B) and coreness (c) can give clues as to whether a graph contains important nodes, such as influential spreaders [11] for example. For the graphs constructed with the lower thresholds, the maximum betweenness value indicates that between 4 and 9 percent of all the shortest paths cross at the most central node. This value increases with an increasing threshold while the link density increases. This indicates that, as the largest component shrinks, some players play an increasingly important role in facilitating short paths.

V. Formalism-Based Match Recommendations

We focus in this section on the use of our formalism for extracting graphs for improving typical game functionality. We propose as an example an algorithm that can assist in matchmaking decisions. A good matchmaking system can ensure that players in a game have matching profiles and could, therefore, take player clustering information into account. In this section, we analyse how matches are formed by players and propose a graph-inspired matchmaking algorithm that leads to much stronger social ties than random matchmaking, and is slightly better than the real algorithms used by the communities from which we collected our datasets.

The two datasets under study represent communities that deploy in practice different matchmaking algorithms. For Dota-League, players who want to play a match first join a waiting queue. When there are 10 or more players in the waiting queue, the matchmaking algorithm will form teams that are balanced in terms of the skill levels of the players. Although this matchmaking algorithm enforces balanced matches, it does not take into account the social ties of the players. Sometimes players synchronise themselves out-of-game via instant messaging tools to join the waiting queue at the same time and thus increase the probability to end up in the same match as their preferred players, but there is no guarantee that they will play on the same team. In contrast, for DotAlicious, each game server has a number of open matches waiting for players to join, and each arriving player can select which match to join and on which team.

To study the effect of a matchmaking algorithm, we present a scoring of matches that reflects the utility derived by players from a game. Following McGonigal [2], we assume that matches played by players with strong social ties are enjoyed to a higher extent than those played amongst players that have

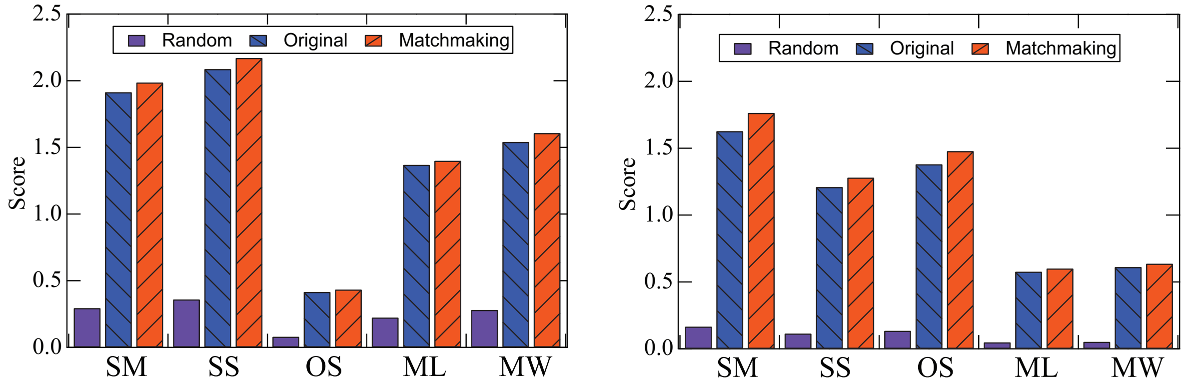

Fig. 8: Average match scores when performing matchmaking for DotAlicious (left) and Dota-League (right).

weak or no social ties. As we have shown in Section IV-B3, a large part of the DotA players are grouped into relatively small, high link-weight clusters. This indicates that these players enjoy their gaming experience best when they play with or against a small set of other players. Since there may be multiple ways to rank matches, our goal is not to propose a unique scoring methodology, but rather to show how to compare and possibly improve matchmaking based on one such socially-aware scoring system.

For each match, we assign a score to a match based on the connected components, referred to as *cluster*, to which the players of the match belong. Cluster membership is derived for the different mappings (excluding PP) described in Section III as follows: first, a threshold value of $n = 100$ is used to filter out weaker links between players. Then, we identify each connected component (cluster) in the extracted graph. The resulting clusters are then numbered, and each cluster member (player) is labelled with its cluster number. In our analysis, we assign a *score* to every match based on the overlap of cluster membership amongst players. Intuitively, we design the score to reward the matches in which many players from the same cluster participate. Specifically, a match receives one score point for every player in clusters who are represented in the game by two or more players.

Consider the match schematically represented in Figure 9. Team 1 consists of players “a” to “e”, as can be seen in the column labelled “Player”; team 2 consists of players “f” to “j”. The column labelled “Cluster” records the cluster number for each player. In total, this match is assigned 7 points, as follows. First, 2 points are given for each of cluster 1, which is represented in this game by players “a” and “c”, and cluster 2 (players “b” and “f”). Then, 3 points are given for players “d”,“h”, and “j” (cluster 3). Players “e”,“g”, and “i” have no fellow cluster members in the match and will be assigned 0 points.

Team 1

Player	Cluster
a	1
b	2
c	1
d	3
e	4

Team 2

Player	Cluster
f	2
g	5
h	3
i	6
j	3

Fig. 9: Example of a match.

To favour the small clusters that lead to novel human emotions [2], and which are shown to be prevalent in our datasets in Section IV-B3, we design the scoring system such that the largest cluster is not considered when assigning points. This effectively avoids biasing the matchmaking algorithm to the dominant set of links, which is akin to recommending a well-known tune, not because it is similar to the requester’s tastes, but because it is present in most other people’s playlists.

We now propose a socially-aware matchmaking algorithm, which works as follows. First, for each 10-minute time interval (sliding windows), the algorithm builds a list of all the players who are online. In practice, the algorithm can build this list from the online information provided by the waiting rooms of each DotA community; to obtain this list from the real (raw) datasets (see Section II), we define a player as being online during an interval if the player has joined at least one match during the interval. Second, the algorithm computes the cluster membership for each player. Third, from the largest online players’ cluster to the smallest, all online players from the same cluster are assigned to new matches if size permits; otherwise, the cluster will be divided into two parts and players from one part will be assigned into new scheduled matches. In practice, the third step can be executed whenever the number of players exceeds the number needed for a game (10), at the end of the 10-minute interval, or periodically every few minutes; when simulating the third step for this study, we use the timestamps when matches become active, as observed in the real (raw) dataset. The rationale behind obtaining all the online players in a time interval is that players who are aware that their friends are online are likely prepared to wait for a short duration in order to play together. This behaviour is consistent with our experience as DotA players and with reports from experienced DotA players.

The matchmaking results are shown in Figure 8; higher scores are better. For comparison purposes, we have also scored the matches obtained via randomly matching players that are online during each time interval (the "Random" matchmaking algorithm in Figure 8) and the matches observed in the real (raw) datasets (the "Original" matchmaking algorithm). The "Random" algorithm indeed scores the lowest, as it does not take into account social ties. Our proposed matchmaking algorithm (the "Matchmaking" algorithm in Figure 8) improves the scores of matches, by a large margin when compared with the "Random" algorithm and by a small but non-negligible margin when compared to the "Original" (real) algorithm. We attribute the better scores achieved by our admittedly simple "Matchmaking" algorithm to the difficulty of seeing whether friends are online in the studied systems, due to shortcomings in the offered service. Without properly displayed information, some players may not be aware that some of their friends are online and thus join other matches. Indeed, we find that the improvement is smaller for DotAlicious than for Dota-League, and attribute this to players in DotAlicious having more freedom and tools for selecting whom to play with. We conclude that, by using a graph perspective instead of manual off-line synchronization, even a simplistic matchmaking algorithm can reach higher match utility scores by leveraging cluster information obtained after a fairly high threshold, that is, after processing only a small part of the original graph. By using a more complex multi-faceted graph extraction rule that, for instance, would include information on friend and foe relationships, we expect that an even better matchmaking could be achieved.

VI. Related Work

Social network analysis and complex networks theory have received increasing attention in the past few years, which has readily resulted in a significant body of related research papers. We refer to [12], [13] for an overview of research on complex networks and to [14], [15] for some excellent overviews of the developments and state of the art in social network analysis. Most research on social network analysis, however, only considers or defines one network for one type of link. Instead, we consider the influence of different (social) link definitions and combinations of link definitions on the emerging (complex) network.

Closest to this study, our previous work [16], [17] on graph extraction and analysis for OSGs investigates several extraction strategies, but does not propose a formalism as comprehensive as we propose in this work, and does not conduct a thorough study of the impact of mapping functions and thresholds on the characteristics of the resulting graphs.

Within the application domain of online social games, few studies use network metrics to divide players into different classes. For example, Kirman and Lawson [18] extract a network from an online game by creating unweighted and undirected links between players that ever exchanged information in the game. They define three types of players based on the successive removal of the highest-degree nodes until the largest connected component falls apart. Shim et al. [3] define implicitly a network that is used to predict future player performance based on the relationship between mentors and apprentices. This analysis could be extended by looking at network-wide properties instead of only local properties. The prediction of the success in games can also be applied to real-world games as is done by Vaz de Melo et al. [19], where a complex network approach is used to predict the performance of basketball teams. The authors propose a network-based ranking of players as a replacement for current statistics such as assists and points scored to predict the future success of a team.

Other studies use different definitions of links to create different networks from the same dataset. Szell and Turner [20] study a detailed dataset of interactions and friend/enemy relations spanning three years in the online game Pardus, which is much less popular than the communities we investigate in this study. They use the game as a substitute of the real-world and test several hypotheses in the field of social dynamics such as social balancing, network densification and triadic closure. They study three different networks extracted from the dataset: the network of communication between players, the network of friends as indicated by players in the game, and the network of enemies as indicated by players in the game. Although the authors present a detailed analysis of the networks for each type of interaction, and especially contribute to existing work by analysing the network of enemy relations, their links are either interaction based or explicitly indicated by the players.

A related study [21] on guild members in World of Warcraft also investigates the differences between networks formed based on different types of interaction. In this work social network analysis is used to explore the network structure of interactions between guild members in the online game World of Warcraft. The authors studied 76 players that formed a single guild and extracted networks by creating links between players that communicate amongst each other. Different types of interaction were classified into seven different categories such as asking for help or group management to form seven different networks. An analysis of these networks in terms of reciprocity and topological structure indicates that the different types of interaction lead to different networks. This study focusses, however, on a rather small group of players, which enables the authors to analyse and classify the communication between players. In our study, we analyse large datasets by applying different mappings, instead of analysing the dataset to find the mappings.

VII. Conclusion

In this paper we have analysed in detail the first step in many online social network studies: the mapping of a dataset on a graph. To investigate the influence of the expert's choice of what type of relationship between nodes should create a link and how strong that relationship must be, in other words, mapping and thresholding, we studied the properties of graphs extracted using six different mappings. As an example application we analysed the gaming relations in two large Online Social Games, Dota-League and DotAlicious.

Our comparison of the graphs obtained from the two datasets through various mappings defined in our formalism has revealed clear differences in the social relations between the players of the respective games. For example, players of DotAlicious are more likely to play with the same group of players, whereas such a preference cannot be seen among Dota-League players. Setting a relatively high threshold value shows that strong social ties between players divide the network into hundreds of small groups. We find that the largest component, which is often the sole focus of network analysis, is no longer the dominant component of the social network. Moreover, even slight variations in the threshold value can completely change the graph metrics for the largest component, also for lower values of the threshold. This gives strong evidence that our formalism offers a complementary view to the simplistic approach based on the lowest possible threshold value, which has been prevalent in earlier studies.

Finally, we have proposed an application that uses the knowledge of implicit social relations in the gaming communities of Dota-League and DotAlicious. Albeit simplistic, our proposed matchmaking algorithm could serve as an example for system designers on how to strengthen or leverage the social ties between players to increase their experience and to attract more players. Our experimental results show that, in comparison to both a random algorithm and the algorithms used by the real communities we have studied, our matchmaking algorithm improves the social cohesion of the community.

For the future, we plan to investigate more mapping strategies, especially multi-thresholded, and apply them and the strategies proposed in this work to the datasets collected in the Game Trace Archive [22]. We also intend to explore other practical applications for the formalism we have proposed in this study.

ACKNOWLEDGMENT

This research has been partly supported by the EU FP7 Network of Excellence in Internet Science EINS (project no. 288021), the STW/NWO Veni grant @larGe (11881), a joint China Scholarship Council and TU Delft grant, and by the National Basic Research Program of China (973) under Grant No.2011CB302603.

REFERENCES

[1] Y. Guo, S. Shen, O. Visser, and A. Iosup, "An analysis of online match-based games," in *Proc. of the Int. Workshop on Massively Multiuser Virtual Environment (MMVE)*, 2012.

[2] J. McGonigal, *Reality is Broken: Why Games Make us Better and How They Can Change the World*. Jonathan Cape, 2011.

[3] K. J. Shim, K.-W. Hsu, and J. Srivastava, "Modeling player performance in massively multiplayer online role-playing games: The effects of diversity in mentoring network," in *Proc. of the Int. Conf. on Advances in Social Networks Analysis and Mining (ASONAM)*, 2011, pp. 438–442.

[4] N. Ducheneaut, N. Yee, E. Nickell, and R. J. Moore, "The life and death of online gaming communities: a look at guilds in world of warcraft," in *Proc. of the Conf. on Human Factors in Computing Systems (CHI)*, 2007, pp. 839–848.

[5] A. Nazir, S. Raza, and C.-N. Chuah, "Unveiling facebook: a measurement study of social network based applications," in *Proc. of the Conf. on Internet Measurement (IMC)*, 2008, pp. 43–56.

[6] J. Rossignol, *This Gaming Life: Travels in Three Cities*. U. Michigan Press, 2008.

[7] J. Leskovec, L. Backstrom, and J. M. Kleinberg, "Meme-tracking and the dynamics of the news cycle," in *Proc. of the Int. Conf. on Knowledge Discovery and Data Mining (KDD)*, 2009, pp. 497–506.

[8] L. d. F. Costa, F. A. Rodrigues, G. Travieso, and P. R. Villas Boas, "Characterization of complex networks: A survey of measurements," *Advances in Physics*, vol. 56, no. 1, pp. 167–242, 2007.

[9] M. E. J. Newman, "The Structure and Function of Complex Networks," *SIAM Review*, vol. 45, no. 2, pp. 167–256, 2003.

[10] D. J. Watts and S. H. Strogatz, "Collective dynamics of small-world networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[11] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley, and H. A. Makse, "Identification of influential spreaders in complex networks," *Nature Physics*, vol. 6, no. 11, pp. 888–893, 2010.

[12] S. H. Strogatz, "Exploring complex networks," *Nature*, vol. 410, no. 6825, pp. 268–276, 2001.

[13] D. Easley and J. Kleinberg, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge: University Press, 2010.

[14] J. Scott, "Social network analysis: developments, advances, and prospects," *Social Network Analysis and Mining*, vol. 1, no. 1, pp. 21–26, 2011.

[15] S. P. Borgatti, A. Mehra, D. J. Brass, and G. Labianca, "Network analysis in the social sciences," *Science*, vol. 323, no. 5916, pp. 892–895, 2009.

[16] V. Posea, M. Balint, A. Dimitriu, and A. Iosup, "An analysis of the bbo fans online social gaming community," in *Proc. of the Int. Conf. RoEduNet*, 2010, pp. 1–6.

[17] M. Balint, V. Posea, A. Dimitriu, and A. Iosup, "An analysis of social gaming networks in online and face to face bridge communities," in *Proc. of the Int. workshop on Large-scale System and Application Performance (LSAP)*, 2011.

[18] B. Kirman and S. Lawson, "Hardcore classification: Identifying play styles in social games using network analysis," in *Proc. of the Int. Conf. on Entertainment Computing (ICEC)*, 2009, pp. 246–251.

[19] P. O. Vaz de Melo, V. A. Almeida, and A. A. Loureiro, "Can complex network metrics predict the behavior of nba teams?" in *Proc. of the Int. Conf. on Knowledge Discovery and Data Mining (KDD)*, 2008, pp. 695–703.

[20] M. Szell and S. Thurner, "Measuring social dynamics in a massive multiplayer online game," *Social Networks*, vol. 32, no. 4, pp. 313–329, 2010.

[21] C. Ang, "Interaction networks and patterns of guild community in massively multiplayer online games," *Social Network Analysis and Mining*, pp. 1–13, 2011.

[22] Y. Guo and A. Iosup, "The game trace archive," in *ACM Annual Workshop on Network and Systems Support for Games (NETGAMES)*, 2012, pp. 1–6.