DEVIANTART IN SPOTLIGHT: A NETWORK OF ARTISTS

Almila Akdag Salah[1], Albert Ali Salah[2], Bart Buter[3], Nick Dijkshoorn[3], Davide Modolo[3], Quang Nguyen[3], Sander van Noort[3], Bart van de Poel[3] [1]Media Studies, UvA; VACD, Sabanci Uni. E-mail: <a.a.akdag@uva.nl> [2]Dept. of Comp. Engineering, Bogazici Uni. E-mail: <salah@boun.edu.tr> [3]Informatics Institute, Uni. of Amsterdam

<Unedited author's proof - do not distribute>

Abstract

deviantArt (dA) is the largest online community of user-generated artworks. So far, a scholarly study of dA has been missing. The main goal of this paper is to describe several tools for the network analysis of this community and to propose future research directions for understanding this collaborative and autonomous art venue.

Launched in 2000, deviantArt is one of the largest online communities showcasing user-made artworks. With its 19 million members, 100 million images and 45 million monthly visitors (all as of May 2012), surpassing any real or virtual museum, dA offers a genuine virtual space for disseminating art. It generates a platform free of institutional and governmental politics, democratizing the way arts are generated, shared and enjoyed, mainly through the underlying social network that allows distributed valorization of arts. Unlike Flickr [1-2], which focuses on photographs only, dA hosts a variety of genres, offering (and even enforcing) a delicate category structure to its users. Thus, all artwork is organized according to a comprehensive category structure, from photographs to various digital and traditional art forms. Each member of the site has their own webpage featuring a gallery, a journal, a favorites section, as well as a basic information box highlighting statistics such as number of visitors, number of comments, number of downloads etc. These statistics build up the main evaluation system of the dA community; a member with a large number of visitors/ comments is seen as successful. The information box also contains demographic data (gender/ geographic location/ age), and details about membership. This rich background information allows us to study the dynamics of dA via network analysis.

In this paper, we highlight two aspects: 1) the social structure, where we characterize dA in terms of artists watching each other, 2) the category structure, where each category is represented with a node. The number of artists publishing in any two categories is converted to weighted edges between the categories, showing the practical affinity between categories.

A Visualization Tool

In order to supplement network analysis, we have developed a visualization tool that allows us to depict galleries of artists or artworks of a category [3]. There exist several tools to visualize large image collections, but these tools are geared towards similarity-based image search or content-based image retrieval. Our tool is designed for the analysis of similarities between artists and categories, and for discovering artworks with unexpected visual qualities. It projects a large number of images onto selected feature spaces (about a hundred features implemented), but it is also able to suggest the most discriminative feature space, given two sets of images. Our experiments with this tool show that dealing with the whole network for any given task is too unwieldy. For this purpose, we have extracted a representative core network.

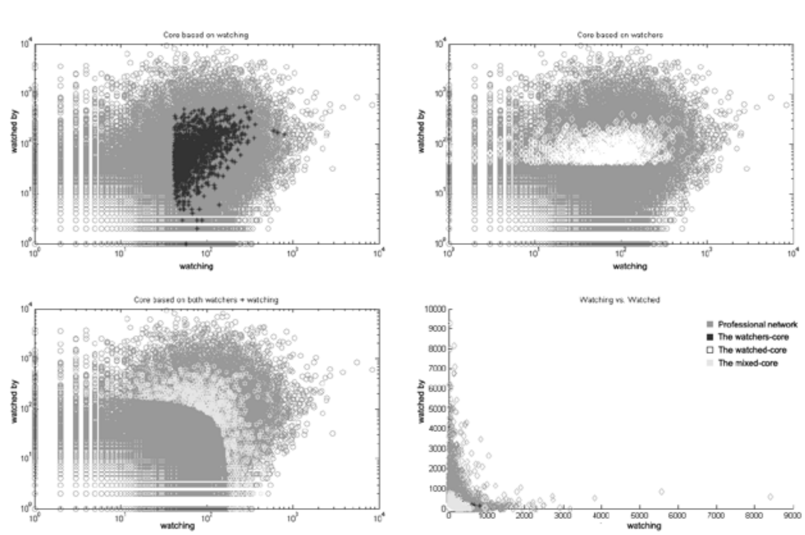
The core of deviantART

The dA network consists of 13 million members, but some of the members are passive users. In order to get to the vibrant core of the dA network, we have used a number of assumptions that weed out most of the members. This helped us reach a manageable and relevant set of users. The first heuristic we used is the subscription status; the paying members of the site are more serious users and have access to more services. These can be automatically determined through scraping. Our first data reduction followed these members, and we thus obtained a network with 103.663 vertices and about 4,5 million arcs, the latter representing a user being 'watched' by another user (average degree is 43,25). This is referred to as the *member network* in the rest of the paper. We have used the member network in the analysis of the category structure, subject of the next section.

Watchers get notifications about the activities of the members they are following. Thus, if a member has a high number of watchers, he/she is able to reach out to a bigger audience. This property guides us in capturing the core of dA. From the member network, we have recursively removed nodes (and all connecting arcs) that had only a few watchers. Each iteration of this k-core procedure peels off one shell from the peripheries of the network, leaving us finally with a densely connected graph [4]. Fig. 1 shows nodes of the member network according to the number of outgoing (watched) and incoming (watchers) arcs, in a log-log plot. The nodes in the core are shown in a different color, depending on the removal.

The three different core networks are artists that are power-watchers, popular artists and lastly a mix of both. The statistics are shown in Table 1. L_G denotes the characteristic path length (average shortest path length between vertices). C_G denotes the directional clustering coefficient, which indicates social grouping. L_{random} and C_{random} denote these statistics for a typical

Fig. 1. ***Member Network*** **and the** ***Core Networks*** **superposed (watchers core, watched core, and the mixed core). (© Almila Akdag Salah Copyright Holder.)**



random graph of the same size.

Table 1. Core Network Statistics

Statistic	Watchers	Watched	Mixed core
Nr. of nodes	1701	1471	1099
Nr. of arcs	139.285	127.837	166.244
Avg. degree	81,88	86,90	151,27
L_G	2,15	2,27	2,14
L_{random}	1,69	1,63	1,40
C_G	0,200	0,220	0,200
C_{random}	0,048	0,059	0,140

The Category Structure

We have scraped the galleries of the member network, for a total of about 13 million artworks. Members have to assign each of their artworks to one of the pre-defined categories. The hierarchically organized category structure is important for two reasons. Firstly, it serves as a ground truth for any explorative method on discovering community-structure in the network, as the artists publishing in distinct categories create their own sub-communities naturally. Secondly, the combination of categorical labels with the social network paves way to a clear picture of the relation between different styles and their adherents.

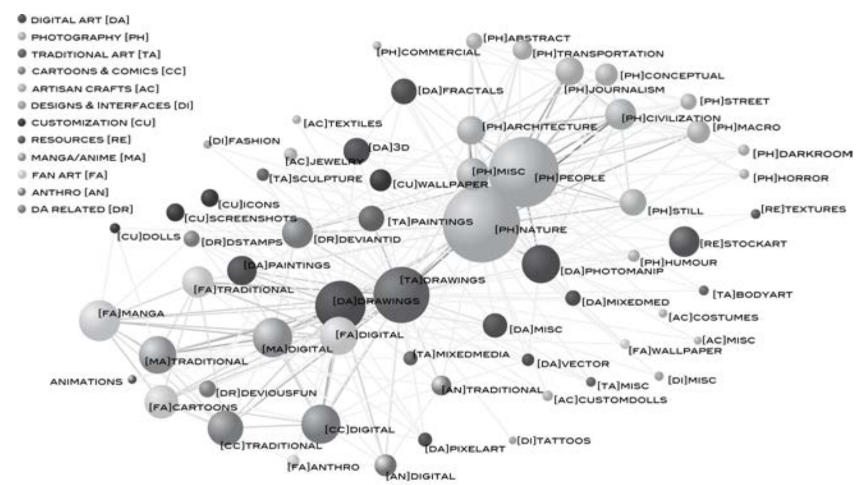


Fig. 2. The category structure of the dA member network. (© Almila Akdag Salah Copyright Holder.)

The category graph [Fig. 2] depicts categories as vertices, and their relations as edges. It is obtained by assigning artists to categories; each artist with a minimum number of works in a particular category (empirically set to 10) is deemed an artist of that category. The vertex sizes are proportional to the number of works in the category. (To give an idea, the Photography-Nature category is the largest vertex with 1.647.425 artworks. Next to it is the Traditional-Paintings category, with 180.204 works.) Links between categories indicate the number of artists (darker means higher) that have produced in both categories. A free energy based optimization [5] is used to spread the vertices on a 2D space, where categories drawn together have more overlapping artists than others. Naturally, the larger categories are placed centrally in this representation, and smaller categories are pushed towards the periphery.

The spatial clustering reveals that there are strong links between subcategories; the producers of natural photographs are also the producers of photographs with people. Thus, the technique is a major determinant here, as much as the content.

The most popular category is *Photography*, with approximately 5 million works. Its subcategories are excessively grouped together, building a densely connected cluster that has strong ties with the category *Resources*. This latter category mainly entails stock-images that are freely available for the dA community. The rest of the network reflects an interesting mix between subcategories of different headings, indicating that genre itself is overridden by the pertinence of the technique. Thus, we see that the subcategories *Digital Art – Drawing* and *Traditional Art – Drawing* are strongly connected. The same type of intermingling is especially observable among the main headings of *Cartoon & Comics, Fan Art*, and *Manga/Anime*. These headings, instead of building their own clusters, are grouped together according to their subheadings. For example, a grouping is visible between traditional and digital cartoons. The cartoon artists do not use the traditional or digital medium exclusively, but produce in both. The boundaries of some categories are fuzzy; the *Fan Art/Cartoon* subcategory is indeed very close to *Cartoon/Traditional* and *Cartoon/Digital Art* subcategories.

Future Directions

The category structure of the dA member network reveals an unexpected picture: Aside from the category of *Photography*, most categories do not create clusters of their sub-categories, but rather form mixed clusters according to production techniques. The network visualization makes it clear that *Photography* should be further analyzed as a separate cluster, whereas artworks in other categories could be analyzed in comparison to each other. Among the main headings, *Fan Art* is the most interesting group, as it covers all the important subcategories (*Digital Art, Traditional Art, Cartoon, Manga*), and is close to the main categories with most works.

The next step in our work plan is to apply comparative image analysis on chosen categories, and determine whether we can identify patterns and styles that play a role across category boundaries. The visual comparison of categories may conform to the structure we depict in Fig.2, or it may reveal visually similar but socially distant groups. Another worthwhile comparison is between the watcher/watched graph and the overlapping artist graph.

Finally, we have collected data that pertains to the temporal growth of the dA network. For each artist, we can plot a social growth trajectory in time, to characterize growth patterns, as well as to interpolate and predict the growth of the entire underlying social network.

References and Notes

1. M. Cha, A. Mislove, and K.P. Gummadi, "A Measurement-driven Analysis of Information Propagation in the Flickr Social Network," *Proc 18th Int.Conf. on WWW* (ACM, 2009) pp. 721–730.

2. J. Leskovec, L. Backstrom, R. Kumar, and A. Tomkins, "Microscopic Evolution of Social Networks," *Proc. 14th ACM SIGKDD* (2008) pp. 462–470.

3. B. Buter, N. Dijkshoorn, D. Modolo, Q. Nguyen, S. van Noort, B. van de Poel, B. A.A. Salah, A.A. Akdag Salah, "Explorative visualization and analysis of a social network for arts: the case of deviantART," *Journal of Convergence* (2011), 2(1), pp.87-94.

4. S.B. Seidman, "Network structure and minimum degree," *Social Networks* (1983), 5, pp. 269-287.

5. T. Kamada, and S. Kawai, "An Algorithm for Drawing General Undirected Graphs" *Information Processing Letters* (1989) 31, pp. 7-15.

6. This research is funded by De Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NOW) Veni grant.