

An insight to the analysis of roles in IMDb

Alberto Lumbreras, James Lanagan
Technicolor
{alberto.lumbreras, james.lanagan}@technicolor.com

Bertrand Jouve, Julien Velcin
ERIC Laboratory, Université Lumière Lyon 2
{bertrand.jouve, julien.velcin}@univ-lyon2.fr

Role theory, the branch of Sociology that studies roles in communities, has been around since the first half of the 20th century. Yet, the definition of a *role* is still a matter of discussion. Sociology lacks a unified role theory that includes both structural (objective) and psychological (subjective) variables [1]. Since the emergence of online communities, Computer Science has tried to construct a practical understanding of a role by capturing measurable variables such as structural position or observed interactions of actors within a network.

Two approaches can be followed when finding roles in a community. Top-down approaches take an *a priori* definition of one or several roles, and examine the community to find persons that match patterns. Bottom-up approaches look for behavioral patterns among users to obtain a descriptive definition of roles [2] [3]. The former approach lacks flexibility: new roles can live in the community. The later lacks of universality: two different communities often lead to different role descriptions. A third approach would be to iteratively mix these two approaches in order to converge on role definitions refined by observed behaviors: explore the community, redefine our *a priori* assumptions, refine the exploration, and so forth, following a cycle that connects our abstract *a priori* assumptions to the observed behaviors [4]. We think this third approach can offer the most fruitful results.

Social roles are, by definition, community based: people play roles as a result of norms, beliefs and preferences (of themselves and of other persons in their community)[1]. Online communities are often grouped around common interests and their temporal stability allows them to generate certain properties such as social norms that explain taboos, vocabulary, friendship patterns and roles. The study of roles and communities must therefore be intertwined.

Our aim is to understand roles within the specific context of IMDb forums as a particular case of online forums, to study its particularities - new roles- and its commonalities with other forums. The first task that we face is to determine whether there are consistent communities in these forums and, if there are, whether they are restricted to threads, the movie, the genre, or maybe the whole IMDb. We start from almost none *a priori* knowledge of what roles we can find [2], and explore IMDb in search of patterns that could give us insights into what is going on. In this first study, we focus on observed interactions between users.

Our dataset consists on 117.000 users that commented on 1.115.616 posts about 706 films. There are 8.453 users that commented on at least 10 films. Films consist on the top 250 and the bottom 100 in terms of votes at September 2012. Films from the top 1000 with more than 75000 votes were included as well. This can introduce a bias as communities could be stronger in popular films.

We observed three major groups of users in terms of the number of movies they commented on: users that occasionally participated

in a forum — some *lurkers* [2] are amongst these, but also *loners* that did not interact with anyone in the forum, *regular* users that commented about different movies and interacted with some other users; and a small *elite* of users that commented in hundreds of movies interacting with a large set of users.

The group of regular users (we included here users with more than 10 films) might contain a variety of roles which deserve further study. In this group, people participate in a sparse set of forums and the mean *co-occurrence* (two users discussing in the same movie) is low even between users that occasionally converse (9.2%). That can be explained because the collection of movies is large and it would be rare that two users like *and* participate in the same collection of movies even if they have some co-occurrence. Moreover, IMDb is not used – or designed – like other social networks, such as Facebook or Twitter, where users tend to form stable groups. In IMDb, most users seem to use it individually and there is no apparent community behavior. The probability that these occasional friends talk to each other is naturally even lower (3.7%).

However, if two users interacted somewhere in the past, its likely (68%) that they will interact if they both participate on another common forum, as if a bond was created after the first time they talked. That indicates that a user follows the activity of his favorites, or that when one of them, by chance, sees the other *old friend* in a new forum she will probably reply to her post.

In future work we further analyze the communities in IMDb. Following [5], we will cluster users based on their interaction behavior and we will apply community detection algorithms to the users graph. These hopefully will give us more insight to understand the nature of roles over IMDb.

A clear role definition will have to be provided. While some authors have provided good definitions [2][4], they lack of temporal dimension. Including this temporal or dynamic dimension in the definition of role will allow us to explicit roles that has an intrinsic dynamic behavior.

REFERENCES

- [1] B. Biddle, "Recent Developments in Role Theory," *Annual Review of Sociology*, vol. 12, no. 1, pp. 67–92, Jan. 1986.
- [2] S. Golder and J. Donath, "Social roles in electronic communities," *Internet Research*, vol. 5, 2004.
- [3] E. M. Airoldi, D. M. Blei, S. E. Fienberg, and E. P. Xing, "Mixed Membership Stochastic Blockmodels," *The Journal of Machine Learning Research*, vol. 9, pp. 1981–2014, Jun. 2008.
- [4] E. Gleave, H. Welser, T. M. Lento, and M. A. Smith, "A Conceptual and Operational Definition of 'Social Role' in Online Community," in *2009 42nd Hawaii International Conference on System Sciences*. IEEE, 2009, pp. 1–11.
- [5] N. Anokhin, J. Lanagan, and J. Velcin, "Social Citation: Finding Roles in Social Networks. An Analysis of TV-Series Web Forums," *Second International Workshop on Mining Communities and People Recommenders - COMPER '12*, 2012.