Temporal Evolution of Scientific Communities *

[Deadline February 22, 2013] †

Bruno Leite Alves
UFMG
Belo Horizonte, Brazil
bruno.leite@dcc.ufmg.br

Fabrício Benevenuto
UFMG
Belo Horizonte, Brazil
fabricio@dcc.ufmg.br

Alberto H. F. Laender
UFMG
Belo Horizonte, Brazil
laender@dcc.ufmg.br

ABSTRACT

Categories and Subject Descriptors

H.4 [**Social Network**]: Temporal Analisys; J.4. [**Computer Applications**]: Social and behavioral sciences Miscellaneous

General Terms

Human Factors, Measurement.

Keywords

communities, scientific communities, core community, evolution

1. INTRODUCTION

Since its beginning, society has been organizing itself into communities, which are groups of people with common interests. Particularly, the proliferation of new communication technologies based on the Internet has facilitated the rapid formation and growth of online communities. Communities exhibit a wide range of characteristics and serve a variety of purposes, from small groups engaged in tightly niche topics such as a very specific scientific community, to millions of users linked by an interest such as a community related to a sport team or fans of a celebrity.

Often, individuals who are socially connected in a community tend to share interests and similarities. Although, there are many factors that might determine a community formation and its growth, there are two main driven forces used to explain similarity in a community formation: influence and homophily. On one hand, influence posits that individuals change to become more similar to their friends in the

community. On the other hand, homophily postulates that individuals create social connections within a community precisely because they are already similar. Recent efforts have provided quantitative evidences of both forces [1–4] and several existing theories [**?**, 17, 22], models [9, 10], and approaches [19, 23] rely on identifying a group of influential individuals with the power to affect not only the underlying network structure of a community, but also to interfere on the spread and flow of information within a community.

In this paper, we take a different perspective and study a complementary problem. Here, we focus on studying the roles that community leaders play and how they can impact intrinsic evolving properties of communities. From a sociological perspective, it helps us understanding how segments of society evolve as well as answering longstanding questions related to the interaction among different types of participants. From the computer science perspective, understanding these aspects is critical not only for link prediction as well as the designing better recommendation systems, but it is also a necessary step for viral marketing strategies and social campaigns. Such a study, however, has been difficult as essential components like human connections and a proper definition of leadership is hard to be reproduced at a large scale within the confines of a research laboratory.

We focus on studying the role of leaderships in scientific communities. When prolific research leaders decide to join or leave communities, they take with them resources, experience, students and possibly influence other authors to do the same, which makes scientific communities very suitable for this kind of study. We used data from DBLP to identify scientific communities, represented by the main ACM SIGs conferences. Then, we propose a strategy to infer the community core, the leaders of a given scientific community in a given period of time. Finally, we investigate how aspects of the core might impact on the community structure. Our results show thatâĂę (TO DO).

The rest of this paper is organized as follows. Next section surveys related efforts. Then, Section 3 describes our strategy and dataset used to construct the connections around scientific communities and analyzes the main evolving properties of this communities. Section 4 describes our strategy to compute the community core and Section 5 investigate the main properties of these sets of authors within their communities. Finally, Section 6 concludes the paper and provides directions for future work.

2. RELATED WORK

Studies has been done to understand the structure of so-