Computer Communications 80 (2016) 72–81

Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

Analysis of topology dynamics for unstructured P2P networks

Baoqun Yin[a], Xiaonong Lu[a,b,*], Jing Huang[a], Yu Kang[a]

[a] Department of Automation, University of Science and Technology of China, Hefei 230027, China
[b] School of Management, Hefei University of Technology, Hefei 230009, China

ARTICLE INFO

Article history:
Received 13 January 2015
Revised 9 January 2016
Accepted 15 January 2016
Available online 2 February 2016

Keywords:
Evolution model
Unstructured P2P network
Simulation
Degree distribution

ABSTRACT

In this paper, we propose an evolution model for unstructured Peer-to-Peer (P2P) file-sharing systems, which tries to capture the dynamic properties of P2P application networks. The model consists of four most important events in the evolution of network topology: node addition, node departure, connection establishment, and edge deletion. Based on the model, we derive a difference equation of degree distribution, which can be solved analytically. By analyzing the solutions, we find that there are two factors affecting the degree distribution. The first one represents the internal dynamics of network topology, while the second one represents the external influence. We find that the internal dynamics makes the degree distribution follow power law while the external influence makes it deviate from power law. User behaviors and resource distributions are the main reasons for this deviation.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Peer-to-Peer file-sharing systems, such as those supported by the well-known Gnutella protocol, have received a lot of attention in the last decade as they offer high availability without the significant cost of highly redundant and central entity. The topology of a P2P network is a graph generated by considering P2P hosts and application level connections as "nodes" and "edges" respectively. It plays an important role in the proper function of P2P network, because a good topology makes file transmission more efficient and network more robust. For example, searching on networks with lower diameter will be faster, and networks, which have an approximately power-law vertex degree distribution, are highly robust against random removal of nodes, but are relatively fragile to the specific removal of the most highly connected nodes [1].

Technical means have been devoted to the research of P2P network topology, for example, [2–5] explore the P2P networks by crawlers. The crawler captures the snapshots of the system during the crawl. Examination of individual snapshots reveals topological properties of the system (e.g., the size and diameter of the network and node degree distribution), while the comparison of sequential snapshots identifies dynamics of such properties as a function of time. The accuracy of the conclusions in these researches relies on the accuracy of the crawler [4]. Specifically, the duration of the crawl determines the time granularity between compared back-to-back snapshots, the dynamics of topological properties obtained based on the crawler may be not accurate if the crawl takes a long time. Based on the results obtained by a network crawler, Jovanović et al. [2] reported that the Gnutella topology exhibits "small-word" properties and the out degree distribution obeys power law. However, in 2008, Xie et al. [3] discovered that the current Gnutella network deviates from the earlier power laws. By analyzing the topology information obtained by the crawler developed by them, Xie et al. observed that the power law property of the rank distribution and degree distribution has become weak, some even invalid. The different observations about the topology of the same P2P file-sharing network in different years indicate that modeling P2P as a static network [2,6] may not be accurate. These static models cannot capture the time evolution properties of P2P topology.

According to the topology, the current P2P networks can be classified to three types, centralized, structured, and unstructured. In a centralized P2P network, such as Napster, all objects index items are kept in a centralized server. Peer only needs to search in the server for its wanted resources. The topology of a centralized P2P is a star-like network. The structured P2P network usually emerges when using Distributed Hashing Table (DHT) to organize the topology. Search is accomplished by looking up key-value pairs in the DHT. The topology of this type of network is structured graph, such as mesh [7,8], d-dimension torus [9] or ring [10]. The unstructured P2P network is completely decentralized. All the peers in the network operate independently, and the search is executed hop-by-hop through the network until success/failure or

* Corresponding author. Tel.: +86 13615600212.
E-mail addresses: bqyin@ustc.edu.cn (B. Yin), lxn520@mail.ustc.edu.cn, xnlu@hfut.edu.cn (X. Lu), xjhuang@mail.ustc.edu.cn (J. Huang), kangduyu@ustc.edu.cn (Y. Kang).

http://dx.doi.org/10.1016/j.comcom.2016.01.009
0140-3664/© 2016 Elsevier B.V. All rights reserved.

timeout. The topology of such network is often random and unstructured mesh. Gnutella protocol 0.4 is a typical unstructured P2P network. In this manuscript, we only focus on the unstructured P2P network as a case-study, so when we use the word “P2P network”, we actually refer to unstructured P2P network.

In this paper, we present a new evolution model for the dynamic of P2P network topology. There are three aims in the modeling. (1) Unlike the previous researches, we try to propose an analytic model for the dynamic of P2P topology, the time granularity may be much small in the model. (2) In order to capture the time varying properties of P2P topology, the model involves a time evolution process. In each time unit, different events may happen with probabilities. (3) Due to the different protocols of P2P networks, their topologies may differ significantly. The factors influence the topological properties include overlay construction algorithms, user behaviors, etc. The model can be used as an analytical framework to figure out the detailed influences of these factors.

The remainder of the paper is organized as follows. In Section 2, some related works of network evolution model are presented and discussed. Section 3 introduces a new evolution model for the unstructured P2P network, and four kinds of events are described. In Section 4, the difference equation of degree distribution is developed and its solution is given. In Section 5, we conduct simulation experiments to demonstrate the effectiveness of our presented model. Finally, conclusions are drawn in Section 6.

2. Related works

The idea of evolution model for P2P is mainly inspired by the research in the evolution model of complex networks. Many evolution models [11–14] have been proposed to describe the general topology dynamic process for generalized networks (by which we mean the network model with consideration of the common properties of all networks). These models mainly focus on generating networks with some of their statistical properties, such as degree distributions and weight distributions, to be consistent with real networks. And some kind of node addition and deletion mechanisms are used to simulate the procedure happening in real networks.

As a result, the generated networks do have the desired statistical properties, such as power law distributions of degree, weight, and strength. However, in the practical P2P networks, in addition to node addition and deletion, the events of edge deletion and connection establishment also happen frequently. From this point of view, these models may not be applicable to describe the realistic P2P networks.

In recent years, significant works have been focused on exploring the topology characteristics of practical networks. Many measurement methods and analytical models are proposed to measure the network. In [15] and [16], Crespelle et al. designed a method for the measurement of the neighborhood of internet core routers, which estimates the degree distribution of the IP-level router topology, referred to as the Neighborhood Flooding method. And the simulation results showed experimentally that Neighborhood Flooding method is free from the bias highlighted in the classical approach and is able to observe properly the exact degree of a vast majority of nodes in the core of the network. In [17], Arakawa et al. developed a modeling method to construct ISP router-level topologies. The presented ISP topologies had a quite different characteristic from other topologies attained with conventional modeling methods. In [18] Liu and Ansari proposed a shared user based affinity measurement between websites and built an affinity graph to represent the topology of observed mobile websites. This method is effective in identifying hidden website communities in mobile internet. In [19], Li et al. utilized the theories of social networks and economic incentives to investigate the formation of P2P networks, and developed an analytical model to evaluate the impacts of system parameters on the emergence of self-organized P2P network structures. A framework with multilevel P2P formation dynamics, incorporating individual sharing decisions, group admission, and group interconnections is presented. In [20], Su et al. presented a measurement study on the topology characteristics of a real bit-torrent (BT) network. A hybrid BT measurement system with a combination of active and passive approaches and explored the BT peer exchange extension to collect connection information among peers, discovered that a steady-stage BT network has a Gaussian-like peer-degree-frequency distribution, especially for a large scale swarm. In [21], the heterogeneous characteristics of Internet applications and traffic are investigated from a complex network perspective. On the basis of the analysis of the flow records, a complex network model is developed for the flow graph growth. The results show that the model matches well the flow graphs in terms of both degree distribution and entropy dynamics.

In this paper, we attempt to develop a more complete evolution model for P2P by additionally considering connection establishment and edge deletion. The model consists of four events, node addition, node deletion, connection establishment and edge deletion. These four events can capture the most important dynamics in the evolution of network topology. Therefore, by exploring the dynamic process of topology evolution, we derive a difference equation about P2P network degree distribution. We obtain the analytical solution for the difference equation, and discover some interesting results. This model may ease the analysis of the interplay between P2P topology and user behaviors. Furthermore, we believe that the model presented in this paper can be extended to describe a class of generalized networks in which the edge deletion and connection establishment happen frequently.

3. The new evolution model

In this section, we will give a detailed description of the evolution model. Generally, there are four events that will affect the topology of a P2P application network, which are **node addition & neighbor selection, node deletion, connection establishment, and edge deletion**. Although may not realistic, we assume these events happen randomly. The happening of these events will cause the addition and deletion of both nodes and edges. We start this section by giving a general description of these events.

At any time, a user of a P2P network may execute the P2P program, which logs into the P2P network and finds a number of application level neighbors for the user. In the P2P topology model, we can mimic this by adding a new node and establishing connections to nodes already in the network. We call such a process **node addition & neighbor selection**. With high probability, a user already in the network is willing to exchange some resources with others; he will find the resource that seems interesting to him and establish connections to some of the peers possessing the resource. We call this process **connection establishment**. When the connection between the requesting node and the requested one is established, the transmission of resource will start. The time that is needed to finish the transmission depends on the available bandwidth between the two nodes. If the transmission is finished or the user decides to interrupt that transmission, the connection between the requesting node and the requested one is terminated. We model this process by deleting the edge between these two nodes, and call it **edge deletion**. In addition, the user may close the P2P program at any time for various reasons. If the P2P program is closed, the node representing the user in the P2P topology will be deleted and the connections associated with it will be removed too. We call this process **node deletion**. All of the above four events happen randomly in a P2P network, and the nodes and/or connections will be added or deleted accordingly.

Download English Version:

https://daneshyari.com/en/article/447630

Download Persian Version:

https://daneshyari.com/article/447630

Daneshyari.com