Journal of Computer and System Sciences 81 (2015) 1088–1109

Contents lists available at ScienceDirect

Journal of Computer and System Sciences

www.elsevier.com/locate/jcss

Distributed agreement in dynamic peer-to-peer networks ☆

John Augustine [a,*,1], Gopal Pandurangan [b,c,2], Peter Robinson [b,3], Eli Upfal [c]

[a] Department of Computer Science and Engineering, Indian Institute of Technology Madras, Chennai 600036, TN, India
[b] Division of Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore
[c] Department of Computer Science, Brown University, Box 1910, Providence, RI 02912, USA

ARTICLE INFO

Article history:
Received 3 March 2013
Received in revised form 4 April 2014
Accepted 26 June 2014
Available online 12 January 2015

Keywords:
Dynamic networks
Distributed computation
Fault-tolerance
P2P networks
Agreement
Churn
Randomization

ABSTRACT

Motivated by the need for robust and fast distributed computation in highly dynamic Peer-to-Peer (P2P) networks, we present first-known, fully-distributed algorithms for the fundamental distributed agreement problem in dynamic networks that experience heavy node *churn* (i.e., nodes join and leave the network continuously over time). Our algorithms guarantee *stable almost-everywhere agreement* with high probability even under high adversarial churn and run in time that is polylogarithmic in n (which is the stable network size). Our first algorithm can tolerate a churn of up to εn per time step, sends only polylogarithmic number of bits per node per time step, and works under an adversary that is oblivious to the algorithm's random choices. Our second algorithm, designed for the more challenging adaptive adversary, can tolerate a churn of up to $\varepsilon\sqrt{n}$. Being easy to implement, our algorithms could serve as building blocks for other non-trivial distributed computation in dynamic networks.

1. Introduction

Peer-to-peer (P2P) computing is emerging as one of the key networking technologies in recent years with many application systems, e.g., Skype, BitTorrent, Cloudmark etc. However, many of these systems are not truly P2P, as they are not fully decentralized — they typically use hybrid P2P along with centralized intervention. For example, Cloudmark [25] is a large spam detection system used by millions of people that operates by maintaining a hybrid P2P network; it uses a central authority to regulate and charge users for participation in the network. A key reason for the lack of fully-distributed P2P systems is the difficulty in designing highly robust algorithms for large-scale dynamic P2P networks. Indeed, P2P networks are highly dynamic networks characterized by high degree of node *churn* — i.e., nodes continuously join and leave the network. Connections (edges) may be added or deleted at any time and thus the topology changes very dynamically.

☆ A preliminary version of this paper appeared in the Proceedings of the ACM/SIAM Symposium on Discrete Algorithms (SODA), 2012, 551–569.

* Corresponding author.

E-mail addresses: augustine@cse.iitm.ac.in (J. Augustine), gopalpandurangan@gmail.com (G. Pandurangan), peter.robinson@ntu.edu.sg (P. Robinson), eli@cs.brown.edu (E. Upfal).

[1] Work done while at the Division of Mathematical Sciences, Nanyang Technological University, Singapore 637371.

[2] Work supported in part by the following grants: Nanyang Technological University grant M58110000, Ministry of Education – Singapore (MOE) Academic Research Fund (AcRF) Tier 2 grant MOE2010-T2-2-082, US NSF grant CCF-1023166, and grant 2008348 from the United States–Israel Binational Science Foundation (BSF).

[3] Work supported by the Nanyang Technological University grant M58110000.

http://dx.doi.org/10.1016/j.jcss.2014.10.005
0022-0000/

In fact, measurement studies of real-world P2P networks [33,40,64,65] show that the churn rate is quite high: nearly 50% of peers in real-world networks can be replaced within an hour. (However, despite a large churn rate, these studies also show that the total number of peers in the network is relatively *stable*.) We note that peer-to-peer algorithms have been proposed for a wide variety of computationally challenging tasks such as collaborative filtering [19], spam detection [25], data mining [28], worm detection and suppression [55,67], and privacy protection of archived data [38]. However, all algorithms proposed for these problems have no theoretical guarantees of being able to work in a network with a dynamically changing topology and a linear churn rate per round. This is a major bottleneck in implementation and wide-spread use of these algorithms.

In this paper, we take a step towards designing robust algorithms for large-scale dynamic peer-to-peer networks. In particular, we study the fundamental distributed agreement problem in P2P networks (the formal problem statement and model is given in Section 2). An efficient solution to the agreement problem can be used as a building block for robust and efficient solutions to other problems as mentioned above. However, the distributed agreement problem in P2P networks is challenging since the goal is to guarantee *almost-everywhere* agreement, i.e., almost all nodes[4] should reach consensus, even under high churn rate. The churn rate can be as much as linear *per time step (round)*, i.e., up to a constant fraction of the stable network size can be replaced per time step. Indeed, until recently, almost all the works known in the literature (see e.g., [32,44–46,66]) have addressed the almost-everywhere agreement problem only in static (bounded-degree) networks and these approaches do not work for dynamic networks with changing topology. Such approaches fail in dynamic networks where both nodes *and* edges can change by a large amount in *every* round. For example, the work of Upfal [66] showed how one can achieve almost-everywhere agreement under up to a *linear* number — up to εn, for a sufficiently small $\varepsilon > 0$ — of Byzantine faults in a bounded-degree expander network (n is the network size). The algorithm required $O(\log n)$ rounds and polynomial (in n) number of messages; however, the local computation required by each processor is exponential. Furthermore, the algorithm requires knowledge of the global topology, since at the start, nodes need to have this information "hardcoded". The work of King et al. [47] is important in the context of P2P networks, as it was the first to study scalable (polylogarithmic communication and number of rounds) algorithms for distributed agreement (and leader election) that are tolerant to Byzantine faults. However, as pointed out by the authors, their algorithm works only for static networks; similar to Upfal's algorithm, the nodes require hardcoded information on the network topology to begin with and thus the algorithm does not work when the topology changes. In fact, this work [47] raises the open question of whether one can design agreement protocols that can work in highly dynamic networks with a large churn rate.

1.1. Our main results

Our first contribution is a rigorous theoretical framework for the design and analysis of algorithms for highly dynamic distributed systems with churn. We briefly describe the key ingredients of our model here. (Our model is described in detail in Section 2.) Essentially, we model a P2P network as a bounded-degree expander graph whose topology — both nodes and edges — can change arbitrarily from round to round and is controlled by an adversary. However, we assume that the total number of nodes in the network is stable. The number of node changes *per round* is called the *churn rate* or *churn limit*. We consider a churn rate of up to some εn, where n is the stable network size. Note that our model is quite general in the sense that we only assume that the topology is an expander at every step; no other special properties are assumed. Indeed, expanders have been used extensively to model dynamic P2P networks[5] in which the expander property is preserved under insertions and deletions of nodes (e.g., [52,59]). Since we do not make assumptions on how the topology is preserved, our model is applicable to all such expander-based networks. (We note that various prior works on dynamic network models make similar assumptions on preservation of topological properties — such as connectivity, expansion etc. — at every step under dynamic *edge* insertions/deletions — cf. Section 1.3. The issue of how such properties are preserved are abstracted away from the model, which allows one to focus on the dynamism. Indeed, this abstraction has been a feature of most dynamic models e.g., see the survey of [20].)

We study stable, almost-everywhere, agreement in our model. By "almost-everywhere", we mean that almost all nodes, except possibly $\beta c(n)$ nodes (where $c(n)$ is the order of the churn and $\beta > 0$ is a suitably small constant — cf. Section 2) should reach agreement on a common value. (This agreed value must be the input value of some node.) By "stable" we mean that the agreed value is preserved subsequently after the agreement is reached.

Our main contribution is the design and analysis of randomized distributed algorithms that guarantee stable almost-everywhere agreement with high probability (i.e., with probability $1 - 1/n^{\gamma}$, for an arbitrary fixed constant $\gamma \geqslant 1$) even under high adversarial churn in a polylogarithmic number of rounds. Our algorithms also guarantee stability once agreement has been reached. In particular, we present the following results (the precise theorem statements are given in the respective sections below):

---

[4] In sparse, bounded-degree networks, an adversary can always isolate some number of non-faulty nodes, hence almost-everywhere is the best one can hope for in such networks [32].

[5] Expander graphs have been used extensively as candidates to solve the agreement and related problems in bounded degree graphs even in static settings (e.g., see [32,44–46,66]). Here we show that similar expansion properties are beneficial in the more challenging setting of dynamic networks.