Computer Networks 94 (2016) 307–317

Contents lists available at ScienceDirect

Computer Networks

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

Parameterized maximum and average degree approximation in topic-based publish-subscribe overlay network design☆

Melih Onus[a,*], Andréa W. Richa[b]

[a] Department of Computer Engineering, Ted University, Ankara, Turkey
[b] Computer Science and Engineering, School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, AZ 85287, Arizona, United States

ARTICLE INFO

Article history:
Received 15 January 2015
Revised 16 October 2015
Accepted 26 October 2015
Available online 10 November 2015

Keywords:
peer to peer
multicast
publish/subscribe
optimization problems
overlay networks

ABSTRACT

Publish/subscribe communication systems where nodes subscribe to many different topics of interest are becoming increasingly more common. Designing overlay networks that connect the nodes subscribed to each distinct topic is hence a fundamental problem in these systems. For scalability and efficiency, it is important to keep the degree of the nodes in the publish/subscribe system low. Ideally one would like to be able not only to keep the average degree of the nodes low, but also to ensure that all nodes have equally the same degree, giving rise to the following problem: Given a collection of nodes and their topic subscriptions, connect the nodes into a graph with low average and maximum degree such that for each topic t, the graph induced by the nodes interested in t is connected. We present the first polynomial time parameterized sublinear approximation algorithm for this problem.

We also propose a heuristic for constructing topic-connected networks with low average degree and diameter 2 and validate our results through simulations.

1. Introduction

In publish/subscribe (pub/sub) systems, publishers and subscribers interact in a decoupled fashion. They use logical channels for delivering messages according to the nodes subscription to the services of interest. Publishers publish their messages through logical channels that deliver the messages to the nodes that subscribed to the respective services.

Pub/sub systems can be either topic-based or content-based. In a *topic-based* pub/sub system, messages are published to “topics”, where each topic is uniquely associated with a logical channel. Subscribers in a topic-based system will receive all messages published to the topics to which they subscribe. The publisher is responsible for defining the classes of messages to which subscribers can subscribe. In a *content-based* system, messages are only delivered to a subscriber if the attributes of those messages match constraints defined by the subscriber; each logical channel is characterized by a subset of these attributes. The subscriber is responsible for classifying the messages.

Given their simplicity and wide applicability, we have seen many implementations of those systems in recent years (see e.g., [1–4,6–10,17,19,24,25,27–29]), as well as many applications built on top of them, such as stock-market monitoring engines, RSS [28] feeds [20], on-line gaming and several others. For a survey on pub/sub systems, see [16].

Publish/subscribe overlays with the topic connected overlay property are employed in a number of industrial solutions at Google, Yahoo, and Spotify [14,26]. PNUTS is a parallel and geographically distributed system for Yahoo!s web applications. PNUTS uses asynchronous replication to ensure

☆ A preliminary version containing some of the results in this paper appeared at IEEE ICDCSÆ10. This research was supported in part by NSF grants CCF-1116368 and CCF-0830704.

* Corresponding author. Tel.: +905302863985.
E-mail addresses: melih.onus@tedu.edu.tr (M. Onus), aricha@asu.edu (A.W. Richa).

http://dx.doi.org/10.1016/j.comnet.2015.10.023
1389-1286/

low latency updates. PNUTS uses Yahoo! message broker, a publish-subscribe system developed at Yahoo!, both as its replacement for a redo log and as its replication mechanism [14]. Spotify is a worldwide popular peer-assisted music streaming service. Spotify uses publish/subscribe system for matching, disseminating and persisting billions of publications everyday [26].

We will implement a topic-based pub/sub system by designing a connected (peer-to-peer) overlay network for each pub/sub topic: more specifically, for each topic t, we will enforce that the subgraph induced by the nodes interested in t will be connected. This translates into a *fully decentralized* topic-based pub/sub system since any given topic-based overlay network will be connected and thus nodes subscribed to a given topic do not need to rely on other nodes (agents) for forwarding their messages. Such an overlay network is said to be *topic-connected*.

Low node degrees are desirable in practice for scalability and also due to bandwidth constraints. Nodes with a high number of adjacent links will have to manage all these links (e.g., monitor the availability of its neighbors, incurring in heartbeats and keep-alive state costs, and connection state costs in TCP) and the traffic going through each of the links, without being able to take great advantage of aggregating the traffic (which would also reduce the number of packet headers, responsible for a significant portion of the traffic for small messages). See [12,21–23] for further motivation.

The node degrees and number of edges required by a topic-connected overlay network will be low if the node subscriptions are well-correlated. In this case, by connecting two nodes with many coincident topics, one can satisfy connectivity of many topics for those two nodes with just one edge. Several recent empirical studies suggest that correlated workloads are indeed common in practice [20,28].

In this paper, we focus on building overlay networks with low (average *and* maximum) node degrees. The importance of minimizing both the maximum and average degree has been recognized in some network domains, such as that of survivable network design [18] and that of establishing connectivity in wireless networks [15]. To the best of our knowledge, minimizing both the maximum and the average degree in topic-connected pub/sub overlay network design had not been directly addressed prior to this work.

As in many other systems, a space-time trade-off exists for pub-sub systems: On one hand, one would like the total time taken by a topic-based broadcast (which directly depends on the diameter of each topic-connected subnetwork) to be as small as possible; on the other hand, for memory and node bandwidth considerations, one would like to keep the total degree of a node small. Those two measures are often conflicting.

Most of the current solutions adopted in practice actually fail at maintaining *both* the diameter and the node degrees low. Take the naive, albeit popular, solution to topic connected-overlay network design to construct a cycle connecting all nodes interested in a topic independently for each given topic [29]: This construction results both in very large diameter for each topic-connected network (proportional to the total number of nodes subscribed to a topic) and in node degrees proportional to the nodes' subscription sizes, whereas a more careful construction, taking into account the

Table 1
Summary of known results on overlay network construction for publish/subscribe communication (n: number of nodes, t: number of topics, k is any parameter between 1 and n).

	Avg. degree	Max degree
Chockler et. al. [12]	$O(\log(n \cdot t))$	$\theta(n)$
Onus and Richa [21,23]	$\theta(n)$	$O(\log(n \cdot t))$
This paper	$O(k \cdot \log(n \cdot t))$	$O((n/k) \cdot \log(n \cdot t))$
Lower bound	$\Omega(\log n)$	$\Omega(\log n)$

correlations among the node subscription sets might result in much smaller node degrees (and total number of edges) and topic-based diameters. Even the more recent advances on approximating the average or maximum degree alone have been made [12,21,23] still fail at approximating the diameter well.

Whereas in the main contribution of our work (see Section 3), we completely neglect the diameter of the networks constructed, in Section 6, we propose a new heuristic for constructing topic-connected networks with low average degree and diameter 2. In fact, the results in Section 6 are a refinement of the preliminary results presented in [21–23]. Section 6 constitutes the major extension in this work over the results in [22].

1.1. Related work

The results in this paper are an improved and extended version of our results in [22]. The major extensions in this work are the following. We introduced an improved heuristic, 2-Diameter Overlay Design Algorithm(2D-ODA), and corresponding simulations(Section 6.3), which provide a much more thorough approach to the problem and improves previous heuristics. We proved previously presented heuristics may produce overlay networks of high average degree. The approximation ratio on the average degree obtained by the previously presented algorithms may be as bad as $\Theta(n)$ (Section 6.2). We introduced a new case and our intuition for the new heuristic and compared the heuristics using this case (Section 6.2).

Chockler et al. [12] introduced the MinAv-TCO problem [In the original paper, this problem was called Min-TCO.], which aims at minimizing the average degree alone of a topic-connected overlay network. They present an algorithm, called *GM*, which achieves a logarithmic approximation on the minimum average degree of the overlay network. The GM algorithm follows the greedy approach described below:

The Greedy-Merge (GM) algorithm [12]: The GM algorithm greedily adds the edge which maximally reduces the total number of topic-connected components at each step of the algorithm (initially we have the set of nodes V and no edges between the nodes).

While minimizing the average degree is a step forward towards improving the scalability and practicality of the pub/sub system, their algorithm may still produce overlay networks of very uneven node degrees where the maximum degree may be unnecessarily high. In [21,23], it is shown that GM algorithm may produce a network with maximum degree $|V|$ while a topic-connected overlay network of constant degree exists for the same configuration of I (See Table 1).