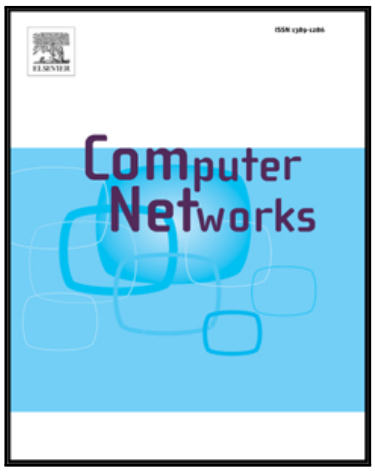

The Operational Cost Minimization in Distributed Clouds via Community-Aware User Data Placements of Social Networks

Qiufen Xia, Weifa Liang, Zichuan Xu

PII: S1389-1286(16)30388-7
DOI: 10.1016/j.comnet.2016.11.012
Reference: COMPNW 6054

To appear in: *Computer Networks*

Received date: 1 April 2016
Revised date: 14 September 2016
Accepted date: 13 November 2016

Please cite this article as: Qiufen Xia, Weifa Liang, Zichuan Xu, The Operational Cost Minimization in Distributed Clouds via Community-Aware User Data Placements of Social Networks, *Computer Networks* (2016), doi: 10.1016/j.comnet.2016.11.012

The Operational Cost Minimization in Distributed Clouds via Community-Aware User Data Placements of Social Networks

Qiufen Xia†, Weifa Liang†, and Zichuan Xu¶

† *Research School of Computer Science, Australian National University, Canberra, ACT 2601, Australia*
¶ *Department of Electronic and Electrical Engineering, University College London, London, UK*
Email: qiufen.xia@anu.edu.au, wliang@cs.anu.edu.au, z.xu@ucl.ac.uk

Abstract

With the increasing popularity of Online Social Networking (OSN) and public cloud platforms, cloud service providers such as Facebook and Google desperately need efficient placements of large-volume user data of social networks into their distributed clouds to enable the placed user data to be not only easily accessed and updated but also highly available, reliable and scalable, in order to minimize their operational costs of accommodating various social networks. In this paper, we investigate the problem of user data placements of social networks into a distributed cloud with the aim to minimize the operational cost of a cloud service provider, where the distributed cloud consists of multiple datacenters located at different geographical regions and interconnected by Internet links. We first devise a fast yet scalable algorithm for the user data placement problem. The key ingredient of this algorithm is the use of the community concept, by grouping users of a social network into different communities and placing the master replicas of user data of the users in the same community to a datacenter, and replicating their slave replicas of the user data into nearby datacenters. We then deal with the dynamic maintenance of the placed user data in an evolving social network, where new users can join in the network and existing users can leave from the network at any time, or existing users can change their read and update rates over time. We finally conduct extensive experiments to evaluate the efficiency of the proposed algorithms through simulations, using three real social network datasets: Facebook, Twitter and WikiVote. Experimental results demonstrate that the proposed algorithms significantly outperform state-of-the-arts in terms of the operational cost, yet run much faster.

Keywords: Community identification, distributed clouds, user data placements, optimization algorithms, online social networks, community maintenance

1. Introduction

Today's Online Social Networking (OSN) has many features to assist people socializing, allowing different scientific communities to expand their knowledge bases and helping individual researchers keep updated activities of peer colleagues. It is estimated that the number of worldwide OSN users will reach 2.5 billion by 2018, around one third of Earth's entire population [27]. Such an intense use of OSNs has generated huge