International Journal of Web Services Research
Volume 17 • Issue 1 • January-March 2020

A Decentralized PageRank Based Content Dissemination Model at the Edge of Network

Xin Zhang, Institute of Acoustics, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Huairou, China
Jiali You, Institute of Acoustics, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Huairou, China
Hanxing Xue, Institute of Acoustics, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Huairou, China
Jinlin Wang, Institute of Acoustics, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Huairou, China

ABSTRACT

In the era of Internet of Things, cloud services are difficult to meet the real-time transmission requirements of users for the data generated in the edge of network especially for the Internet video services. Utilizing the devices at the edge of network, such as an intelligent router, to achieve nearby content services for users can effectively reduce backbone traffic and enhance service performance. This article proposes a decentralized PageRank-based content dissemination model at the edge of network, in which a suitable node selection algorithm is designed to distribute the content evenly in the network. Each node can quickly obtain data from neighbor nodes, thereby reducing the cloud load as well as the network bandwidth and improving the service response performance. The simulation shows that, compared with the other two dissemination algorithms, the content is distributed more even, which means every node has more opportunity to obtain the data from neighbors; and the service rejection rate can be decreased by an average of 5.2% in the case of high concurrent requests.

KEYWORDS

Content Dissemination, Edge Caching, Edge Services, Overlay Neighborhood, PageRank

INTRODUCTION

With the development of the Internet, mobile Internet and the Internet of Things (Xu, He, & Li, 2014), terminal devices with certain processing and data transmission capabilities are expanding on the network. The user-centric content dissemination (Delmastro, Arnaboldi, & Conti, 2016) has become the trend of the network information dissemination. With such a rapid growth in user resources, it remains a great challenge how to organize and manage user resources to provide elastic services. The linear growth of centralized cloud computing capabilities cannot match the explosive growth of the massive edge data. The transmission of the massive data from the cloud center to the network edge

DOI: 10.4018/IJWSR.2020010101

device causes the drastically increase of the load of network bandwidth, which may result in high network latency (Chang, Hari, Mukherjee, & Lakshman, 2014).

The total resources of terminal devices at the edge of network are enormous, such as the computing and storage resources owned by devices including but not limited to smartphones, set-top boxes, routers, and PCs, especially IoT terminals such as a variety of smart sensors. But these resources are still idle and are not properly utilized (Ma & Feng, 2015). In contrast, the total amount of resources deployed by service providers is still relatively small, which is one of the typical causes for network bottlenecks. Therefore, in order to improve the quality of content services, content is usually placed at the edge of the network which is close to users, such as CDN. In recent years, with the enhancement of edge device capabilities, it is hoped that more and more content will be placed on the wider edge including user devices and operator devices, such as set-top boxes, APs. Nevertheless, the traditional distribution methods are mostly based on centralized management. As the scale of edge devices growing larger, it will have a negative impact on the dissemination efficiency. Therefore, how to disseminate content at the edge of the network in a decentralized manner is an important issue.

The edge network (Shi, Cao, Zhang, Li, & Xu, 2016) has two main features. Firstly, it is highly heterogeneous. The hardware, operating system, network, management domain and security policies for terminal devices are all different. Secondly, the capabilities of various terminals, such as resource sharing and coordinating with other terminals, are quite different. Terminals with strong service capabilities, high impact and near-user characteristics can be considered as potential nodes providing extra services.

For some applications, such as video service and information distribution in the sensor network, the critical problem is to identify the most potential nodes to cache content by means of approximation of global knowledge in a decentralized manner. In this paper, our main contribution is to create a mechanism for automatic content dissemination in a distributed manner. The core algorithm in this mechanism is a PageRank-based node selection algorithm, which can disseminate the content to the properly selected nodes in the network. Each node can quickly obtain data from its neighbors or surrounding nodes, which can reduce the cloud load as well as the network bandwidth and improve the service response performance. This algorithm is executed in each node, by which each neighbor node is assigned a score. Replicas for content are cached at neighbor nodes according to their ranks until all nodes reach the default content coverage requirement among neighbors or no more quotas are available. The users can quickly get data from its neighbors or surrounding nodes without having to cache them. Therefore, the service performance is the ultimate goal concerned about. Finally, the proposed schemes are evaluated in consideration of different configurations at the edge network scenarios.

RELATED WORKS

Although the content caching has been extensively investigated in the past, the rise of IoT and edge computing has renewed interest in this topic (Anjum, Karamshuk, Shikh-Bahaei, & Sastry, 2017). Here is a brief overview of existing information diffusion models and content distribution solutions.

Diffusion Models

Diffusion models were originally used in social networks to model the spread of influence in a network. The Susceptible-Infected-Recovered (SIR) model is the most classic information diffusion model. In these models each node is either active or inactive. Over iterations an inactive node becomes active as more of its neighbors become active. The two most popular diffusion models are the linear threshold model and the independent cascade model.

In the independent cascade (IC) model, each edge has an activation probability and the influence is propagated by activated nodes independently activating their inactive neighbors based on the edge activation probabilities. In the linear threshold (LT) model, each edge has a weight, each vertex has

Related Content