Available online at www.sciencedirect.com
SciVerse ScienceDirect
Procedia Engineering 29 (2012) 3727 – 3731
Procedia Engineering
www.elsevier.com/locate/procedia

2012 International Workshop on Information and Electronics Engineering (IWIEE)

Gossip-Based Indexing Ring Topology for 2-Dimension Spatial Data in Overlay Networks

Z. Q. Zou[a,b,c], Y. Wang[a], B. Hu[a]*, R. C. Wang [a,b,c],

[a]*College of Computer, Nanjing University of Posts and Telecommunications,Nanjing, Jiangsu 210003, China*
[b]*Jiangsu High Technology Research Key Laboratory for Wireless Sensor Networks, Nanjing, Jiangsu 210003, China*
[c]*Key Lab of Broadband Wireless Communication and Sensor Network Technology (Nanjing University of Posts and Telecommunications),Ministry of Education Jiangsu Province, Nanjing, Jiangsu 210003, China*
[d]*Key Laboratory of Virtual Geographic Environment; Ministry of Education, Nanjing Normal University, Nanjing 210046, China*

Abstract

Overlay networks are used widely in the Internet, such as retrieval and share of files, multimedia games and so on. However, in distributed system, the retrieval and share of 2-dimension spatial data still have some difficult problems and can not solve the complex retrieval of 2-dimension spatial data efficiently. This article presents a new indexing overlay networks, named 2D-Ring, which is the ring topology based on gossip for 2-dimension spatial data. The peers in our overlay networks exchange the information periodically and update each local view by constructing algorithm. 2-dimension spatial data is divided by quad-tree and mapped into control points, which are hashed into 2D-Ring by SHA-1 hash function. In such way, the problem of 2-dimension spatial data indexing is converted to the problem of searching peers in the 2D-Ring. A large of extensive experiments show that the time complexity of constructing algorithm of 2D-Ring can reach convergence logarithmically as a function of the network size and hold higher hit rate and lower query delay.

Keywords: Overlay networks; 2-Dimension Spatial Data; Indexing; Gossip

1. Introduction

There are many successful applications based on overlay networks, such as files sharing, streaming multimedia, network games, and so on [1],[2],[7]. However, there exist many challenges in network process of spatial information, which is multi-dimension data, such as 2-dimension spatial data.

* Corresponding author,B. Hu. Tel.: +0-86-13813878460; fax: +0-86-25-85866151.
E-mail address: hb_hubin@126.com.

doi:10.1016/j.proeng.2012.01.561

Furthermore, the corresponding process task is different and difficult, especially in requiring complex spatial region range data with the massive amount of concurrent users [3],[4]. There are some distinguishing characteristics of 2-dimension spatial data under internet. First, it needs a kind of mapping function to map 2-dimension spatial data into one dimension. Second, this application maybe undergoes a large amount of users to query concurrently. In order to meet this demand, we propose a new indexing overlay networks that is the ring topology based on gossip for 2-dimension spatial data.

Related work with ours includes two main aspects. One is about the process to spatial data under distributed environment. The other one is concerned with the topology of overlay networks. First, as for one dimension data under overlay network, such as T-Chord [5] and Chord [9], we can directly get the required data with its unique ID by one hash function. However, the process method of 2-dimension spatial data is different from the before since the user's query often deals with the complex retrieval to region range, which maybe includes multi-ID of multi-sub-region [3]. One way is to use the distributed quad-tree (DQT) to express 2-dimension spatial data [3],[4],[8]. Second, as for the overlay network, there are some worthy of studies [1],[2],[5]. Jinu K. presents a survey of recent advancements in application-layer overlay networks, such as overlay topology design and enhanced routing performance [1]. Mawji A. proposes a computationally feasible game-theoretic heuristic algorithm on mobile ad hoc networks [2]. The classical theories based on gossip spread protocol, as described in [5] are closest to our work, which are often used to build and maintain the structure of large-scale overlay network for reliable transmission of data.

Different existing work, our main contribution with respect to related work is three-fold. First, we introduce a new algorithm based on gossip, which can construct a spatial data indexing overlay networks, named 2D-Ring. Second, we design a large-scale spatial information indexing method. Third, we made a large of extensive experiments to evaluate the performance of 2D-Ring on various aspects, including convergence, rate of indexing hit, and so on.

The rest of paper is organized as follows: section 2 introduces our theoretical model of 2D-Ring overlay network and its algorithm. Section 3 presents large-scale spatial information indexing method under 2D-Ring. In section 4, we will present theoretical and experimental results to characterize key properties of this 2D-Ring overlay network and performance analysis of our algorithm. The last section presents conclusion and future work.

2. Theoretical model of 2D-Ring overlay network

Theoretical model of 2D-Ring overlay network includes two main part: one is the Gossip-Based Communication algorithm and the other one is the rank(i) function. The communication algorithm is the fundamental of the information exchange in peers while the rank(i) function is the core of the algorithm.

2.1. Gossip-Based Communication algorithm

The gossip-based communication algorithm is extended based on the newscast model of computation [5] by using the value of metric in next Equation (1) as the standard of distance in rank(i) function. In this algorithm, all peers periodically exchange their descriptors with other peers, thereby constantly improving the set of peers they know — their partial views. The basic idea of the algorithm is as follows 7 steps:
(1) initial local views in each peer;(2) wait a constant time units periodically;(3) each peer selects a random sample peer X in list;(4) update each local view by adding the view of peer X;(5) send the updated view to peer X;(6) receive the view from peer X;(7) goto step (2) until obtain the perfect target.

Next is the explanation for some key steps in algorithm:

- step (5) and step (6) are to be run in thread way, which supports massive peers concurrency;

- the list in step (3) is the ordered peer list, which is ordered by the value of metric in Equation (1). Each local view includes multiple waited-select peers.

2.2. Introduction of rank(i) Function

The list in step (3) of the above algorithm needs to be ordered. The order uses rank(i) function instead of simple method based on Euclidean distance. In 2D spatial indexing, as we will show in next section, two region's distance is not computed by the distance between two points in one dimension. The rank(i) function can sort a set of nodes (potential neighbors) according to the ''taste" of a given base node. Here "taste" means semantic similarity of two control points, which is corresponding to their 2D spatial data.

$$\mathrm{Sim}(CPx, CPy) = \underset{T \in Set(CPx,CPy)}{Max}[-\log(P(T))] = P(-\log(NSCP(CPx, CPy))) = -\log(CPNum / N) \quad (1)$$

In Equation (1), *CPx* and *CPy* present the control points of corresponding to their 2D spatial data. In this section, we present a concise explanation for Equation (1) (For details of computing procedure, please refer to [4, 8]). The basic idea of Equation (1) is that the semantic similarity is bigger while the distance is smaller.

Next we focus on the construction of 2D-Ring overlay network that supports the efficient spatial data indexing under distributed environment. Based on the Equation (1), according to the ''taste" of peer i as a based peer, an arbitrary peer i can sort the peers in the view of itself. And then, each peer builds its finger table and leaves from the ordered list. Allowing for the characteristic of ring, we adopt the following Equation (2) to compute the semantic distance. Each peer adopts Equation (2) as ranking function and each peer in 2D-Ring ranks its own view according to it. And then, the sort ascending list can be gotten as its neighborhood. After hash mapping, each peer has a unique ID in the 2D-Ring.

$$dist\_sim(i, j) = \min(N - |i - j|, |i - j|) \quad (2)$$

3. Large-scale spatial information indexing method

One of the main different features of 2D spatial data from the one dimension is that the region range data can not be divided by linear method. In order to indexing these 2D spatial data, it needs to map the two dimensions into one dimension. So we introduce the DQT to divide the 2D spatial data recursively based on our previous work [4].

3.1. 2D spatial data indexing method

Here we use DQT to divide the 2D spatial data and map it into control points. Firstly, the 2D spatial data is abstracted into BBox (Bounding Box) and the across point of BBox is acted as control point. Then we divided it recursively. And then, we encode this DQT and let each node has unique ID. By using SHA-1 hash function, this ID can be mapped into the 2D-Ring. So far, the problem of 2-dimension spatial data indexing is converted to the problem of searching peers in the 2D-Ring.

3.2. Routing table of 2D-Ring

We assume that N in Equation (2) is 2^t and t is the bit number of ID. When it meets Equation (3), we define that peer i follows peer j, otherwise peer i precedes peer j. There are two key parts in the peer's routing table as follows: children and fingers in 2D-Ring. The children part is defined as a short hop to four children in the DQT while the fingers part is defined as a long hop to other peer in the DQT.

$$(i - j + 2^t) \bmod 2^t < 2^{t-1} \quad (3)$$

4. Performance Evaluation

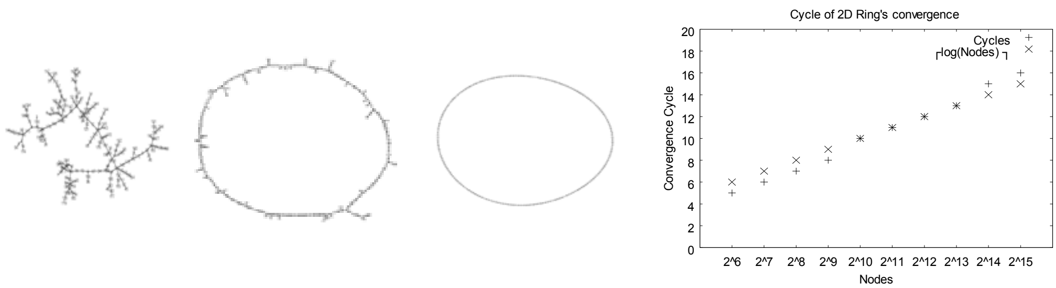
This section we focus on evaluating the key properties. All the experiments are performed using PEERSIM [6], which extended with the class: QuadTree, TWireStrategy and Chordidcomparator.

4.1. Experiment of 2D-Ring's convergence

In 2D-Ring's initialization, all peers are connected randomly. After running the algorithm in the section 1, it will converge into a ring as shown in Fig.1(a)(b)(c). The relationship between the cycle of convergence and the size of network N is presented in Fig.1(d), which suggests a function O(log N).

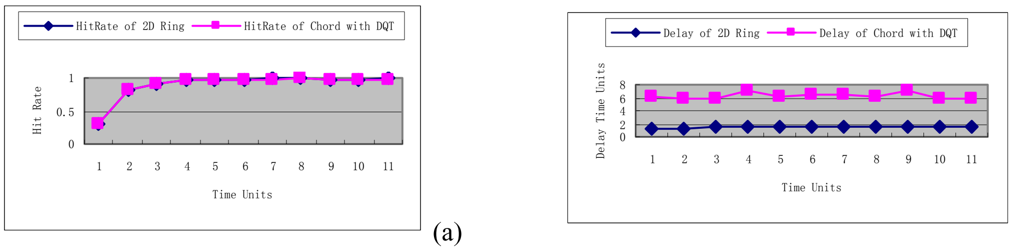
4.2. Experiment of 2D-Ring's indexing performance

Next experiment is to evaluate the indexing performance under this overlay networks from two aspects: hit rate of indexing and query delay under different topology. And then, we compare 2D-Ring with Chord with DQT [3]. As Fig.2 (a), and Fig.2 (b) shown, 2D-Ring topology is similar to Chord with DQT at the aspect of Hit Rate while superiority to it at the aspect of query delaying.



(a) step-a (b) step-b (c) step-c (d) convergence relationship

Fig.1. Converging procedure (a)step-a;(b)step-b;(c)step-c;(d) convergence cycle as a function of Nodes' number.



(a) (b)

Fig.2. (a) Hit Rate of 2D-Ring and Chord with DQT; (b) Delay of 2D-Ring and Chord with DQT

5. Conclusions

In this paper, a gossip based overlay network 2D-Ring is constructed. And 2D spatial data can be indexed in 2D-Ring efficiently. Simulation experiments show that 2D-Ring has desirable properties. In the future research, we will focus on reducing the burden of indexing on 2D-Ring further.

Acknowledgements

This work is supported by NSFC (No.40801149, No.60903181, and No.61003040);China Planned Projects for Postdoctoral Research Funds (20100471356);Open Research Fund from Key Laboratory of Computer Network and Information Integration in Southeast University; Ministry of Education, China, (No. K93-9-2010-06), and a Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (yx002001).

References

[1] Jinu K, Kamil S. A survey on the design,applications,and enhancements of application-layer overlay networks.*ACM Computing Surveys (CSUR)*; 2010, 43(1), 5.1-5.34.

[2] Mawji A, Hassanein H, Zhang XY.Peer-to-peer overlay topology control for mobile ad hoc networks.*Pervasive and Mobile Computing*; 2011, 7(4), 467-478.

[3] Egemen Tanin, Aaron Harwood, Hanan Samet. Using a distributed quadtree index in peer-to-peer networks. *The VLDB Journal;* 2007, 16(2),165–178.

[4] Zou ZQ, Jiang N, Wu JG, Hu B. A P2P Spatial Data Index Mechanism based on Semantic Clustering Distributed Quad-Tree.*Geomatics and Information Science of Wuhan University*; 2011,36(1):76-81.

[5] Jelasity M, Montresor A, Babaoglu O. T-MAN: Gossip-based fast overlay topology construction.*ACM Transactions on Computer Networks*; 2009, 53:2321–2339.

[6] PEERSIM. *http://peersim.sourceforge.net/.*

[7] Hu SY, Huang TH, Chang SC, et al. FLoD- A Framework for Peer-to-Peer 3D Streaming. *INFOCOM 2008, The 27th Conference on Computer Communications IEEE, Phoenix,AZ,USA*; 2008:1373-1381.

[8] Wu J., Jiang N, Zou ZQ, etc. HPSIN: A New Hybrid P2P Spatial Indexing Network. *The Journal of China Universities of Posts and Telecommunications*; 2010,17(3):66-72.

[9] Ratnasamy S, Francis P, Handley M, Karp R, Shenker S.:A scalable content-addressable network. In: Proceedings of the ACM SIGCOMM'01, San Diego, CA;2001:161–172.