A configuration method for structured P2P overlay network considering delay variations

Tomoya Kitani
Division of Global Research Leaders
Shizuoka University
Hamamatsu, Shizuoka, JAPAN
t-kitani@ieee.org

Yoshitaka Nakamura
Graduate School of Information Science
Nara Institute of Science and Technology
Ikoma, Nara, JAPAN
y-nakamr@is.naist.jp

***Abstract*—P2P networks can achieve high scalability since they distribute service contents/resources to multiple nodes in the network. In a P2P network, it is necessary to search the resource location on the network when we use some contents/resources. Space filling curve is known as technique to map information of the multi-dimensional space such as the location information onto the one-dimensional space such as ID. In this paper, we propose a novel space filling curve for configuring structured P2P overlay network considering delay based on the geographic information of each node. By using the proposed space filling curve, we can convert geographic information of nodes into their ID (label) of P2P network. Through the numerical evaluation, we confirmed that the proposed curve is more suitable for handling hierarchical-spread nodes than the conventional curves.**

***Keywords*-P2P; Overlay Network; Space Filling Curve; Delay Variation; Node Labeling**

I. Introduction

In recent years, with the development of the communication technology, mobile terminals such as cellular telephones or the note PCs come to be connected frequently to the network with wireless communication. As a result, the user of mobile terminals become possible to receive e-mails or the multimedia contents such as the image, and movie, etc. without depending at the specific location and time. Moreover, a huge amount of terminals such as sensors deployed in the living space or terminals installed by cars are expected to compose the network in coming ubiquitous computing society.

It is difficult to keep a scalability to offer high quality service for such large amount of terminals because the load to the server concentrates in server-client type system mainly used in current computer network. Therefore, the Peer-to-Peer (P2P) network system attracts attention. In the P2P network system, each terminal manages data autonomously and distributedly, and communicates directly with other terminals. Gnutella[1] and BitTorrent[2] are typical P2P network technologies. P2P network can provide network services to millions of terminals by distributing the load of the terminals and network unlike the server-client type architecture.

On the other hand, in the ubiquitous computing society, it becomes available a location-aware service that selects appropriate services by using user's geographic information, and a contents-aware service that selects appropriate services considering of the state of sensors in the neighborhood. Current cellular phone terminals and many of in-vehicle navigation systems are equipped with GPS (Global Positioning System) and can easily acquire the geographical location information of each user. LL-net[3] is one of the configuration method of the P2P overlay network using geographic information. LL-net is a structured P2P overlay network where the overlay link is configured by hierarchically dividing the area on the map. However, there is some problem that it is necessary to assume the existence of special nodes for the area management or that the search efficiency turns worse according to the node distribution.

Space filling curve is known as technique to map information of a multi-dimensional space such as location information onto the one-dimensional space such as ID. There is a trade-off between the inefficiency of space filling in simple method such as Z-Ordering (Lebesgue curve)[4] and high calculation cost of a curve with good filling efficiency.

In this paper, we propose a novel space filling curve for configuring structured P2P overlay network considering the delay variation based on the geographic information of each node. The proposed method gives each node a label (ID) from the space coordinate of the node based on its geographic information and the link delay between the nodes. And the proposed curve can let the space coordinate be converted into the one-dimensional space efficiently.

II. Related Work

In Peer-to-Peer(P2P) technology, each peer constructs application layer overlay network, searches contents or resources over this network by direct communication between peers without a server, and distributes and shares the resources. This application layer network (overlay network) is composed independently of the physical layer network, and classified from its structure into the structured overlay networks and the non-structured overlay networks. Earlier P2P technology used the non-structured overlay networks. Gnutella[1] and BitTorrent[2] are typical technologies using the non-structured overlay networks. On the other hand, the structured overlay networks can search a resource on the P2P networks, and many technologies are proposed in recent years. Especially, there are many technologies that use Distributed Hash Table(DHT). Chord[5], CAN (content addressable network)[6], Pastry[7], and Tapestry[8] are known well. In the conventional non-structured P2P overlay networks, search of information uses flooding queries to peers. On the other hand, the search success rate can be achieved almost 100% in the structured P2P overlay networks, because it is possible to search efficiently with the structure of the networks.

About the configuration of the structured overlay networks, some researches try to improve the efficiency of the overlay networks with the state of the lower physical links. Reference[9] proposes the framework that separates an overlay network from failures in its physical layer by using a measurement based on the heuristic method for an arrangement of the overlay nodes in an ISP and the selection of ISPs. Laptop[10] improves efficiency, the scalability and the robustness of the overlay network by using the cache of the routing information and maintaining only the connection between each node and the parent of the node. LTM[11] constructs an efficient overlay network with cutting an inefficient connection, selecting nodes that exist physically at near positions as the logical neighbor nodes based on IP address. Mithos[12] achieves to connect with neighbor nodes efficiently by using the shortest path information with DHT, and to minimize the overhead of the message forwarding. In [13], it is described that the update of an overlay routing table is made more efficient in measuring the neighborhood of the node by using *Network Coordinates*(NC). DDAPS[14] uses delay between end hosts for an evaluation criterion of pruning the flooding to improve the routing efficiency in the overlay network.

On the other hand, in order to enable the range search which treats the consecutive quantity, some techniques which do not use DHT for decentralization of data and search queries have been proposed. SkipNet[15] is one method for configuring a structured P2P overlay network. In this method, the range search is enabled by using SkipGraph[16] that uses consecutive values for ID instead of DHT that uses a hash function. SkipGraph achieves the search efficiency of the logarithmic order by hierarchically grouping one-dimensional node array using the balanced tree structure. However, the range specification search in SkipNet might need to search wider range than other methods, because SkipNet constructs a hierarchical overlay network by using single search key, without regard to the geographical location and multiple search keys. LL-net[3] is structured P2P overlay network where the area on the map is hierarchically divided into four sub areas, and in each hierarchy the overlay links should be the different length links. In LL-net, the scalability of system is lacked so that the existence of a special node that manages each area may be assumed, and the management cost of the overlay network may increase. Moreover, if the distribution of the node is not uniform, it might become impossible to converse from geographic coordinates into node ID easily, the hierarchical structure might become biased, and the search efficiency might turn worse. In Mill[17], the range search is enabled by connecting nodes with the ring structure by applying ID which based on geographical location information. Hereby, the search efficiency can be realized $O(logN)$ without special nodes.

Vivaldi[18] is a technique to calculate the coordinate system autonomously by considering both the proximity of the physical network and real communication delay on a P2P network. This technique leads virtual coordinate system by gradually correcting the difference between Euclidean distance and the actual measurement delay between nodes using the spring model after each participation node decides its virtual coordinates autonomously.

III. Proposed Method

In this paper, we research a new configuration method for structured P2P overlay network considering delay variations between nodes. We propose a new space filling curve to convert coordinate (geographic) information of nodes in the two-dimensional space into a one-dimensional ID sequence.

The search efficiency can be improved when given an ID which has small distance from IDs of the physically neighbor nodes. Therefore, it can be good for labeling a node ID to use geographical information of nodes. Space filling curve[19] (Peano curve) is known as a technique to convert information on multi-dimensional space like location information into information on one-dimensional space like ID. We propose a new space filling curve with following features that (1) it is easily convertible on node ID on the P2P network from the geographic information, (2) information search that specifies the range is efficiently executable, and (3) it is easy to construct a hierarchical structure to give the scalability.

A. Conventional Space Filling Curves

Some space filling curves have been proposed so far.

The typical space filling curves are Lebesgue curve (Z-Ordering, Fig. 1), Hilbert curve (Fig. 2), Sierpinski curve (H-indexing), and $\beta\Omega$-indexing, etc. Hilbert curve, Sierpinski curve, and $\beta\Omega$-indexing fill the space so that the neighbor nodes on the two-dimensional space is converted to the close ID on the one-dimensional space. In these curves, the communication delay between logical neighbor nodes is suppressed small, and the range search can be searched efficiently. The performance of the Lebesgue curve is not better than that of the above curves, because the terminal

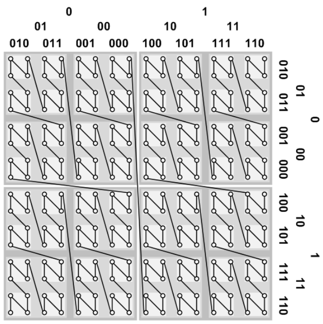

Figure 1. Lebesgue Curve (Z-Ordering)

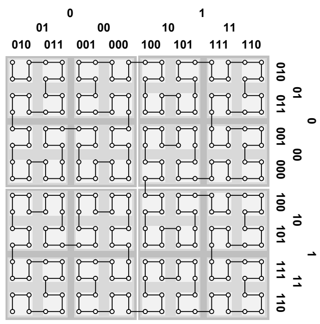

Figure 2. Hilbert Curve

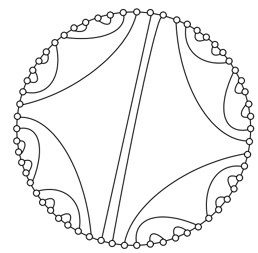

Figure 3. Hierarchical Chordal Ring Network ($N = 62$)

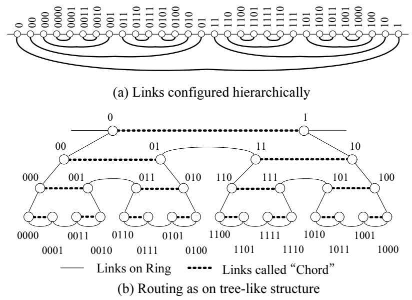

Figure 4. A cluster of HCRN

nodes of the link that connects between clusters are a long way each other on the two-dimensional space (see Fig. 1). However, Lebesgue has simple structure and two dimensional coordinates can be easily converted into the corresponding IDs on it. Therefore, Lebesgue curve is often used for ID conversion in P2P networks, especially for mobile ones where the computing power is limited and it is required to reduce overhead of complex processing. Our space filling curve aims to be able to convert ID as easily as the Lebesgue curve and to give an ID the correlation with the geographical proximity of the node which has the next ID such as Hilbert curve.

B. Consideration on the Curves and Proposed Curve

The conventional space filling curves such as Hilbert curve and $\beta\Omega$-indexing have self-similarity because they are constructed by hierarchical processes. But they are not considered to design to hierarchically spread nodes. In the curves, it can be possible to change the hierarchical level with respect to each sub area, for example there are 2^4 nodes to be traveled (the hierarchical level is 2) in an sub area and there are 2^6 nodes to be traveled (the hierarchical level is 3) in another sub area. However, in the sub areas, all nodes should be in the same hierarchical level and they should be relabeled when the number of nodes increased. It is difficult for the labeling scheme of the curves to be assigned hierarchically-spread nodes efficiently. In order to achieve scalability for information storage and retrieval on P2P networks based on the curves, service providers need to adopt a hierarchical structure in an upper service layer. The curve to be proposed can be flexibly assigned and label hierarchically-spread nodes in each area. It can travel all nodes on the curve easily. The curve can handle hierarchical-spread nodes and achieve scalability at the labeling scheme.

Only Lebesgue curve could be assigned such nodes by enhancing the curve to assign hierarchical nodes at the middle point of each edge which goes across an area vertically. In the evaluation section, we compare our curve with the enhanced Lebesgue curve.

C. The Constitution of the Space Filling Curve

The proposed space filling curve is composed based on Hierarchical Chordal Ring Network (HCRN)[20] shown in Fig. 3 by enhancing HCRN to two dimensions.

HCRN constructs the tree structure topology on the ring by connecting the link on the node on a cluster edge after clustering by recursive division of node clusters on the ring (Fig. 4). Here, each node's ID is expressed by a variable-length gray code according to the hierarchy where the node is located, and each node queues up in the order of the gray code in the hierarchy. The binary number and the gray code can be easily converted by using the following exclusive-OR operation and the simple algorithm. If a certain n digit binary number is assumed to be $\mathbf{b}$ and the gray code corresponding to it is assumed to be $\mathbf{g}$, they are possible to be converted to each other.

Algorithm Binary2Gray

$$\mathbf{g} = \mathbf{b} \oplus (\mathbf{b} >> 1)$$

Algorithm Gray2Binary

Given a gray code $\mathbf{g} = (g_1, g_2, \cdots, g_n)$
$flag_{rev} \leftarrow$ false
for $i \leftarrow 1$ **to** n **step** 1
  if $g_i =$ '1' **then**
    if $flag_{rev} =$ false
    then $b_i \leftarrow$ '1' **else** $b_i \leftarrow$ '0' **endif**
    $flag_{rev} \leftarrow$ **not** $flag_{rev}$
  else
    if $rev_f lag =$ false
    then $b_i \leftarrow$ '0' **else** $b_i \leftarrow$ '1' **endif**
endfor
return the binary code $\mathbf{b} = (b_1, b_2, \cdots, b_n)$

In Lebesgue curve, given a geographical position $(\mathbf{x_i}, \mathbf{y_i})$ of a node n_i $(\mathbf{x_i} = (x_1, x_2, \cdots, x_H), \mathbf{y_i} = (y_1, y_2, \cdots, y_H))$, the node's geographical label $\mathbf{p_i}$ is given as $\mathbf{p_i} = (x_1, y_1, x_2, y_2, \cdots, x_H, y_H)$. Here, H denotes both the length of geographical information and the maximum number of hierarchical level.

In our curve, first, in order to let the curve be a closed curve, we define that a new geographical information of a node n_i is given as $\mathbf{p_i} =$

Binary2Gray$(x_1, y_1, \overline{x_2}, \overline{y_2}, \cdots, x_{2j-1}, y_{2j-1}, \overline{x_{2j}}, \overline{y_{2j}}, \cdots)$. Here, $\overline{x}$ means bit inversion and $j \in \mathbb{N}$. This definition provides a result that the nearer two nodes is located, the smaller the Hamming distance between their node labels is.

Next, to handle hierarchically-spread nodes, we define the following algorithm to label ID to each node. Each node is labeled by this scheme in the chronological order of the nodes.

Given

$\mathbf{p_i} = f(\mathbf{x_i}, \mathbf{y_i})$: The geographical label of each node n_i
H : Maximum hierarchical level

Procedure

```
Given p_i = (p_1, p_2, p_3, ..., p_2H)
for j <- 2 to 2H step 2
    if Label (p_1, ..., p_j) has not been used yet
    then Label l_i <- (p_1, ..., p_j) and goto 7.
    endif
endfor
end.
```

The enhanced Lebesgue curve is obtained by connecting nodes in the numerical order of the label $\mathbf{l_i}$. Note, however, suppose $(l_1, l_2, \cdots, l_m, 0, 1) < (l_1, l_2, \cdots, l_m) < (l_1, l_2, \cdots, l_m, 1, 0)$ when the length of the labels are diffrent.

The proposed curve is obtained by connecting nodes in the numerical order of Gray2Binary($\mathbf{l_i}$). Note, however, suppose $(l_1, \cdots, l_{m-1}, 0) < (l_1, \cdot, l_{m-1}, 0, l_{m+1}, \cdots, l_n)$ and $(l_1, \cdots, l_{m-1}, 1, l_{m+1}, \cdots, l_n) < (l_1, l_2, \cdots, l_{m-1}, 1)$ when the length of the labels are different.

An example of the proposed curve is shown in Fig. 5. Figure 6 shows the lowest level nodes and their connections out of the curve in order to let it be easier to compare with the conventional space filling curves such as Hilbert curve.

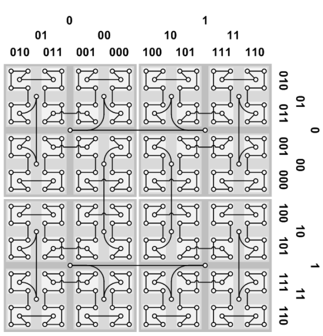

Figure 5. Proposed curve

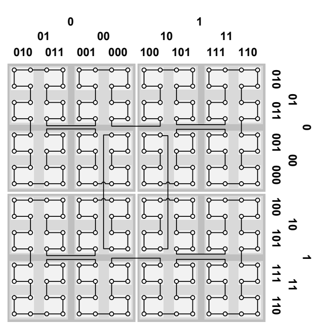

Figure 6. The links of the lowest level in the proposed curve

IV. Numerical Evaluation

The advantage of the proposed method is evaluated by the following points.

- The index that shows the physical neighbor node is also near on space filling curve

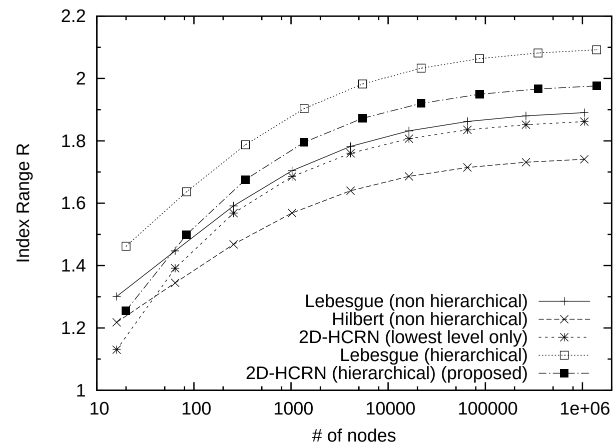

Figure 7. Average of logarithmic index range R_{avg}^{curve}

- The communication delay which necessary for travelling all over the filling curve

First, we compared the proposed curve with Lebesgue curve and Hilbert curve form the standpoint of the proximity index. Next, we compared it with the enhanced Lebesgue curve. Nodes on the proposed curve and the enhanced Lebesgue curve can belong to various hierarchical levels as shown in Fig. 5 although nodes on Hilbert curve belong to the same hierarchical level.

A. Proximity with Neighbor Nodes on the Curves

First of all, we evaluated how far the nodes which is physically in four neighborhoods which are on the filling curves, using logarithmic index range R_{avg}^{curve} shown in Ref.[19]. Suppose that each node n_i has a sequence number $seq(i)$ on the curve and it is located on coordinate $pos(i) = (x_i, y_i)$. Proximity index R_{avg}^{curve} can be calculated by the following expressions.

$$\begin{aligned}
r_1^{curve}(i) &= \log(|seq(i) - seq(pos^{-1}(x-1, y))| + 1) \\
r_2^{curve}(i) &= \log(|seq(i) - seq(pos^{-1}(x+1, y))| + 1) \\
r_3^{curve}(i) &= \log(|seq(i) - seq(pos^{-1}(x, y-1))| + 1) \\
r_4^{curve}(i) &= \log(|seq(i) - seq(pos^{-1}(x, y+1))| + 1) \\
R_{avg}^{curve} &= \text{avg}_i\{r_j^{curve}(i)\}, \quad (j \in \{1, 2, 3, 4\})
\end{aligned}$$

The evaluation results are shown in Fig. 7.

A physically neighbor node can be assigned to nearer seqence number on the proposed curve comparing with the Lebesgue curve. It can be a remarkable improvement, especially considering that R_{avg}^{curve} is the log scale. Hilbert curve is one of the most efficient filling curves and it shows better result. However, it is difficult to use Hilbert curve when the nodes location changes dynamically, because it is difficult for the Hilbert curve to construct hierarchical structure for the geographical uneven distribution of the node efficiently.

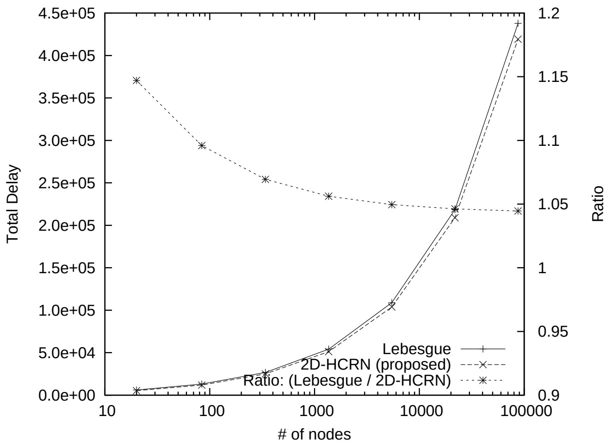

Figure 8. Overall delay on the curves with regularly distributed nodes

B. *Length of Curves*

Next, we evaluated the communication delay to travel all nodes on the filling curves. We assume that each delay between any two nodes are proportional to the distance between the nodes.

Figure 8 shows the total delay for travelling all nodes on the curves when the number of nodes are decided so that the nodes can fill all grid positions like Figs. 1, 2, and 5. As shown in the figure, the proposed method is able to reduce total delay by at least about 5% and up to 15%.

Next, we evaluated the delay for various numbers of nodes with computer simulations. In the simulations, we assume that the number of nodes are changed from 50 to 10,000, and each result is the average of 30 trials.

Figure 9 shows the result when the position of the node is generated according to the uniform distribution, and Fig. 10 shows the result when the geographical position of the node is generated according to Zipf Law (also known as Power Law). From each result, the performance of the proposed curve has improved about 3-6% more than Lebesgue curves by whole areas.

Figures 11 and 12 show the square sum of delay between each node and each next node on the curves where the geographical position of nodes are generated by the uniform distribution and Zipf law, respectively. As a result, the proposed curve has been improved by as much as 30-60% against Lebesgue curve. Form the results, it can be said that the distance between any adjacent nodes on the proposed curve are always small without large bias than the enhanced Lebesgue curve.

V. Conclusion

In this paper, we proposed the method which configure small delay structured P2P overlay network by assigning a label considering the geographic information of each terminal node that became peer. The proposed space filling curve to configure the P2P network converts geographic information into node ID (label) of P2P network, and it

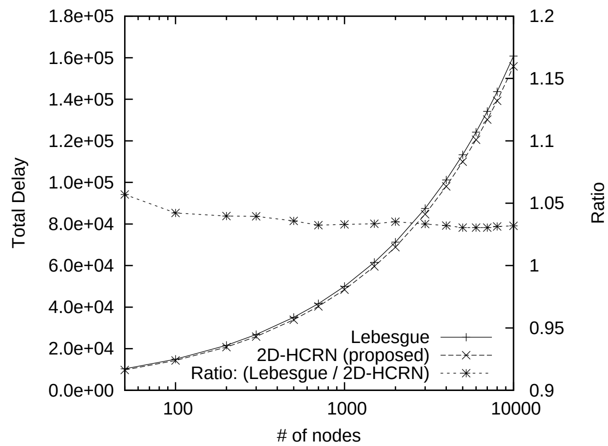

Figure 9. Overall delay on the curves with randomly distributed nodes

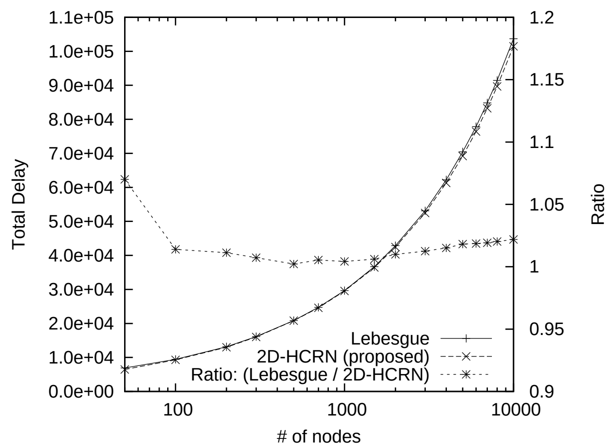

Figure 10. Overall delay on the curves with nodes distributed under Zipf's law

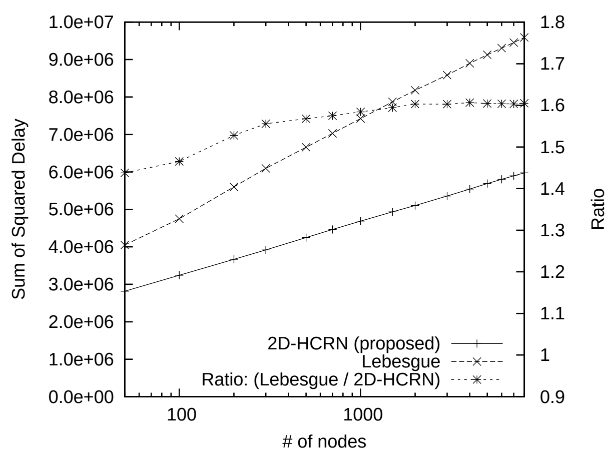

Figure 11. Sum of squared delays on the curves with randomly distributed nodes

can search information over specified range efficiently, and can construct hierarchical connection structure for scalability. Through the numerical evaluation, we confirmed that the proposed curve is more suitable for hierarchical-spread nodes than the conventional curves.

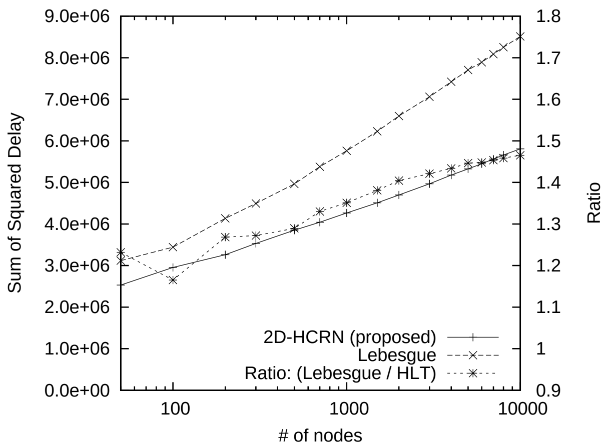

Figure 12. Sum of squared delays on the curves with nodes distributed under Zipf's law

For future work, we need the performance evaluation of the proposed space filling curve and the P2P overlay network in the real network environment especially in node distribution with bias, and the reexamination of the routing entry of each node to improve the performance of this method. In order to improve it and apply to real networks, we have a plan to adopt Vivaldi[18] that was the technique for constructing low delay network by using the spring model with the measurement result of the real communication delay.

REFERENCES

[1] "Gnutella," http://gnutella.wego.com.

[2] "Bittorrent," http://bittorrent.com.

[3] Y. Kaneko, K. Harumoto, S. Fukumura, S. Shimojo, and S. Nishio, "A location-based peer-to-peer network for context-aware services in a ubiquitous environment," in *Proceedings of Interjatinoal Symposium on Applications and the Internet (SAINT 2005)*, February 2005, pp. 208–211.

[4] J. Orenstein and T. Merrett, "A class of data structures for associative searching," in *Proceedings of the 3rd ACM SIGACT-SIGMOD Symposium on Principles of Database Systems*, April 1984, pp. 181–190.

[5] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan, "Chord: A scalable peer-to-peer lookup service for internet applications," in *Proceedings of the ACM SIGCOMM 2001*, August 2001, pp. 149–160.

[6] S. Ratnasamy, P. Francis, M. Handley, R. Karp, and S. Schenker, "A scalable content-addressable network," in *Proceedings of the ACM SIGCOMM 2001*, August 2001, pp. 161–172.

[7] A. I. T. Rowstron and P. Druschel, "Pastry: Scalable, decentralized object location, and routing for large-scale peer-to-peer systems," in *Proceedings of the IFIP/ACM International Conference on Distributed Systems Platforms Heidelberg (Middleware 2001)*. Springer-Verlag LNCS 2218, November 2001, pp. 329–350.

[8] B. Y. Zhao, J. Kubiatowicz, and A. D. Joseph, "Tapestry: An infrastructure for fault-tolerant wide-area location and routing," Computer Science Division, University of California at Berkeley, Tech. Rep., 2001.

[9] J. Han, D. Watson, and F. Jahanian, "Topology aware overlay networks," in *Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE INFOCOM 2005)*, vol. 4, March 2005, pp. 2554–2565.

[10] C.-J. Wu, D.-K. Liu, and R.-H. Hwang, "A location-aware peer-to-peer overlay network," *International Journal of Communication Systems*, vol. 20, no. 1, pp. 83–102, January 2007.

[11] Y. Liu, X. Liu, L. Xiao, L. M. Ni, and X. Zhang, "Location-aware topology matching in P2P systems," in *Proceedings of the 23rd Annual Joint Conference of the IEEE Computer and Communications Societies(IEEE INFOCOM 2004)*, vol. 4, March 2004, pp. 2220–2230.

[12] M. Waldvogel and R. Rinaldi, "Efficient topology-aware overlay network," *ACM SIGCOMM Computer Communication Review*, vol. 33, no. 1, pp. 101–106, January 2003.

[13] P. Pietzuch, J. Ledlie, M. Mitzenmacher, and M. Seltzer, "Network-aware overlays with network coordinates," in *Proceedings of the 26th IEEE International Conference on Distributed Computing Systems (ICDCS 2006)*, July 2006, p. 12.

[14] D. Zhang and C. Lin, "Efficient delay aware peer-to-peer overlay network," in *Proceedings of the 6th International Conference on Web-Age Information Management (WAIM 2005)*. Springer-Verlag LNCS 3739, October 2005, pp. 682–687.

[15] N. J. A. Harvey, M. B. Jones, S. Saroiu, M. Theimer, and A. Wolman, "Skipnet: A scalable overlay network with practical locality properties," in *Proceedings of the 4th USENIX Symposium on Internet Technologies and Systems (USITS '03)*, March 2003, pp. 113–126.

[16] J. Aspnes and G. Shah, "Skip graphs," in *Proceedings of the 14th annual ACM-SIAM Symposium on Discrete Algorithms (SODA 03)*, January 2003, pp. 384–393.

[17] S. Matsuura, K. Fujikawa, and H. Sunahara, "Mill: A geographical location oriented overlay network managing data of ubiquitous sensors," *IEICE TRANSACTIONS on Communications*, vol. E90-B, no. 10, pp. 2720–2728, October 2007.

[18] F. Dabek, R. Cox, and R. Morris, "Vivaldi: A decentralized network coordinate system," in *Proceedings of the ACM SIGCOMM 2004*, August 2004, pp. 15–26.

[19] J.-M. Wierum, "Logarithmic path-length in space-filling curves," in *Proceedings of the 14th Canadian Conference on Computational Geometry*, August 2002, pp. 22–26.

[20] T. Kitani, N. Funabiki, H. Yamaguchi, and T. Higashino, "Hierarchical logical topology in wdm ring networks with limited adm," in *Proceedings of the IFIP Networking 2008 (Networking 2008)*. Springer-Verlag LNCS 4982, May 2008, pp. 326–337.