A distributed measurement method for reducing measurement conflict frequency in overlay networks

Dinh Tien Hoang
Graduate School of Information
Science and Technology, Osaka University
1-5, Yamadaoka, Suita,
Osaka, 565-0871 Japan
Email: d-hoang@ist.osaka-u.ac.jp

Go Hasegawa
Cybermedia Center
Osaka University
1-32, Machikaneyama, Toyonaka,
Osaka, 560-0043 Japan
Email: hasegawa@cmc.osaka-u.ac.jp

Masayuki Murata
Graduate School of Information
Science and Technology, Osaka University
1-5, Yamadaoka, Suita,
Osaka, 565-0871 Japan
Email:murata@ist.osaka-u.ac.jp

***Abstract*—In overlay networks, in order to obtain accurate measurement results, it is important to solve the measurement conflict problem : the measurement tasks for overlapping paths conflict with each other. In this paper, we propose a measurement method which reduces the number of measurement conflicts, without centralized control in measurement tasks scheduling. In this method, each overlay node uses `traceroute` to get path information to other overlay nodes and exchanges it with nearby overlay nodes to estimate path overlaps. Based on the number of overlapping paths, the overlay node calculates an appropriate measurement frequency and a measurement timing to minimize the probability of measurement conflicts occurring among overlapping paths. Furthermore, the overlay node exchanges measurement results with a small number of overlay nodes to statistically obtain more accurate measurement results. Simulation results show that when the density of the overlay nodes (the ratio of the number of overlay nodes to the number of routers) is greater than 0.3, our method can run twice as many measurement tasks as existing methods without measurement conflict.**

I. INTRODUCTION

Recently, overlay networks have attracted much attention as a technology that enables early deployment of new network services without standardization processes. In general, the overlay network is a logical network constructed on the underlay network. In this paper, we consider the application-level overlay networks constructed on the IP network.

In overlay networks, the overlay nodes are often installed on end hosts as an application program. In this case, routing and traffic control at the overlay level is conducted at the end hosts, and such controls cannot be activated inside the network. On the other hand, the overlay routing inside the network becomes possible by installing overlay nodes on routers in the network. In this paper, in order to realize efficient routing control by overlay networks, we consider an overlay network in which the overlay nodes are deployed on the routers.

The overlay network should obtain the network resource information of the underlay network (including available bandwidth, propagation delay and packet loss ratio) for maintenance and improvement of the performance of network service. Several network measurement methods have been proposed in the literature [1]–[3], [5]–[7]. In [1], the authors assume that paths between all overlay nodes are measured. Therefore, the measurement overhead becomes $O(n^2)$, where n is the number of overlay nodes. Thus, [8] pointed out that the number of overlay nodes that can be applied is up to around 50.

Many solutions have been proposed to reduce measurement overhead [2], [3], [5]–[7]. In [2], the overlay path is divided into some segments based on the overlapping information of the overlay paths. Next, the minimum cover set of paths that covers all of the segments is found by using a heuristic algorithm. In general, because the overlay paths often overlap each other, the number of overlay paths in the minimum cover set is much less than the number of the entire overlay paths. Therefore, measurement overhead is reduced to $O(n \log n)$. The resource information of the segments is approximately calculated from the resource information of the overlay paths of the minimum cover set, and then the resource information of the entire overlay paths is approximately calculated from the resource information of segments. Therefore, the accuracy of measurement results obtained by this technique is not high.

In [3], the resource information of a path is expressed as a linear equation of the resource information of the links included in the route of the path. Some paths that include all links will be measured, and the result will be used to calculate the resource information of all links and then infer the resource information of all paths. The number of links in a general network is much less than the number of paths, so the overhead can be decreased to $O(n \log n)$. Although this method can get more accurate measurement results than the method in [2], the overhead to calculate the resource information of all links is greater than the overhead to calculate the resource information of segments, because the number of links is greater than the number of segments.

The authors in [5] propose a measurement technique called BRoute that can measure the available bandwidth of overlay paths. BRoute reduces measurement overhead to $O(n)$, based on two characteristics of overlay networks constructed over the Internet: (1) the bottleneck link exists from both ends of the overlay path in roughly four hops or less, and (2) the path overlapping often exists near both ends of the overlay path. Therefore, the available bandwidth of a segment near both ends of each overlay path can be used to get the available bandwidth of the entire path, so that measurement overhead can be reduced greatly. However, this technique cannot be applied to the measurement of delay and packet loss ratio. Furthermore, the characteristics (2) can not be satisfied in the router-based overlay networks.

The number of overlapping paths often increases when the density of the overlay node rises. Measurement conflicts happen when measuring overlapping paths at the same time and causes an additional error in measurement results. However, this problem is not considered in the methods mentioned above [1]–[3], [5]. To avoid measurement conflicts, [6] proposes a technique that schedules the timing of the measurement tasks of the overlay paths, based on topology information aggregated at a master node. In this technique, the execution time and period are defined for each measurement task, and then the tasks are divided into several groups, each one has the same execution time and period. Then task groups that can be executed simultaneously are obtained using a heuristic algorithm. As a result, measurement conflicts can be avoided completely. However, sometimes the measurement task cannot be divided. For example, in an available bandwidth measurement tool like PathChirp [9], the transmission interval of the measurement probes is adjusted to infer available bandwidth, so the measurement task cannot be divided.

Moreover, all methods mentioned above [2], [3], [6] require a master node to aggregate topology information, decide measurement timing, and give instructions to each overlay node. Therefore, the amount of time and network traffic for aggregation of topology information and instructions is large, and the performance of overlay networks decreases when some changes occur in the underlay network or overlay network.

[7] proposes a measurement system for available bandwidth, called ImSystemPlus that can avoid measurement conflicts without using a master node. In ImSystemPlus, the measurement timing of overlapping paths are decided randomly based on the number of hops of overlapped parts to avoid measurement conflicts. However, this method requires complete topology knowledge of the IP network at each overlay node.

In this paper, we propose an overlay network measurement method that does not use a master node, and that does not require complete topology information of the IP network. In the proposed method, each overlay node decides the measurement timing of paths whose source node is itself to avoid measurement conflicts. The overlay node first gets the topology information of the paths whose source node is itself, and exchanges the topology information with a small number of other overlay nodes to infer the overlap between paths whose source node is this node and those whose source nodes are the other overlay nodes. The overlay node then measures overlapping paths whose source node is itself sequentially. The measurement timing is decided randomly to avoid measurement conflicts with overlapping paths whose source nodes are the other nodes. After measuring those overlay paths, the overlay node exchanges measurement results with nearby overlay nodes and use statistical analysis for data obtained by information exchange to improve measurement accuracy.

The remainder of this paper is organized as follows. In Section II, we explain the method for detecting the overlapping of overlay paths. Section III describes the overlay path measurement technique for reducing the frequency of measurement conflict and explains the scheme of exchanging measurement results for improving measurement accuracy. In Section IV, we describe the measurement system based on the proposed technique. In Section V, the proposed method is evaluated with both mathematical analysis and numerical examples. We conclude and discuss future work in Section VI.

II. DETECTING OVERLAPPING PATHS

Figure 1 shows a classification of the overlapping state of overlay paths. In this paper, we classify overlapping states into three types, as follows.

- Complete overlapping : One path completely includes another path.
- Half-overlapping : Two paths share a route from the source overlay node to a router that is not an overlay node.
- Partial overlapping : Two paths share a route that does not include overlay nodes.

For example, in Figure 1, path AB and path AC are in the relation of complete overlapping, path AB and path AE are in the relation of half-overlapping, path AB and path DE are in the relation of partial overlapping.

We assume that every overlay node knows the IP addresses of the other overlay nodes. Then complete overlapping and half-overlapping can be detected by the source overlay node of the overlay path using `traceroute` as described in [10]. For example, in Figure 1, when the overlay node A issues `traceroute` to B and C, complete overlapping of path AB and path AC can be detected. Similarly, the shared route from A to router R_2 by path AB and path AE can be detected when A issues `traceroute` to B and E. On the other hand, partial overlapping cannot be precisely detected only by `traceroute`. Therefore, this paper proposes a detection procedure for partial overlapping as follows.

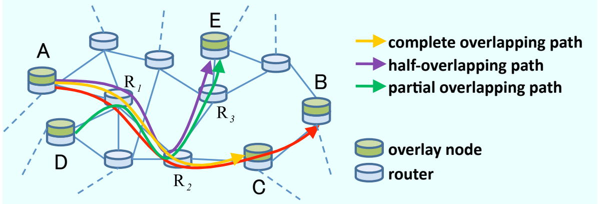


Fig. 1. Classification of path overlapping

1) Infer partial overlapping using `traceroute` results
   For example, in Figure 1, overlay node A issues `traceroute` to other overlay nodes and detects half-overlapping of path AB and path AD, and that of path AB and path AE. Then it is possible for path AB and path DE to partially overlap.
2) Exchange topology information between overlay nodes to confirm partial overlapping
   Overlay node A exchanges topology information with node D, the source node of the path that may partially overlap with path AB, and A can judge whether path AB and path DE are actually in the relation of partial overlapping.

By exchanging topology information with a small number of nearby overlay nodes, each overlay node can detect partial overlapping without using a master node.

III. MEASUREMENT METHOD FOR OVERLAY PATHS

In this section, we propose a method for reducing the frequency of measurement conflict. Because of space limitations, we explain the proposed method by showing a detailed behavior for overlay path AB in Figure 1. First, node A detects overlapping paths of path AB by using the method described in Section II. If path AB has no overlapping paths, it is unnecessary to consider a method that will avoid measurement conflicts. Therefore, we are only concerned with the case when path AB overlaps with other overlay paths.

We consider the following two cases of overlapping states.

1) When path AB completely includes other overlay paths, overlay path AB is not measured because path AB and the overlapping path are in the relation of complete overlapping.
2) When path AB does not include other overlay paths, we can reduce the probability of measurement conflicts between path AB and its overlapping paths by adjusting the frequency and timing of measurements.

The detailed mechanisms for the above two cases are described in Subsections III-A1 and III-A2, respectively. Furthermore, we improve the accuracy of the measurement results by exchanging the measurement results between overlay nodes and doing statistical analysis with aggregated data.

A. Avoiding measurement conflicts

1) Complete overlapping: In this case, the longer overlay path is not measured directly. That is, the measurement result is estimated based on the measurement results of shorter overlay paths included in the longer path.

We use the example of Figure 1 to explain this method. As shown in Figure 1, path AB completely includes path AC. When A issues `traceroute` to B, the `traceroute` packet goes through C. Then C can know that it itself exists on path AB. Then, C measures path CB, and transmits the result to A. A can also know that C exists on the path AB, based on the `traceroute` result. Therefore, A does not measure path AB directly, it only measures path AC and A then estimates the measurement result of path AB from the measurement result of path AC and that of path CB received from C. Refer to [10] for details on this method.

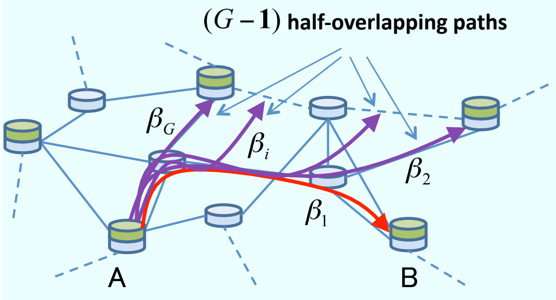


Fig. 2. Half-overlapping paths of AB

2) Half-overlapping and partial overlapping: In general, path AB has more than one half-overlapping path and more than one partial overlapping path. Here, we assume that AB has $G-1$ ($G \geq 1$) half-overlapping paths, as shown in Figure 2. For simplicity, we denote path AB as path 1, and its half-overlapping paths as paths i ($2 \leq i \leq G$) respectively. Furthermore, we assume that, by using the inferring method described in Section II to detect partial overlapping paths, path i has $K_i - 1$ ($1 \leq i \leq G$) partial overlapping paths respectively.

Overlay node A can avoid the measurement conflicts of path AB and its half-overlapping paths, by measuring them sequentially. On the other hand, because the source nodes of the partial overlapping paths of AB are not A, measurement conflicts between them cannot be avoided completely. Therefore, we propose a technique combining a sequential measurement scheme for half-overlapping paths and a random measurement scheme for partial overlapping paths.

In general, because of measurement conflicts and the changes in the network condition, the measurement results of a path fluctuate by time. Therefore, it is necessary to adjust measurement frequency according to the degree of fluctuation of measurement results. In detail, it is necessary to increase the measurement frequency when the fluctuation of measurement results is large, and we may reduce the measurement frequency if the fluctuation in measurement results is small.

We define the measurement frequency as follows. First, we assume that the time required for each measurement task is identical for all overlay paths and we denote it τ. We also assume that a measurement of the path is done in the time duration of T_{path} ($T_{path} \geq \tau$). Then the measurement frequency of the path is defined by τ / T_{path}. Note that T_{path} is changed by the source node according to the fluctuation of measurement results.

Assume that based on the degree of fluctuation of measurement results, the measurement frequency of path AB is calculated as β_1. Similarly, the measurement frequencies of half-overlapping paths of path AB are calculated as β_i ($2 \leq i \leq G$). Note that the calculation method of β_i is omitted due to space limitation. To avoid measurement conflicts of path AB and its half-overlapping paths, the sum of their measurement frequencies should be equal to or less than one ($\sum_{i=1}^{G} \beta_i \leq 1$). Therefore, when the sum of their measurement frequencies is greater than one ($\sum_{i=1}^{G} \beta_i > 1$), we should reduce the measurement frequencies of the measurement paths. The reduced measurement frequencies x_i for path i should satisfy the following equations.

$$\begin{gathered} x_1 \leq \beta_1, x_2 \leq \beta_2, ..., x_G \leq \beta_G \\ \frac{\beta_1 - x_1}{x_1} = \frac{\beta_2 - x_2}{x_2} = ... = \frac{\beta_G - x_G}{x_G} \\ \sum_{i=1}^{G} x_i = 1 \end{gathered} \tag{1}$$

The inequalities in the first line of (1) guarantees that all of the measurement frequencies are reduced, while the equation in the second line means that the reduction rates of the measurement frequencies are the same, and the equation in the third line keeps the reduction of frequencies as small as possible.

It can be confirmed that solution x_i that satisfies (1) is given by Equation (2).

$$x_i^* = \frac{\beta_i}{\sum_{i=1}^{G} \beta_i}, \quad i = 1, ..., G \tag{2}$$

In the proposed method, the value derived by Equation (2) is used. If $\sum_{i=1}^{G} \beta_i \leq 1$, we set $x_i^* = \beta_i$.

The following theorem can be utilized to arrange the measurement timing of path AB and its half-overlapping paths with the measurement frequency of x_i^*.

Theorem 1. *A method of arranging the measurement timing of path AB and its half-overlapping paths, with the measurement frequency of x_i^*, exists.*

Proof: Let $L_i = 1/x_i^*$. We can assume that $L_1 \leq L_2 \leq ... \leq L_G$ without loss of generality. Since $\sum_{i=1}^{G} 1/L_i = \sum_{i=1}^{G} x_i^* = 1$, there exists $1 \leq l \leq G$, such that $L_1 \leq ... \leq L_l \leq G \leq L_{l+1} \leq ... \leq L_G$.

We call the time of one measurement a `measurement time slot` and the time to measure path AB and its half-overlapping paths a `measurement cycle`. We consider the following measurement scheme.

1) Decide the measurement order of path i ($1 \leq i \leq G$) at one measurement circle randomly, and allocate the measurement time slot for each path.
2) For path i that $i > l$, we measure the path with the probability of G/L_i, at the measurement time slot allocated to it. In other words, this path is not measured at probability of $1 - G/L_i$. When path i ($i > l$) is not measured, the measurement time slot is used to measure path j ($j \leq l$).
3) For path j that $j \leq l$, we measure the path at the measurement time slot allocated to it. Therefore, the measurement frequency of path j becomes $1/G$ ($< 1/L_j$), and it is smaller by $1/L_j - 1/G$ than the measurement frequency $1/L_j$ for path j.

Hereafter, we prove that the measurement frequency of path i becomes $1/L_i$ ($i = 1, ..., G$) by using the above scheme.

- For path i ($i > l$)
  At the measurement time slot allocated to the path, the probability that measurement is carried out is G/L_i. Because we have G slots in one measurement circle, the probability that measurement is carried out becomes $1/L_i$.
- For path j ($j \leq l$)
  The measurement is carried out at the measurement time slot of path j, so the measurement frequency of path j becomes $1/G$, and is $1/L_j - 1/G$ smaller than the measurement frequency $1/L_j$ that we want to set for path j. Therefore, the sum of the measurement frequencies of all

paths j ($j \le l$) is $\sum_{j=1}^{l} (1/L_j - 1/G)$ smaller than the sum of the desired measurement frequencies.

On the other hand, at the measurement time slot of path i ($i > l$), the probability that the measurement of path i is not carried out is $1 - G/L_i$, so in one measurement circle, the probability that measurement is not carried out becomes $(1 - G/L_i) * 1/G = 1/G - 1/L_i$.

For all path i ($i > l$), the sum of probability that measurement is not carried out becomes $\sum_{i=l+1}^{G} (1/G - 1/L_i)$, and we can use these measurement time slots to measure path j ($j \le l$) when measurement of path i ($i > l$) is not carried out.

Since $\sum_{i=1}^{G} 1/L_i = \sum_{i=1}^{G} x_i^* \le 1$, we have $\sum_{j=1}^{l} (1/L_j - 1/G) \le \sum_{i=l+1}^{G} (1/G - 1/L_i)$. Therefore, we can say the measurement frequencies of path i ($i \le l$) can be set to $1/L_i$.

■

We next explain the statistical method of avoiding the measurement conflicts of path AB and $K_1 - 1$ partial overlapping paths. As mentioned above, because $K_1 - 1$ partial overlapping paths are measured by the overlay nodes different from A, the measurement conflicts of path AB and the partial overlapping paths cannot be completely avoided. Therefore, we propose a method of reducing the probability of measurement conflicts as follows.

When path AB has $K_1 - 1$ partial overlapping paths, the measurement frequency of path AB is set to a value not greater than $1/K_1$. In other words, the measurement frequency of path AB is decided according to the following equation.

$$y_1^* = \min\{x_1^*, 1/K_1\} \tag{3}$$

The measurement frequency of path AB is finally set to y_1^*.

B. Statistical method for improving the accuracy for measurement results

As mentioned in Section III-A, because it is not possible to avoid measurement conflicts completely, the accuracy of measurement results decreases due to measurement conflicts. Therefore, in the proposed method, the measurement results are exchanged between overlay nodes, and the confidence interval of measurement results is evaluated by using statistical analysis for aggregated measurement results. This section briefly explains the measurement, information exchange, and statistical methods.

1) Measurement method for partial overlapping paths: We explain the measurement method of path AB that has partial overlapping paths. Here, we assume the measuring metric is delay.

Generally, there are many overlapped parts of path AB and its partial overlapping paths. We assume that the overlapped parts are divided by routers R_1, R_2, ..., R_l. In the proposed method, the delay measurements are individually conducted for overlapping parts $\mathrm{R}_1\mathrm{R}_2$, $\mathrm{R}_2\mathrm{R}_3$, ..., $\mathrm{R}_{l-1}\mathrm{R}_l$ as well as for end-to-end path AB. Figure 3 shows an example of path AB and routers R_1, R_2, ..., R_l.

In the following, we explain the procedure for A to measure the delay of path AB.

A issues `traceroute` to B, and detects the routers on the path AB. Next, A measures the delays to these routers and calculates the delay of AR_1, $\mathrm{R}_1\mathrm{R}_2$, ..., $\mathrm{R}_{l-1}\mathrm{R}_l$ and $\mathrm{R}_l\mathrm{B}$. The delay of AR_1, AR_2, ..., AR_l, AB are denoted as t_{A,R_1}, t_{A,R_2}, ..., t_{A,R_l}, $t_{A,B}$ respectively. These values can be got by using ping from A to each router. The delays t_{A,R_1}, t_{R_1,R_2}, ..., t_{R_{l-1},R_l}, $t_{R_l,B}$ of AR_1, $\mathrm{R}_1\mathrm{R}_2$, ..., $\mathrm{R}_{l-1}\mathrm{R}_l$, $\mathrm{R}_l\mathrm{B}$ are then calculated as follows.

$$\begin{array}{lcll} t_{R_i,R_{i+1}} & = & t_{A,R_{i+1}} - t_{A,R_i} & ,i = 1,...,l-1 \\ t_{R_l,B} & = & t_{A,B} - t_{A,R_l} & \end{array} \tag{4}$$

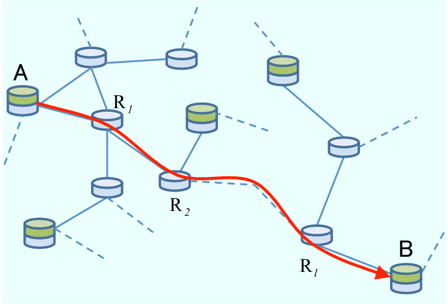


Fig. 3. Routers on the path AB

2) Exchange of measurement results: The measurement results and the route information of path AB are sent to the source overlay node of its partial overlapping paths. A also receives similar information from those overlay nodes.

3) Statistical method: Finally, we use statistical analysis for data obtained by information exchange and calculate the measurement result of path AB. Due to space limitation, we omit a detailed explanation of the statistical methods.

IV. MEASUREMENT SYSTEM

In this section, we explain the composition of the measurement system based on the measurement method described in the above sections. Because of space limitation, in the following we explain the procedure of the measurement of overlay node A in Figure 1.

- Job 1 : Investigation overlaps by using `traceroute`
  Node A first executes `traceroute` to other overlay nodes and investigates the overlapping state of overlay paths whose source node is A.
- Job 2 : Path measurement, information exchange, and measurement results calculation
  - Job 2.1 : Path measurement
    Node A measures each path based on the method proposed in Section III-A.
  - Job 2.2 : Information exchange with other overlay nodes
    Node A exchanges path information and the measurement results with other overlay nodes as described in Subsection III-B.
  - Job 2.3 : Calculation of measurement results
    Node A does statistical processing on the measurement results obtained by its own measurements and those from other overlay nodes, based on the method explained in Section III-B, and calculates the measurement results.

  Note that A can concurrently execute measurement of one path and exchange measurement results of the other paths as long as these paths do not overlap each other. So the above three procedures can be executed concurrently.

Because the configuration of the underlay network may change, Job 1 is executed at regular intervals. In general, because the frequency of the change in the underlay network is smaller than the frequency of the change in the measurement results, the execution frequency of Job 1 is set smaller than the execution frequency of Job 2.

V. PERFORMANCE EVALUATION

A. Measurement overhead analysis

In [2], [3], [6], all overlay nodes should send the network topology information, to the master node. So when n is the number of overlay nodes, then the amount of traffic of transferring topology information becomes $O(n^2)$. On the other hand, the proposed method does not aggregate all topology information to one node. Instead, an overlay node only exchanges topology information of a small number of partial overlapping paths with nearby overlay nodes. So we can expect that the traffic overhead is small.

In the method in [2], the number of measurement paths is reduced to $O(n \log n)$. On the other hand, because our method does not measure complete overlapping paths, the number of measurement path is also reduced. In the following, we use some results from [11] to estimate the number of measurement paths, and compare it to the number of measurement tasks of the full mesh measurement method. We denote the maximum length of the overlay path (the number of IP links that the path traverses) as L_{max}, the ratio of the number of paths whose length is r as $P(r)$, the number of routers as N, and the density of the overlay nodes as d ($d = n/N$). Because the total number of overlay paths is $Nd(Nd-1)$, so the number of overlay paths whose length is r becomes $P(r)Nd(Nd-1)$. We estimate the number of measurement paths whose length is r as follows. First, note that an overlay path becomes a measurement path when it does not contain overlay nodes between source node and destination node. An overlay path whose length is r contains $r-1$ routers between source node and destination node. So the probability that this overlay path becomes a measurement path is equal to the probability that all of these $r-1$ routers are not overlay node. We denote these routers R_1,...,R_{r-1}, started from the router next to source node. The probability that the router R_1 is not an overlay node is $1 - \frac{Nd-2}{N-2}$, because there are $Nd-2$ overlay nodes among $N-2$ routers except source node and destination node. Similarly, if we have already known that the routers R_1,...,R_{i-1} ($1 < i < r-1$) are not overlay node, then the probability that the router R_i is also not an overlay node is $1 - \frac{Nd-2}{N-1-i}$, because there are $Nd-2$ overlay nodes among $N-1-i$ routers except source node, destination node and routers R_1,...,R_{i-1}. So the probability that all of these $r-1$ routers are not overlay node becomes $\prod_{i=1}^{r-1}\left(1 - \frac{Nd-2}{N-1-i}\right)$. Therefore, the number of measurement paths whose length is r becomes $P(r)Nd(Nd-1)\prod_{i=1}^{r-1}\left(1 - \frac{Nd-2}{N-1-i}\right)$, and the total number of measurement paths is given by Equation (5) [11].

$$M = \sum_{r=1}^{L_{max}} \left(P(r)Nd(Nd-1) \prod_{i=1}^{r-1} \left(1 - \frac{Nd-2}{N-1-i}\right) \right) \quad (5)$$

On the other hand, the number of measurement paths of the full mesh measurement methods is given by Equation (6).

$$M_{full} = Nd(Nd-1) \quad (6)$$

Therefore, the ratio of the number of measurement paths of the proposed method to the number of measurement paths of the full mesh measurement method is given from Equation (7).

$$\mu = \frac{M}{M_{full}} = \sum_{r=1}^{L_{max}} \left(P(r) \prod_{i=1}^{r-1} \left(1 - \frac{Nd-2}{N-1-i}\right) \right) \quad (7)$$

Let i_0 be the maximum value of integer i satisfies $d \leq \frac{2}{i+1}$.

- when $i \leq i_0$
  In this case, when N is large enough, Equation (8) establishes.

$$\frac{Nd-2}{N-1-i} \approx d \quad (8)$$

- when $i > i_0$
  In this case, Equation (9) holds.

$$\frac{Nd-2}{N-1-i} > d \quad (9)$$

Therefore, when N is large enough, the value of μ is approximated by Equation (10).

$$\begin{aligned} \mu &\leq \sum_{r=1}^{L_{max}} \left(P(r) \prod_{i=1}^{r-1} (1-d) \right) \\ &= \sum_{r=1}^{L_{max}} \left(P(r)(1-d)^{r-1} \right) \end{aligned} \quad (10)$$

The right side of Equation (10) becomes smaller when d approaches to 1, because $\sum_{r=1}^{L_{max}} P(r) = 1$, $(1-d)^{r-1} < 1, \forall r > 1$, and the more d approaches 1 the smaller it becomes.

Therefore, when d is large, the number of measurements of the proposed method is greatly reduced compared with the full mesh measurement methods. In [11], this result is verified by simulation evaluations, and it is confirmed that in some cases, the number of measurement paths is reduced up to $1/4000$ compared to full mesh measurement methods.

B. Measurement accuracy analysis

Here we present numerical evaluation results of the proposed method. We compare the results with those of the measurement method proposed in [6].

In [6], a scheduling method is proposed so that measurement tasks do not conflict with each other. To do that, a heuristic algorithm is used to divide the measurement tasks into some groups of tasks that can run concurrently. The tasks of different groups are not executed at the same time, so measurement conflicts can be avoided completely. However, the number of measurement tasks that can be scheduled is limited when the number of required measurements is large. In contrast, our method does not avoid measurement conflicts completely but can run as many measurement tasks as possible. In this subsection we compare the proposed method and the method in [6] by using the following metrics.

Definition 1. *Measurement execution ratio*
Measurement execution ratio is defined as the average value of measurement frequencies of all overlay paths.

Definition 2. *Measurement success ratio*
Measurement success ratio is defined as the average value of frequency of measurement tasks that do not conflict with other measurement tasks.

Note that in the method proposed in [6], because the measurement conflicts are completely avoided, the measurement execution ratio is equal to the measurement success ratio.

We use underlay network topologies based on the BA model [12] and the Waxman model [13] that are generated by BRITE [14]. We also utilized the network topology based on AT&T's actual network topology obtained from [4]. Each topology has 523 routers and 1304 links. The density of the overlay nodes is set to values from 0.05 to 0.4 with steps of 0.05. We utilized ten generated topologies for the BA model and the Waxman model. Furthermore, for averaging the results, we chose overlay nodes randomly 10 times for the BA model and the Waxman model and 100 times for the AT&T model.

In Figure 4 we plot the average value and 95% confidence interval of the measurement execution ratio and the measurement success ratio. The green curve shows the measurement execution ratio of the method proposed in [6], the blue curve and red curve show the measurement execution ratio and measurement success ratio of our method, respectively. We can see that in our method, the measurement success ratio is much greater than those of the methods in [6], especially when the density of overlay nodes is large. In addition, the measurement success ratio of the proposed method is up to twice those of the methods in [6]. This is because when the density of the overlay nodes becomes large, the number of half-overlapping paths and partial overlapping paths increases, so the methods of [6] create more task groups. As a result, the measurement execution ratio decreases. We can also see that the measurement execution ratio in the Waxman model is greater than those in the BA and AT&T models. This is because the number of half-overlapping paths and partial overlapping paths of the BA and AT&T models are greater than those in Waxman model. To verify this, in Figure 5 we plot the average value and 95% confidence interval of the total number of half-overlapping paths and partial overlapping paths of one overlay path in BA model, Waxman model and AT&T model networks. We can see that the number of half-overlapping paths and partial overlapping paths of AT&T model is greater than those of BA model, and those of BA model is greater than those of Waxman model.

VI. CONCLUSION

In this paper, we proposed a distributed overlay network measurement method that reduces the probability of measurement conflicts, by inferring the overlapping of paths and adjusting the measurement frequency and the measurement timing of paths. We also proposed a method to improve measurement accuracy by exchanging measurement results between a small number of nearby overlay nodes.

Future tasks include the evaluation of measurement accuracy for comparison with existing methods.

ACKNOWLEDGMENT

This work was supported in part by the National Institute of Information and Communications Technology of Japan (NICT).

REFERENCES

[1] D. Andersen, H. Balakrishnan, M.Kaashoek, and R. Morris, “Resilient overlay networks,” in *Proceedings of SOSP 2001*, Oct. 2001.

[2] C. Tang and P. McKinley, “On the cost-quality tradeoff in topology-aware overlay path probing,” in *Proceedings of ICNP 2003*, Nov. 2003.

[3] Y. Chen, D. Bindel, H. Song, and R. Katz, “An algebraic approach to practical and scalable overlay network monitoring,” in *Proceedings of SIGCOMM 2004*, Aug. 2004.

[4] N. Spring, R. Mahajan, and D. Wetherall, “Measuring ISP Topologies with Rocketfuel,” in *Proceedings of SIGCOMM 2002*, Aug. 2002.

[5] N. Hu and P. Steenkiste, “Exploiting internet route sharing for large scale available bandwidth estimation,” in *Proceedings of the IMC 2005*, Oct. 2005.

[6] M. Fraiwan and G. Manimaran, “Scheduling algorithms for conducting conflict-free measurements in overlay networks,” *Computer Networks*, vol. 52, pp. 2819–2830, Oct. 2008.

[7] C. L. T. Man, G. Hasegawa, and M. Murata, “Monitoring overlay path bandwidth using an inline measurement technique,” *IARIA International Journal on Advances in Systems and Measurements*, vol. 1, no. 1, pp. 50–60, 2008.

[8] A. Nakao, L. Peterson, and A. Bavier, “Scalable routing overlay networks,” *ACM SIGOPS Operating Systems Review*, vol. 40, pp. 49–61, Jan. 2006.

[9] V. J.Ribeiro, R. H. Riedi, R. G. Baraniuk, J. Navratil, and L. Cottrell, “pathchirp: Efficient available bandwidth estimation for network paths,” in *Proceedings of Passive and Active Measurement Workshop 2003*, Apr. 2003.

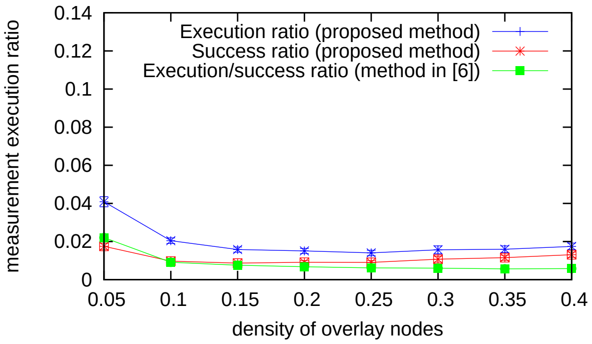
[10] G. Hasegawa and M. Murata, “Scalable and density-aware measurement strategies for overlay networks,” in *Proceedings of ICIMP 2009*, pp. 21–26, May 2009.

[11] H. Mori, G. Hasegawa and M. Murata, “Reducing measurement overhead using path overlapping in overlay networks,” *IEICE technical report(ICM2008-67)*, vol. 108, pp. 53–58, Mar. 2009.

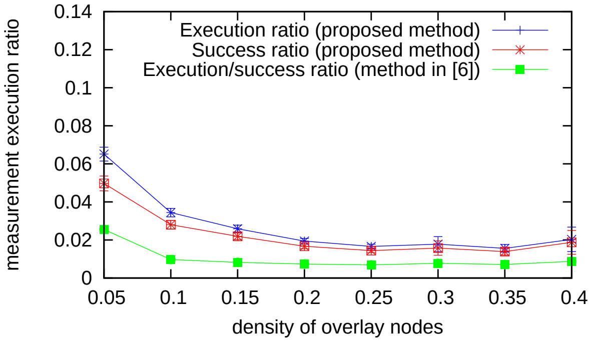
[12] A. Barabasi and R. Albert, “Emergence of scaling in random networks,” *Science*, vol. 286, pp. 509–512, Oct. 1999.

[13] B. M. Waxman, “Routing of multipoint connections,” *IEEE journal on selected areas in communications*, vol. 6, pp. 1617–1622, Dec. 1988.

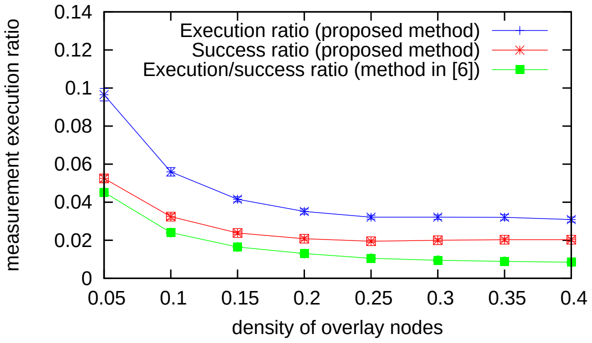
[14] BRITE: Boston university Representative Internet Topology gEnerator. available at http://www.cs.bu.edu/brite/index.html.



(a) AT&T model



(b) BA model



(c) Waxman model

Fig. 4. Measurement execution/success ratio

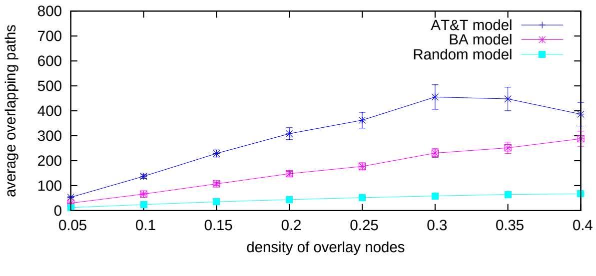


Fig. 5. Average overlapping paths