MOBILITY MODEL BASED ON INCOMING AND OUTGOING NODES TO AN AREA

S. Aihara[a], **M. Sasabe**[b], **H. Nakano**[b]

[a]Graduate School of Information Science and Technology, Osaka University
1-5 Yamadaoka, Suita-shi, Osaka-fu 565-0871, Japan - s-aihara@ist.osaka-u.ac.jp
[b]Cybermedia Center, Osaka University,
1-32 Machikaneyama-cho, Toyonaka-shi, Osaka-fu 560-0043, Japan - (m-sasabe, nakano)@cmc.osaka-u.ac.jp

KEY WORDS: mobility model, node distribution, incoming and outgoing nodes to an area, combinatorial optimization, Bayesian estimation

ABSTRACT:

In this paper, we propose a mobility model which estimates a node distribution in a service area from node flows on the boundary of the area. Our model targets new wireless communication techniques, such as Bluetooth and ZigBee, whose transmission ranges, called cells, are relatively small. The proposed model reproduces the mobility of nodes in an area by the observation of the incoming and outgoing nodes to the area. At first, we focus on a person as a node and observe an area where people actually pass through. From the observation results, we show that a node moves along with an approximately straight line with a velocity following a normal distribution. Furthermore, we propose two methods that estimate correspondences between incoming and outgoing events to the area; combinatorial optimization based method (COM) and Bayesian estimation based method (BEM). As a result, the reproduction ratio of COM is 0.736, and that of BEM is 1.

1 INTRODUCTION

In this paper, we propose a mobility model which estimates a node distribution in a service area from node flows on the boundary of the area. Our model targets new wireless communication techniques, such as Bluetooth (Bray and Sturman, 2001) and ZigBee (ZigBee Alliance, 2007), whose transmission ranges, called cells, are relatively small. These techniques can be available for everyone unlike mobile phone networks which belong to specific communication carriers. Therefore, if system administrators in companies or universities use them appropriately, they can construct a new wireless infrastructure. If they arrange the deployment of access points according to the distribution of users in a service area, communication quality for users will be improved. Though the distribution of users can be derived by observing the movements of users, it is unwise to observe the whole area from the view point of monitoring overheads and costs. If we can model the mobility of users in an area, we can estimate the distribution of users in the area while suppressing the observation costs.

The most popular mobility model is random waypoint model (RWP) (Camp et al., 2002), which randomly determines the node traveling direction, velocity, and residence time. Because of its simplicity, RWP is often referred to represent node mobility in wide-range areas such as mobile phone networks. There have been studied on other mobility models. For example, RWP is modified so that the probability of the node continuing to follow the same direction is higher than that of the node changing directions (Camp et al., 2002). Baochun Li and K.H. Wang propose a model which groups nodes based on their velocities or traveling directions (Li and Wang, 2003). In another study (Ashtiani et al., 2003), an area is divided into subregions and the node mobility in an area is regarded as a multi class Jackson network between the subregions. All of them exclusively focus on node mobility in an area. If the size of an area becomes small, we cannot neglect the effect of the node flows on the boundary of the area and the obstacles which disturb the node mobility in the area. These factors, however, are not considered in the traditional mobility models.

In this paper, we propose a mobility model based on the incoming and outgoing nodes to an area. The proposed model reproduces the mobility of nodes in an area by monitoring the node incoming and outgoing at gates located on the boundary of the area. We firstly focus on a person as a node and observe the mobility of people in a certain area in Osaka University. Then, we model the trajectory of nodes and velocity distribution in the area from observation results. Furthermore, we propose two methods that estimate correspondences between incoming and outgoing events to the area; combinatorial optimization (Korte and Vygen, 2005) based method (COM) and Bayesian estimation (Fox et al., 2003) based method (BEM). Finally, we validate the accuracy of the proposed method by comparing with the observation results.

The rest of the paper is organized as follows. In Section 2, we explain the knowledge that derived from observation and shows validity of some assumptions used in the proposed model. We introduce the proposed model in Section 3, and evaluate the model in Section 4. Finally, our conclusion and future works are stated in Section 5.

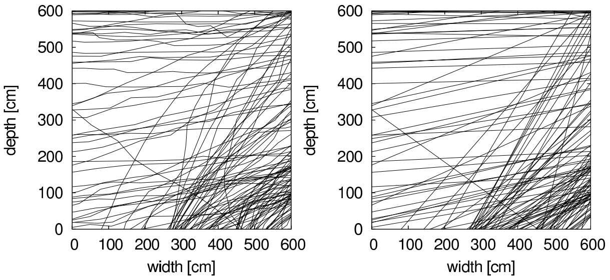
2 OBSERVATION

2.1 Overview

We took a movie of people flows in front of Cybermedia Center in Osaka University from 7th floor of the building. Next, we extracted static images from the movie with a sampling of interval 500 [ms]. Then, we arranged an area of 6×6 [m^2] at the center of each image. We calculated the trajectory and incoming and outgoing positions for each node by using a measurement tool developed by Visual C++. Although the proposed model estimates the node distribution in an area from observations at the boundary of the area, we also observed node trajectories in the area in order to evaluate the accuracy of the proposed model. Since we could not observe the area from the right above, the extracted images did not correctly represent the real world. We used projective transformation (Ito, 2007) to transform coordinates on each extracted image to those on a real coordinate system. Note that we assume that there is no aberration of a lens. Table 1 shows details of the observation environment.

Table 1: Observation environment

Date	2006/11/22
Time	14:30-14:40
Weather	Fine
Sampling Interval	500 [ms]
Area Size	600×600 [cm]
Number of Passing Nodes	125



(a) Observation result

(b) Straightly approximated lines each of which is drawn from the incoming position to the outgoing position of a node

Figure 1: Node trajectories

To evaluate the monitoring overheads quantitatively, we assume that the area is divided into $n \times n$ grids. Accordingly, node incoming and outgoing to the area occurs at $n \times 4$ gates on the boundary. Note that node incoming and outgoing positions are compensated to the center of the corresponding gate. We can derive a distribution of node residence time by calculating sum of node passing time for each grid. Then, we obtain the normalized distribution of node residence time. Finally, we get a probability density distribution of nodes by dividing each grid value by its area.

2.2 Node Trajectories

Figure 1(a) shows the actual node trajectories, and Fig. 1(b) shows straightly approximated lines each of which is drawn from the incoming position to the outgoing position of a node. Figure 2 illustrates a histogram of the destance errors obtained by comparing the approximation with the observation result. 90 % of the distance errors is not over 30 [cm]. Since this value seems to be smaller than the width a person occupies, the straight approximation is valid.

2.3 Distribution of Node Velocities

Figure 3 depicts a histogram of the node velocities. We also show a normal distribution $N(1.31 \times 10^2, 7.38 \times 10^2)$ multiplied by a correction factor C = (the number of samples of the histogram × data interval). Since they present a similar tendency, we assume that node velocities can be approximated by a normal distribution.

2.4 Number of Nodes

Table 2 shows the frequency ratio of the number of nodes that concurrently exist in the area. The observation environment has relatively a sparse density of people since there are not simultaneously over two people in the area at about 80 % of the observation. As future work, we plan to evaluate our proposed model in people-denser environments.

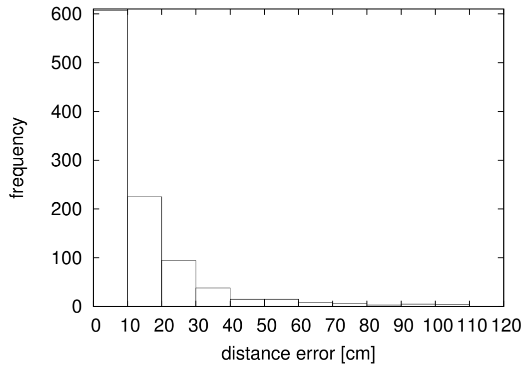


Figure 2: Histogram of distance errors caused by approximation

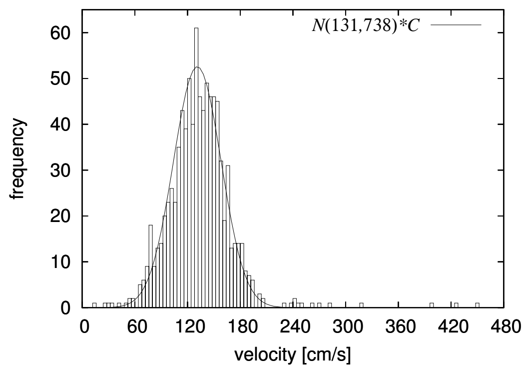


Figure 3: Distribution of node velocities (mean: 133 [cm/s])

3 MOBILITY MODEL BASED ON INCOMING AND OUTGOING NODES TO AN AREA

3.1 Overview

The proposed model reproduces the mobility of nodes in the area by observing the incoming and outgoing events to the area. Firstly, from the knowledge obtained in section 2, we suppose a node moves along with a straight line at a constant speed. Furthermore, node velocities follow a normal distribution. Then, we estimate the correspondences between incoming and outgoing events to the area. Ideally, the estimation can yield the trajectories in Fig. 1(b).

3.2 Correspondences between Incoming and Outgoing Events

Whenever a node enters (leaves) the area, we obtain the time and position of the corresponding incoming (outgoing) event from gates located on the boundary of the area. Because the proposed model does not monitor the inner area, it cannot know the gate out of which an incoming node gets. To derive an accurate distribution of nodes, estimation accuracy of the correspondences between incoming and outgoing events is an important factor. In this paper, we propose combinatorial optimization (Korte and Vygen, 2005) based method (COM) and Bayesian estimation (Fox et al., 2003) based method (BEM).

3.2.1 Combinatorial Optimization Based Method (COM)

Whenever the detected number of outgoing events reaches N, we choose an incoming event for each outgoing event. In what follows, we give the details of scheme to determine the correspondences between incoming and outgoing events. Examples in the

Table 2: Frequeny ratio of the number of nodes that simultaneously exist

Number of Nodes	0	1	2	3	4	5	6	7	8
Frequency Ratio	0.6583	0.1817	0.0958	0.0300	0.0100	0.0133	0.0058	0.0033	0.0017

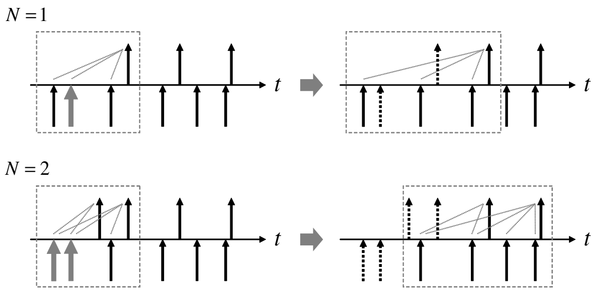


Figure 4: Overview of COM ($N = 1, 2$)

case of $N = 1, 2$ are shown in Fig. 4. Horizontal arrows denote the time axis, vertical arrows under the time axis denote incoming events, and opposite ones denote outgoing events. In the left side of the figure, we choose N incoming events which will correspond to N outgoing events in a dashed square. As the result, if the outgoing events are corresponded to the incoming events (bold arrows), then the next correspondences are chosen as in the right of the figure. We should note here that an incoming event that has already matched an outgoing event, that is the dashed arrow, is excluded in the following estimation.

Here, we explain how to choose an appropriate pair of incoming and outgoing events. Let A be a set of gates on the boundary of the area. $in(t, a)$ and $out(t, a)$ denote incoming and outgoing events on a gate $a \in A$ at time t, respectively. Additionally, let $D(a, a')$ be the distance between a and a', and $\bar{v}$ be the average velocity of nodes. Now, if $in(t, a)$ corresponds to $out(t', a')$, the time that the node passed through the area can be estimated as follows.

$$t' - t \simeq \frac{D(a, a')}{\bar{v}}$$

The left side of the equation is obtained from the observation, and the right side of that is derived from the estimation. Since $\bar{v}$ may be different from the actual velocity in practical situations, both sides of the equation may not be equal. Absolute differences of them becomes as follows.

$$Er(in(t, a), out(t', a')) = \left| (t' - t) - \frac{D(a, a')}{\bar{v}} \right|$$

This equation is expected to be minimized when $in(t, a)$ corresponds to $out(t', a')$. In case $N \geq 2$, we calculate sum of Er which is derived from each N pair as follows.

$$S_{Er}(N) = \sum_{i=1}^{N} Er(in(t_i, a_i), out(t'_i, a'_i))$$

We conduct paring of incoming and outgoing events by minimizing $S_{Er}(N)$ for $N \geq 1$.

This combinatorial optimization problem can be calculated within a practical time by reducing the combination patterns under the following considerations: an outgoing event must occur after the corresponding incoming event, and N does not need to exceed the maximum number of nodes exist in an area at the same time.

3.2.2 Bayesian Estimation Based Method (BEM) Whenever an outgoing event is detected, we estimate the corresponding incoming event probabilistically. In this section, we describe how to derive the probability that $out(t_{out}, a_{out})$ corresponds to $in(t_{in}, a_{in})$ based on Bayesian estimation.

According to Bayes' theorem, we first obtain the relationship between the conditional and marginal probabilities of stochastic events $in(a_{in})$ and $out(a_{out})$. Here, $in(a_{in})$ is an incoming event occurred at gate a_{in} and $out(a_{out})$ is an outgoing event occurred at gate a_{out}.

$$\begin{aligned} & P(in(a_{in})|out(a_{out})) \\ & \quad = \frac{P(out(a_{out})|in(a_{in}))P(in(a_{in}))}{P(out(a_{out}))} \end{aligned} \tag{1}$$

By extending Eq. (1) that take into account time relationship, we derive a probability that $out(t_{out}, a_{out})$ corresponds to $in(t_{in}, a_{in})$ as follows.

$$\begin{aligned} & P(in(t_{in}, a_{in})|out(t_{out}, a_{out})) \\ & \quad = P(out(t_{out}, a_{out})|in(t_{in}, a_{in})) \\ & \quad\quad \cdot \frac{P(in(t_{in}, a_{in}))}{P(out(t_{out}, a_{out}))} \end{aligned} \tag{2}$$

In a stationary state of the system, we assume that the rate that incoming and outgoing events happen does not depend on time. As a result, the right side of Eq. (2) is equal to

$$\frac{P(out(t_{out}, a_{out})|in(a_{in})) \cdot P(in(a_{in}))}{P(out(a_{out}))} \tag{3}$$

$P(out(t_{out}, a_{out})|in(a_{in}))$ can be denoted by the products of probability distribution of the passing time between a_{in} and a_{out}, $P(D(a_{in}, a_{out})/v)$, and a maximum probability of node traveling from a_{in} to a_{out}. Consequently, Eq. (3) can be transformed to

$$\begin{aligned} & P(D(a_{in}, a_{out})/v) \cdot P_{max}(out(a_{out})|in(a_{in})) \\ & \quad\quad \cdot \frac{P(in(a_{in}))}{P(out(a_{out}))} \end{aligned} \tag{4}$$

where v denotes the node velocity. From section 2.3, node velocity v follows a normal distribution with mean $\bar{v}$ and variance σ^2. With this distribution, we derive a probability distribution of the passing time t that a node requires to travel on distance l as

$$g(t) = \frac{l}{t^2} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\frac{l}{t} - \bar{v})^2}{2\sigma^2} \right) \tag{5}$$

From Eqs. (4) and (5), we derive the probability that $in(t_{in}, a_{in})$ corresponds to $out(t_{out}, a_{out})$ as

$$\begin{aligned} & P(in(t_{in}, a_{in}))|out(t_{out}, a_{out})) \\ & \quad = g(t_{out} - t_{in}) P_{max}(out(t_{out}, a_{out})|in(a_{in})) \\ & \quad\quad \cdot \frac{P(in(a_{in}))}{P(out(a_{out}))} \end{aligned} \tag{6}$$

where $\frac{P_{max}(out(t_{out}, a_{out})|in(a_{in})) \cdot P(in(a_{in}))}{P(out(a_{out}))}$ means the probability that a node moves from a_{in} to a_{out}. We define this probability as $w(a_{in}, a_{out})$. $w(a_{in}, a_{out})$ is updated based on the observation result appropriately.

Figure 5 shows an example of correspondence between incoming

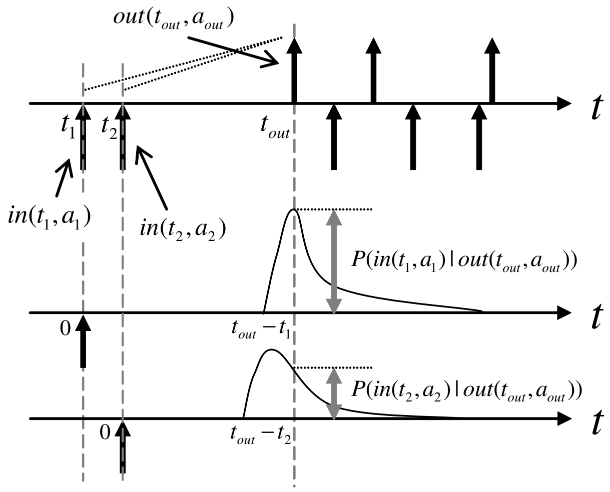


Figure 5: Overview of BEM

and outgoing events based on Eq. (6). Assume that when outgoing event $out(t_{out}, a_{out})$ is detected, there are $in(t_1, a_1)$ and $in(t_2, a_2)$ as candidates for incoming events. By using the distance from the incoming position to the outgoing position and the velocity following a normal distribution, $P(in(t_1, a_1)|out(t_{out}, a_{out}))$ becomes as follows.

$$P(in(t_1, a_1)|out(t_{out}, a_{out})) = w(a_1, a_{out})g(t_{out} - t_1)$$

Similarly, the probability that $in(t_2, a_2)$ corresponds to $out(t_{out}, a_{out})$ is derived by

$$P(in(t_2, a_2)|out(t_{out}, a_{out})) = w(a_2, a_{out})g(t_{out} - t_2)$$

From Fig. 5, we can see $P(in(t_2, a_2)|out(t_{out}, a_{out})) < P(in(t_1, a_1)|out(t_{out}, a_{out}))$. Therefore, we choose $in(t_1, a_1)$ as a correspondence of $out(t_{out}, a_{out})$.

4 EVALUATION

In this section, we evaluate the validity of the proposed model based on COM and BEM by comparing with the observation results in section 2. As an evaluation criterion, we define a reproduction ratio as the ratio of the number of pairs successfully estimated to the whole number of pairs. We set n to 128.

4.1 Reproduction ratio of COM

Reproduction ratio of COM is shown in Fig. 6. We change N from 1 to 8 according to the knowledge in section 2.4. $\bar{v}$ is set to 133 [cm/s], which is the average velocity derived from observation. We also use half and double of the average velocity, that is 65 [cm/s] and 130 [cm/s], as $\bar{v}$.

We expected the reproduction ratio was improved with increase of N. As shown in Fig. 6, however, the ratio oscillates rather than monotonically increases. We further find that the reproduction ratio of $\bar{v}$=266 is lower than others. This is caused by the estimation error of the average velocity. However, there are almost no differences between the results of $\bar{v}$=133 and those of $\bar{v}$=65.

In actual situations, it may be difficult for a system administrator to know $\bar{v}$ at the start of monitoring. To reduce the initial estimation error of $\bar{v}$, we introduce a mechanism to adapt $\bar{v}$ with

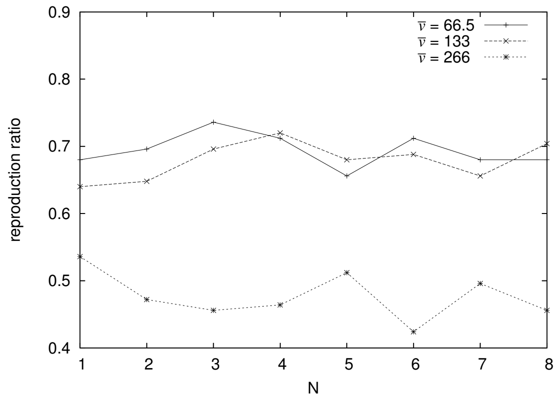


Figure 6: Reproduction ratio of COM without velocity modification

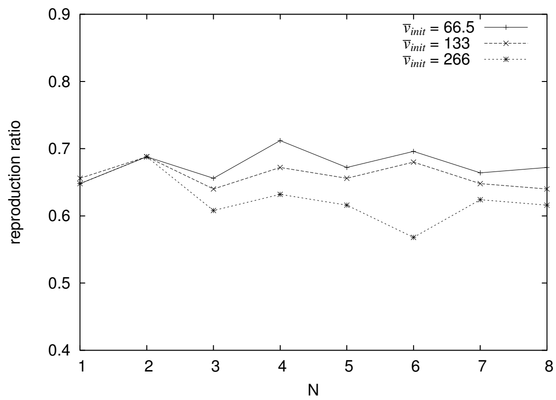


Figure 7: Reproduction ratio of COM with velocity modification

the observation result. Figure 7 shows the result of the reproduction ratio with this modification. $\bar{v}_{init}$ denotes the initial value of $\bar{v}$. As we expected, the reproduction ratio is improved with the modification when $\bar{v}_{init}$=266. However, this improvement does not appear in the case of $\bar{v}_{init}$=65, 133. This means that the velocity approximation with only the average is not sufficient to reproduce the node behavior. We tackle this problem in BEM by using the velocity following a normal distribution.

4.2 Reproduction ratio of BEM

By obtaining the average velocity $\bar{v}$, the variance σ^2, and $w(\cdot, \cdot)$ from the observation results, we could accomplish the reproduction ratio of 1, which means BEM can estimate the correspondences perfectly. To compare BEM with COM fairly in terms of the initially available information, we also evaluated BEM with uniform $w(\cdot, \cdot)$. In this case, the reproduction ratio was 0.664 that was close to the result of COM. Moreover, we can improve the reproduction ratio 0.336 by appropriately estimating $w(\cdot, \cdot)$ from the observation. We expect that it is relatively easier to calculate $\bar{v}$, σ^2, and $w(\cdot, \cdot)$ with a constant interval than to measure the velocity of each node. In summary, BEM is also favorable from the view point of observation overheads.

4.3 Node distribution

In this paper, we assume that the area is divided into $n \times n$ grids. Since we can't reproduce the node trajectories precisely with small n, we set n to a large number of 128. On the other hand, a resolution $m \times m$ ($m \leq n$) that is actually required depends on applications. In this paper, we set m to 16 and the grid

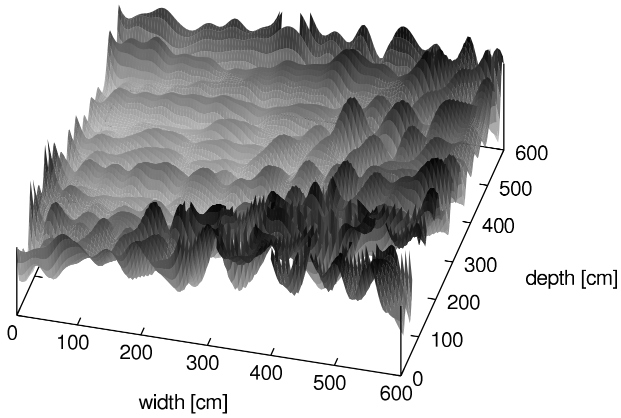


Figure 8: Node distribution based on observation

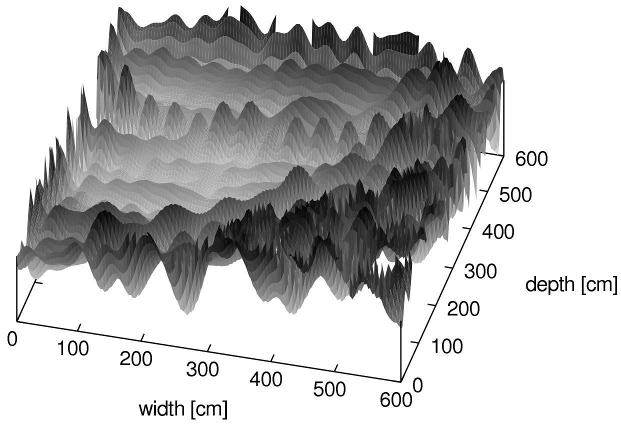


Figure 9: Node distribution based on COM

size to 37.5×37.5 [cm^2] which corresponds to a minimum area that a person occupies. We first divide the area $m \times m$ grids and then smooth each grid. For the smoothing, firstly we perform Fourier transformation to the original node distribution. Then, we make a filtering operation of $\frac{sin(x)}{x}$ to the transformed distribution. Finally, we derive a smoothed distribution by performing inverse Fourier transformation to the distribution. These operations are equal to tolerating the trajectory errors by expanding the width of the node trajectories to some extent. By lowering the resolution of the area with adjustment of m, we can alleviate the distance errors in section 2.2.

Figure 8 shows a smoothed distribution obtained by the observation. Figure 9 and Fig. 10 show those based on COM and BEM, respectively. We set N=3, $\bar{v}$=66.5, and no velocity modification in COM. Although the distributions based on COM and BEM are relatively similar to that obtained by the observation, BEM seems to be a little closer to the observation result.

5 CONCLUSIONS

In this paper, we proposed the mobility model based on the node incoming and outgoing to an area. Firstly, we showed a node moved approximately along with a straight line and the node velocities followed a normal distribution. Moreover, we proposed the COM and BEM which estimated the correspondences between incoming and outgoing node events. As a result, the reproduction ratio of COM was 0.736 while that of BEM was 1.

As future work, we plan to add a mechanism to change $w(\cdot,\cdot)$ in BEM dynamically based on the observation result and evaluate

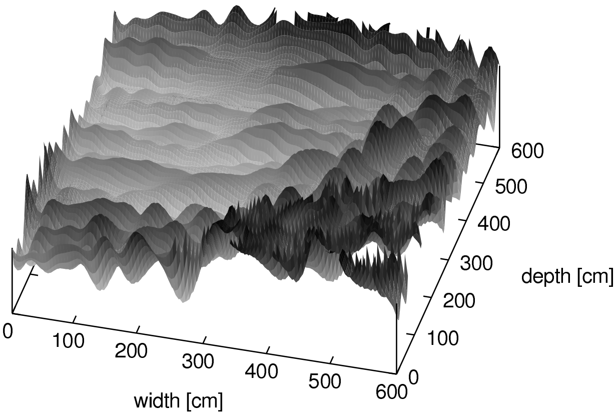


Figure 10: Node distribution based on BEM

the relationship between the reproduction ratio and the update interval. Furthermore, we have to evaluate the models at other locations and explore the feasible area of the proposed models from the view point of the number of nodes that simultaneously exist in the area.

REFERENCES

Ashtiani, F., Salehi, J. and Aref, M., 2003. Mobility Modeling and Analytical Solution for Spatial Traffic Distribution in Wireless Multimedia Networks. IEEE Journal on Selected Areas in Communications 21(10), pp. 1699–1709.

Bray, J. and Sturman, C., 2001. Bluetooth 1.1: Connect Without Cables. Prentice Hall PTR Upper Saddle River, NJ, USA.

Camp, T., Boleng, J. and Davies, V., 2002. A Survey of Mobility Models for Ad Hoc Network Research. Wireless Communications and Mobile Computing 2(5), pp. 483–502.

Fox, D., Hightower, J., Liao, L., Schulz, D. and Borriello, G., 2003. Bayesian Filtering for Location Estimation. IEEE Pervasive Computing 2, pp. 24–33.

Ito, T., 2007. Projective Transformation and Mobius Transformation. "`http://www1.kcn.ne.jp/~iittoo/us27d_infi.htm`".

Korte, B. and Vygen, J., 2005. Combinatorial Optimization: theory and algorithms. Springer.

Li, B. and Wang, K., 2003. NonStop: Continuous Multimedia Streaming in Wireless Ad Hoc Networks with Node Mobility. IEEE Journal on Selected Areas in Communications 21(10), pp. 1627–1641.

ZigBee Alliance, 2007. "`http://www.zigbee.org/`".