Submitted version

Continuous-Time Analysis of the Simple Averaging Scheme for Global Clock Synchronization in Sparsely Populated MANETs

Masahiro Sasabe *Member, IEEE,* and Tetsuya Takine, *Member, IEEE*

***Abstract*—In sparsely populated mobile ad hoc networks (MANETs), mobile nodes are chronically isolated each other and they meet very occasionally. Global clock synchronization among nodes in such networks is a challenging problem because reference clock information cannot be disseminated promptly over nodes due to the lack of stable connections among nodes. In recent years, distributed global clock synchronization based on consensus algorithms has been studied. In this paper, we conduct the continuous-time analysis of the simplest consensus-based clock synchronization, where two mobile nodes exchange their local clock times when they meet and adjust their own clocks to the average of them. Through the analysis and simulation experimennts, we reveal how the clock accuracy of nodes and meeting rates among them affect the rate of convergence to the steady state and the accuracy of clock synchronization in steady state.**

***Index Terms*—Continuous-time analysis, averaging scheme, global clock synchronization, sparsely populated MANETs.**

I. INTRODUCTION

In challenged networks such as sparsely populated mobile ad hoc networks (MANETs) and mobile wireless sensor networks (WSNs), mobile nodes are chronically isolated each other and they meet very occasionally. Such a kind of networks typically arise in deep-space exploration, wildlife tracking, underwater networking, and emergency networking in disaster areas [7]. To realize effective networking in those situations, global clock synchronization is one of key issues [3], [17].

The local clock time $c(t)$ of a node at real time t ($t \geq 0$) can be expressed to be [16]

$$c(t) = \rho t + \phi, \tag{1}$$

where ρ and ϕ are called the *clock rate* and *clock offset*, respectively. The clock rate ρ is equivalent to the first derivative $dc(t)/dt$ of $c(t)$ and ideally, it should be equal to one. In practice, however, ρ differs from one in the range of $[10^{-4}, 10^{-6}]$, due to crystal inaccuracies and short- and long-term environmental variations such as temperature and aging [16]. The difference $\rho - 1$ of the clock rate from one is called *clock skew* (or drift). On the other hand, the clock offset $\phi = c(0)$ represents the initial error of the local clock time when the clock was initialized, which is not equal to zero in general.

The authors are with the Department of Information and Communications Technology, Graduate School of Engineering, Osaka University, Suita 565-0871, Japan (e-mail: {sasabe, takine}@comm.eng.osaka-u.ac.jp).

This research was supported in part by Grant-in-Aid for Young Scientists (B) of JSPS (No. 22700070) and Grant-in-Aid for Scientific Research (C) of JSPS (No. 22500056).

Manuscript submitted November 28, 2012.

There are extensive studies on global clock synchronization in multi-hop wireless networks and the surveys are given in [6], [14]. If the network is static or stable, the simplest way is to form a hierarchical topology rooted by a reference node and to broadcast the clock time of the reference node to all other nodes along with the topology. This category of global time synchronization schemes includes Network Time Protocol (NTP) [11] and its extension [17], tree-based approach [9], [15], and cluster-based approach [5]. These approaches, however, do not work well in challenged networks due to the following reasons: 1) Making and maintaining the hierarchical topology is difficult due to sparse node density, node mobility, and node failures, and 2) estimation errors increase with the number of hops from the reference node.

To tackle these problems, there are several studies on distributed global clock synchronization based on consensus/agreement algorithms [3], [10], [13]. The consensus algorithm enables a large number of distributed nodes to reach consensus on a common value, e.g., the global average among their local values, in a fully distributed manner [1]. Li and Rus first proposed consensus-based clock synchronization for offset compensation [10], where each node periodically calculates the average of clock times among neighbors. In recent years, more sophisticated consensus-based clock synchronization schemes have also been proposed to both drift compensation and offset compensation [3], [13].

The original consensus algorithm has a good convergence property under some assumptions. However, there is a significant difference between the original consensus problem and the global clock synchronization problem. More specifically, the local clock time at each node varies due to not only the consensus algorithm but also clock itself. The existing schemes [3], [10], [13], however, implicitly ignore the clock-driven effect and the consensus-based clock synchronization is modeled as a discrete-time system, where each node performs averaging operations at least once at each time step. As a result, the convergence of the local clock times is proved in a way similar to the original consensus problem.

In actual systems, the clock changes continuously, as shown in (1). Furthermore, the meeting rate (which is determined by communication intervals) between a pair of nodes depends on those nodes' mobility. To the best of our knowledge, there is no study on revealing how the clock drift, clock offset, and meeting rate affect the rate of convergence to the steady state and the accuracy of clock synchronization in steady state. In this paper, we clarify these fundamental characteristics by conducting the continuous-time analysis of the simplest consensus-based clock synchronization called the *simple averaging scheme*, where two mobile nodes exchange their local clock times when they meet and adjust their own clocks to be the average of them.

The rest of this paper is organized as follows. Section II describes the simple averaging scheme for global clock synchronization. Section III provides the continuous-time analysis of the simple averaging scheme. Section IV provides some numerical illustrations with simulation experiments and demonstrate the fundamental characteristics of the simple averaging scheme and the practicality of the analytical results using real trace data. Finally, conclusion is provided in section V.

II. Simple Averaging Scheme for Global Clock Synchronization

Suppose there exist N ($N \geq 2$) mobile nodes, labeled 1 to N, in a closed region. Let $\mathcal{N}$ denote $\{1, 2, \dots, N\}$. The clock rate and offset time of node k ($k \in \mathcal{N}$) are denoted by ρ_k and ϕ_k, respectively. We define $c_k(t)$ ($k \in \mathcal{N}$, $t \geq 0$) as the local clock time of node k at real time t. Thus $c_k(0) = \phi_k$ ($k \in \mathcal{N}$), and when the clock of node k is not adjusted during a time interval $(t_1, t_2]$ ($0 \leq t_1 < t_2$),

$$c_k(t_2) = \rho_k(t_2 - t_1) + c_k(t_1). \tag{2}$$

We assume that node k ($k \in \mathcal{N}$) has an estimated value $\hat{\rho}_k$ of its clock rate ρ_k and utilizes it for estimating real time t from its local clock time $c_k(t)$ in the following way. Let $\hat{c}_k(t)$ ($k \in \mathcal{N}$, $t \geq 0$) denote the estimated time of node k at time t. Also let $\tau_k(t)$ ($k \in \mathcal{N}$, $t \geq 0$) denote a time instant at which the clock of node k was adjusted last time before time t, where $\tau_k(t) = 0$ if the clock has never been adjusted during $(0, t]$. Note that whenever the clock of node k is adjusted, the estimated time is also set to be the same time, i.e., $\hat{c}_k(\tau_k(t)) = c_k(\tau_k(t))$. Node k generates the estimated time $\hat{c}_k(t)$ by

$$\hat{c}_k(t) = \frac{c_k(t) - c_k(\tau_k(t))}{\hat{\rho}_k} + c_k(\tau_k(t)), \qquad t \geq 0. \tag{3}$$

It then follows from (2) and (3) that the estimated time $\hat{c}_k(t)$ at time t is given by $\hat{c}_k(t) = \rho_k^*(t - \tau_k(t)) + \hat{c}_k(\tau_k(t))$, where ρ_k^* ($k \in \mathcal{N}$) denotes the adjusted clock rate of node k, i.e., $\rho_k^* = \rho_k/\hat{\rho}_k$. Note that the default value of $\hat{\rho}_k$ is equal to one. In what follows, we assume that $c_k(t)$ and $\hat{c}_k(t)$ are right-continuous and have left-limits, and we denote the left-limits of $c_k(t)$ and $\hat{c}_k(t)$ at time t by $c_k(t-)$ and $\hat{c}_k(t-)$, respectively.

Suppose node k and node j ($k, j \in \mathcal{N}$, $k \neq j$) meet at time $t = \tau$. They instantaneously exchange their estimated clock times and adjust their clock to be the average of them.

$$c_k(\tau) = c_j(\tau) = \hat{c}_k(\tau) = \hat{c}_j(\tau) = \frac{\hat{c}_k(\tau-) + \hat{c}_j(\tau-)}{2}. \tag{4}$$

It then follows from (4) that $\hat{c}_k(\tau) + \hat{c}_j(\tau) = \hat{c}_k(\tau-) + \hat{c}_j(\tau-)$, so that the sum of estimated times does not change before and after meetings of nodes. Note also that the sum of estimated times of all nodes increase at constant rate $\rho_1^* + \rho_2^* + \cdots + \rho_N^*$ unless the estimated clock rate $\hat{\rho}_k$ is updated. Therefore we have

$$\begin{aligned}\sum_{k \in \mathcal{N}} \hat{c}_k(t) &= \left(\sum_{k \in \mathcal{N}} \rho_k^*\right) t + \sum_{k \in \mathcal{N}} \hat{c}_k(0) \\ &= \left(\sum_{k \in \mathcal{N}} \rho_k^*\right) t + \sum_{k \in \mathcal{N}} \phi_k,\end{aligned} \tag{5}$$

for all t ($t \geq 0$), because (3) implies $\hat{c}_k(0) = c_k(0) = \phi_k$ ($k \in \mathcal{N}$).

Remark 1: In the above formulation, we ignore the propagation delay between two nodes because it will be negligible compared with the achievable granularity of clock synchronization. If the estimated propagation delay is available (cf. [12]), however, it can be incorporated into the simple averaging scheme.

We now define the *reference time* $c^*(t)$ ($t \geq 0$) as

$$c^*(t) = \frac{1}{N} \sum_{k \in \mathcal{N}} \hat{c}_k(t) = \overline{\rho} t + \overline{\phi}, \tag{6}$$

where

$$\overline{\rho} = \frac{1}{N} \sum_{k \in \mathcal{N}} \rho_k^*, \quad \overline{\phi} = \frac{1}{N} \sum_{k \in \mathcal{N}} \phi_k.$$

From (5) and (6), we have the following theorem.

Theorem 2: The sum of differences between the estimated time $\hat{c}_k(t)$ of each node and the reference time $c^*(t)$ is always zero.

$$\sum_{k \in \mathcal{N}} (\hat{c}_k(t) - c^*(t)) = 0, \quad \forall t \geq 0.$$

III. Continuous-Time Analysis of Simple Averaging Scheme

In this section, we analyze the performance of the simple averaging scheme under the assumption that the adjusted clock rate ρ_k^* ($k \in \mathcal{N}$) of node k is fixed. Let s_k ($k \in \mathcal{N}$) denote the difference between the adjusted clock rate ρ_k^* of node k and the clock rate $\overline{\rho}$ of the reference time.

$$s_k = \rho_k^* - \overline{\rho}.$$

Hereafter, we call s_k the *relative clock skew* of node k. Further, we define S as a random variable representing the relative clock skew of a randomly chosen node, i.e., $\Pr(S = s_k) = 1/N$ ($k \in \mathcal{N}$). Note here that by the definition of $\overline{\rho}$,

$$\sum_{k \in \mathcal{N}} s_k = 0, \tag{7}$$

and therefore $\mathrm{E}[S] = 0$.

Let $X_k(t)$ ($k \in \mathcal{N}$, $t \geq 0$) denote the difference between the estimated time $\hat{c}_k(t)$ of node k and the reference time $c^*(t)$ at time t.

$$X_k(t) = \hat{c}_k(t) - c^*(t).$$

Hereafter, we call $X_k(t)$ the *relative time difference* of node k at time t. It follows from Theorem 2 that

$$\sum_{k \in \mathcal{N}} X_k(t) = 0, \tag{8}$$

for all $t \geq 0$. Thus $X_k(t)$'s ($k \in \mathcal{N}$) are dependent random variables. We then define $\boldsymbol{X}(t)$ ($t \geq 0$) as an N-dimensional random vector of $X_k(t)$ ($k \in \mathcal{N}$), i.e., $\boldsymbol{X}(t) = (X_1(t), X_2(t), \dots, X_N(t))$, and let $F(\boldsymbol{x}, t)$ ($t \geq 0$) denote the joint distribution function of $X_k(t)$ ($k \in \mathcal{N}$),

$$F(\boldsymbol{x}, t) = \Pr(X_k(t) \leq x_k; k \in \mathcal{N}),$$

where $\boldsymbol{x} = (x_1, x_2, \dots, x_N)$. By definition, $F(\boldsymbol{x}, 0)$ represents the joint distribution of relative offset differences $X_k(0) = \phi_k - \overline{\phi}$ ($k \in \mathcal{N}$), which are treated as random variables in our formulation. We assume that $\boldsymbol{X}(t)$ is ergodic and let X_k ($k \in \mathcal{N}$) denote the limiting random variable of $X_k(t)$.

$$X_k = \lim_{t \to \infty} X_k(t).$$

We also define $X(t)$ ($t \geq 0$) as the relative time difference of a randomly chosen node at time t, i.e., $\Pr(X(t) = X_k(t)) = \frac{1}{N}$ ($k \in \mathcal{N}$). Let X denote the limiting random variable of $X(t)$.

$$X = \lim_{t \to \infty} X(t).$$

A. General N-Node Systems

We assume that there exist N nodes ($N \geq 2$). To make things tractable, we assume that the sequence of time instants at which node k and node j ($k, j \in \mathcal{N}$, $k \neq j$) meet forms an independent Poisson process with rate $\lambda_{k,j}$. Thus the sequence of time instants at which node k ($k \in \mathcal{N}$) encounters other nodes forms an independent Poisson process with rate λ_k, where λ_k is given by

$$\lambda_k = \sum_{\substack{j \in \mathcal{N} \\ j \neq k}} \lambda_{k,j}, \qquad k \in \mathcal{N}.$$

Let $\boldsymbol{\Lambda}$ denote an $N \times N$ matrix whose kth ($k \in \mathcal{N}$) diagonal element is given by $-\lambda_k$ and the (k, j)th ($k, j \in \mathcal{N}, k \neq j$) off-diagonal element is given by $\lambda_{k,j}$. Note that $\boldsymbol{\Lambda}$ is considered as an infinitesimal generator of a continuous-time Markov chain with finite state space $\mathcal{N}$. We assume that $\boldsymbol{\Lambda}$ is irreducible. Let $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_N)$ denote a $1 \times N$ invariant probability vector of $\boldsymbol{\Lambda}$, which is determined uniquely by $\boldsymbol{\pi}\boldsymbol{\Lambda} = \boldsymbol{0}$ and $\boldsymbol{\pi}\boldsymbol{e} = 1$, where $\boldsymbol{e}$ denotes an $N \times 1$ vector whose elements are all equal to one. By definition, $\lambda_{k,j} = \lambda_{j,k}$, so that $\boldsymbol{\Lambda}$ is a symmetric matrix. Thus we have $\boldsymbol{\pi} = (1/N) \cdot \boldsymbol{e}^T$, where $\cdot^T$ stands for the transpose operator. It then follows that $\pi_k \lambda_{k,j} = \pi_j \lambda_{j,k}$ ($k, j \in \mathcal{N}, k \neq j$), and therefore the finite state Markov chain defined by $\boldsymbol{\Lambda}$ is reversible (cf. Chap. 1 in [8]).

Let $f^*(\boldsymbol{\eta}, t)$ ($t \geq 0$) denote the joint characteristic function of $X_k(t)$ ($k \in \mathcal{N}$).

$$f^*(\boldsymbol{\eta}, t) = \mathrm{E}\left[\prod_{k \in \mathcal{N}} e^{\imath \eta_k X_k(t)}\right],$$

where $\boldsymbol{\eta} = (\eta_1, \eta_2, \dots, \eta_N)$ and $\imath$ denotes the imaginary unit.

Theorem 3: The joint characteristic function $f^*(\boldsymbol{\eta}, t)$ ($t \geq 0$) satisfies

$$\frac{\partial}{\partial t} f^*(\boldsymbol{\eta}, t) = \sum_{k \in \mathcal{N}} \left(\imath s_k \eta_k - \frac{1}{2}\lambda_k\right) \cdot f^*(\boldsymbol{\eta}, t) + \sum_{\substack{k,j \in \mathcal{N} \\ k<j}} \lambda_{k,j} f^*(\boldsymbol{\eta}_{k,j}, t), \tag{9}$$

where $\boldsymbol{\eta}_{k,j}$ ($k, j \in \mathcal{N}$, $k < j$) is given by

$$\boldsymbol{\eta}_{k,j} = \left(\eta_1, \dots, \eta_{k-1}, \frac{\eta_k + \eta_j}{2}, \eta_{k+1}, \dots, \eta_{j-1}, \frac{\eta_k + \eta_j}{2}, \eta_{j+1}, \dots, \eta_N\right).$$

The proof of Theorem 3 is given in Appendix A. Unfortunately, it seems to be hard to obtain the solution for $F(\boldsymbol{x}, t)$ from (9). Thus we consider moments of $X_k(t)$ ($k \in \mathcal{N}$) below.

We first discuss the first moment of the relative time difference $X_k(t)$. Let $f_k^*(\eta, t)$ ($k \in \mathcal{N}$, $t \geq 0$) and $f_{k,j}^*(\eta_k, \eta_j, t)$ ($k, j \in \mathcal{N}$, $t \geq 0$) denote the marginal characteristic function of $X_k(t)$ and the marginal joint characteristic function of $X_k(t)$ and $X_j(t)$, respectively.

$$f_k^*(\eta, t) = \mathrm{E}\left[e^{\imath \eta X_k(t)}\right],$$
$$f_{k,j}^*(\eta_k, \eta_j, t) = \mathrm{E}\left[e^{\imath \eta_k X_k(t)} e^{\imath \eta_j X_j(t)}\right].$$

Substituting $\eta_j = 0$ for all j ($j \in \mathcal{N}$, $j \neq k$) in (9) yields

$$\frac{\partial}{\partial t} f_k^*(\eta, t) = \imath s_k \eta f_k^*(\eta, t) - \lambda_k f_k^*(\eta, t) + \sum_{\substack{j \in \mathcal{N} \\ j \neq k}} \lambda_{k,j} f_{k,j}^*\left(\frac{\eta}{2}, \frac{\eta}{2}, t\right). \tag{10}$$

Let $\boldsymbol{s}$ denote an $N \times 1$ vector whose kth ($k \in \mathcal{N}$) element is given by s_k. We define $\boldsymbol{x}^{(n)}(t)$ ($n = 1, 2$, $t \geq 0$) as an $N \times 1$ vector whose kth ($k \in \mathcal{N}$) element $x_k^{(n)}(t)$ is given by $\mathrm{E}[X_k^n(t)]$. By definition,

$$x_k^{(1)}(t) = \mathrm{E}[X_k(t)] = -\imath \lim_{\eta \to 0} \frac{\partial}{\partial \eta} f_k^*(\eta, t).$$

Theorem 4: $\boldsymbol{x}^{(1)}(t)$ is expressed to be

$$\boldsymbol{x}^{(1)}(t) = \boldsymbol{x}^{(1)} + \exp\left(\frac{1}{2}\boldsymbol{\Lambda} t\right) \left(\boldsymbol{x}^{(1)}(0) - \boldsymbol{x}^{(1)}\right), \tag{11}$$

where $\boldsymbol{x}^{(1)} = \lim_{t \to \infty} \boldsymbol{x}^{(1)}(t)$ is given by

$$\boldsymbol{x}^{(1)} = 2\left(\frac{1}{N}\boldsymbol{e}\boldsymbol{e}^T - \boldsymbol{\Lambda}\right)^{-1} \boldsymbol{s}. \tag{12}$$

The proof of Theorem 4 is given in Appendix B.

Recall that $\boldsymbol{\Lambda}$ is considered as a symmetric infinitesimal generator of a finite-state reversible Markov chain. Therefore all eigenvalues of $\boldsymbol{\Lambda}$ are real, the maximum eigenvalue is equal to 0, and other eigenvalues are strictly negative. Therefore, from (11), we observe that the influence of relative offset times $x_k^{(1)}(0)$ ($k \in \mathcal{N}$) decays exponentially and the relaxation time is given by the reciprocal of the absolute value of the second largest eigenvalue of $\boldsymbol{\Lambda}/2$, which is independent of relative clock skews.

Remark 5: The second largest eigenvalues of infinitesimal generators of reversible Markov chains have been studied in depth, e.g., see [4] and references therein.

Next we consider the time-dependent second moment $\mathrm{E}[X_k^2(t)]$ $(k \in \mathcal{N})$. Let $f^*_{k,j,i}(\eta_k, \eta_j, \eta_i, t)$ $(k, j, i \in \mathcal{N},\ t \geq 0)$ denote

$$f^*_{k,j,i}(\eta_k, \eta_j, \eta_i, t) = \mathrm{E}\left[e^{\imath \eta_k X_k(t)} e^{\imath \eta_j X_j(t)} e^{\imath \eta_i X_i(t)}\right].$$

Substituting $\eta_i = 0$ for all i $(i \in \mathcal{N},\ i \neq k, j)$ in (9) yields

$$\begin{aligned}\frac{\partial}{\partial t} f^*_{k,j}(\eta_k, \eta_j, t) &= \imath s_k \eta_k f^*_{k,j}(\eta_k, \eta_j, t) + \imath s_j \eta_j f^*_{k,j}(\eta_k, \eta_j, t) \\ &\quad - (\lambda_k + \lambda_j - \lambda_{k,j}) f^*_{k,j}(\eta_k, \eta_j, t) \\ &\quad + \lambda_{k,j} f^*_{k,j}\left(\frac{\eta_k + \eta_j}{2}, \frac{\eta_k + \eta_j}{2}, t\right) \\ &\quad + \sum_{\substack{i \in \mathcal{N} \\ i \neq k,j}} \lambda_{k,i} f^*_{k,j,i}\left(\frac{\eta_k}{2}, \eta_j, \frac{\eta_k}{2}, t\right) \\ &\quad + \sum_{\substack{i \in \mathcal{N} \\ i \neq k,j}} \lambda_{j,i} f^*_{k,j,i}\left(\eta_k, \frac{\eta_j}{2}, \frac{\eta_j}{2}, t\right). \end{aligned} \tag{13}$$

We define $x^{(1,1)}_{k,j}(t)$ $(k, j \in \mathcal{N})$ as $\mathrm{E}[X_k(t) X_j(t)]$. By definition,

$$x^{(1,1)}_{k,j}(t) = -\lim_{\eta_k \to 0} \lim_{\eta_j \to 0} \frac{\partial}{\partial \eta_k} \frac{\partial}{\partial \eta_j} f^*_{k,j}(\eta_k, \eta_j, t),$$
$$x^{(2)}_k(t) = \mathrm{E}[X_k^2(t)] = -\lim_{\eta \to 0} \frac{\partial^2}{\partial \eta^2} f^*_k(\eta, t).$$

Theorem 6: $\mathrm{E}[X_k^2(t)]$ $(k \in \mathcal{N})$ is given by

$$\mathrm{E}[X_k^2(t)] = -\left(\sum_{j=1}^{k-1} x^{(1,1)}_{j,k}(t) + \sum_{j=k+1}^{N} x^{(1,1)}_{k,j}(t)\right),$$

where $x^{(1,1)}_{k,j}(t)$ $(1 \leq k < j \leq N)$ is the solution of the following system of $N(N-1)/2$ linear ordinary differential equations.

$$\begin{aligned}\frac{d}{dt} x^{(1,1)}_{k,j}(t) &= \sum_{l=1}^{k-1} \left\{ \left(\frac{\lambda_{l,j}}{2} + \frac{\lambda_{k,j}}{4}\right) x^{(1,1)}_{l,k}(t) + \left(\frac{\lambda_{l,k}}{2} + \frac{\lambda_{k,j}}{4}\right) x^{(1,1)}_{l,j}(t) \right\} \\ &\quad + \sum_{l=k+1}^{j-1} \left\{ \left(\frac{\lambda_{l,j}}{2} + \frac{\lambda_{k,j}}{4}\right) x^{(1,1)}_{k,l}(t) + \left(\frac{\lambda_{k,l}}{2} + \frac{\lambda_{k,j}}{4}\right) x^{(1,1)}_{l,j}(t) \right\} \\ &\quad + \sum_{l=j+1}^{N} \left\{ \left(\frac{\lambda_{j,l}}{2} + \frac{\lambda_{k,j}}{4}\right) x^{(1,1)}_{k,l}(t) + \left(\frac{\lambda_{k,l}}{2} + \frac{\lambda_{k,j}}{4}\right) x^{(1,1)}_{j,l}(t) \right\} \\ &\quad - \frac{1}{2}(\lambda_k + \lambda_j - 2\lambda_{k,j}) x^{(1,1)}_{k,j}(t) \\ &\quad + s_k x^{(1)}_j(t) + s_j x^{(1)}_k(t), \\ &\qquad 1 \leq k < j \leq N, \end{aligned} \tag{14}$$

which can be solved numerically.
The proof of Theorem 6 is given in Appendix C.

Note that all coefficients of unknowns $x^{(1,1)}_{k,j}(t)$ in (14) are independent of relative clock skews s_k. Thus the convergence speed of the second moments to the limiting values is determined only by $\lambda_{k,j}$ $(k, j \in \mathcal{N},\ k \neq j)$, and it is independent of relative clock skews.

It is easy to see that the limiting second moments $\mathrm{E}[X_k^2]$ $(k \in \mathcal{N})$ are given by

$$\mathrm{E}[X_k^2] = -\left(\sum_{j=1}^{k-1} x^{(1,1)}_{j,k} + \sum_{j=k+1}^{N} x^{(1,1)}_{k,j}\right),$$

where $x^{(1,1)}_{k,j} = \lim_{t\to\infty} x^{(1,1)}_{k,j}(t)$ $(k, j \in \mathcal{N},\ k < j)$ denotes the solution of a system of linear simultaneous equations, which is obtained by taking limit $t \to \infty$ on both sides of (14), where $\lim_{t\to\infty} dx^{(1,1)}_{k,j}(t)/dt = 0$.

Because the above formulation for the general case provides us only with limited insight, we consider two special cases below.

B. Special Case 1: Homogeneous Meetings

We now assume that all nodes are homogeneous in terms of their meetings.

Corollary 7: Suppose $\lambda_{k,j} = \lambda$ for all k, j $(k, j \in \mathcal{N},\ k \neq j)$. We then have for $k \in \mathcal{N}$,

$$\mathrm{E}[X_k(t)] = \mathrm{E}[X_k] + (\mathrm{E}[X_k(0)] - \mathrm{E}[X_k])\, e^{-\frac{N\lambda}{2} t}, \tag{15}$$

where

$$\mathrm{E}[X_k] = \lim_{t\to\infty} \mathrm{E}[X_k(t)] = \frac{2 s_k}{N\lambda}, \qquad k \in \mathcal{N}.$$

The proof of Corollary 7 is postponed till section III-C.

In the homogeneous meeting case, the relaxation time γ is given by $\gamma = 2/(N\lambda)$, which is independent of relative clock skews, as shown in the general case. Note that when N is large, $\gamma \approx 2 \times \{(N-1)\lambda\}^{-1}$, which is approximately equal to the mean length of time which a specific node takes to encounter other nodes twice. Thus the relaxation time is fairly short and we expect the fast convergence to steady state. Also the mean relative time difference $\mathrm{E}[X_k(t)]$ of node k converges to $\mathrm{E}[X_k] = \gamma s_k$ as t goes to infinity, and this limiting value depends on the relative clock skews of other nodes through the constraint of (7).

Theorem 8: Suppose $\lambda_{k,j} = \lambda$ for all k, j $(k, j \in \mathcal{N},\ k \neq j)$. We then have for $k \in \mathcal{N}$,

$$\mathrm{E}[X_k^2] = \lim_{t\to\infty} \mathrm{E}[X_k^2(t)] = \frac{8}{3N^2\lambda^2}\left(2 s_k^2 + \mathrm{E}[S^2]\right), \tag{16}$$

$$\mathrm{E}[X^2] = \lim_{t\to\infty} \mathrm{E}[X^2(t)] = \frac{8\mathrm{E}[S^2]}{N^2\lambda^2}, \tag{17}$$

and

$$\begin{aligned}\mathrm{E}[X_k^2(t)] &= \mathrm{E}[X_k^2] + (2C_k + D_k - E_k) e^{-\frac{3N\lambda}{4} t} \\ &\quad + (A - 3C_k + E_k) e^{-\frac{N\lambda}{2} t} + B t e^{-\frac{N\lambda}{2} t}, \end{aligned} \tag{18}$$

$$\mathrm{E}[X^2(t)] = \mathrm{E}[X^2] + A e^{-\frac{N\lambda}{2} t} + B t e^{-\frac{N\lambda}{2} t}, \tag{19}$$

where

$$A = \mathrm{E}[X^2(0)] - \mathrm{E}[X^2],$$

$$B = \frac{2}{N} \sum_{j \in \mathcal{N}} s_j (\mathrm{E}[X_j(0)] - \mathrm{E}[X_j]),$$
$$C_k = \mathrm{E}[X_k^2] - \mathrm{E}[X^2],$$
$$D_k = \mathrm{E}[X_k^2(0)] - \mathrm{E}[X^2(0)],$$
$$E_k = 4 \left(\mathrm{E}[X_k]\mathrm{E}[X_k(0)] - \frac{1}{N} \sum_{j \in \mathcal{N}} \mathrm{E}[X_j]\mathrm{E}[X_j(0)] \right).$$

The proof of Theorem 8 is given in Appendix D.

As observed in the general case, the speed of convergence is independent of relative clock skews. We also observe that the second moment $\mathrm{E}[X_k^2]$ of the limiting relative time difference of node k is composed of two factors, one of which is the common factor $\mathrm{E}[X^2]/3$ determined by overall system parameters. The other factor is given by $4(\mathrm{E}[X_k])^2/3$ (see Corollary 7), which is proportional to the square s_k^2 of the relative clock skew of node k. Thus even if $s_k = 0$, node k is influenced by other nodes with non-zero relative clock skews.

Next we discuss the transient behavior of the second moments. From (19), we observe that the second moment $\mathrm{E}[X^2(t)]$ of the relative time difference of a randomly chosen node has two decay terms containing factors $\exp(-t/\gamma)$ and $t\exp(-t/\gamma)$, respectively. Note that the second decay term stems from the transient dynamics of $\mathrm{E}[X_k(t)]$ ($k \in \mathcal{N}$). In fact, if $\mathrm{E}[X_k(0)] = \mathrm{E}[X_k]$ for all $k \in \mathcal{N}$, the second term vanishes away and $\mathrm{E}[X_k(t)] = \mathrm{E}[X_k]$ for all $t \geq 0$. Even when $\mathrm{E}[X_k(0)] = \mathrm{E}[X_k]$ for all $k \in \mathcal{N}$, however, the first decay term remains, so that it is the essential term representing the transient behavior of the second moment $\mathrm{E}[X^2(t)]$ of the relative time difference of a randomly chosen node.

We then turn our attention to the time-dependent second moment $\mathrm{E}[X_k^2(t)]$ of the relative time difference of node k. From (18) and (19), we have

$$\begin{aligned} &\mathrm{E}[X_k^2(t)] - \mathrm{E}[X_k^2] \\ &= \mathrm{E}[X^2(t)] - \mathrm{E}[X^2] \\ &\quad + (2C_k + D_k - E_k)e^{-\frac{3N\lambda}{4}t} - (3C_k - E_k)e^{-\frac{N\lambda}{2}t}, \quad (20) \end{aligned}$$

which shows that the transient dynamics of $\mathrm{E}[X_k^2(t)]$ is composed of two factors, one of which is determined by overall system parameters. We also observe from (20) that $\mathrm{E}[X_k^2(t)]$ has two additional terms decaying exponentially at rates $1/\gamma$ and $3/(2\gamma)$, respectively. Note here that C_k, D_k, and E_k in those terms satisfy

$$\sum_{k \in \mathcal{N}} C_k = \sum_{k \in \mathcal{N}} D_k = \sum_{k \in \mathcal{N}} E_k = 0,$$

so that they do not appear in $\mathrm{E}[X^2(t)]$ of (19). Thus they are considered as decay terms stemming from the diversity of respective nodes.

C. Special Case 2: N-Node System with One Active Node

Next we consider a situation where meetings of nodes are not homogeneous: node 1 moves around actively and encounter other nodes frequently, while interactions among other $N-1$ nodes are homogeneous. More specifically, the sequence of meeting times of node k and node j ($k, j \in \mathcal{N}$, $k, j \neq 1$, $k \neq j$) is assumed to be an independent Poisson process with rate λ_1. We also assume that the sequence of meeting times of node 1 and node k ($k \in \mathcal{N}$, $k \neq 1$) forms an independent Poisson process with rate $\lambda^* = \lambda_1 + \lambda_2$. Note that the model is reduced to Special Cases 1 when $\lambda_2 = 0$ and we are mainly interested in the case of $\lambda_2 > 0$, even though the following analytical results are valid for all $\lambda_2 > -\lambda_1$ (i.e., $\lambda^* > 0$).

Corollary 9: Suppose $\lambda_{k,j}$ ($k, j \in \mathcal{N}$, $k \neq j$) is given by

$$\lambda_{k,j} = \begin{cases} \lambda^* = \lambda_1 + \lambda_2, & k = 1 \text{ or } j = 1, \\ \lambda_1, & \text{otherwise.} \end{cases}$$

We then have

$$\mathrm{E}[X_1(t)] = \mathrm{E}[X_1] + (\mathrm{E}[X_1(0)] - \mathrm{E}[X_1])\, e^{-\frac{N\lambda^*}{2}t}, \quad (21)$$

$$\begin{aligned} \mathrm{E}[X_k(t)] &= \mathrm{E}[X_k] + (\mathrm{E}[X_k(0)] - \mathrm{E}[X_k])\, e^{-\frac{N\lambda_1 + \lambda_2}{2}t} \\ &\quad + \frac{\mathrm{E}[X_1(0)] - \mathrm{E}[X_1]}{N-1} \left(e^{-\frac{N\lambda_1 + \lambda_2}{2}t} - e^{-\frac{N\lambda^*}{2}t} \right), \\ &\qquad k \neq 1, \quad (22) \end{aligned}$$

where $\mathrm{E}[X_k] = \lim_{t \to \infty} \mathrm{E}[X_k(t)]$ ($k \in \mathcal{N}$) is given by

$$\mathrm{E}[X_k] = \begin{cases} \dfrac{2s_1}{N\lambda^*}, & k = 1, \\ \dfrac{2s_k}{N\lambda_1 + \lambda_2} + \dfrac{\lambda_2}{N\lambda_1 + \lambda_2} \cdot \dfrac{2s_1}{N\lambda^*}, & k \neq 1. \end{cases} \quad (23)$$

The proof of Corollary 9 is given in Appendix E.

It is interesting to observe that the time-dependent mean relative time difference $\mathrm{E}[X_1(t)]$ and the limiting mean relative time difference $\mathrm{E}[X_1]$ of node 1 take the same forms as those in Special Case 1 of the homogeneous meeting model. Note that node 1 encounters other nodes randomly with equal probabilities and the rate of meetings of node 1 to a specific node is given by λ^*. Therefore this result suggests that the mean relative time difference of a specific node is determined mainly by the frequency of meetings to other nodes and their relative clock skews, and interactions among other nodes have little impact on the specific node.

Next we consider node k ($k \neq 1$). Because Special Case 2 for $N = 2$ is equivalent to Special Case 1, we assume $N \geq 3$. The relaxation time $\gamma_k = 2/(N\lambda_1 + \lambda_2)$ of node k ($k \neq 1$) is longer (resp. shorter) than the relaxation time $\gamma_1 = 2/(N\lambda^*)$ of node 1 when $\lambda^* > \lambda_1$ (resp. $\lambda^* < \lambda_1$). This suggests that when meeting frequencies among nodes are different, the relaxation time $\max_{k \in \mathcal{N}} \gamma_k$ of the system is determined by the node with the least meeting frequency.

In the system with one active node 1 (i.e., $\lambda_2 > 0$), meetings of node k ($k \in \mathcal{N}$, $k \neq 1$) to other nodes can be divided into two sorts. One is the homogeneous meeting to other nodes and the meeting rate to each pair of nodes is given by λ_1. The other is the addition meeting to node 1 with rate λ_2. As a result, $\mathrm{E}[X_k(t)]$ ($k \neq 1$) has a term involving $\mathrm{E}[X_1(0)] - \mathrm{E}[X_1]$, which is common for all k ($k \neq 1$). Similarly, $\mathrm{E}[X_k]$ is composed of two terms, one is the product of the relaxation time γ_k and the relative clock skew s_k, and the other is a term involving $\mathrm{E}[X_1]$, which is common for all k ($k \neq 1$).

Proofs of Corollary 7 Corollary 7 is obtained by setting $\lambda_2 = 0$ in Corollary 9. ■

Theorem 10: Suppose $\lambda_{k,j}$ $(k, j \in \mathcal{N},\ k \neq j)$ is given by

$$\lambda_{k,j} = \begin{cases} \lambda^* = \lambda_1 + \lambda_2, & k = 1 \text{ or } j = 1, \\ \lambda_1, & \text{otherwise.} \end{cases}$$

We then have

$$\mathrm{E}[X_1^2] = \frac{8\mathrm{E}[S^2]}{(N\lambda_1 + \lambda_2)(3N\lambda_1 + 4\lambda_2)} + \frac{8}{N\lambda^*(3N\lambda_1 + 4\lambda_2)} \cdot \left(\frac{2(N\lambda_1 + \lambda_2)}{N\lambda^*} - \frac{\lambda_2}{N\lambda_1 + \lambda_2} \right) s_1^2, \quad (24)$$

$$\mathrm{E}[X^2] = \frac{24\mathrm{E}[S^2]}{(N\lambda_1 + \lambda_2)(3N\lambda_1 + 4\lambda_2)} - \frac{8\lambda_2}{N\lambda^*(3N\lambda_1 + 4\lambda_2)} \left(\frac{2}{N\lambda^*} + \frac{3}{N\lambda_1 + \lambda_2} \right) s_1^2, \quad (25)$$

and for $k \neq 1$,

$$\mathrm{E}[X_k^2] = \frac{1}{|-4\boldsymbol{B}|} \Big[2N\lambda_1 \{(2N-1)\lambda_1 + N\lambda_2\} \mathrm{E}[X^2] + [2\lambda_2\{(2N-1)\lambda_1 + N\lambda_2\} + 2\lambda_2(\lambda_2 - \lambda_1)] \mathrm{E}[X_1^2] + 8\lambda_2 s_k \mathrm{E}[X_1] + 8[\lambda_2 s_1 + 2\{(2N-1)\lambda_1 + N\lambda_2\} s_k] \mathrm{E}[X_k] \Big], \quad (26)$$

where

$$|-4\boldsymbol{B}| = 6(N\lambda_1 + \lambda_2)\{(2N-1)\lambda_1 + N\lambda_2\} + 2\lambda_2\lambda^*.$$

The proof of Theorem 10 is given in Appendix F. It is easy to verify that when $\lambda_2 = 0$, Theorem 10 is reduced to the result in Theorem 8.

Compared with the mean relative time difference, the second moment is quite complicated, especially, for $X_k(t)$ $(k \neq 1)$. It is interesting to observe that when $\mathrm{E}[S^2]$ is fixed and $\lambda_2 > 0$, the limiting second moment $\mathrm{E}[X^2]$ of the relative time difference of a randomly chosen node is a decreasing function of $|s_1|$. Therefore $\mathrm{E}[X^2]$ is minimized when a node with the largest absolute value of the relative clock skew moves around actively. In other words, it is maximized when a node with the minimum absolute value of the relative clock skew moves around actively. This fact might be counterintuitive. Note that the high meeting frequency of a node leads to frequent adjustments of its clock. Thus, by letting the node with the largest relative clock skew move around, we can eliminate the influence of such a node to some extent and this leads to a small variation in the limiting relative time difference.

IV. Numerical Illustrations with Simulation Experiments

In this subsection, we provide some numerical illustrations obtained by simulation experiments, in order to demonstrate the fundamental characteristics of the simple averaging scheme visually. We first consider Special Case 1 (the homogeneous meeting model), where $N = 20$, the unit of time is chosen to be the relaxation time $\gamma = 2/(N\lambda)$, and the relative clock skew s_k and offset time $\phi_k = \mathrm{E}[X_k(0)]$ of node k are set to be

$$(s_k, \phi_k) = \begin{cases} (1, 1000), & k = 1, 2, \ldots, 10, \\ (-1, -1000), & k = 11, 12, \ldots, 20. \end{cases}$$

In this setting, a specific node encounters other nodes $(N-1)\lambda \times \gamma = 2(N-1)/N = 1.9$ times per unit time on average and the average of the total number of node meetings per unit time is equal to 19 $(= N-1)$. Also, it follows from Corollary 7 that $\mathrm{E}[X_k(t)] = s_k[1 + 999\exp(-t)]$ and $\mathrm{E}[X_k] = s_k$ $(k \in \mathcal{N})$. Thus the system is nearly in steady state at time $t = 10$ (cf. $\mathrm{E}[X_k(10)] - \mathrm{E}[X_k] \approx 0.045 s_k$).

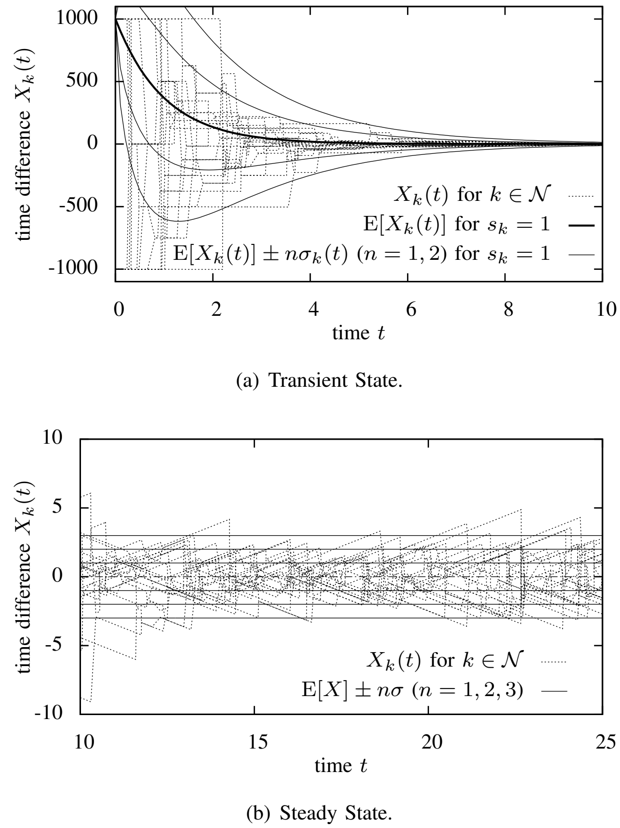
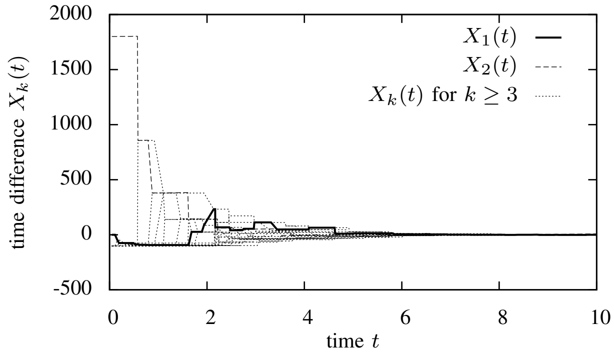


Fig. 1. Transient State and Steady State in Special Case 1.

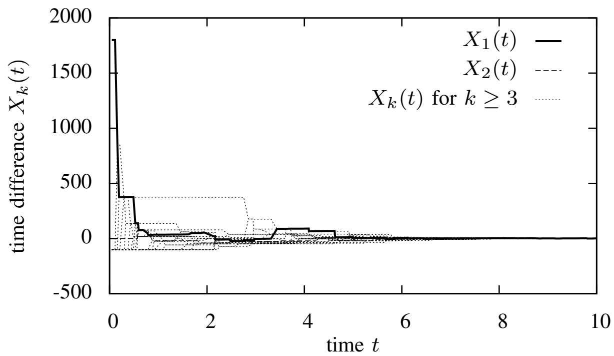
Figures 1(a) and 1(b) plot all sample paths of $X_k(t)$ $(k \in \mathcal{N})$ in transient state and steady state, respectively, with dotted lines. For reference, we also plot $\mathrm{E}[X_k(t)] \pm n\sigma_k(t)$ $(n = 0, 1, 2)$ for $k \in \mathcal{N}$ such that $s_k = 1$ in Fig. 1(a), where $\sigma_k(t) = \sqrt{\mathrm{E}[X_k^2(t)] - \mathrm{E}[X_k(t)]^2}$ denotes the time-dependent standard deviation of $X_k(t)$. Further, in Fig. 1(b), we plot $\mathrm{E}[X] \pm n\sigma = 0 \pm n$ $(n = 1, 2, 3)$, where $\sigma = 1$ denotes the limiting standard deviation of $X(t)$. We observe that the influence of huge initial offsets vanish away rapidly, and all $X_k(t)$'s remain in a certain range when t is large enough.

A. Fundamental Characteristics

Next we consider Special Case 2, where node 1 moves around actively. We set $N = 20$, $\lambda_1 = \lambda_2$, the unit of time is

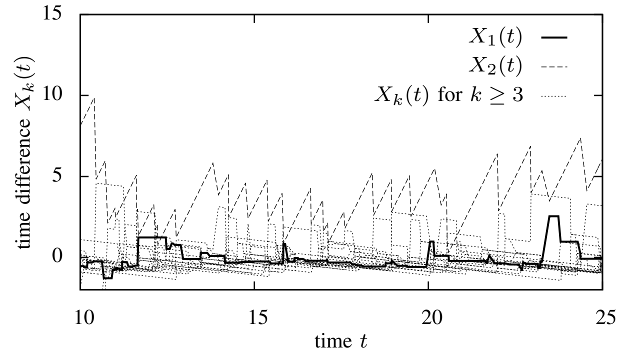


(a) $s_1 = 0$.

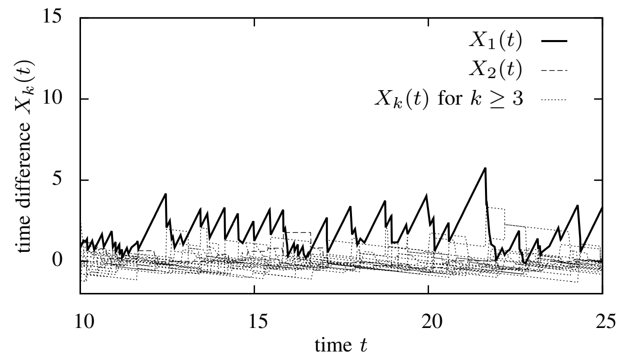


(b) $|s_1| = \max_{k \in \mathcal{N}} |s_k|$.

Fig. 2. Transient State in Special Case 2.



(a) $s_1 = 0$.



(b) $|s_1| = \max_{k \in \mathcal{N}} |s_k|$.

Fig. 3. Steady State in Special Case 2.

chosen to be the relaxation time $2/(N\lambda_1 + \lambda_2)$, $\mathrm{E}[S^2] = 1$, and for all $k \geq 3$,

$$(s_k, \phi_k) = (-\sqrt{\frac{N}{(N-1)(N-2)}}, -100) \approx (-0.24, -100).$$

Within this setting, we consider two cases, Case (a) and Case (b). In Case (a), we set

$$(s_1, \phi_1) = (0, 0),$$
$$(s_2, \phi_2) = (\sqrt{\frac{N(N-2)}{N-1}}, 1800) \approx (3.08, 1800),$$

so that the node with an exact clock moves around actively. On the other hand, in Case (b), we interchange the roles of node 1 and node 2, i.e., we set

$$(s_1, \phi_1) = (\sqrt{\frac{N(N-2)}{N-1}}, 1800) \approx (3.08, 1800),$$
$$(s_2, \phi_2) = (0, 0),$$

so that the node with the largest relative clock skew (and the largest offset time) moves around actively. Sample paths in those two cases are identical in terms of meeting epochs and a pair of node IDs at each meeting epoch.

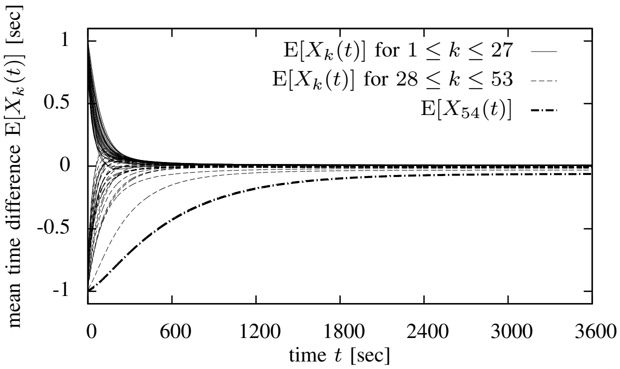
Figures 2 and 3 show all sample paths of $X_k(t)$ in transient state and steady state, respectively, where the sample path of node 1 is plotted with a solid line and the rest is plotted with dotted lines. From Fig. 2, we observe that $X_k(t)$'s in both cases converge to steady state rapidly and there is little qualitative difference between those two cases. From Fig. 3, however, we observe that Cases (a) and (b) exhibit different characteristics in steady state. At a glance, variation in Case (a) is much larger than that in Case (b). In fact, it follows from (25) that $\mathrm{E}[X^2] \approx 1.97$ in Case (a) and $\mathrm{E}[X^2] \approx 0.71$ in Case (b). Compared with Case (b), node 2 with the largest relative clock skew in Case (a) has less opportunities to encounter other nodes, so that the variation of the relative time difference of this node gets large. Further, when it encounters another nodes, its incorrect clock time propagates to the other. These facts lead to the large variation in Case (a).

B. Evaluation Using Real Trace Data

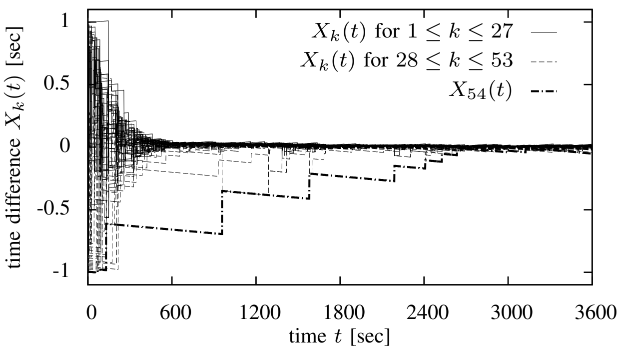
In this subsection, we examine the practicality of the analytical results using the real trace data of meeting epochs among participants of IEEE Infocom'06 [2]. We use the trace data composed of meeting epochs among 54 mobile nodes (i.e., $N = 54$) from 9:30am to 10:30am on April 24, 2006. The average rate of meeting epochs between a pair of nodes becomes 7.96×10^{-4} [1/sec]. We rearranged the node IDs in descending order of $\lambda_k = \sum_{j \in \mathcal{N}, j \neq k} \lambda_{k,j}$, where $\lambda_{k,j}$ ($k, j \in \mathcal{N}$, $k \neq j$) is estimated by the number of meetings divided by 3,600 [sec]. We set s_k and ϕ_k to be

$$(s_k, \phi_k) = \begin{cases} (10^{-4}, 1 - 0.01 * (27 - k)), & 1 \leq k \leq 27, \\ (-10^{-4}, -1 + 0.01 * (54 - k)), & 28 \leq k \leq 54, \end{cases}$$

in order to clarify: 1) the influence of clock offset and drift, and 2) transition of each node's clock.



(a) Analytical results of $\mathrm{E}[X_k(t)]$.



(b) Trace data results of $X_k(t)$.

Fig. 4. Results for real trace data.

Figure 4(a) shows the analytical results of $\mathrm{E}[X_k(t)]$ based on Theorem 4. We observe that the global clock synchronization is almost achieved at 3,000 [sec] under the assumption of Poisson meeting epochs. Note that the relaxation time γ is given by 513 [sec]. Thus, the system converges to the steady state after about 6γ [sec]. Note here that the relaxation time is inherently determined by the meeting rate of the node (i.e., node #54) with the least meeting opportunities. Except for node #54, the time synchronization is achieved around 1,500 [sec]. Fig. 4(b) illustrates the sample paths of $X_k(t)$ based on the meeting epochs in the trace data. We observe that result for the trace data is very similar to the analytical results in Figure 4(a) and the global clock synchronization is almost achieved at 3,000 [sec]. We observe that the relative time difference of node #54 becomes about half every time it meets other nodes, so that node #54 has to meet other nodes ten times or more before the relative time difference shrinks sufficiently.

V. CONCLUSION

This paper considered the simple averaging scheme for global clock synchronization in sparsely populated MANETs. Under the assumption that the estimated clock rates $\hat{\rho}_k$ are fixed, we discussed the fundamental characteristics of the scheme through the analysis and simulation results. Recall that the performance was evaluated in terms of two metrics: the speed of convergence to the steady state and the variation of relative time differences in steady state. Roughly speaking, the former is determined by the meeting frequency, independent of relative clock skews. On the other hand, the variation of relative time differences in steady state is influenced directly by relative clock skews.

APPENDIX A
PROOF OF THEOREM 3

Consider the dynamics of $\boldsymbol{X}(t)$ in a time interval $[t, t+\delta t]$. Any two nodes do not meet with probability

$$1 - \sum_{\substack{k\in\mathcal{N}\\ k<j}} \lambda_{k,j}\delta t + o(\delta t) = 1 - \frac{1}{2}\sum_{k\in\mathcal{N}} \lambda_k \delta t + o(\delta t),$$

and in this case, $\boldsymbol{X}(t+\delta t) = \boldsymbol{X}(t) + (s_1, s_2, \dots, s_N)\cdot \delta t$. Also, with probability $\lambda_{k,j}\delta t + o(\delta t)$ ($k, j \in \mathcal{N}$, $k \neq j$), node k and node j meet and those two nodes synchronize their clocks. Thus, when node k and node j meet at time $t+z$ ($0 < z < \delta t$), we have for $i \in \mathcal{N}$,

$$X_i(t+\delta t) = \begin{cases} X_i(t) + s_i\delta t, & i \neq k, j, \\ \frac{1}{2}(X_k(t) + s_k z + X_j(t) + s_j z) & \\ \quad + s_i(\delta t - z), & i = k, j. \end{cases}$$

The probability of any other event is of $o(\delta t)$.

Therefore, noting $(X_k(t) + s_k z + X_j(t) + s_j z)/2 + s_i(\delta t - z) = (X_k(t) + X_j(t))/2 + o(1)$ and $\exp(\imath s_k \eta_k \delta t) = 1 + \imath s_k \eta_k \delta t + o(\delta t)$, we have

$$\begin{aligned} f^*(\boldsymbol{\eta}, t+\delta t) &= \left(1 - \frac{1}{2}\sum_{k\in\mathcal{N}} \lambda_k \delta t\right) f^*(\boldsymbol{\eta}, t) \prod_{k\in\mathcal{N}} (1 + \imath s_k \eta_k \delta t) \\ &\quad + \sum_{\substack{k\in\mathcal{N}\\ k<j}} \lambda_{k,j}\delta t f^*(\boldsymbol{\eta}_{k,j}, t) + o(\delta t) \\ &= \left(1 + \sum_{k\in\mathcal{N}} \imath s_k \eta_k \delta t - \frac{1}{2}\sum_{k\in\mathcal{N}} \lambda_k \delta t\right) f^*(\boldsymbol{\eta}, t) \\ &\quad + \sum_{\substack{k\in\mathcal{N}\\ k<j}} \lambda_{k,j}\delta t f^*(\boldsymbol{\eta}_{k,j}, t) + o(\delta t), \end{aligned}$$

from which the theorem follows. ∎

APPENDIX B
PROOF OF THEOREM 4

Differentiating both sides of (10) once with respect to η, substituting 0 into η, and rearranging terms, we obtain

$$\frac{d}{dt} x_k^{(1)}(t) = -\frac{\lambda_k}{2} x_k^{(1)}(t) + \sum_{\substack{j\in\mathcal{N}\\ j\neq k}} \frac{\lambda_{k,j}}{2} x_j^{(1)}(t) + s_k,$$

or equivalently

$$\frac{d}{dt}\boldsymbol{x}^{(1)}(t) = \frac{1}{2}\boldsymbol{\Lambda}\boldsymbol{x}^{(1)}(t) + \boldsymbol{s}. \tag{27}$$

It is easy to verify that the solution of (27) takes a form:

$$\boldsymbol{x}^{(1)}(t) = \boldsymbol{x}^{(1)} + \exp\left(\frac{1}{2}\boldsymbol{\Lambda} t\right)(\boldsymbol{x}^{(1)}(0) - \boldsymbol{x}^{(1)}),$$

where $\boldsymbol{x}^{(1)} = \lim_{t\to\infty} \boldsymbol{x}^{(1)}(t)$ satisfies

$$-\boldsymbol{\Lambda}\boldsymbol{x}^{(1)} = 2\boldsymbol{s}. \tag{28}$$

Adding $\boldsymbol{e}\boldsymbol{\pi}\boldsymbol{x}^{(1)}$ to both sides of (28) and noting $\boldsymbol{e}\boldsymbol{\pi} - \boldsymbol{\Lambda}$ is nonsingular, we have $\boldsymbol{x}^{(1)} = 2(\boldsymbol{e}\boldsymbol{\pi}-\boldsymbol{\Lambda})^{-1}\boldsymbol{s}+\boldsymbol{e}\boldsymbol{\pi}\boldsymbol{x}^{(1)}$. Because $\boldsymbol{\pi} = N^{-1}\boldsymbol{e}^T$ and (8), we have $\boldsymbol{\pi}\boldsymbol{x}^{(1)} = \boldsymbol{0}$, from which (12) follows. ■

Appendix C
Proof of Theorem 6

Differentiating both sides of (13) twice with respect to η_k, substituting 0 into η_k and η_j, and rearranging terms, we obtain

$$\frac{d}{dt}x_k^{(2)}(t) = 2s_k x_k^{(1)}(t) + \frac{1}{4}\sum_{i\in\mathcal{N}} \lambda_{k,i} x_i^{(2)}(t) + \frac{1}{2}\sum_{i\in\mathcal{N}} \lambda_{k,i} x_{i,k}^{(1,1)}(t), \tag{29}$$

where we use $x_{k,i}^{(1,1)}(t) = x_{i,k}^{(1,1)}(t)$ and for $k \in \mathcal{N}$, $\lambda_{k,k} = -\lambda_k$ and $x_{k,k}^{(1,1)}(t) = x_k^{(2)}(t)$. Also, differentiating both sides of (13) with respect to η_k and η_j ($k, j \in \mathcal{N}$, $k \neq j$) once each, substituting 0 into η_k and η_j, and rearranging terms, we obtain

$$\begin{aligned}\frac{d}{dt}x_{k,j}^{(1,1)}(t) &= s_k x_j^{(1)}(t) + s_j x_k^{(1)}(t) + \frac{\lambda_{k,j}}{2} x_{k,j}^{(1,1)}(t) \\ &\quad - \frac{\lambda_{k,j}}{4}(x_k^{(2)}(t) + x_j^{(2)}(t)) + \frac{1}{2}\sum_{i\in\mathcal{N}} \lambda_{k,i} x_{i,j}^{(1,1)}(t) \\ &\quad + \frac{1}{2}\sum_{i\in\mathcal{N}} \lambda_{j,i} x_{i,k}^{(1,1)}(t).\end{aligned} \tag{30}$$

Note here that (8) implies

$$\sum_{k\in\mathcal{N}} x_k^{(1)}(t) = \sum_{k\in\mathcal{N}} \mathrm{E}[X_k(t)] = 0, \tag{31}$$

$$\begin{aligned}\sum_{\substack{j\in\mathcal{N}\\ j\neq k}} x_{k,j}^{(1,1)}(t) &= \mathrm{E}\left[X_k(t) \sum_{\substack{j\in\mathcal{N}\\ j\neq k}} X_j(t)\right] \\ &= \mathrm{E}[X_k(t)(-X_k(t))] = -x_k^{(2)}(t).\end{aligned} \tag{32}$$

Summing up both sides of (30) for all j ($j \in \mathcal{N}$, $j \neq k$) and using (31) and (32), we can obtain (29), which implies that (29) is redundant. With (32) and $x_{k,j}^{(1,1)}(t) = x_{j,k}^{(1,1)}(t)$, (30) is rewritten to be (14). Finally, $\mathrm{E}[X_k^{(2)}(t)] = x_k^{(2)}(t)$ is obtained with (32). ■

Appendix D
Proof of Theorem 8

When $\lambda_{k,j} = \lambda$ for all k, j ($k, j \in \mathcal{N}$, $k \neq j$), (29) is reduced to

$$\frac{d}{dt}x_k^{(2)}(t) = 2s_k x_k^{(1)}(t) - \frac{3N\lambda}{4} x_k^{(2)}(t) + \frac{\lambda}{4}\sum_{i\in\mathcal{N}} x_i^{(2)}(t), \tag{33}$$

where we use (32). We define $\overline{x}^{(2)}(t)$ as

$$\overline{x}^{(2)}(t) = \frac{1}{N}\sum_{k\in\mathcal{N}} x_k^{(2)}(t) = \frac{1}{N}\sum_{k\in\mathcal{N}} \mathrm{E}[X_k^2(t)] = \mathrm{E}[X^2(t)].$$

Summing up both sides of (33) for all k ($k \in \mathcal{N}$), dividing the both sides by N, and rearranging terms yield

$$\frac{d}{dt}\overline{x}^{(2)}(t) = \frac{2}{N}\sum_{k\in\mathcal{N}} s_k x_k^{(1)}(t) - \frac{N\lambda}{2}\overline{x}^{(2)}(t),$$

and using (15), we obtain

$$\begin{aligned}\frac{d}{dt}\overline{x}^{(2)}(t) &= -\frac{N\lambda}{2}\overline{x}^{(2)}(t) + \frac{4\mathrm{E}[S^2]}{N\lambda} \\ &\quad + 2\left(\frac{1}{N}\sum_{k\in\mathcal{N}} s_k \mathrm{E}[X_k(0)] - \frac{2\mathrm{E}[S^2]}{N\lambda}\right) e^{-\frac{N\lambda}{2}t}.\end{aligned}$$

Thus we have

$$\begin{aligned}\overline{x}^{(2)}(t) &= \frac{8\mathrm{E}[S^2]}{N^2\lambda^2} + \left(\mathrm{E}[X^2(0)] - \frac{8\mathrm{E}[S^2]}{N^2\lambda^2}\right) e^{-\frac{N\lambda}{2}t} \\ &\quad + 2\left(\frac{1}{N}\sum_{k\in\mathcal{N}} s_k \mathrm{E}[X_k(0)] - \frac{2\mathrm{E}[S^2]}{N\lambda}\right) te^{-\frac{N\lambda}{2}t},\end{aligned} \tag{34}$$

from which (17) and (19) follow. Further, noting

$$\frac{\lambda}{4}\sum_{j=1}^{N} x_j^{(2)}(t) = \frac{N\lambda}{4}\overline{x}^{(2)}(t),$$

substituting (15) and (34) into (33), and rearranging terms, we obtain

$$\begin{aligned}\frac{d}{dt}x_k^{(2)}(t) &= -\frac{3N\lambda}{4}x_k^{(2)}(t) + \frac{2}{N\lambda}(2s_k^2 + \mathrm{E}[S^2]) \\ &\quad + \left[2s_k\mathrm{E}[X_k(0)] + \frac{N\lambda\mathrm{E}[X^2(0)]}{4} - \frac{2}{N\lambda}(2s_k^2 + \mathrm{E}[S^2])\right] e^{-\frac{N\lambda}{2}t} \\ &\quad + \left[\frac{\lambda}{2}\sum_{j=1}^{N} s_j\mathrm{E}[X_j(0)] - \mathrm{E}[S^2]\right] te^{-\frac{N\lambda}{2}t},\end{aligned}$$

from which (16) and (18) follow. ■

Appendix E
Proof of Corollary 9

In Special Case 2, $\boldsymbol{\Lambda}$ can be represented to be

$$\boldsymbol{\Lambda} = -[(N\lambda_1+\lambda_2)\boldsymbol{I} - \lambda_1\boldsymbol{e}\boldsymbol{e}^T + N\lambda_2\boldsymbol{e}_1\boldsymbol{e}_1^T - \lambda_2\boldsymbol{e}\boldsymbol{e}_1^T - \lambda_2\boldsymbol{e}_1\boldsymbol{e}^T], \tag{35}$$

where $\boldsymbol{e}_1$ denotes an $N \times 1$ unit vector whose first element is equal to one. Thus, for any vector $\boldsymbol{x}$ such that $\boldsymbol{e}^T\boldsymbol{x} = 0$, it can be shown that

$$\begin{aligned}\boldsymbol{\Lambda}^n\boldsymbol{x} &= (-(N\lambda_1+\lambda_2))^n\left[\boldsymbol{I} - \frac{N}{N-1}\boldsymbol{e}_1\boldsymbol{e}_1^T + \frac{1}{N-1}\boldsymbol{e}\boldsymbol{e}_1^T\right]\boldsymbol{x} \\ &\quad + (-N(\lambda_1+\lambda_2))^n\left[\frac{N}{N-1}\boldsymbol{e}_1\boldsymbol{e}_1^T - \frac{1}{N-1}\boldsymbol{e}\boldsymbol{e}_1^T\right]\boldsymbol{x}.\end{aligned}$$

Therefore we have

$$\begin{aligned}&\exp\left(\frac{\boldsymbol{\Lambda}}{2}t\right)(\boldsymbol{x}^{(1)}(0) - \boldsymbol{x}^{(1)}) \\ &\quad = e^{-\frac{N\lambda_1+\lambda_2}{2}t}\left[\boldsymbol{I} - \frac{N}{N-1}\boldsymbol{e}_1\boldsymbol{e}_1^T + \frac{1}{N-1}\boldsymbol{e}\boldsymbol{e}_1^T\right]\end{aligned}$$

$$\cdot (\boldsymbol{x}^{(1)}(0) - \boldsymbol{x}^{(1)}) + e^{-\frac{N(\lambda_1+\lambda_2)}{2}t}\left[\frac{N}{N-1}\boldsymbol{e}_1\boldsymbol{e}_1^T - \frac{1}{N-1}\boldsymbol{e}\boldsymbol{e}_1^T\right] \cdot (\boldsymbol{x}^{(1)}(0) - \boldsymbol{x}^{(1)}),$$

from which (21) and (22) follow. On the other hand, (23) immediately follows from (28), (35), and $\boldsymbol{e}^T\boldsymbol{x}^{(1)} = 0$, i.e.,

$$-\boldsymbol{\Lambda}\boldsymbol{x}^{(1)} = [(N\lambda_1+\lambda_2)\boldsymbol{I} + N\lambda_2\boldsymbol{e}_1\boldsymbol{e}_1^T - \lambda_2\boldsymbol{e}\boldsymbol{e}_1^T]\boldsymbol{x}^{(1)} = 2\boldsymbol{s},$$

which completes the proof. ■

Appendix F
Proof of Theorem 10

Note that the special case of $\lambda_2 = 0$ has already been considered in Theorem 8 of section III-B. Therefore we assume $\lambda_2 \neq 0$. For simplicity in description, we define $\overline{x}^{(2)}(t)$ as

$$\overline{x}^{(2)}(t) = \frac{1}{N}\sum_{k\in\mathcal{N}} x_k^{(2)}(t).$$

It then follows from (29) that

$$\frac{d}{dt}\begin{pmatrix} x_1^{(2)}(t) \\ \overline{x}^{(2)}(t)\end{pmatrix} = \begin{pmatrix} -\frac{3N\lambda^*}{4} & \frac{N\lambda^*}{4} \\ -\frac{\lambda_2}{2} & -\frac{N\lambda_1+\lambda_2}{2}\end{pmatrix}\begin{pmatrix} x_1^{(2)}(t) \\ \overline{x}^{(2)}(t)\end{pmatrix} + \begin{pmatrix} 2s_1x_1^{(1)}(t) \\ \frac{2}{N}\sum_{k\in\mathcal{N}} s_kx_k^{(1)}(t)\end{pmatrix}$$
$$= \boldsymbol{A}\begin{pmatrix} x_1^{(2)}(t) \\ \overline{x}^{(2)}(t)\end{pmatrix} + \boldsymbol{b} + e^{-\frac{N\lambda^*}{2}t}\boldsymbol{c} + e^{-\frac{N\lambda_1+\lambda_2}{2}t}\boldsymbol{d},$$

where

$$\boldsymbol{A} = \frac{1}{4}\begin{pmatrix} -3N\lambda^* & N\lambda^* \\ -2\lambda_2 & -2(N\lambda_1+\lambda_2)\end{pmatrix},$$
$$\boldsymbol{b} = 2\begin{pmatrix} s_1\mathrm{E}[X_1] \\ \frac{1}{N}\sum_{k\in\mathcal{N}} s_k\mathrm{E}[X_k]\end{pmatrix},$$
$$\boldsymbol{c} = 2s_1 \cdot \frac{\mathrm{E}[X_1(0)] - E[X_1]}{N-1}\begin{pmatrix} N-1 \\ 1\end{pmatrix},$$
$$\boldsymbol{d} = \frac{2}{N}\left(\sum_{k=2}^N s_k(\mathrm{E}[X_k(0)] - \mathrm{E}[X_k]) - s_1 \cdot \frac{\mathrm{E}[X_1(0)] - \mathrm{E}[X_1]}{N-1}\right)\begin{pmatrix} 0 \\ 1\end{pmatrix}.$$

Because $\boldsymbol{A}$, $\frac{N\lambda^*}{2}\boldsymbol{I} + \boldsymbol{A}$, and $\frac{N\lambda_1+\lambda_2}{2}\boldsymbol{I} + \boldsymbol{A}$ are nonsingular when $\lambda_2 \neq 0$, we have

$$\begin{pmatrix} x_1^{(2)}(t) \\ \overline{x}^{(2)}(t)\end{pmatrix} = (-\boldsymbol{A})^{-1}\boldsymbol{b} - e^{-\frac{N\lambda^*}{2}t}\left(\frac{N\lambda^*}{2}\boldsymbol{I} + \boldsymbol{A}\right)^{-1}\boldsymbol{c} - e^{-\frac{N\lambda_1+\lambda_2}{2}t}\left(\frac{N\lambda_1+\lambda_2}{2}\boldsymbol{I} + \boldsymbol{A}\right)^{-1}\boldsymbol{d}$$
$$+ \exp(\boldsymbol{A}t)\left[\begin{pmatrix} x_1^{(2)}(0) \\ \overline{x}^{(2)}(0)\end{pmatrix} - (-\boldsymbol{A})^{-1}\boldsymbol{b} + \left(\frac{N\lambda^*}{2}\boldsymbol{I} + \boldsymbol{A}\right)^{-1}\boldsymbol{c} + \left(\frac{\lambda_2}{2}\boldsymbol{I} + \boldsymbol{A}\right)^{-1}\boldsymbol{d}\right].$$

It can be readily shown that all eigenvalues of $\boldsymbol{A}$ are negative. Thus we have

$$\lim_{t\to\infty}\begin{pmatrix} x_1^{(2)}(t) \\ \overline{x}^{(2)}(t)\end{pmatrix} = (-\boldsymbol{A})^{-1}\boldsymbol{b},$$

from which (24) and (25) follow. Further, from (29), (30), and (32), we have for k ($k \in \mathcal{N}$, $k \neq 1$),

$$\frac{d}{dt}\begin{pmatrix} x_k^{(2)}(t) \\ x_{1,k}^{(1,1)}(t)\end{pmatrix} = \boldsymbol{B}\begin{pmatrix} x_k^{(2)}(t) \\ x_{1,k}^{(1,1)}(t)\end{pmatrix} + \frac{N\lambda_1}{4}\begin{pmatrix} 1 \\ 0\end{pmatrix}\overline{x}^{(2)}(t) + \frac{1}{4}\begin{pmatrix} \lambda_2 \\ \lambda_2 - \lambda_1\end{pmatrix}x_1^{(2)}(t) + \begin{pmatrix} 0 & 2s_k \\ s_k & s_1\end{pmatrix}\begin{pmatrix} x_1^{(1)}(t) \\ x_k^{(1)}(t)\end{pmatrix},$$

where

$$\boldsymbol{B} = \frac{1}{4}\begin{pmatrix} -3(N\lambda_1+\lambda_2) & 2\lambda_2 \\ -\lambda^* & -2[(2N-1)\lambda_1 + N\lambda_2]\end{pmatrix},$$

and taking the limit $t \to \infty$, we have

$$-\boldsymbol{B}\begin{pmatrix} \mathrm{E}[X_k^2] \\ \mathrm{E}[X_1X_k]\end{pmatrix} = \frac{N\lambda_1}{4}\begin{pmatrix} 1 \\ 0\end{pmatrix}\mathrm{E}[X^2] + \frac{1}{4}\begin{pmatrix} \lambda_2 \\ \lambda_2 - \lambda_1\end{pmatrix}\mathrm{E}[X_1^2] + \begin{pmatrix} 0 & 2s_k \\ s_k & s_1\end{pmatrix}\begin{pmatrix} \mathrm{E}[X_1] \\ \mathrm{E}[X_k]\end{pmatrix},$$

from which (26) follows. ■

References

[1] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Prentice-Hall, Englewood Cliffs, NJ, 1997.

[2] A. Chaintreau, P. Hui, J. Scott, R. Gass, J. Crowcroft, and C. Diot, "Impact of Human Mobility on Opportunistic Forwarding Algorithms," *IEEE Trans. Mobile Computing*, vol. 6, no. 6, pp. 606–620, 2007.

[3] B. Choi, H. Liang, X. Shen, W. Zhuang, "DCS: Distributed Asynchronous Clock Synchronization in Delay Tolerant Networks," *IEEE Trans. Parallel and Distributed Systems*, vol. 23, no. 3, pp. 491–504, 2012.

[4] P. Diacoins and D. Stroock, "Geometric Bounds for Eigenvalues of Markov Chains," *Annals Appl. Probab.*, vol. 1, no. 1, pp. 36–61, 1991.

[5] J. Elson, L. Girod, and D. Estrin, "Fine-Grained Network Time Synchronization Using Reference Broadcasts," *ACM SIGOPS Operat. Sys. Review*, vol. 36, no. SI, pp. 147–163, 2002.

[6] Y. R. Faizulkhakov, "Time Synchronization Methods for Wireless Sensor Networks: A Survey," *Program. Comput. Soft.*, vol. 33, no. 4, pp. 214–226, 2007.

[7] S. Farrel and V. Cahill, *Delay- and Disruption-Tolerant Networking*, Artech House, Boston, MA, 2006.

[8] F. P. Kelly, *Reversibility and Stochastic Networks*, John Willy & Sons, Chichester, 1979.

[9] C. Lenzen, P. Sommer, and R. Wattenhofer, "Optimal Clock Synchronization in Networks," in *Proc. of the 7th ACM Int'l Conf. on Embedded Networked Sensor Sys. (Sensys'09)*, pp. 225–239, 2009.

[10] Q. Li and D. Rus, "Global Clock Synchronization in Sensor Networks," *IEEE Trans. Computers*, vol. 55, no. 2, pp. 214–226, 2006.

[11] D. L. Mills, *Computer Network Time Synchronization: The Network Time Protocol*. CRC Press, Boca Raton, FL, 2006.

[12] K. Römer, "Time Synchronization in Ad Hoc Networks," in *Proc. of the 2nd ACM Int'l Symp. on Mobile Ad Hoc Netw. and Computing (MobiHoc'01)*, pp. 173–182, 2001.

[13] L. Schenato and F. Fiorentin, "Average TimeSynch: A Consensus-Based Protocol for Clock Synchronization in Wireless Sensor Networks," *Automatica*, vol. 47, no. 9, pp. 1878–1886, 2011.

[14] O. Simeone, U. Spagnolini, Y Bar-Ness, and S. H. Strogatz, "Distributed Synchronization in Wireless Networks," *IEEE Sig. Processing Mag.*, vol. 25, no. 5, pp. 81–97, 2008.

[15] W. Su and I. F. Akyildiz, "Time-Diffusion Synchronization Protocol for Wireless Sensor Networks," *IEEE/ACM Trans. Netw.*, vol. 13, no. 2, pp. 384–397, 2005.

[16] J. R. Vig, "Introduction to Quartz Frequency Standards," *Tech. Rep. SLCET-TR-92-1 (Rev. 1) Army Res. Lab. Electron. and Power Sources Directorate*, 1992. available at http://www.ieee-uffc.org/frequency_control/teaching.asp?name=vigtoc.

[17] Q. Ye and L. Cheng, "DTP: Double-Pairwise Time Protocol for Disruption Tolerant Networks," in *Proc. of Int'l Conf. on Distributed Computing Sys.*, pp. 345–352, 2008.

PLACE PHOTO HERE

Masahiro Sasabe (M'06) received the M.E. and Ph.D. degrees from Osaka University, Osaka, Japan, in 2003 and 2006, respectively. He is currently an Assistant Professor with the Department of Information and Communication Technology, Osaka University. From 2003 to 2004, he was a Research Fellow with 21COE-JSPS, Japan. From 2004 to 2007, he was an Assistant Professor with the Cybermedia Center, Osaka University. His research interests include P2P/overlay networking, game-theoretic networking, and ubiquitous networking. Dr. Sasabe is a member of IEICE.

PLACE PHOTO HERE

Tetsuya Takine (M'94) is currently a Professor in the Department of Information and Communications Technology, Graduate School of Engineering, Osaka University. He was born in Kyoto, Japan, on November 28, 1961. He received B.Eng., M.Eng., and Dr.Eng. degrees in applied mathematics and physics from Kyoto University, Kyoto, Japan, in 1984, 1986, and 1989, respectively. In April 1989, he joined the Department of Applied Mathematics and Physics, Faculty of Engineering, Kyoto University, as an Assistant Professor. Beginning in November 1991, he spent one year at the Department of Information and Computer Science, University of California, Irvine, on leave of absence from Kyoto University. In April 1994, he joined the Department of Information Systems Engineering, Faculty of Engineering, Osaka University as a Lecturer, and from December 1994 to March 1998, he was an Associate Professor in the same department. From April 1998 to May 2004, he was an Associate Professor in the Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University. His research interests include queueing theory, emphasizing numerical computation, and its application to performance analysis of computer and communication networks. He is now serving as an associate editor of Queueing Systems, Stochastic Models, and International Transactions in Operational Research. He received Telecom System Technology Award from the Telecommunications Advancement Foundation in 2003 and 2010, and Best Paper Awards from ORSJ in 1997, from IEICE in 2004 and 2009, and from ISCIE in 2006. Dr. Takine is a fellow of ORSJ and a member of IEICE, IPSJ, and ISCIE.