

# SYNCHRONIZATION in SENSOR NETWORKS How does network topology affect behavior?



Ananthram Swami U.S. Army Research Lab a.swami@ieee.org

21 July 2008, SAM 2008

Acknowledgements:

Sergio Barbarossa, Prithwish Basu, Mounir Ghogho, Prabir Barooah, Joao Hespanha, Xiaoli Ma, Brian Sadler, Arnt Salberg, Anna Scaglione

# Sync: An old problem!







Poincare: 1909 Broadcast radio sync

1884 – World Time Conference

#### Sync: An old problem! V.') huygens' clocks 1665. [Fig. 75.]<sup>2</sup>) 22 febr. 1665. lorum ], miut ne altero ciproe cum arent, ab alicari cœpi. ut experimentum caperem turbavi alterius penduli reditus ne fimul incederent fed quadrante horæ poft vel femihora rurfus concordare inveni. SAM 2008 3

# Sync: An older solved problem?



Kaempfer, 1680; Kuaramato, 1984; Mirollo-Strogtaz, 1990

- ~10K male Southeast Asian fireflies congregate in trees and flash in near perfect unison \*
  - flash every 800 to 1600 ms
  - Sync to within 30 ms
  - Only local interaction: leaves, limited perception etc.
  - flash every ~1s, without external stimulus.
- Knowledge of distributed mechanism used to synchronize flashes can be used to develop better ways to synchronize on a network



Photo from presentation by: Ibiso Wokoma, Ioannis Liabotis, Ogngen Prnjat, Lionel Sacks, Ian Marshall – University College of London

SAM 2008

Source: Talk by Dr. John Parmentla, Director for Research and Laboratory Management, US Army, at USMA Network Science Workshop, on 22 Oct 2007.

## Sync: An older solved problem?



Source: Talk by Dr. John Parmentla, Director for Research and Laboratory Management, US Army, at USMA Network Science Workshop, on 22 Oct 2007.

#### **Clock Neurons in the Suprachiasmatic Nucleus (SCN)**







#### Experiment

- 3 watches placed next to each other
- Left for 140 days, time recorded each day



Source: From a CTA talk by Prof Xiaoli Ma, Georgia Tech; taken from `The Science of Timekeeping', HP Application Note 1289

6

SAM 2008





	Accuracy (PPM)	Power	Lifetime with AAA battery 1250 mAh	Notes
Watch clock	200 x 10 <sup>-6</sup>	1 micro W	142 yrs	Temperature, aging
тсхо	6 x 10 <sup>-6</sup>	6 mW	208 hours	>1 PPM
МСХО	3 x 10 <sup>-8</sup>	75 mW	17 hours	large, aging drift
GPS	<b>10</b> <sup>-8</sup> <b>10</b> <sup>-11</sup>	180 mW	7 hours	Outdoors, cost
DARPA CSAC	10-11	30 mW	42 hours	Target
		125 mW	10 hours	prototype

Caveats:

- Battery lifetime depends upon discharge current
- Energy storage depends upon medium
- Drift to temperature, aging  $\rightarrow$  must resync
- Resync time  $\approx$  (Tolerable offset) / (relative accuracy) <sub>SAM 2008</sub>



## **Motivation**

#### **Time-Sync Crucial for**

- Target tracking, ranging, localization
- Distributed MIMO, Collaborative signal processing
- Data fusion
- Feedback control
- Network probing and monitoring
- Cooperative communications
- Energy efficient MAC (e.g., with duty cycling)

#### Local information

- But need to estimate `global' state

Constraints

- Resource constraints (batteries, BW)
- Application constraints (timeliness, desired accuracy)

Robustness

- Link and node failures; asymmetric comms; variable delays

Exemplar for distributed inference



- Accuracy: Worst case (or avg case) pair wise error between onehop neighbors. Performance bounds?
- Resource efficiency: The number of broadcasts necessary to achieve sync and the rate and frequency of messages that need to be exchanged to maintain sync.
- Convergence time: The time taken for all nodes (or a high percentage of nodes) to sync to their neighbors.
- Fault-tolerance: Robustness to failure of critical nodes and/or links, clock jitter and drift, congestion, mobility.
- Scalability with network size: Does the sync-error increase with size? Does the convergence time increase with the diameter of the network? Node density? Clock parameters?

#### Complexity

Impact of variable delays (queue, processing ...)



Lindsey et al, Proc. IEEE, 1985 Bregni, IEEE Comm Mag, 1998

Anceaume, Puaut; INRIA, 1998 Sivrikaya, Yener, IEEE Network 2004 Johannessen, IEEE Contr. Sys Mag, 2004 Sundararaman et al, AHN, 2005 Sadler, Swami, MILCOM 2006 Faizulkhakov, PCS, 2007



# **Slot Synchronization**

Motivation: Enabling energy conserving duty cycling MAC

- Oscillator drifts and duty cycling
  - => misalignment of slot boundaries
  - => loss in throughput

Pair wise slot sync and re-sync

- Requires guard times in each slot; and re-sync
- Strong correlation between
  - · worst case oscillator drift
  - re-sync period
  - slot utilization
  - energy consumption



B has a slower clock than A ( $\theta_{AB} > 0$ )



Transmission from node A to B and A to C

(source: Dr. P. Basu BBN) (C&N CTA)



#### Dimensions: Sync accuracy, Energy efficiency, Convergence Time, Fault tolerance, Scalability, Engineering Simplicity

Advantages	Disadvantages	
Fast convergence after tree computation (can be reused or amortized); Low #local broadcasts for sync maintenance; Provably low worst-case sync-error	Fault-tolerance hard to achieve in duty-cycle mode (tree needs to be recomputed); If root fails, then re-election is hard; Tree maintenance could costs energy	
Sync-trees with better properties hence lower energy for maintaining sync	Repairs may take a long time because of the dependence on LSUs (low frequency)	
Faster convergence because the tree structure is only implicit after root election; Better fault-tolerance than other tree based schemes	Worst-case sync-error may be worse than other tree based schemes; Also energy consumption (#broadcasts) for maintenance may be high	
No root election or sync-tree computation step, hence highly fault-tolerant; Sync maintenance is very efficient; Friendly to network dynamics (join/leave/move); Potentially better scaling properties	Slower convergence than tree-based schemes; Potentially lower accuracy; Determining optimal aggregation function (median? mean? other?) may be hard	
Best of both worlds: achieve fast convergence by tree-based protocol followed by fault-tolerant maintenance of sync in low-energy mode	Two phase protocol – hence may be harder to analyze and implement	
	<ul> <li>Fast convergence after tree computation (can be reused or amortized); Low #local broadcasts for sync maintenance; Provably low worst-case sync-error</li> <li>Sync-trees with better properties hence lower energy for maintaining sync</li> <li>Faster convergence because the tree structure is only implicit after root election; Better fault-tolerance than other tree based schemes</li> <li>No root election or sync-tree computation step, hence highly fault-tolerant; Sync maintenance is very efficient; Friendly to network dynamics (join/leave/move); Potentially better scaling properties</li> <li>Best of both worlds: achieve fast convergence by tree-based protocol followed by fault-tolerant</li> </ul>	



#### Broadcast Protocols

RBS *Elson, Girod, Estrin, 2002:* Beacon-aided TPSN, *Ganeriwal et al, 2003,* Broadcast with hierarchy

#### Distributed Synchronization Protocols

Diffusion-based:Li, Rus, 2004Spatial smoothing:Solis et al, 2006Bio-inspiredHong, Scaglione, 2005Average consensus:Xiao, Boyd 2003Jacobi iterations:Barooah et al, 2007Advection-diffusionBarbarossa, 2008 (here!)

#### Other taxonomies:

- Single-hop vs. multi-hop
- Server initiated vs. client initiated vs. always-on





# **Clock Models**

t = "true" time, local node time is T(t)

• Clock drift  $\rho(t) = dT(t) / dt - 1$ 

• Bounded drift, and clocks not running backwar  $|\rho(t)| < \rho_0$  - 1 <  $\rho(t)$ 

Taylor expansion at the I<sup>th</sup> node

 $T_i(t) = \alpha_i + \beta_i t + \gamma_i t^2 + \dots$ 

offset skew models time variations

- Could model skew as AR process
- Time Sync problem: Estimate α, β, γ, and adjust clock







 $\zeta_{mn} = (\alpha_n - \alpha_m) + \varepsilon_{mn}$ 

Zero mean error

Same formulation in estimating heading, position, skew

Time shifts from

reference clock

Use multiple measurements. Linear least-squares problem:

- Pair-wise sync easy
- ✓ Can easily compute variance CRB, robust estimators confidence intervals etc.
- ✓ Test for drift  $\beta \neq 0$ ?

SAM 2008



 $\blacktriangleright$  Node u transmits a pair of packets spaced by  $\rho$ 

- Extends to multiple measurements.
- Usual techniques to deal with NG noise

20

SAM 2008













 Every agent measures the difference in the heading between itself and nearby agents.

 One (or more) agent's heading is known (reference).

Problem: determine headings (w.r.t. reference headings)

Headings with respect to reference heading

Zero mean error

[source Dr. Prabir Barooah, U. Florida (ICB)]

# **Network time synchronization**



$$\zeta_{mn} = (\alpha_n - \alpha_m) + \varepsilon_{mn}$$

Same formulation in skew-offset estimation heading estimation position estimation



How to go from pair-wise to network-wide sync? Incorporate info from r' reference nodes? Additive ambiguity : r > 0





Measurements Graph Communications Graph G = (V,E) $\zeta = A' \alpha + \varepsilon$  $\zeta = [A_u' A_r'] [\alpha_u \alpha_r] + \varepsilon$  $\Rightarrow L \alpha_u = b$ 



### Estimators:

L = A<sub>u</sub> P<sup>-1</sup>A<sub>u</sub>' is invertible iff every weakly connected component in G has a reference node

Optimal estimate computed by FC with error-free links from all nodes



Covariance of optimal estimator depends on

- distance from reference node (`# hops')
- structure of the network

 $\boldsymbol{\Sigma}_{\alpha} = \mathbf{L}^{-1}, \ \mathbf{L} = \mathbf{A}_{u} \mathbf{P}_{\varepsilon}^{-1} \mathbf{A}_{u}^{T}$ 

Distributed algorithms

- Converges? To optimal ?
- Convergence rate
- Robustness to link failures ?
- Asymmetric communications ?
- With dynamics in topology ?

26



# **Distributed Algorithms: Assumptions**

Assumptions

- Every weakly connected component of G = (V,E) has at least one reference node.
- □ Communication graph  $G^c = (V, E^c)$ : (u,v) in E ⇒ (u,v) and/or (v,u) in E<sup>c</sup>
- No edge in E<sup>c</sup> is directed towards a reference node
- Measurement errors uncorrelated, known variances
- At every t, each node may fail with prob q, and every link with prob q.

Jacobi iteration for every node  $u \in V \setminus V_r$ 

$$\left(\sum_{\boldsymbol{v}\in\mathcal{N}_{\boldsymbol{u}}}\frac{1}{\sigma_{\boldsymbol{u}\boldsymbol{v}}^{2}}\right)\hat{x}_{\boldsymbol{u}}^{(\boldsymbol{k}+1)} = \sum_{\boldsymbol{v}\in\mathcal{N}_{\boldsymbol{u}}}\frac{1}{\sigma_{\boldsymbol{u}\boldsymbol{v}}^{2}}\left[\hat{x}_{\boldsymbol{v}}^{(\boldsymbol{k})} + a_{\boldsymbol{u}\boldsymbol{v}}\zeta_{\boldsymbol{u}\boldsymbol{v}}\right]$$







# **Convergence Results**

$$\begin{split} \left(\sum_{v \in \mathcal{N}_{u}} \frac{1}{\sigma_{uv}^{2}}\right) \hat{x}_{u}^{(k+1)} &= \sum_{v \in \mathcal{N}_{u}} \frac{1}{\sigma_{uv}^{2}} \left[ \hat{x}_{v}^{(k)} + a_{uv} \zeta_{uv} \right] \\ \text{Weighted in-degree} & D_{uu} &= \sum_{v \in \mathcal{N}_{u}} \frac{1}{\sigma_{uv}^{2}} \\ \text{Weighted adjacency} & C_{uv} &= \frac{1}{\sigma_{uv}^{2}}, \text{ if } (v, u) \in E^{c} \\ \text{Submatrices of C & D} & M, \ N \in \mathcal{R}^{n_{u} \times n_{u}} \\ \text{Iteration:} & M\mathbf{x}^{(k+1)} = N\mathbf{x}^{(k)} + \mathbf{b}_{u} \\ \text{Fixed point} \\ \text{exists & is unique if} & L_{c} = M - N = A_{u}^{c}P^{-1}A_{u}^{T} \\ \text{L}_{c} \text{ is invertible} \end{split}$$

SAM 2008

28



# Synchronous update, no node/link failures:

Matrix  $L_c$  is invertible iff there is a directed path in  $G_c$  from at least one reference node to every non-ref node.

# Asynchronous case, with iid failures:

- If  $G_c(t)$  satisfies AS1; there is a directed path in  $G_{init}$  from at least one ref. node to every non-ref node; no communication edge in  $G_{init}$  fails permanently; no edge outside  $G_{init}$  remains active infinitely often. Then the algorithm converges a.s.
- Proof follows Frommer & Syzld, On asynchronous iterations, Journal of computation & applied math, 2000; also Tsitsiklis and Bertsekas



There is a penalty in the asymmetric case: asymptotic covariances:

 $\boldsymbol{\Sigma}_{s} = (\boldsymbol{A}_{u} \boldsymbol{P}_{\varepsilon}^{-1} \boldsymbol{A}_{u}^{T})^{-1}$ 

 $\boldsymbol{\Sigma}_{a} = (\boldsymbol{A}_{u}^{a} \boldsymbol{P}_{\varepsilon}^{-1} \boldsymbol{A}_{u}^{T})^{-1} (\boldsymbol{A}_{u}^{a} \boldsymbol{P}_{\varepsilon}^{-1} \boldsymbol{A}_{u}^{aT}) (\boldsymbol{A}_{u}^{a} \boldsymbol{P}_{\varepsilon}^{-1} \boldsymbol{A}_{u}^{T})^{-T}$ 

where A<sub>u</sub><sup>a</sup> is obtained from A<sub>u</sub> by setting appropriate elements to zero.

(u,e) = 0 if u is a reference node or the comm link e is not directed to u.





Link failures p=0.2 No node failures q=0

Estimate is unbiased Converges rapidly But with large variance compared with centralized BLUE



### One link can be worth a lot!





# **Do more measurements help?**





Var (G1) = 
$$[1, 1, 2.5]$$
  
Var (G2) =  $[2/3, 2/3, 5.3]$ 

SAM 2008

33



Iterations  $x(t+1) = [I + D_t]^{-1} [A_t + I] x(t) = F_t x(t)$ 

AS1  $F_t(i,i) \ge a \ge 0$ ,  $F_t(i,j) = 0$  or in [a,1]; unity row sums AS2 Graph G(t) = (N, E(t))

Graph (N,  $\mathbf{\acute{U}}_{t\leq s} E(s)$ ) is strongly connected for all t >= 0 AS3 Bounded inter-communication interval or symmetry

Then, the iteration converges to a common value

AS4. Bounded delays

AS5 With positive probability, some updates do not occur

Then, asymptotic consensus is achieved

Bondel et al (CDC 2005) / Tsitsiklis 1984:





But Nodes can obtain `new measurements' during the iterations.

• Avg. relative offset with node 
$$\nu$$

$$\xi_{u,v}(i) = \frac{1}{i} \sum_{k=1}^{i} \zeta_{u,v}(i)$$

$$:= \hat{x}_v^{(i)} + \xi_{u,v}(i) \qquad \forall v \in \mathcal{N}(u)$$

• Overall updated estimate 
$$\hat{x}_u^{(i)} = (1 - \beta_i)\hat{x}_u^{(i-1)} + \beta_i \frac{1}{d_u} \sum_{v \in \mathcal{N}_u} w_{u,v} \hat{x}_u^{(i-1)}(v)$$

 $\hat{x}_{u}^{(i)}(v)$ 

• In matrix form, for all nodes  $\hat{\mathbf{x}}(i) = J\hat{\mathbf{x}}(i-1) + B\bar{\boldsymbol{\xi}}(i-1)$ 

where

$$J = I - \beta M^{-1} L_b, \qquad B(i) = \beta M^{-1} A_b W,$$
  
$$\bar{\xi}(i) := \xi^{(i)} - A_r^T \mathbf{x}_r, \qquad \xi^{(i)} = [\xi_1^{(i)}, \dots, \xi_m^{(i)}]^T,$$

Barooah & Swami: MILCOM 2008 (submitted)

SAM 2008

## **RJA – Convergence Results**



$$\hat{x}_{u}^{(i)} = \alpha \hat{x}_{u}^{(i-1)} + \beta \frac{1}{d_{u}} \sum_{v \in \mathcal{N}_{u}} w_{u,v} \hat{x}_{u}^{(i-1)}(v),$$

- If initial estimates x̂<sup>(0)</sup><sub>u</sub>, u ∈ V for the iterative update algorithm (8) are unbiased, the estimates x̂<sup>(i)</sup>, u ∈ V are unbiased at every iteration i.
- 2) If the measurement noise {e<sup>(i)</sup>, i = 1,...} is a wide sense stationary white noise sequence with R(i) = R<sub>0</sub>, i = 1,..., then P(i) → P<sub>∞</sub> as i → ∞, where the steady state covariance matrix P<sub>∞</sub> is the unique positive definite solution of the Lyapunov equation:

$$JP_{\infty}J^{T} - P_{\infty} + BR_{0}B^{T} = 0.$$
 (12)

3) The spectra gap, assuming weak connectivity, is:

$$g(J) = 1 - \rho(J) = \beta \lambda_{\min} \left( M^{-1/2} L_b M^{-1/2} \right)$$

Barooah & Swami : Milcom 2008 (submitted)

 $L_b := \bar{A}_b W A_b^T$ 

 $M = diag (L_b)$ 

$$J = \alpha I + \beta M^{-1}N,$$
  
$$B = \beta M^{-1}A_bW,$$



Fig. 1. The tradeoff in choosing larger  $\beta$  - it decreases the spectral radius (faster convergence rate) but increases the trace of the steady state covariance matrix  $P_{\infty}$  (poorer estimation accuracy).



### **RJA – Simulation Example**



Fig. 1. The measurement graph for the 25 node simulation. All communi cations are bidirectional.



Fig. 2. The trace of the estimation error covariance matrix P(i) as a function of the iteration counter *i* for the network shown in Figure 1. The covariance P(i) is computed from the recursive relationship (12), with initial condition chosen as the identity matrix. Time evolution of the variances for two values of  $\beta$  are shown.



### **RJA – Simulation Example (2)**



Fig. 3. Five sample runs of the estimates at node 13 (upper right hand corner node in Figure 1) as a function of time, with  $\beta = 0.9$ . The solid line shows the true value of node 13's time offset with respect to the reference.



Fig. 4. The temporal evolution of the variance of the estimation at node #13 (shown in Figure 1). The legend "empirical" refers to the empirically estimated error variance of the node's estimates obtained by the proposed algorithm, with  $\beta = 0.3$ . The estimates were averaged over 100 sample runs. "Jacobi" refers to the steady-state variance the estimates produced by the Jacobi algorithm in [8] and its variants in [5], [10]



### **RJA – Simulation Example (3)**



Fig. 5. Five sample runs of the estimates at node 13 (upper right hand corner node in Figure 1) as a function of time in the presence of random communication faults (probability of failure is 0.3), with  $\beta = 0.9$ . The solid line shows the true value of node 13's time offset with respect to the reference.



Fig. 6. Evolution of node 13's estimation error variance with random communication faults. At every iteration, every communication fails independently of all other edges, with a probability of 0.3. The legend "empirical" refers to the empirically estimated (from 100 sample runs) error variance of the node's estimates produced by the proposed algorithm. "Jacobi" refers to the variance the estimates produced by the Jacobi algorithm in [8] and its variants in [5], [10] achieve upon convergence without communication faults.



### **Convergence for grid graphs**

Theorem 2: When all the edge weights are chosen as 1 (i.e.,  $w_{u,v} = 1$  for every  $(u, v) \in \mathcal{E}$ ) and there is a single reference node  $o \in \mathcal{V}$ , the following statements hold.

1) In a 1-D grid of ntotal nodes, we have

$$g(J) \ge \frac{\beta}{(n-1)(n-3)}$$

 In a N<sub>1</sub> × N<sub>2</sub> 2-D grid with a bounded aspect ratio (i.e., one in which there exist positive constants <u>c</u>, <u>c</u> independent of N<sub>1</sub> and N<sub>2</sub> such that <u>c</u> ≤ N<sub>1</sub>/N<sub>2</sub> ≤ <u>c</u>),

$$g(J) \ge \frac{\beta}{n_{total}(\log n_{total} + \bar{c} + \underline{c} + 6)}$$

where  $n_{total} = N_1 \times N_2$ .

3) In a N<sub>1</sub> × N<sub>2</sub> × N<sub>3</sub> 3-D grid, where there exists positive scalars <u>c</u>, <u>c</u>, <u>d</u>, <u>d</u> such that <u>c</u>N<sub>1</sub> ≤ N<sub>2</sub> ≤ <u>c</u>N<sub>1</sub> and <u>d</u> log N<sub>3</sub> ≤ N<sub>2</sub> ≤ <u>d</u>N<sub>3</sub><sup>2</sup>, we have that

$$g(J) \ge \gamma \frac{\beta}{n_{total}}$$

where  $\gamma$  is a constant independent of the number of nodes  $n_{total} = N_1 \times N_2 \times N_3$ .



Fig. 3. The trend of several estimation error variances of the upper right hand corder node in the graph shown in Figure 2. The legend "empirical" refers to the empirically estimated (by Monte Carlo methods) error variance of the node's estimates produced by the proposed algorithm. "predicted" refers to the same variance, but that computed from the iteration (11). The legend "steady state" refers to the diagonal entry (corresponding to the node) of the steady state covariance matrix  $P_{\infty}$  described in Theorem 1. The trend of the variances corroborate the claims made in the Theorem. "Jacobi" refers to the variance the the estimates produced by the Jacobi algorithm in [6] and its variants in [4, 8] achieve upon convergence.

SAM 2008



- Studied convergence of distributed (consensus) algorithms with asynchronous updates, iid failures, and asymmetric links.
- There is a performance penalty in the asymmetric case
- Variations on consensus algorithms to incorporate new measurements
- Adding measurements may be harmful:
  - Collaborative decision on which measurements should be added
  - > Appropriate protocols
- When reference nodes disagree?



## Anna Scaglione's Sync Video





The video clip was created by Prof. Anna Scaglione (Cornell / UC-Davis) and her group; see: http://www.youtube.com/watch?v=5F7Qhdf9ZJg

SAM 2008



# **QUESTIONS?**

SAM 2008

45