A Martingale Framework for Trust

B. Hajek and Charles Spuckler

Department of Electrical and Computer Engineering
and the Coordinated Science Laboratory
University of Illinois at Urbana-Champaign

January 8, 2010

Outline

Orientation: a martingale framework for trust

- Two examples from ancient Egypt

Orientation: a martingale framework for trust

- Two examples from ancient Egypt
  - Construction of a giant pyramid

Orientation: a martingale framework for trust

- Two examples from ancient Egypt
  - Construction of a giant pyramid

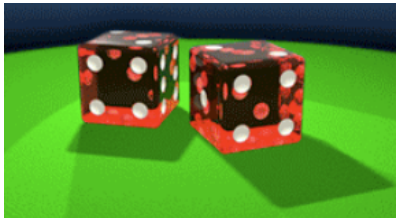
  - Cutting a giant (1,168 ton) obelisk

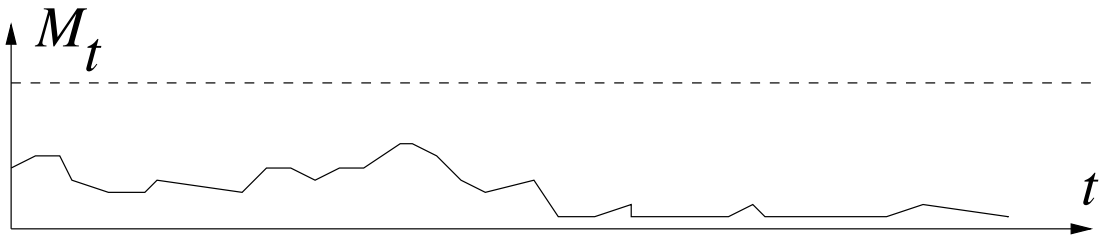
Orientation: a martingale framework for trust

- Two examples from ancient Egypt
  - Construction of a giant pyramid

  - Cutting a giant (1,168 ton) obelisk

- Where do martingales fit in?





Some notation

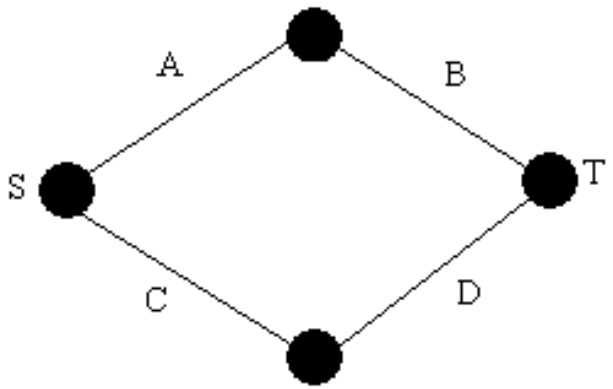
An *instance* of the trust framework is a pair $(E, \mathcal{G})$.

- E is a finite set, elements $e \in E$ are called *links*
- Each link e has a state $x_e \in \{0, 1\}$
- $\mathbf{x} = (x_e : e \in E)$ is called the *network state*
- $\{0, 1\}^E$ denotes the set of all network states
- $\mathcal{G}$ denotes the set of *good* network states, $\mathcal{G} \subset \{0, 1\}^E$.

Models

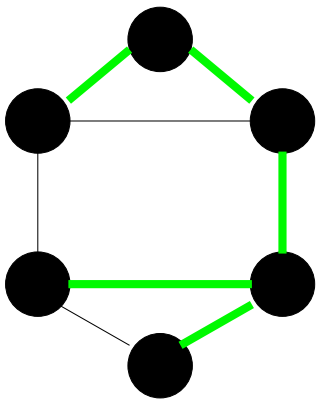
(ways to specify many instances):

- $S - T$ *reachability model*
  - E is the set of links for a graph (V, E) with $S, T \in V$
  - $\mathcal{G}$ is the set of network states with an all one $S - T$ path
  - Example:



  - $\mathcal{G} = \{1100, 0011, 1101, 1110, 0111, 1011, 1111\}$.

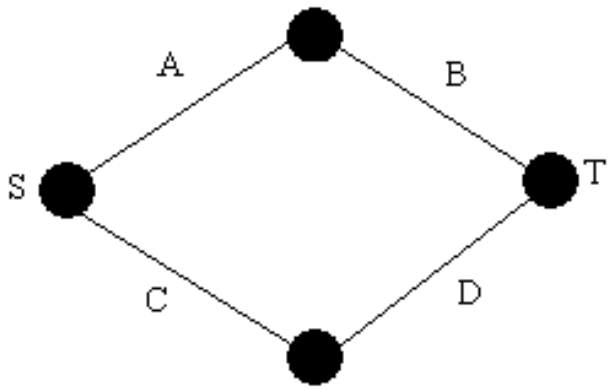
- *Graph connectivity model*
  - E is the set of links for a graph (V, E)
  - $\mathcal{G}$ is the set of network states such that the graph is connected by all one links.
  - Equivalently, $\mathcal{G}$ is set of network states such that a an all one spanning tree exits
  - Example:



- *Majority vote model*
  - E is a finite set.
  - $\mathcal{G}$ is the set of network states $\mathbf{x}$ with at least $\tau = \left\lceil \frac{|E|+1}{2} \right\rceil$ ones
- *Parity model*
  - E is a finite set.
  - $\mathcal{G}$ is the set of network states $\mathbf{x}$ with an odd number of ones.

Coordinate convexity

- If $\mathcal{G}$ is coordinate convex (models above except parity) let:
  - $\mathcal{A}$ be the set of minimal states in $\mathcal{G}$.
  - $\mathcal{B}$ be the set of maximal states in $\mathcal{G}^c$.
- Example: For $S - T$ reachability and



$\mathcal{A} = \{1100, 0011\}$ and $\mathcal{B} = \{0101, 1001, 0110, 1010\}$.

- Example: For an instance of the majority vote model, $\mathcal{A}$ is the set of network states with τ one's, and $\mathcal{B}$ is the set of network states with $\tau - 1$ one's.

Statistical model

Assume random network state $\mathbf{X} = (X_e : e \in E)$ such that:

- X_e is a Bernoulli r.v. with known parameter π_e, for each $e \in E$
- the X_e's are independent.

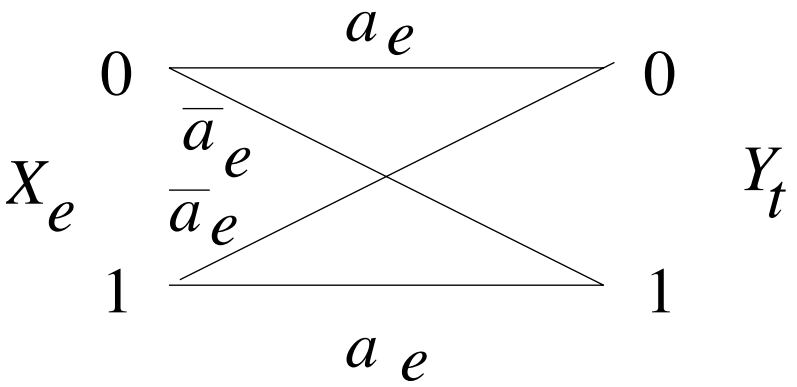
Let $\chi = I_{\{\mathbf{X} \in \mathcal{G}\}}$. Then (use notation $\overline{b} = 1 - b$)

$$E[\chi] = \sum_{\mathbf{x} \in \mathcal{G}} \prod_{e \in E} \pi_e^{x_e} (\overline{\pi}_e)^{\overline{x}_e}$$

Note: $\mathbf{X}$, and therefore χ, are fixed for all time.

Probing strategies

- Decision maker would like to accurately determine χ
- Decision maker probes one link at each integer time $t \geq 1$
- In response, decision maker observes Y_t, generated by BSC with known crossover probability a_e :



- A *probing strategy* determines which link to probe at time $t+1$, based on observations $\mathcal{F}_t = \sigma\{Y_1, \ldots, Y_t\}$ up to time t.

The principle focus of this work is the choice of probing strategies.

Bayesian updates

Let $p_{e,t} = P[X_e = 1|\mathcal{F}_t] = E[X_e|\mathcal{F}_t]$. By Bayes rule, we find that

$$p_{e,t} = \phi(N_{e,t}, K_{e,t}, a_e, \pi_e)$$

where

- $N_{e,t}$ is the number of times link e is probed up to time t,
- $K_{e,t}$ is the number of probes of link e up to time t resulting in an observation of value one
- $\phi(n, k, a, \pi) = \frac{\pi a^k (\overline{a})^{n-k}}{\pi a^k (\overline{a})^{n-k} + \overline{\pi} (\overline{a})^k a^{n-k}}$.

Or, recursively: $p_{e,0} = \pi_e$ and

$$p_{e,t+1} = \begin{cases} p_{e,t} & \text{if link } e \text{ is not probed at time } t+1 \\ \frac{p_{e,t} a_e}{p_{e,t} a_e + \overline{p}_{e,t} \overline{a}_e} & \text{if } e \text{ is probed at time } t+1 \text{ and } Y_{t+1} = 1 \\ \frac{p_{e,t} \overline{a}_e}{p_{e,t} \overline{a}_e + \overline{p}_{e,t} a_e} & \text{if } e \text{ is probed at time } t+1 \text{ and } Y_{t+1} = 0. \end{cases}$$

The process G

Let $G_t = P[\text{network state is good}|\mathcal{F}_t] = E[\chi|\mathcal{F}_t]$. Or

$$G_t = \sum_{\mathbf{x}\in\mathcal{G}} \prod_{e\in E} p_{e,t}^{x_e} (\overline{p}_{e,t})^{\overline{x}_e}.$$

In particular, $G_0 = E[\chi]$.

- The random processes $(p_{e,t} = E[X_e|\mathcal{F}_t] : t \geq 0)$ for $e \in E$ and $(G_t : t \geq 0)$ are martingales.
- If $a_e \neq 0.5$ and e is probed infinitely often, $p_{e,t} \to X_e \quad a.s.$
- Probing strategies thus exist so that $G_t \to \chi \quad a.s.$

A measure of uncertainty

- Let $U_t = h(G_t)$, where h is the binary entropy function or something similar.
- By concavity of h, U is a supermartingale (negative drift)
- If h is the binary entropy function:
  - U_t is the (random) conditional entropy of χ given $Y_1 = y_1, \cdots, Y_t = y_t$
  - $-E[U_{t+1} - U_t | \mathcal{F}_t] = I(\chi; Y_{t+1} | Y_1 = y_1, \cdots, Y_t = y_t)$
- Our focus is to find probing strategies so that U converges to zero as quickly as possible, in some sense to be determined.

Some candidate probing strategies

- *Uniform probing strategy* Repeatedly cycle through the links in a fixed order, probing each link once in each cycle.
- *Greedy probing strategy* Given $\mathcal{F}_t$, compute for each $e \in E$:

$$\gamma_{e,t} = -E[U_{t+1} - U_t | \text{link } e \text{ is probed at } t+1, \mathcal{F}_t].$$

  The greedy probing strategy is to probe a link e at time $t+1$ that maximizes $\gamma_{e,t}$.

- *The k-empirical greedy probing strategy*
  - Follow uniform probing for k rounds
  - Then switch to greedy strategy with estimated $\gamma_{e,t}$ values
  - $\widehat{\gamma}_{e,t}$ is average of $-(U_{t'+1} - U_{t'})$ for k most recent probes of e

- *Best chance probing strategy* This strategy is intended for coordinate convex models.
  - Find $\mathbf{x}^* \in \mathcal{A}$ that maximizes $P[\mathbf{X} \geq \mathbf{x} | \mathcal{F}_t] = \prod_{e:x_e=1} p_{e,t}$ over all $\mathbf{x} \in \mathcal{A}$,
  - Probe a link e with the smallest value of $p_{e,t}$, subject to $x_e^* = 1$.

  Intuitive motivation: Seeking a *certificate* for $\chi = 1$.

Three special cases of the best chance probing strategy:

1. *Best path strategy (best chance strategy for the $S - T$ reachability model)* Probe the worst link of the best path from s to t.
2. *Best tree strategy (best chance strategy for the connectivity model)* Probe the worst link of the best spanning tree.
3. *Middle link strategy (best chance strategy for the majority model)* List the links in order of decreasing $p_{e,t}$ values, and probe the link at (or just past) the middle.

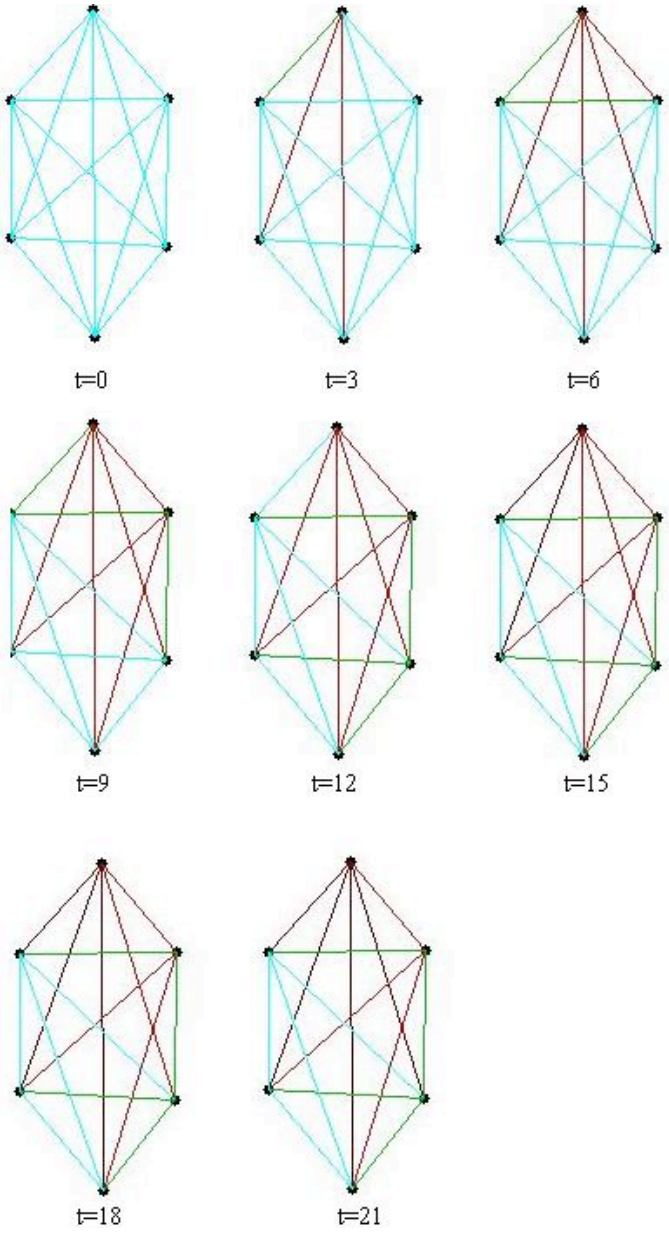


Figure: Best tree strategy: without an all one tree existing

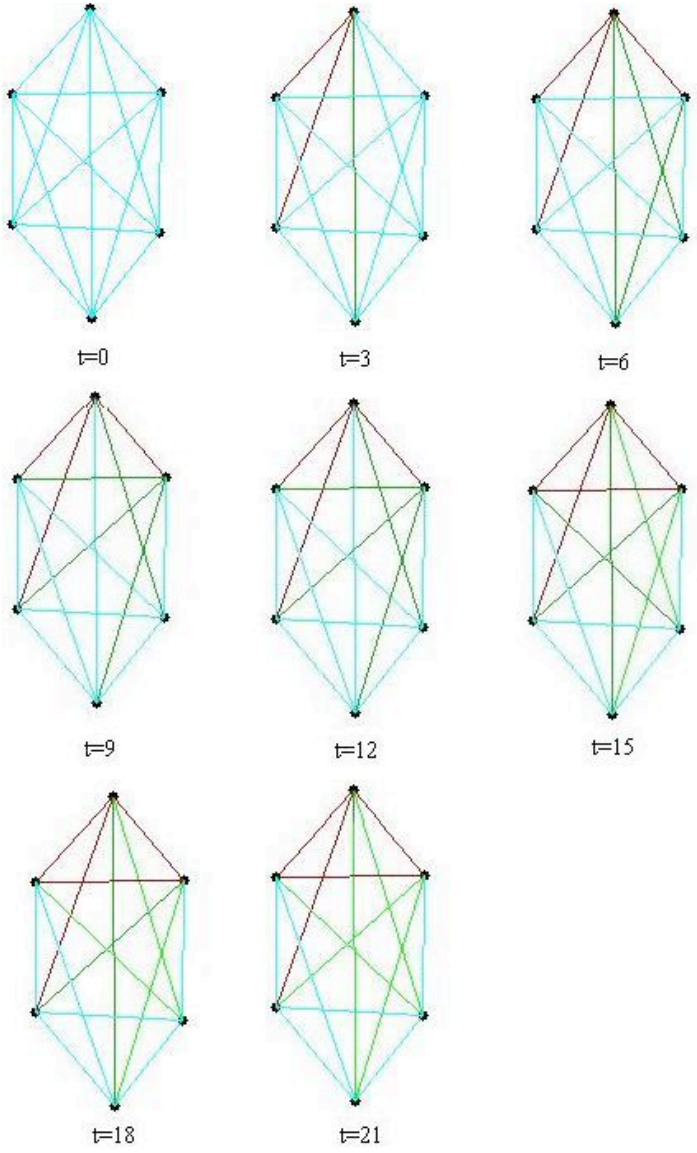


Figure: Best tree strategy: with an all zero tree existing

- *Best-chance worst-block probing strategy* This strategy is also intended for coordinate convex models. The best-chance worst-block probing strategy is to first identify an $x^* \in \mathcal{A}$ that maximizes $P[\mathbf{X} \geq \mathbf{x}|\mathcal{F}_t] = \prod_{e:x_e=1} p_{e,t}$ over all $\mathbf{x} \in \mathcal{A}$, and a $y^* \in \mathcal{B}$ that maximizes $P[\mathbf{X} \leq \mathbf{y}|\mathcal{F}_t] = \prod_{e:y_e=0} \overline{p}_{e,t}$ over all $\mathbf{y} \in \mathcal{B}$. If $P[\mathbf{X} \geq \mathbf{x}^*|\mathcal{F}_t] \geq P[\mathbf{X} \leq \mathbf{y}^*|\mathcal{F}_t]$ then a link with the smallest value of $p_{e,t}$ is probed from among those links e with $x_e^* = 1$. Otherwise, a link with the largest value of $p_{e,t}$ is probed from among those links e with $y_e^* = 0$.
  1. *Best-path worst-cut strategy (best-chance worst-block strategy for the $S - T$ reachability model)*

- *Closest to 0.5 strategy* The closest to 0.5 strategy is proposed specifically for the parity model. For this strategy and any $t \geq 0$, a link e with $p_{e,t}$ closest to 0.5 is probed at time $t + 1$. Intuitively, such a link causes the greatest uncertainty about the parity of the set of links.

Example: One link network

$E = \{e\}$, $\{0,1\}^E = \{0,1\}$, and $\mathcal{G} = \{1\}$.
(Here drop subscript e on the variables a_e, π_e, $N_{e,t}$, and $K_{e,t}$.)
$N_t = t$ and $G_t = p_{e,t} = \phi(t, K_t, a, \pi)$.
For such a network $N_t = t$ and $G_t = p_{e,t} = \phi(t, K_t, a, \pi)$.

$$E[U_t|X_e = 1] = \sum_{j=0}^{t} \binom{t}{j} a^j (\overline{a})^{t-j} h(\phi(t, j, a, \pi)). \qquad (1)$$

$$\text{median}(G_t|X_e = 1) \approx \phi(t, at, a, \pi) = \frac{\frac{\pi}{\overline{\pi}} \left(\frac{a}{\overline{a}}\right)^{(2a-1)t}}{\frac{\pi}{\overline{\pi}} \left(\frac{a}{\overline{a}}\right)^{(2a-1)t} + 1}.$$

$$\text{median}(U_t|X_e = 1) \approx h(\text{median}(G_t|X_e = 1)) \approx h(\phi(t, at, a, \pi)).$$

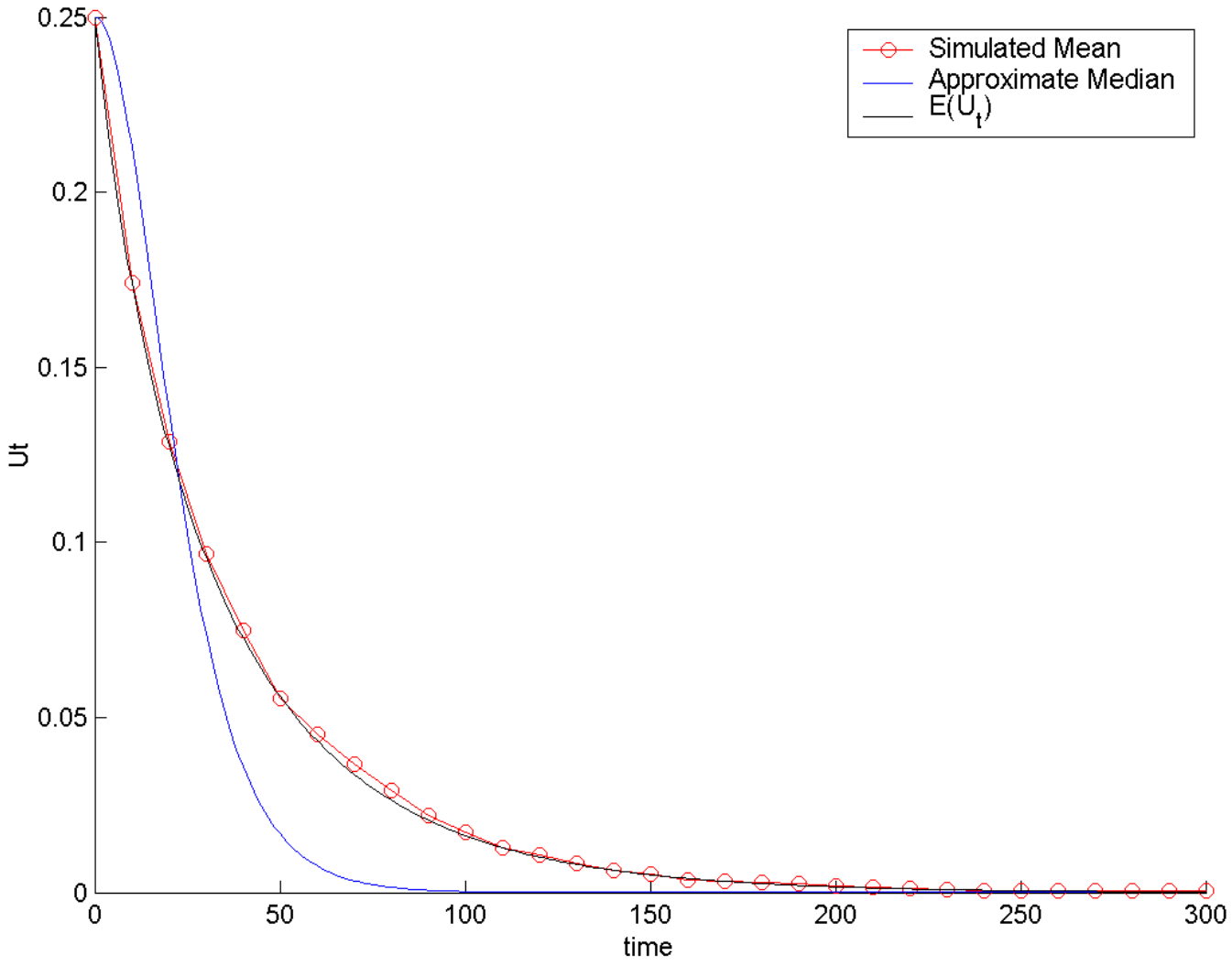


Figure: The mean, approximate median, and observed empirical average of U_t in simulations for the one link network, for $a = 0.6, \pi = 0.5$ and $h(p) = p\overline{p}$.

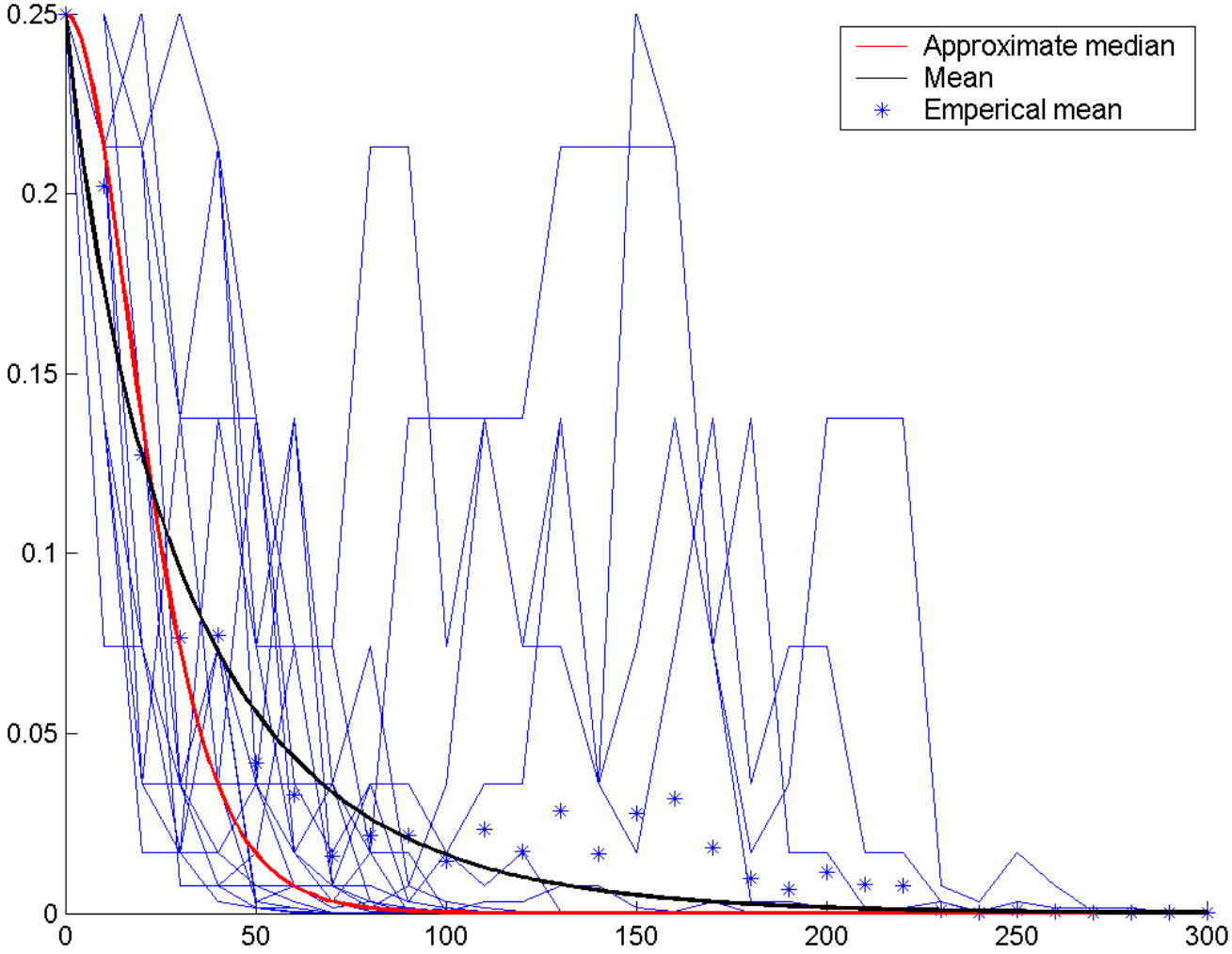


Figure: Twenty sample paths for one link network, $a = 0.6, \pi = 0.5$ and $h(p) = p\overline{p}$.

Extension of median approximation to larger networks

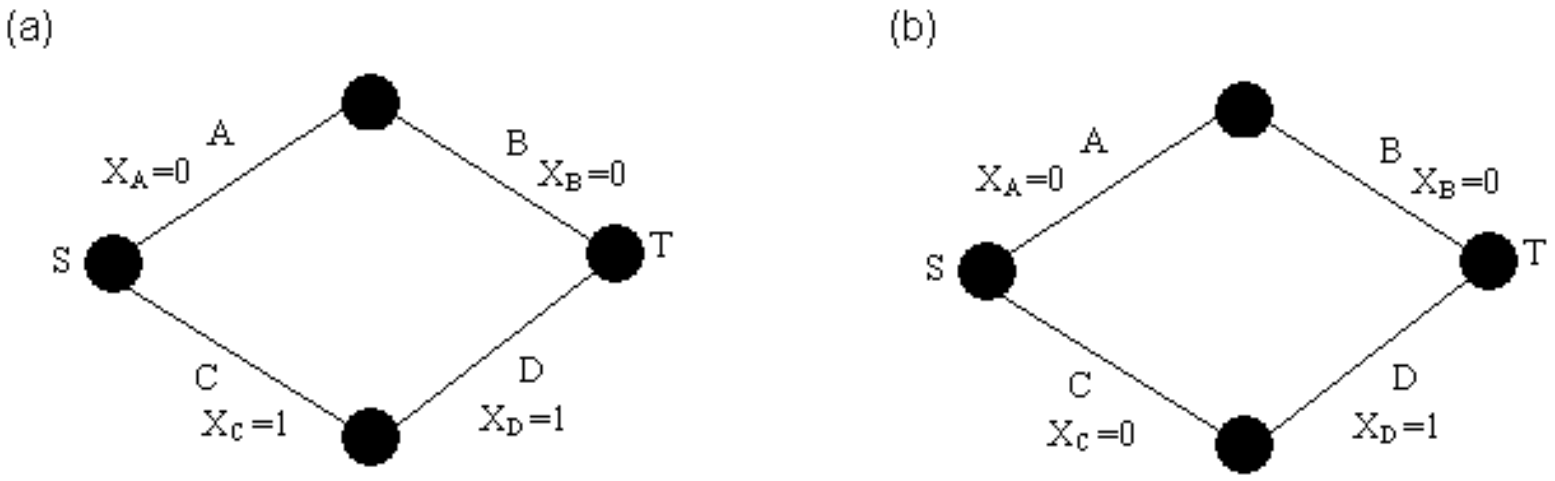


Figure: Example graph and two different network states.

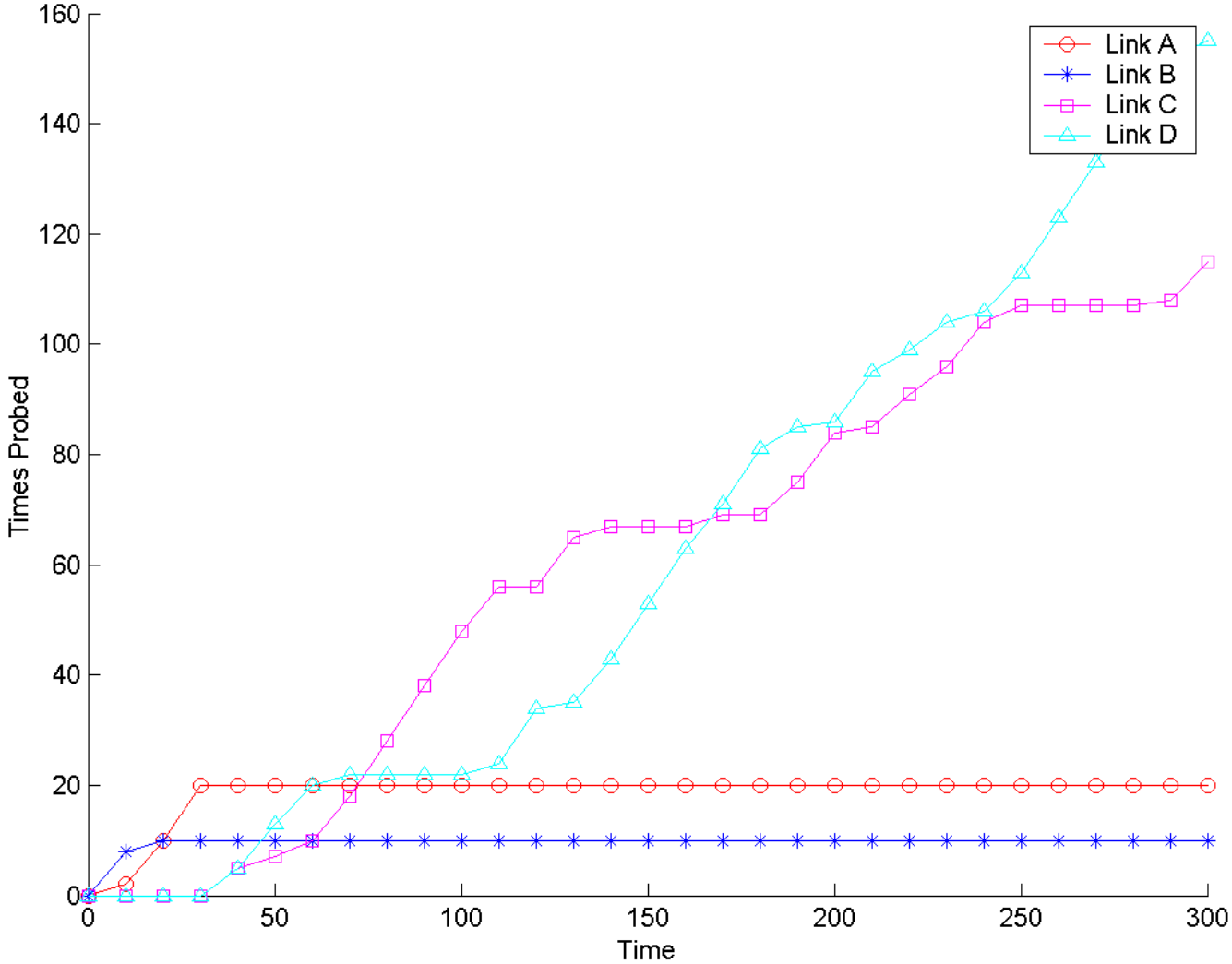


Figure: Number of probes for each link using the greedy strategy for the network state in Figure 5(a)

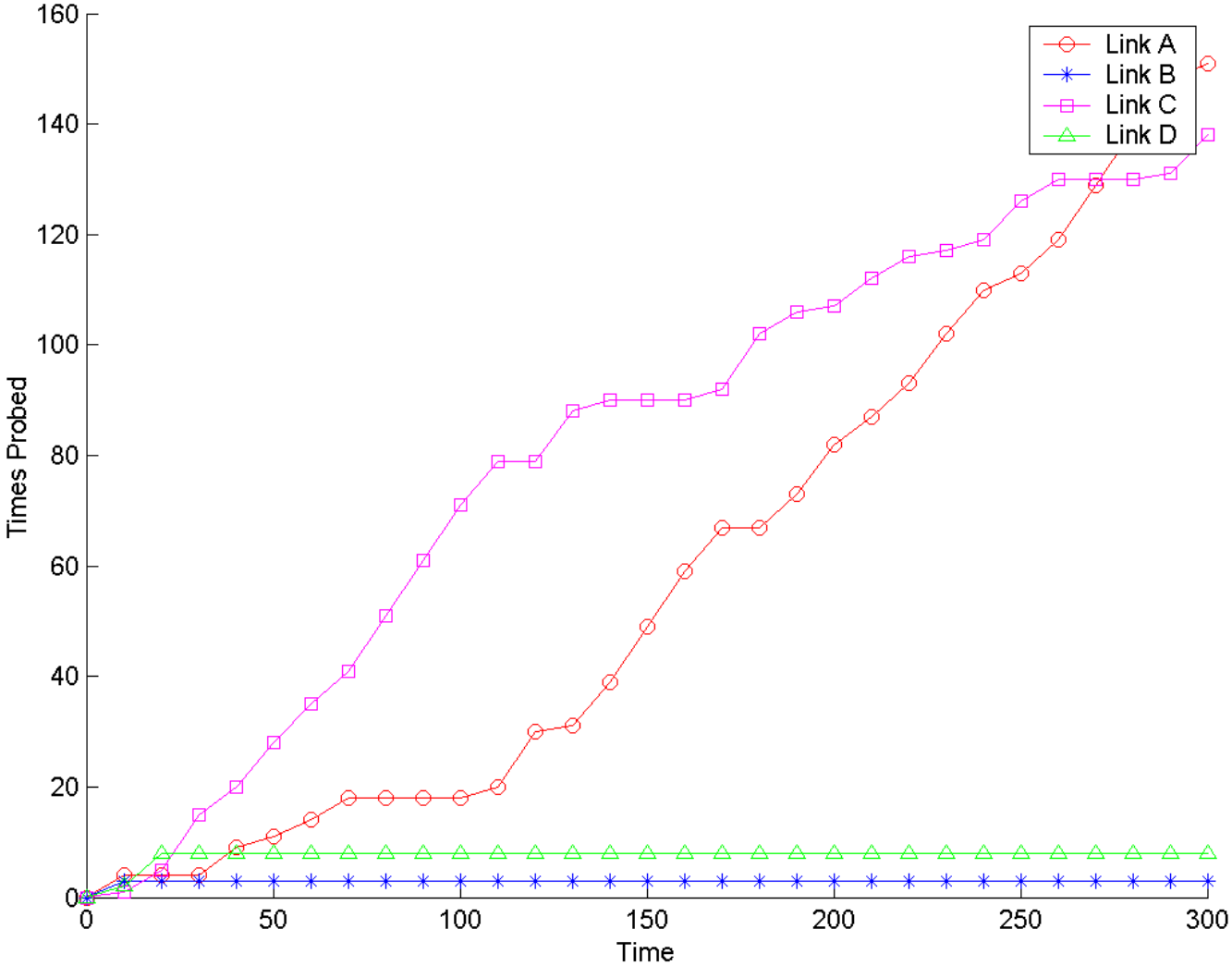


Figure: Number of probes for each link using the best path strategy for the network state in Figure 5(b)

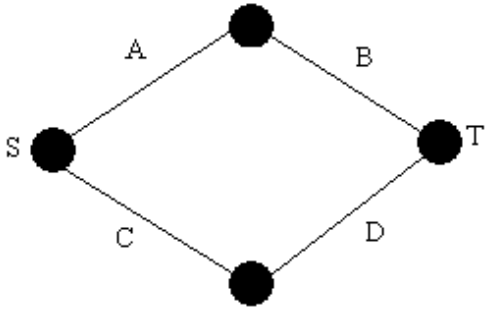
Suggests large time approximation:

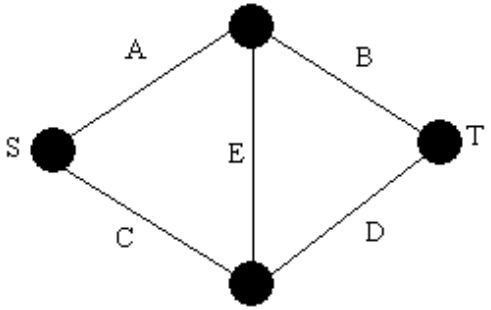
$$\mathrm{median}(U_t | X_E = 1) \approx h(\phi(t/C, at/C, a, \pi)). \tag{2}$$

Expect to hold rather generally, where C is the number of links that are probed infinitely often.

Comparison of Strategies by Simulation

The slides that follow present various simulation results, which won't be covered in this talk for lack of time. They used $\pi_e = 0.5$ and $a_e = 0.6$, for all e. Each simulation covered 300 time slots and the plotted empirical averages of U_t were computed using 1,000 samples. The four and five link networks referred to in the simulation captions are the following:





The curves labeled WorstCut are for the best-path worst-cut strategy. The k-empirical greedy strategy was simulated for $k = 1, 5, 10$.

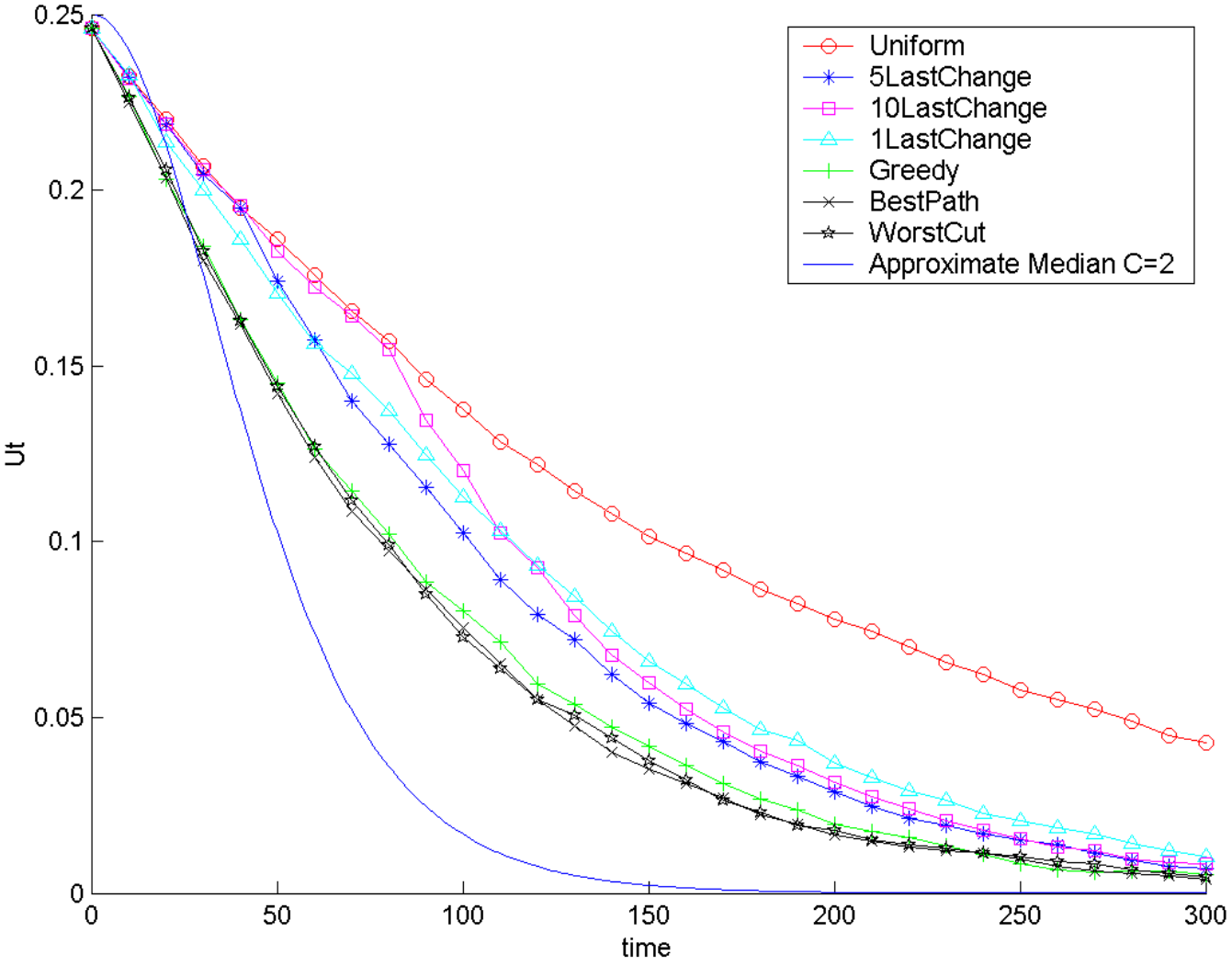


Figure: Averages of U_t for various strategies and the four link graph.

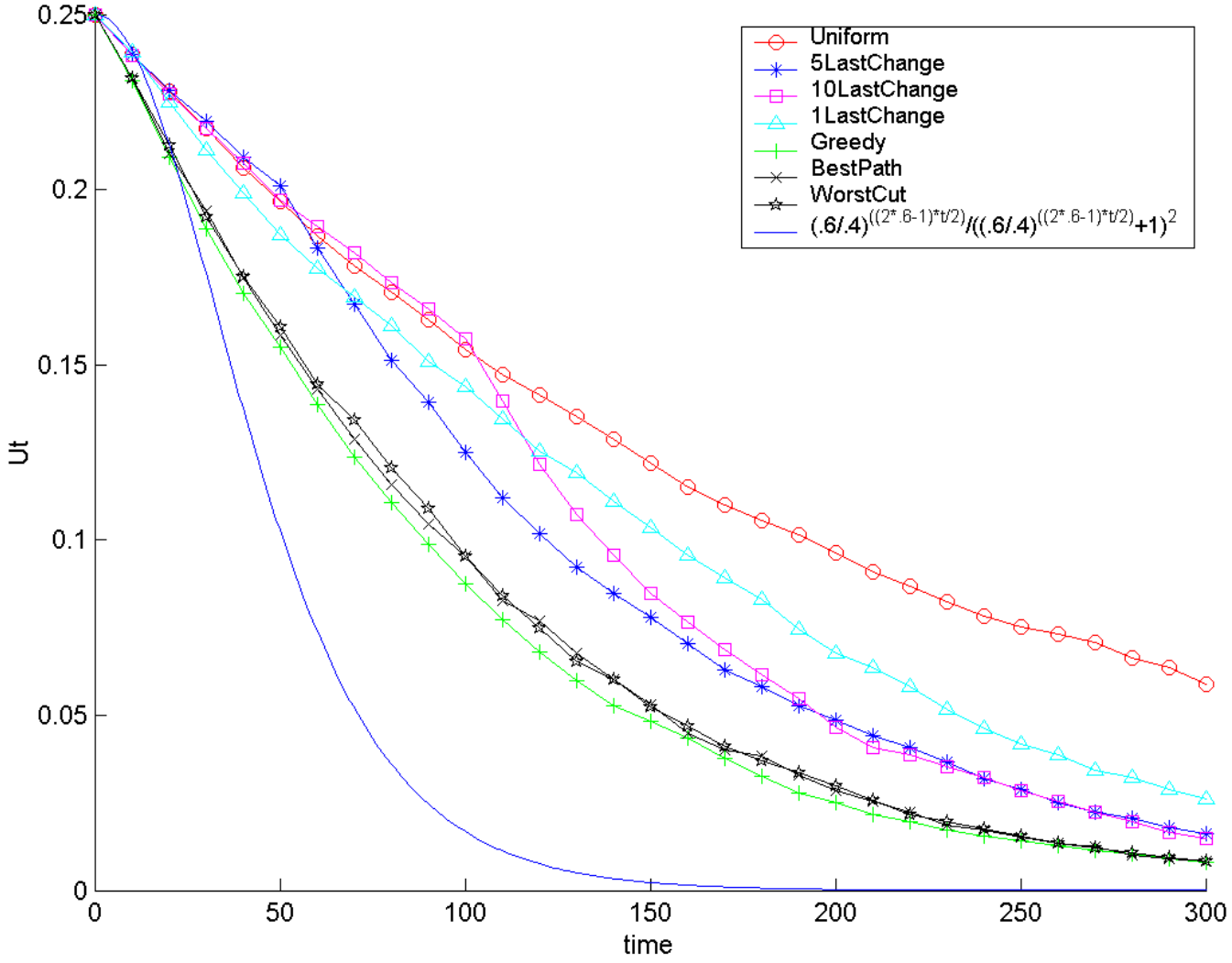


Figure: Average value of U_t for various strategies and the five link graph.

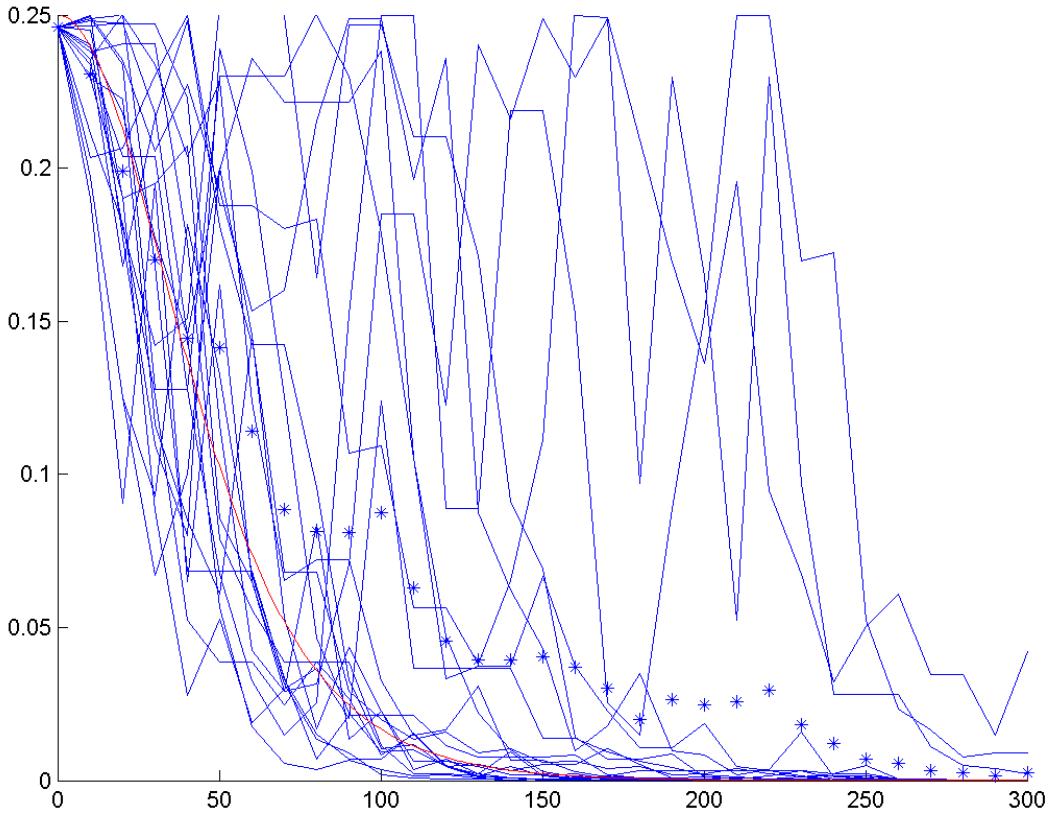


Figure: Twenty iterations of the S-T connectivity model using the greedy strategy with stars placed in the mean value. The solid line is the approximate median

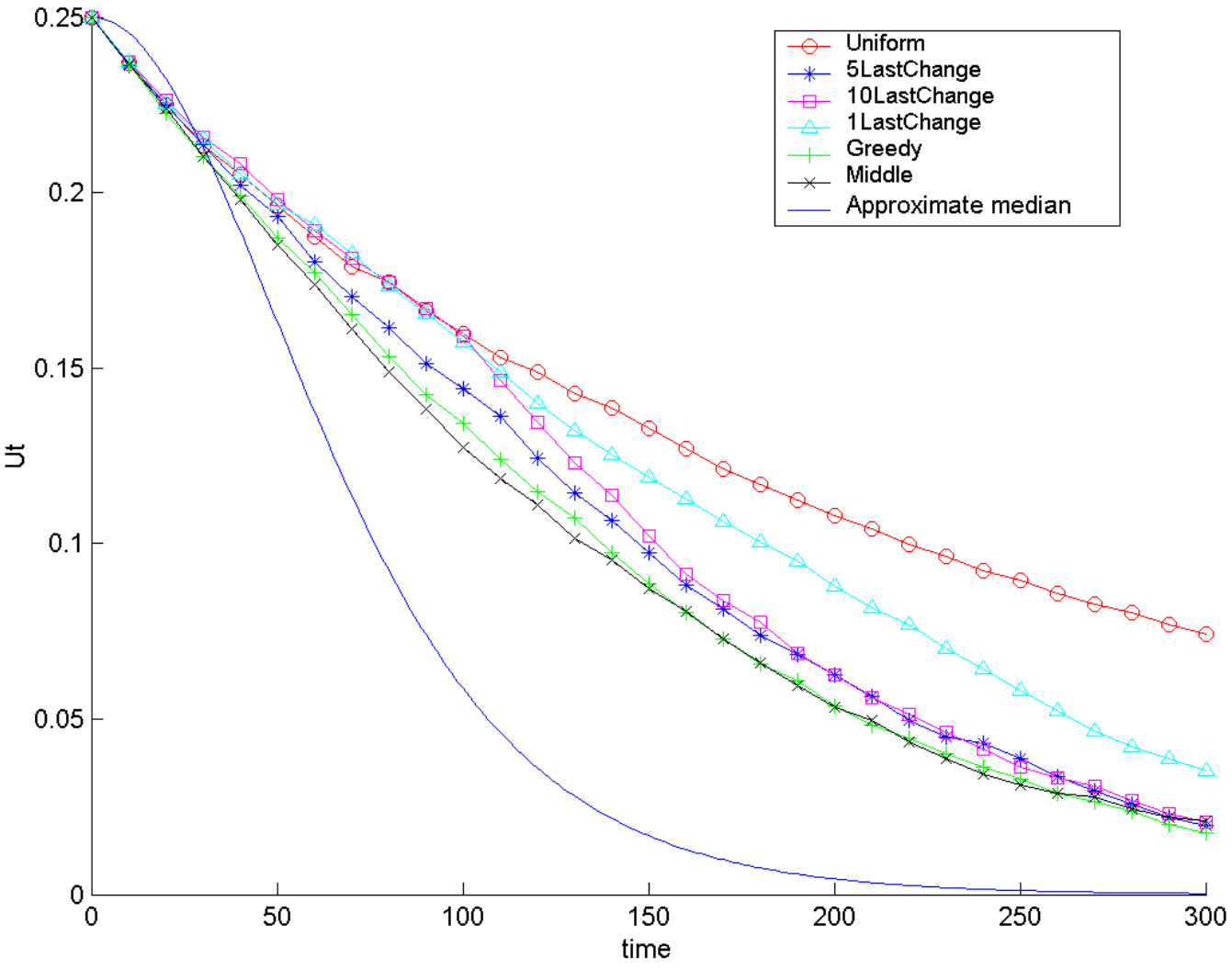


Figure: Average U_t for the majority model using five links

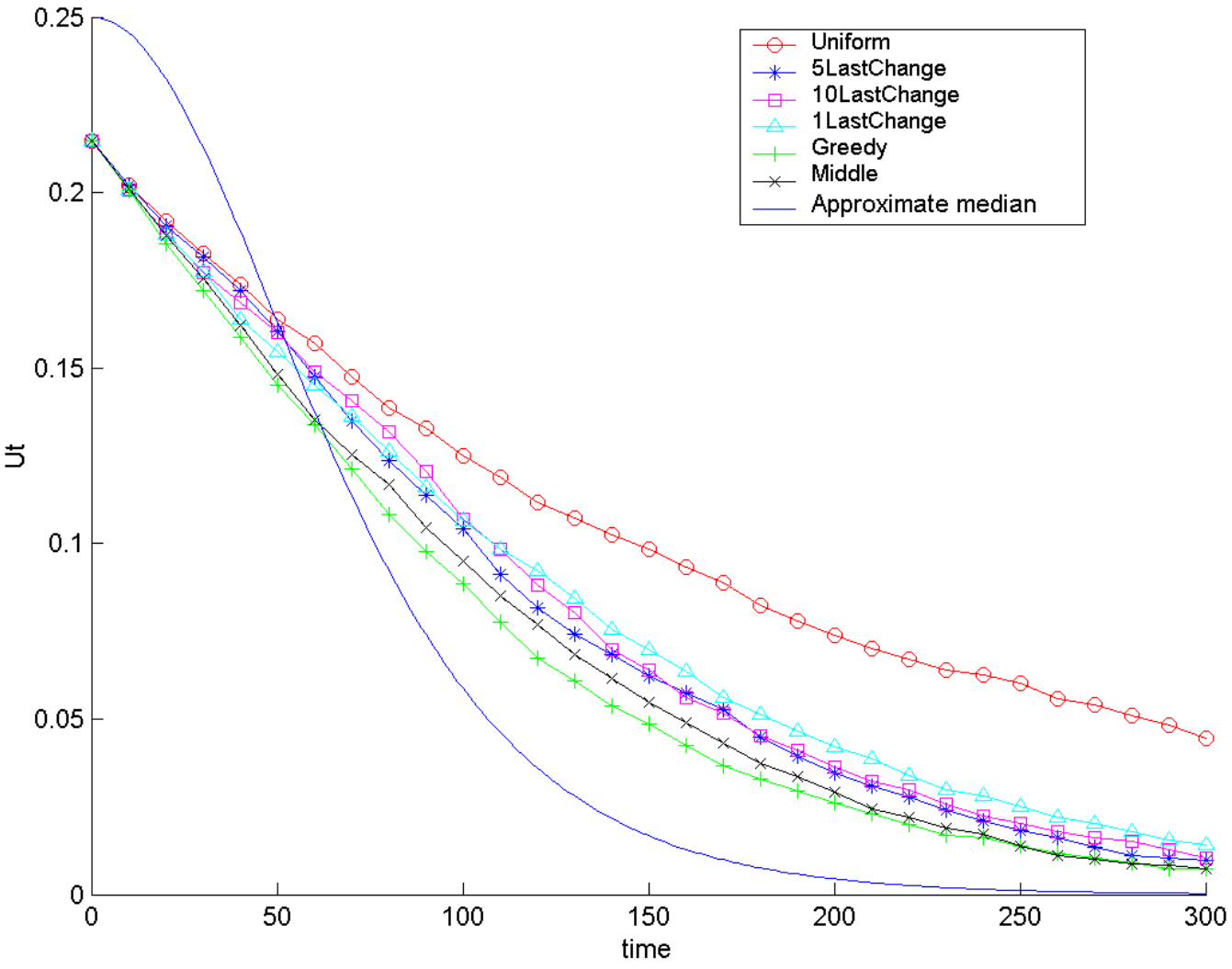


Figure: Average U_t for the majority model using four links

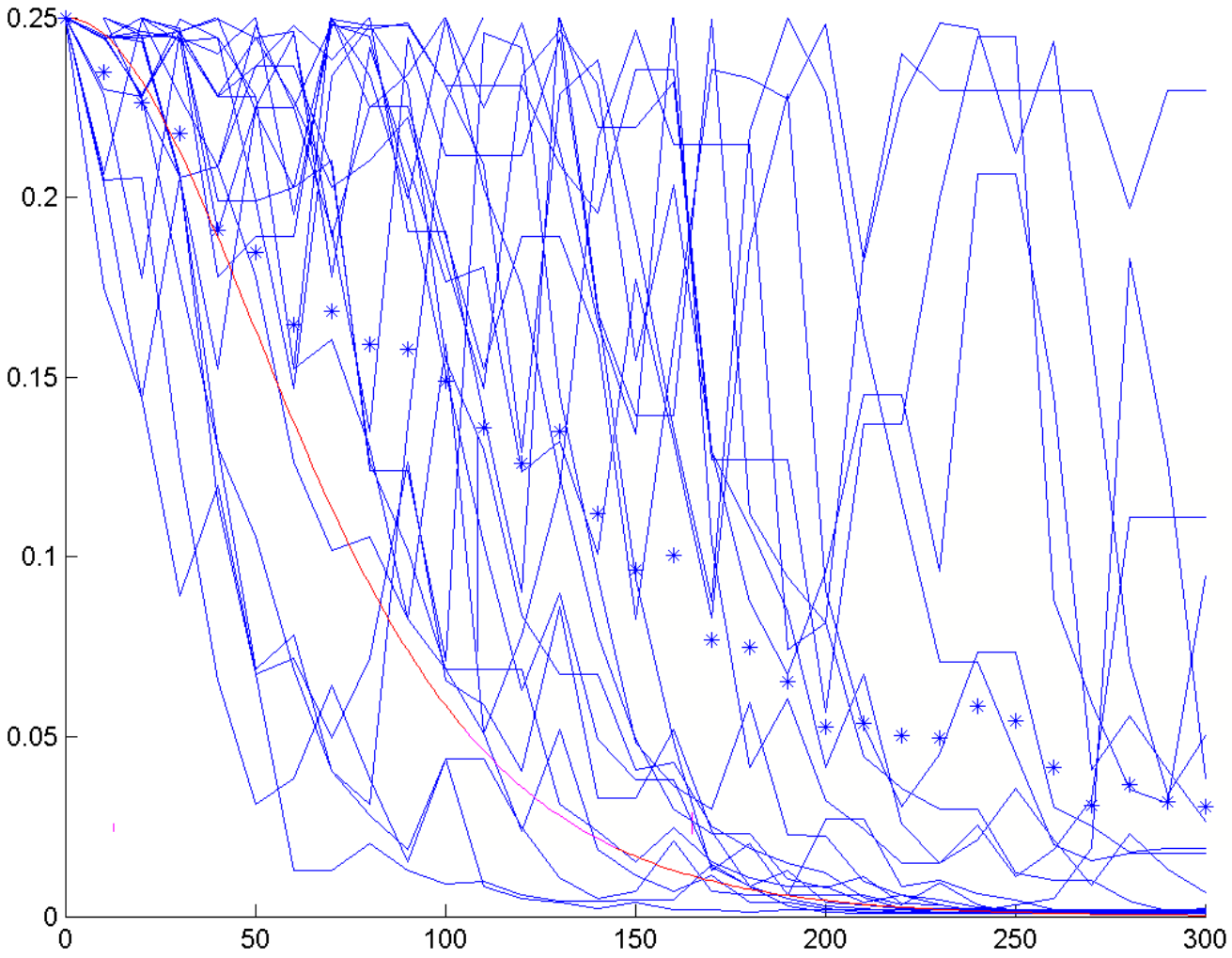


Figure: Twenty iterations of the majority model using the greedy strategy with stars placed in the mean value. The solid line is the approximate median.

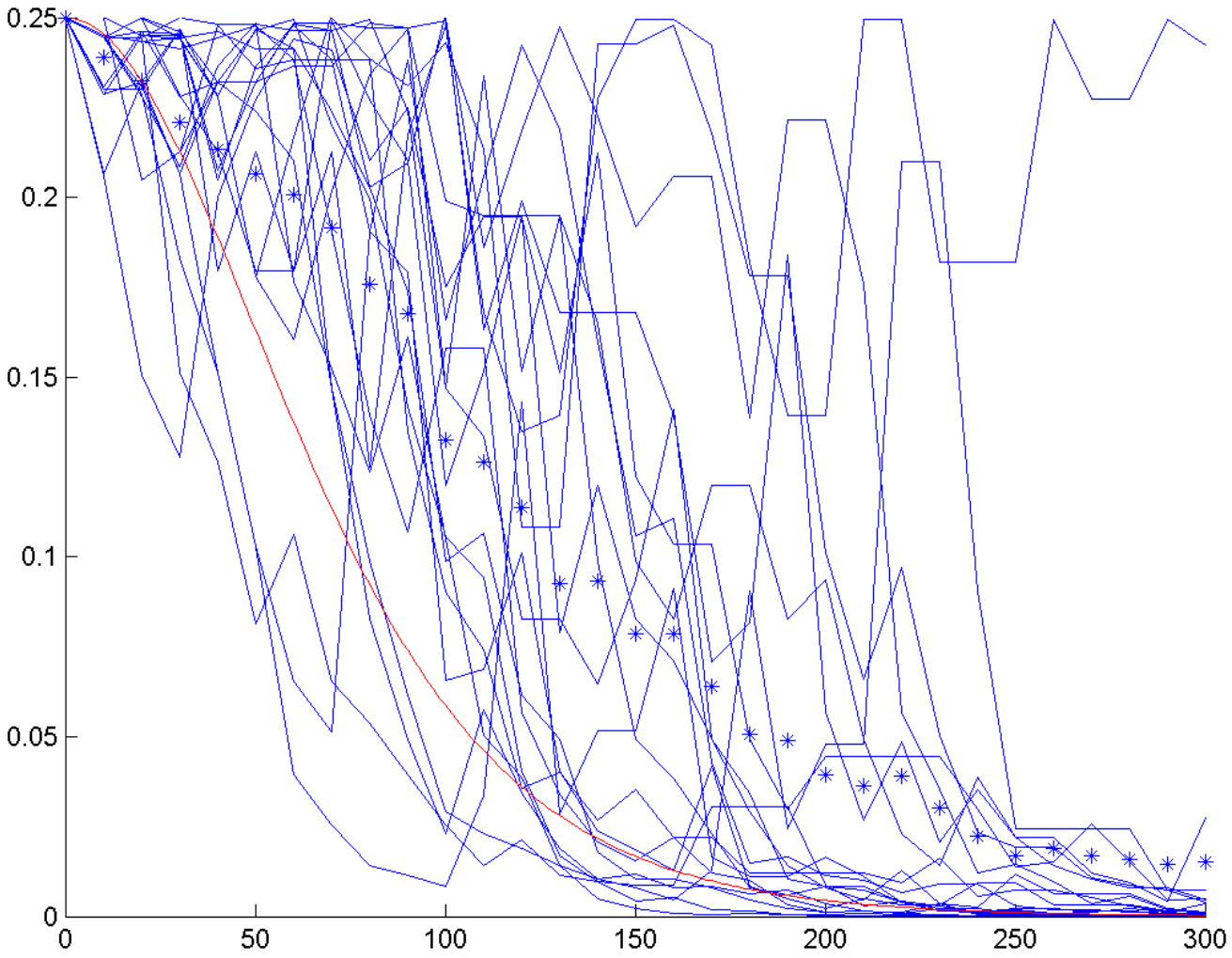


Figure: Twenty iterations of the majority model using the middle strategy with stars placed in the mean value. The solid line is the approximate median.

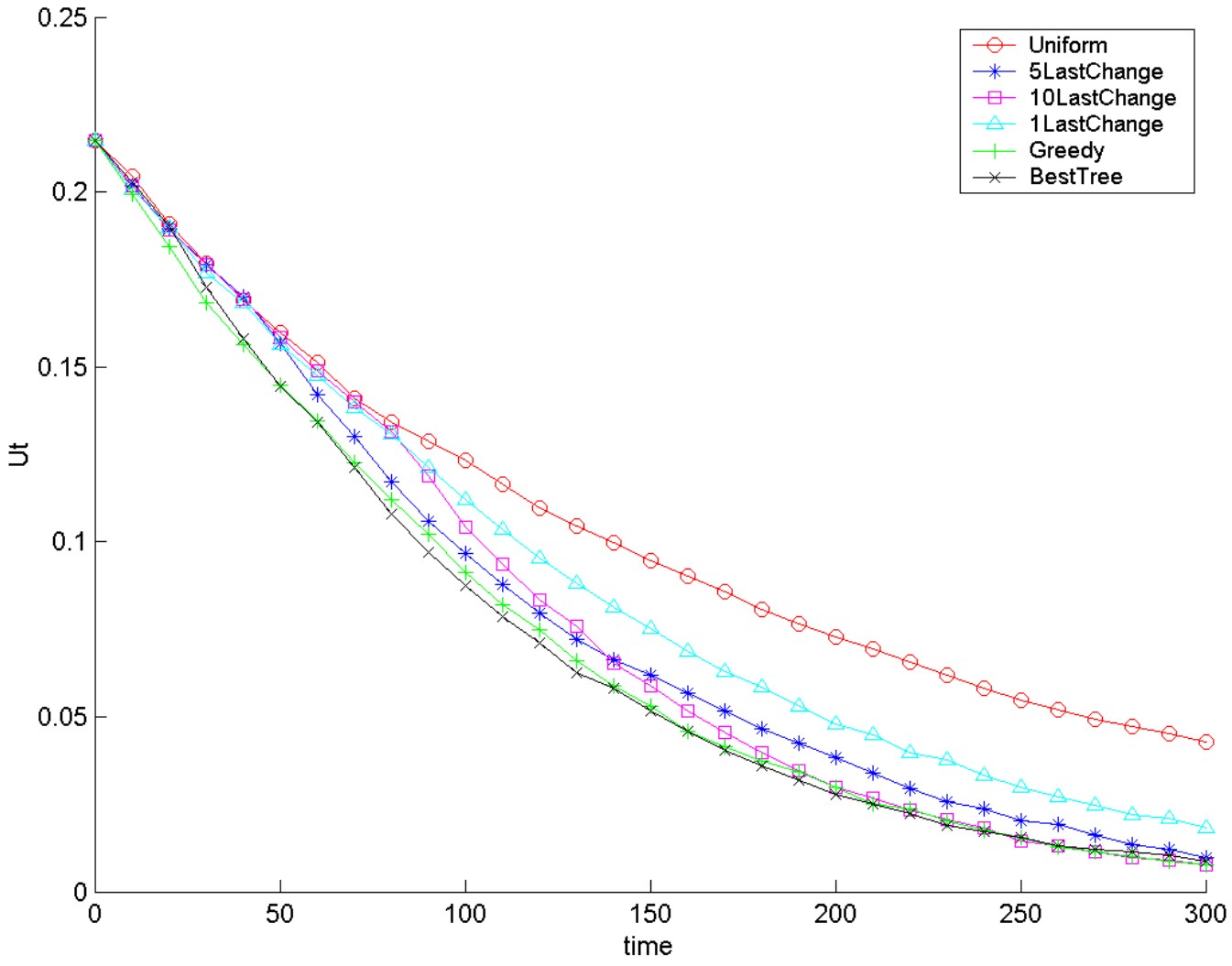


Figure: Average U_t for the connectivity model for the four link graph.

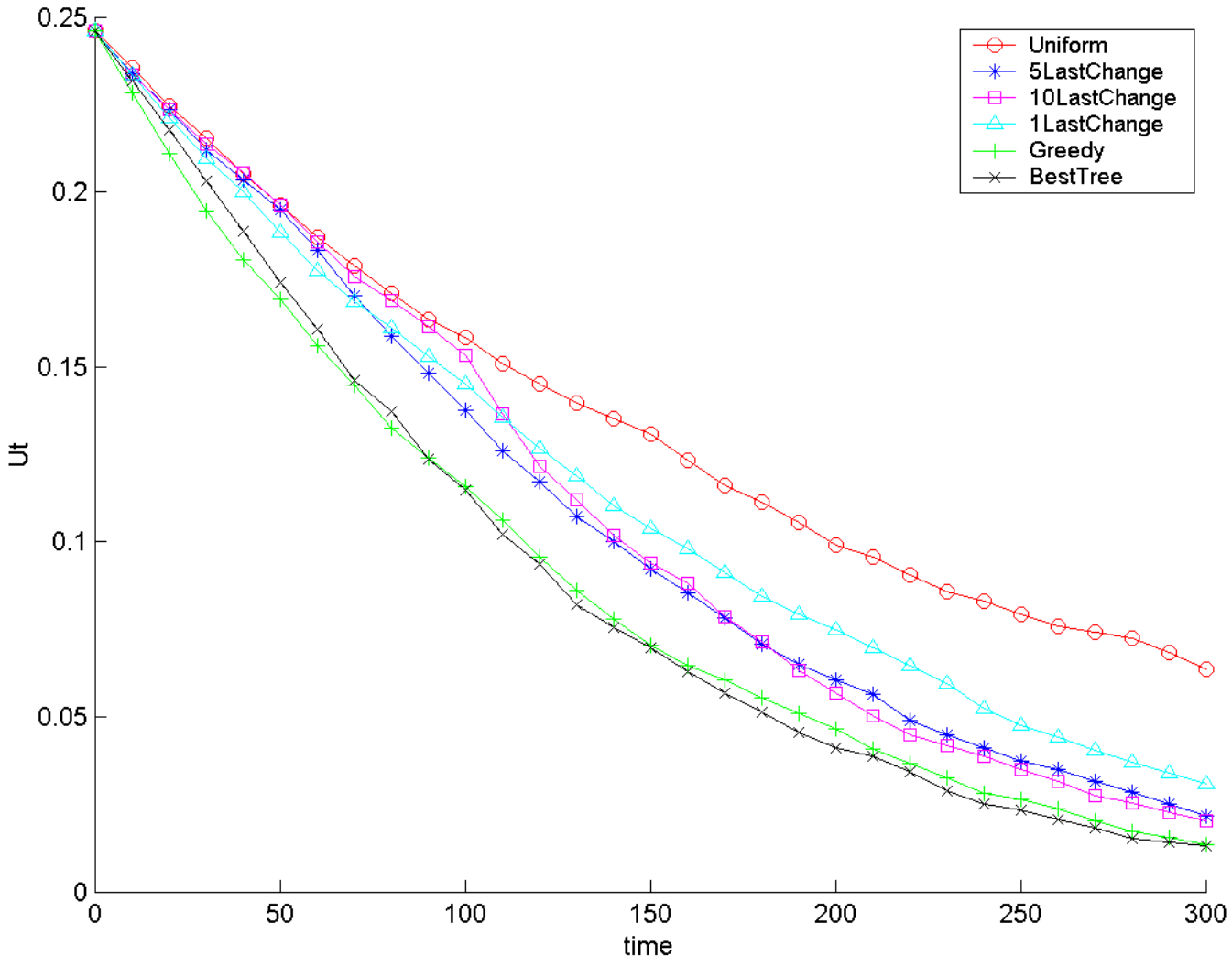


Figure: Average U_t for the connectivity model for the five link graph.

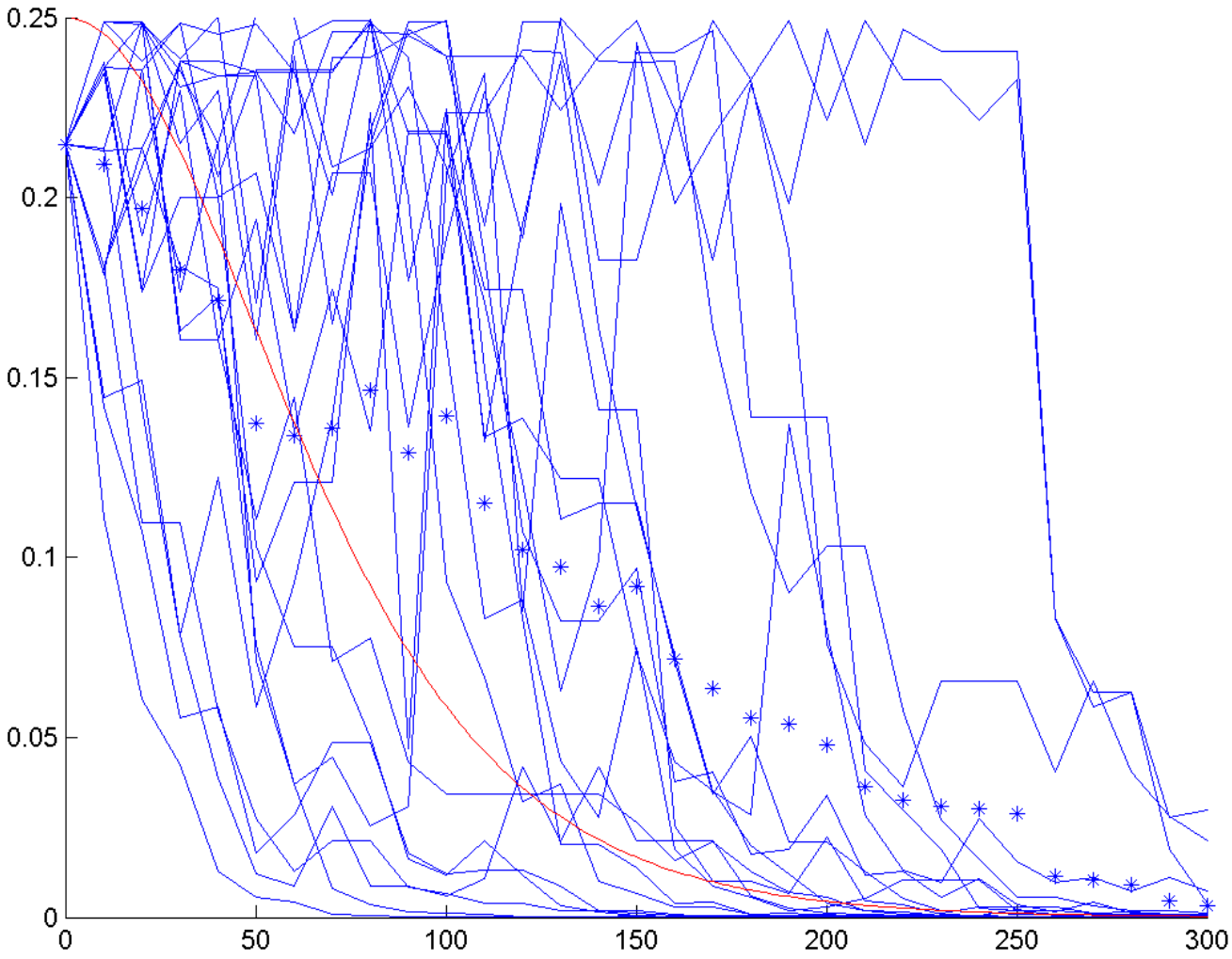


Figure: Twenty sample paths for the connectivity model for the four link graph and the greedy strategy. Stars indicate the sample mean value and the solid line is the approximate median.

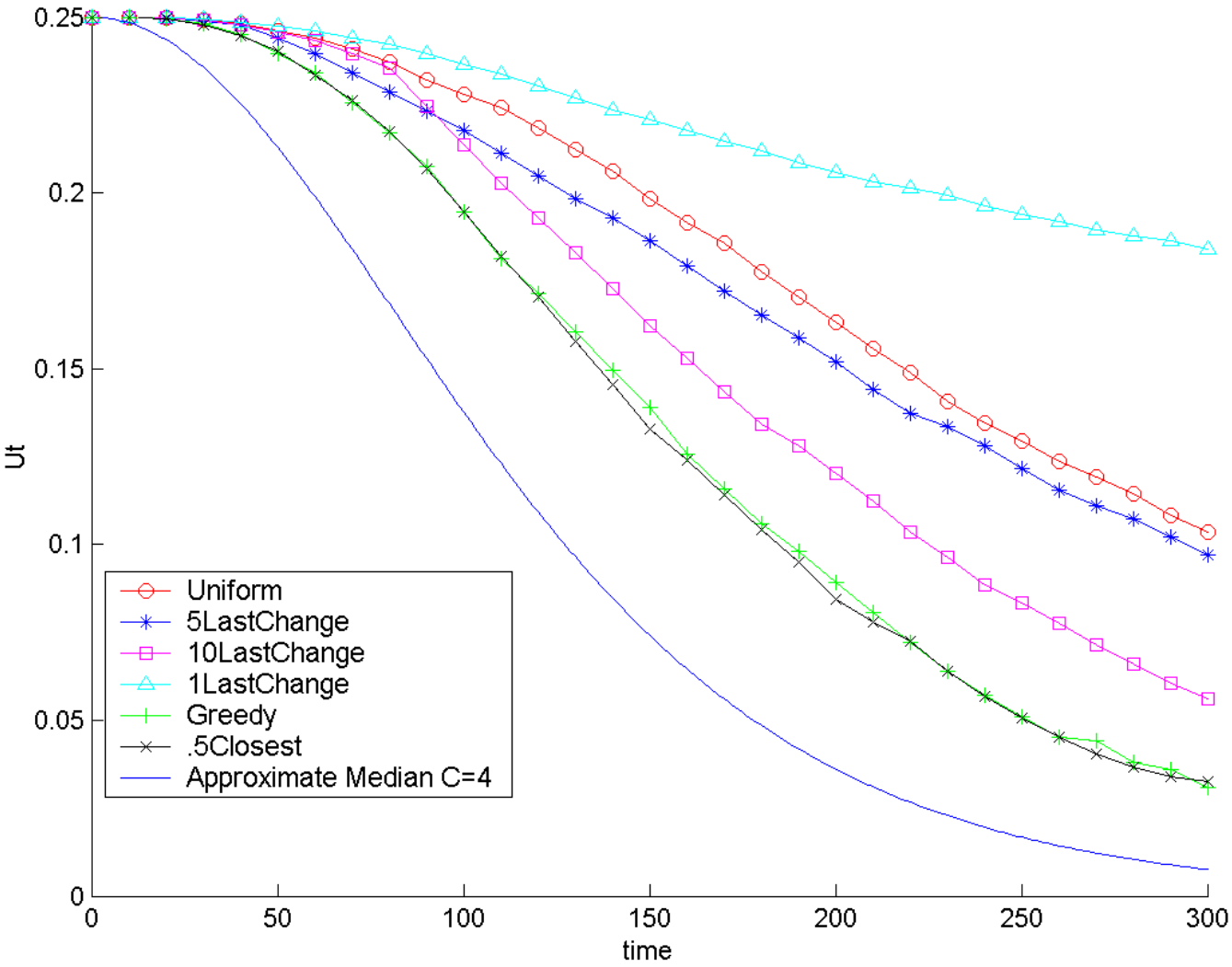


Figure: Average U_t for the parity model for four links

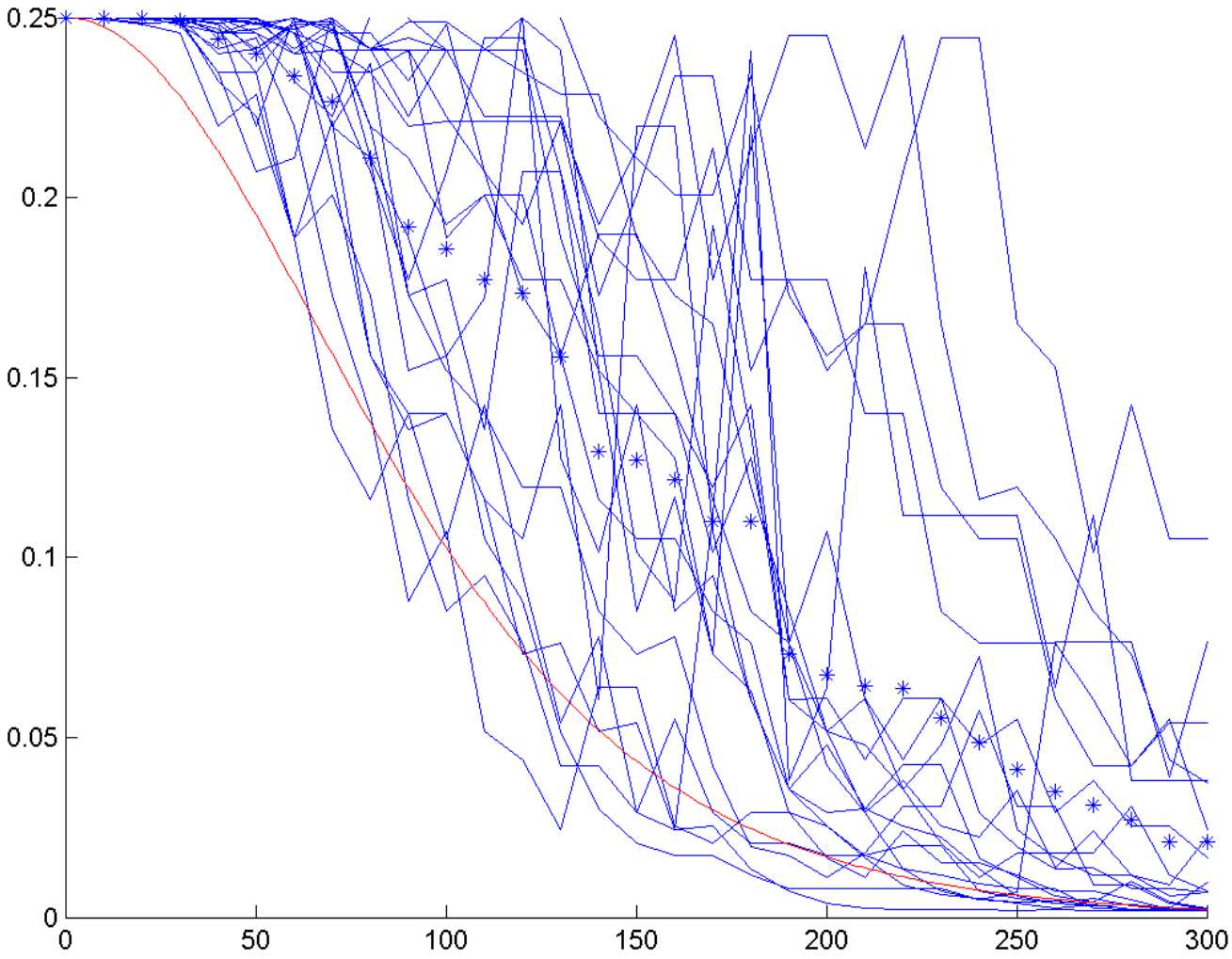
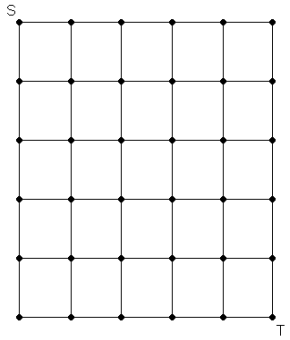


Figure: Twenty iterations of the parity model using the closed to 0.5 strategy. Stars mark sample mean values and the solid line is the approximate median.

Directions to explore

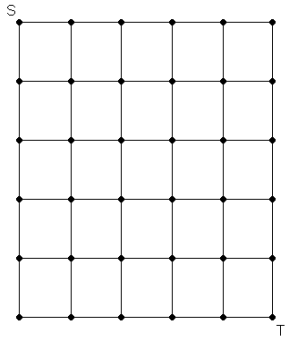
- Three types of results
  - Small networks, exact optimality (e.g. closest to 0.5 for partiy model)
  - Asymptotic behavior for large times
  - Asymptotic behavior for large networks
- Improved probing strategies for various models. For example, for $S - T$-connectivity and a graph such as



  we can be more confident of a path existing in the middle section of the network.
- More complex models: $X_e \in [0, 1]$, or require distributed probing process, accounting for trust in reports.

Directions to explore

- Three types of results
  - Small networks, exact optimality (e.g. closest to 0.5 for partiy model)
  - Asymptotic behavior for large times
  - Asymptotic behavior for large networks
- Improved probing strategies for various models. For example, for $S - T$-connectivity and a graph such as



  we can be more confident of a path existing in the middle section of the network.
- More complex models: $X_e \in [0, 1]$, or require distributed probing process, accounting for trust in reports.

Thanks!