Decentralized routing on stochastic temporal networks

R. Lambiotte
Department of Mathematics
Research group on Comple**x**ity and **N**etworks
Namur Center for Complex Systems
University of Namur, Belgium

Decentralized Routing on Spatial Networks with Stochastic Edge Weights, Till Hoffmann, R.L. and Mason Porter, PRE 2013

From static to temporal networks

Empirical data on interacting systems are more and more often time-resolved: timings at which events take place in the system.

Observations

Node and link dynamics exhibits complex temporal patterns

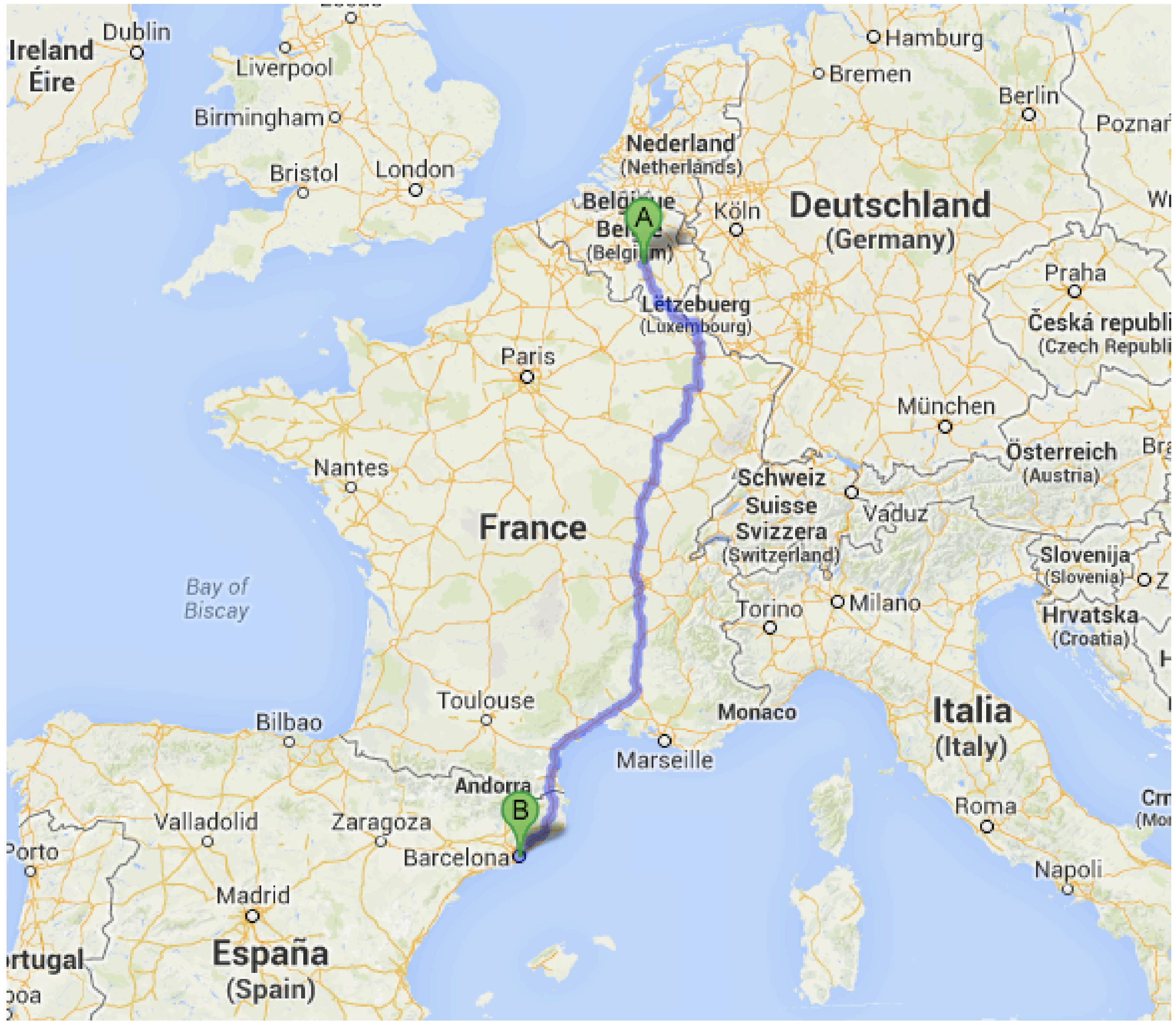
Models

Burstiness and temporal correlations affect spreading

Algorithms?

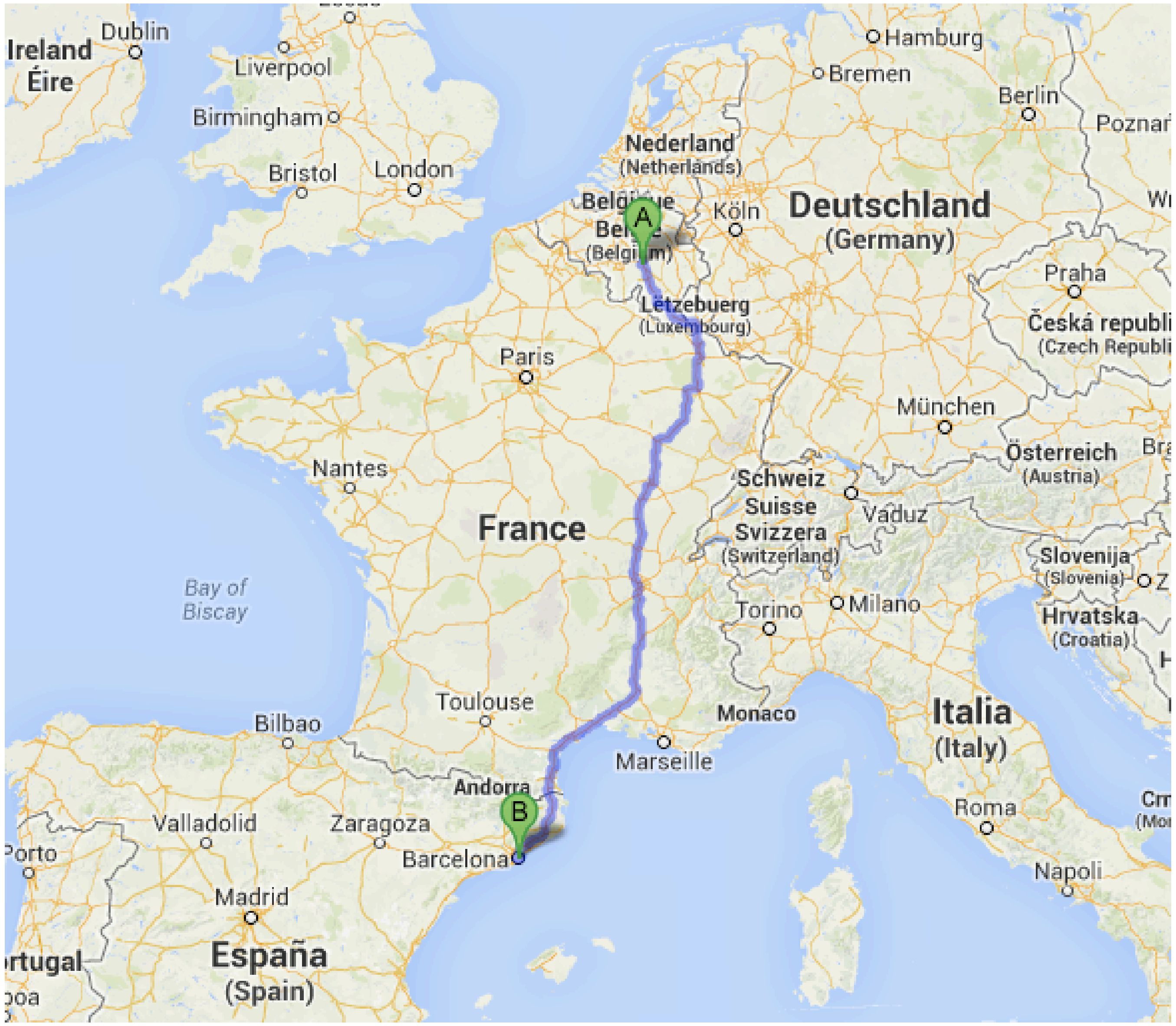
Modularity
?
?

Routing algorithms for static networks



On static networks: finding shortest paths can be done efficiently

Routing algorithms for static networks



On static networks: finding shortest paths can be done efficiently
... but requires complete knowledge of the system.

Routing algorithms for static networks

brief communications

Navigation in a small world

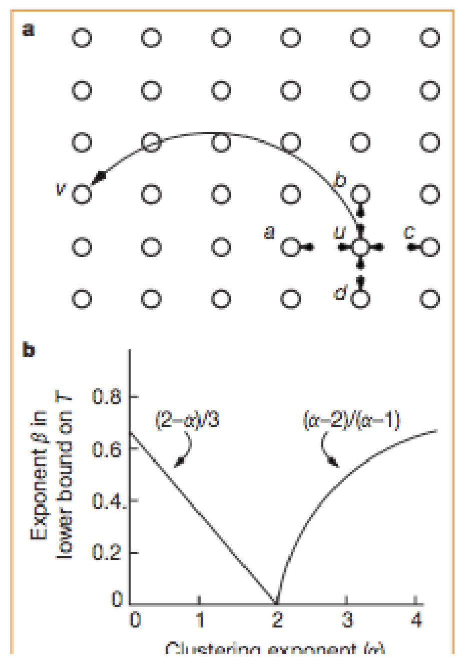
It is easier to find short chains between points in some networks than others.

The small-world phenomenon — the principle that most of us are linked by short chains of acquaintances — was first investigated as a question in sociology[1,2] and is a feature of a range of networks arising in nature and technology[3–5]. Experimental study of the phenomenon[1] revealed that it has two fundamental components: first, such short chains are ubiquitous, and second, individuals operating with purely local information are very adept at finding these chains. The first issue has been analysed[2–4], and here I investigate the second by modelling how individuals can find short chains in a large social network.

I have found that the cues needed for discovering short chains emerge in a very simple network model. This model is based on early experiments[1], in which source individuals in Nebraska attempted to transmit a letter to a target in Massachusetts, with the letter being forwarded at each step to someone the holder knew on a first-name basis. The networks underlying the model follow the 'small-world' paradigm[3]: they are rich in structured short-range connections and have a few random long-range connections.

Long-range connections are added to a two-dimensional lattice controlled by a clustering exponent, α, that determines the probability of a connection between two

tions follow an inverse-square distribution, there is a decentralized algorithm that achieves a very rapid delivery time; T is bounded by a function proportional to $(\log N)^2$. The algorithm achieving this bound is a 'greedy' heuristic: each message holder forwards the message across a con-



nection that brings it as close as possible to the target in lattice distance. Moreover, $\alpha = 2$ is the only exponent at which any decentralized algorithm can achieve a delivery time bounded by any polynomial in $\log N$: for every other exponent, an asymptotically much larger delivery time is required, regardless of the algorithm employed (Fig. 1b).

These results indicate that efficient navigability is a fundamental property of only some small-world structures. The results also generalize to d-dimensional lattices for any value of $d \geq 1$, with the critical value of the clustering exponent becoming $\alpha = d$. Simulations of the greedy algorithm yield results that are qualitatively consistent with the asymptotic analytical bounds (Fig. 1c).

In the areas of communication networks[7] and neuroanatomy[8], the issue of routing without a global network organization has been considered; also in social psychology and information foraging some of the cues that individuals use to construct paths through a social network or hyperlinked environment have been discovered[9,10]. Although I have focused on a very clean model, I believe that a more general conclusion can be drawn for small-world networks — namely that the correlation between local structure and long-range connections pro-

However, some networks can be navigated efficiently by using only local information (see Milgram)
Decentralized algorithms for networks embedded in space.

Routing algorithms for temporal networks

Definition of the **stochastic shortest path** problem when the network is not deterministic

Algorithms to find shortest paths

Approximate, **decentralized** algorithm finding good paths does not require the knowledge of the whole system.

Routing on deterministic networks

We associate a *weight* T_{ij}, representing a cost or travel time, with each edge (i, j). A *path* ℓ with k steps is a sequence of $k+1$ nodes $\ell = \{i_1, \ldots, i_{k+1}\}$ that are connected to one another via edges. The weight T_ℓ of a path ℓ is given by the sum of the weights of its constituent edges:

$$\boxed{T_\ell = \sum_{j=1}^{k} T_{i_j i_{j+1}} .}$$

The shortest-path problem (SPP) aims to determine the path from an *origin* node to a *target* node that has the smallest total weight. In the DSPP, each edge weight T_{ij} is deterministic, and a path with minimal total weight is considered to be *optimal*.

Routing on stochastic networks

The weights are now stochastic variables, distributed with a given probability.
The probability that edge ij has weight t is:

$$\boxed{p_{ij}(t)}$$

This random variable can be:
- a nondeterministic travel time in a transportation networks
- the waiting time of a random walker before an edge appears on a temporal network

In this work, we assume that:
- weights are independent of each other,
- the PDFs do not change during the routing process

=> Distributions are assigned to edges instead of weights (scalar)

Routing on stochastic networks

The probability for a path l to have a certain weight is:

$$\boxed{p_\ell(t) = \left(\mathop{*}_{j=1}^{k} p_{i_j i_{j+1}} \right)(t) \ ,} \qquad (1)$$

where the right-hand side denotes k consecutive convolutions. The probability to traverse the path ℓ and incur a weight $T_\ell \leq t$ is given by the cumulative distribution function (CDF)

$$u_\ell(t) = \int_0^t dt' \, p_\ell(t') \ .$$

What is the shortest path?

In contrast with the deterministic case, there is no longer a unique concept of optimality

Frank defines a path to be optimal if its cdf surpasses a threshold within the shortest time

Fan suggests maximizing the cdf for a given time budget

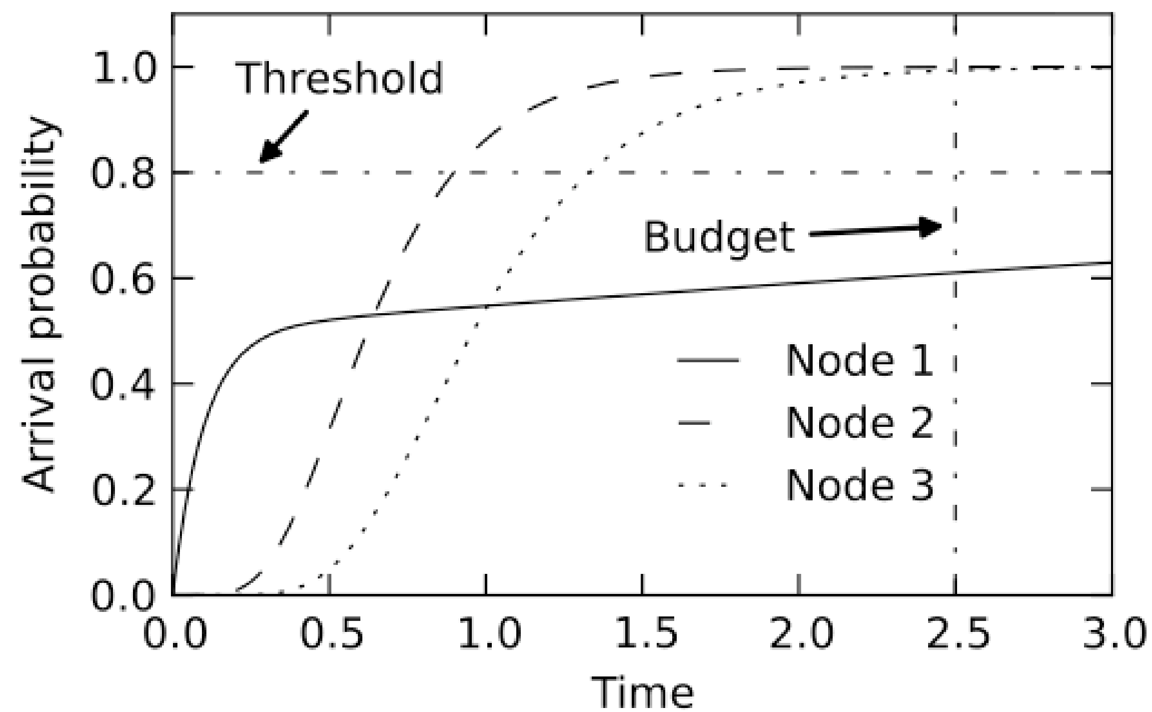


FIG. 1: Comparison of path optimality criteria using CDFs of three paths. Fan et al.'s criterion prefers paths 2 and 3 to path 1 but cannot discriminate between the CDFs of paths 2 and 3. Frank's criterion prefers path 2 to path 3 but cannot be applied to path 1. The joint criterion is applicable to all CDFs and chooses path 2 as the optimal one.

Y. Y. Fan, R. E. Kalaba, and J. E. Moore, *Journal of Optimization Theory and Applications 127, 497(2005).*
H. Frank*, Operations Research 17, pp. 583(1969).*

What is the shortest path?

In contrast with the deterministic case, there is no longer a unique concept of optimality

Frank defines a path to be optimal if its cdf surpasses a threshold within the shortest time **BAD**

Fan suggests maximizing the cdf for a given time budget **GOOD**

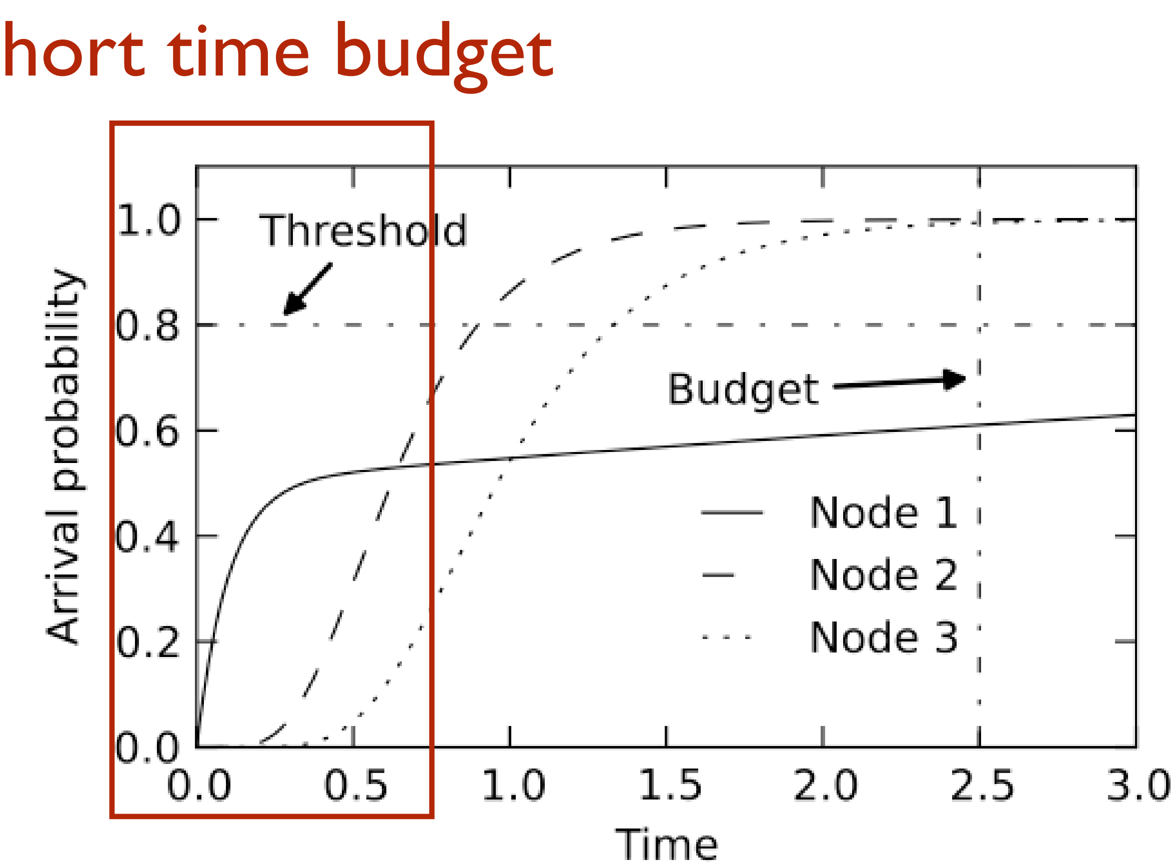


FIG. 1: Comparison of path optimality criteria using CDFs of three paths. Fan et al.'s criterion prefers paths 2 and 3 to path 1 but cannot discriminate between the CDFs of paths 2 and 3. Frank's criterion prefers path 2 to path 3 but cannot be applied to path 1. The joint criterion is applicable to all CDFs and chooses path 2 as the optimal one.

Y. Y. Fan, R. E. Kalaba, and J. E. Moore, *Journal of Optimization Theory and Applications 127, 497(2005).*
H. Frank, *Operations Research 17, pp. 583(1969).*

What is the shortest path?

In contrast with the deterministic case, there is no longer a unique concept of optimality

Long time budget

Frank defines a path to be optimal if its cdf surpasses a threshold within the shortest time

GOOD

Fan suggests maximizing the cdf for a given time budget

BAD

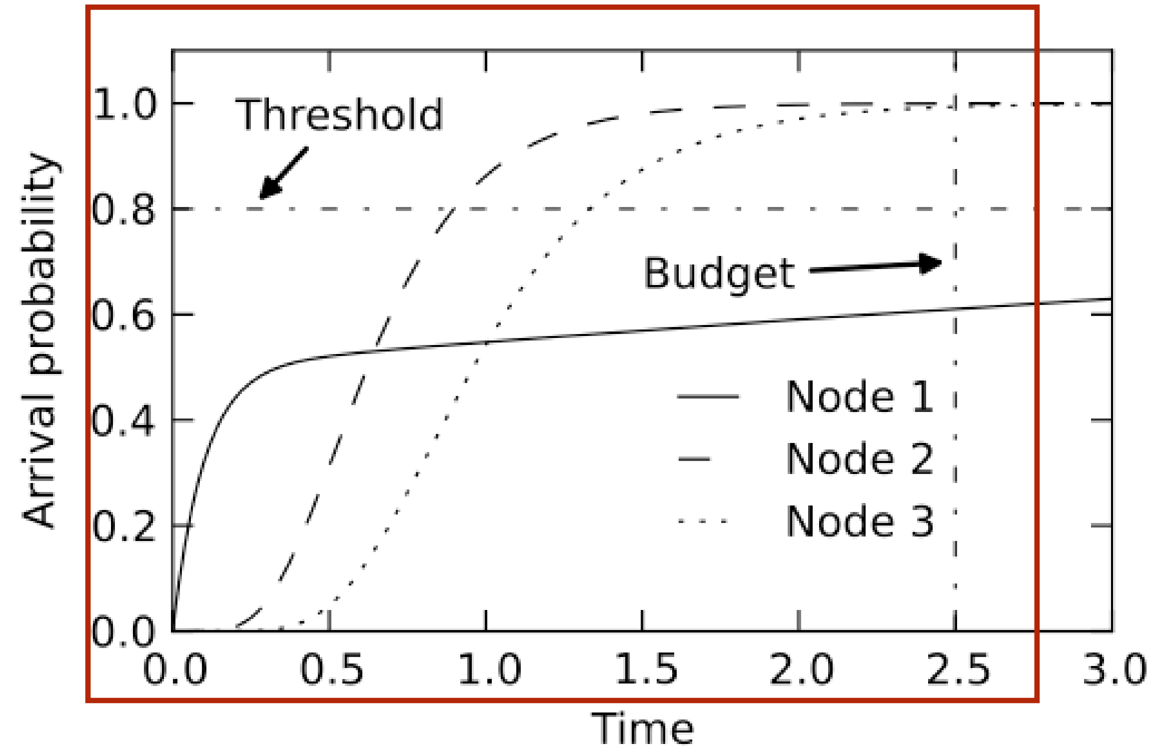


FIG. 1: Comparison of path optimality criteria using CDFs of three paths. Fan et al.'s criterion prefers paths 2 and 3 to path 1 but cannot discriminate between the CDFs of paths 2 and 3. Frank's criterion prefers path 2 to path 3 but cannot be applied to path 1. The joint criterion is applicable to all CDFs and chooses path 2 as the optimal one.

Y. Y. Fan, R. E. Kalaba, and J. E. Moore, *Journal of Optimization Theory and Applications 127, 497(2005).*
H. Frank, *Operations Research 17, pp. 583(1969).*

What is the shortest path?

In contrast with the deterministic case, there is no longer a unique concept of optimality

Frank defines a path to be optimal if its cdf surpasses a threshold within the shortest time

Fan suggests maximizing the cdf for a given time budget

Joint criterion: Frank criterion if it finds a path, otherwise we use Fan

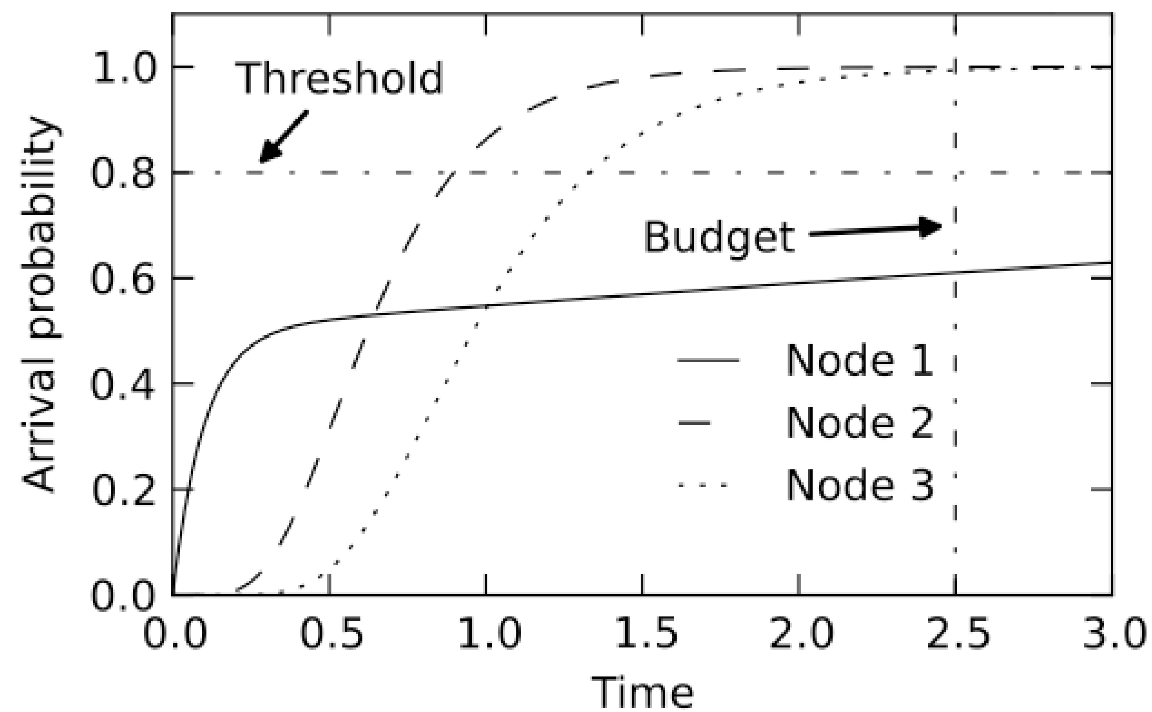


FIG. 1: Comparison of path optimality criteria using CDFs of three paths. Fan et al.'s criterion prefers paths 2 and 3 to path 1 but cannot discriminate between the CDFs of paths 2 and 3. Frank's criterion prefers path 2 to path 3 but cannot be applied to path 1. The joint criterion is applicable to all CDFs and chooses path 2 as the optimal one.

Y. Y. Fan, R. E. Kalaba, and J. E. Moore, *Journal of Optimization Theory and Applications 127, 497(2005).*
H. Frank, *Operations Research 17, pp. 583(1969).*

Finding the best path: A centralized algorithm

Routing table to reach a target r from all other nodes, solution of

$$u_i\left(t\right)=\max_{j\in J_i}\left[\int_0^t p_{ij}\left(t'\right)u_j\left(t-t'\right)\,dt'\right]\,, \qquad (2)$$

$$u_r\left(t\right)=1\,, \qquad (3)$$

$u_i(t)$ is the probability to arrive at node r starting from node i with a total time no longer than t

When on i, the node to chose to optimize the path is:

$$q_i\left(t\right)=\arg\max_{j\in J_i}\left[\int_0^t p_{ij}\left(t'\right)u_j\left(t-t'\right)\,dt'\right]$$

Y. Y. Fan, R. E. Kalaba, and J. E. Moore, *Journal of Optimization Theory and Applications 127, 497(2005).*

Finding the best path: A centralized algorithm

Routing table to reach a target r from all other nodes, solution of

$$u_i(t) = \max_{j \in J_i} \left[\int_0^t p_{ij}(t') \, u_j(t - t') \, dt' \right] , \tag{2}$$

$$u_r(t) = 1 , \tag{3}$$

$u_i(t)$ is the probability to arrive at node r starting from node i with a total time no longer than t

When on i, the node to chose to optimize the path is:

$$q_i(t) = \arg\max_{j \in J_i} \left[\int_0^t p_{ij}(t') \, u_j(t - t') \, dt' \right]$$

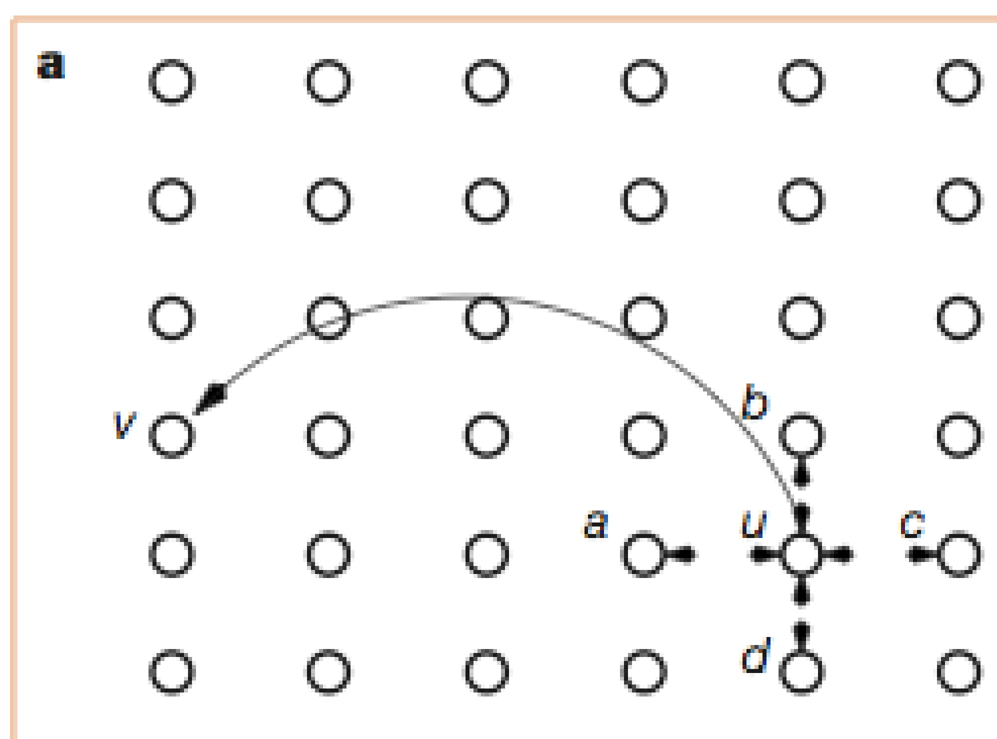
... can be solved recursively, but requires global knowledge of the system because of the recurrence

Y. Y. Fan, R. E. Kalaba, and J. E. Moore, *Journal of Optimization Theory and Applications 127, 497(2005).*

Finding the best path: A decentralized algorithm

We build an estimation function *f(i,j,t)* that estimates the arrival probability between i and j.

1. The network is embedded in a metric space. Let dij the physical distance between two nodes
(proximity will help us to guide travellers)



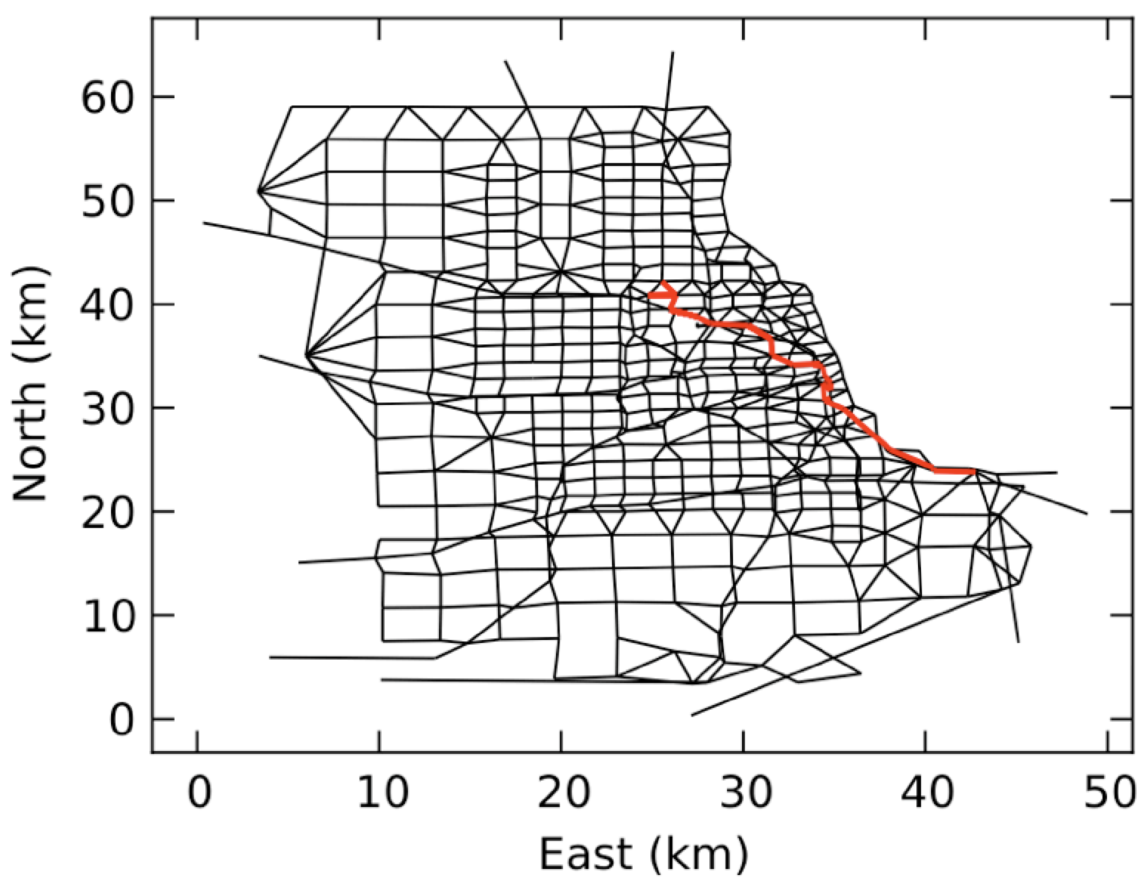
Navigation in small world networks, Kleinberg, Nature 2000

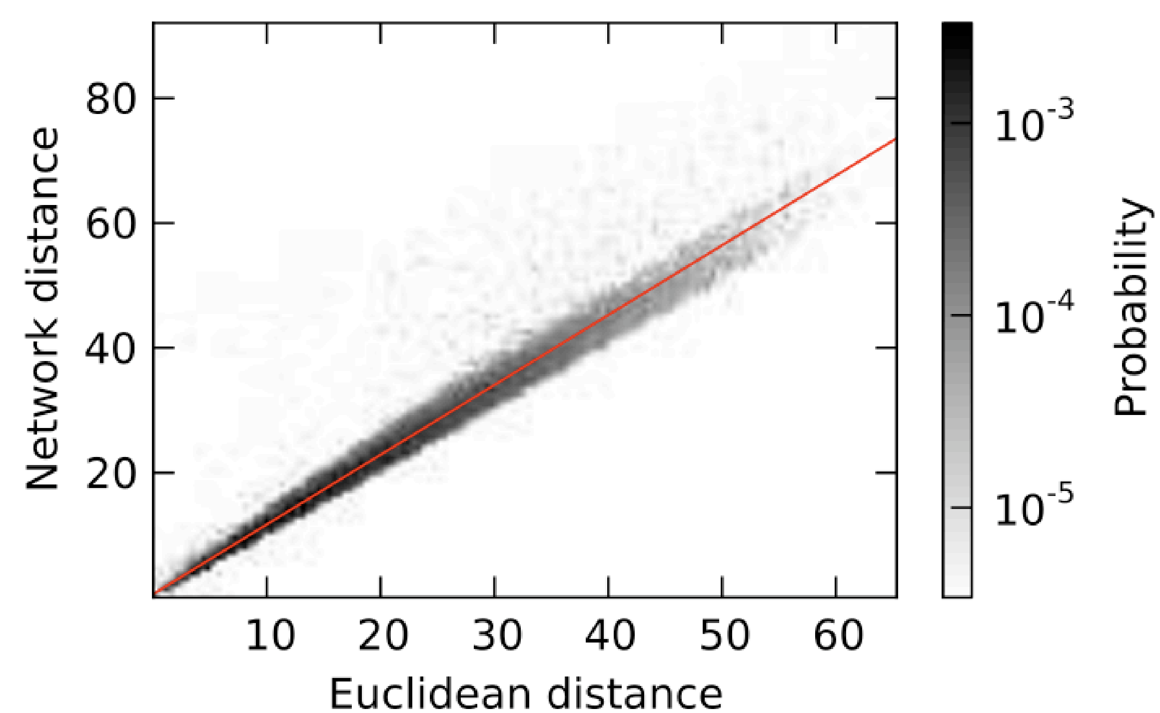
Finding the best path: A decentralized algorithm

We build an estimation function *f(i,j,t)* that estimates the arrival probability between i and j.

1. The network is embedded in a metric space. Let dij the physical distance between two nodes
(proximity will help us to guide travellers)

2. Let g_{ij} be the physical distance of the shortest path between two nodes, and assume that $$g_{ij} \approx h\left(d_{ij}\right)$$





$$h\left(d_{ij}\right) \approx 0.547(8)\ \mathrm{km} + 1.1176(4) \times d_{ij}$$

Finding the best path: A decentralized algorithm

We build an estimation function *f(i,j,t)* that estimates the arrival probability between i and j.

1. The network is embedded in a metric space. Let dij the physical distance between two nodes
(proximity will help us to guide travellers)

2. Let g_{ij} be the physical distance of the shortest path between two nodes, and assume that $g_{ij} \approx h\left(d_{ij}\right)$

3. One estimates the number of steps between i and j by

$$\bar{k} = \left\lceil \frac{g_{ij}}{\lambda} \right\rceil \approx \left\lceil \frac{h\left(d_{ij}\right)}{\lambda} \right\rceil$$

The weight on each step is chosen uniformly at random from the known weights associated with edges

$$\bar{p}\left(t\right) = \frac{1}{m} \sum_{(i,j)\in E} p_{ij}\left(t\right)$$

Finding the best path: A decentralized algorithm

Under these assumptions, we build the estimation function

$$f(i,j;t) = \begin{cases} \int_0^t dt' \left(\overset{\bar{k}}{\underset{k=1}{\ast}} \bar{p} \right)(t'), & \text{if } i \neq j, \\ 1, & \text{if } i = j \end{cases}$$

estimating the probability for a path from i to j to have a weight smaller than t

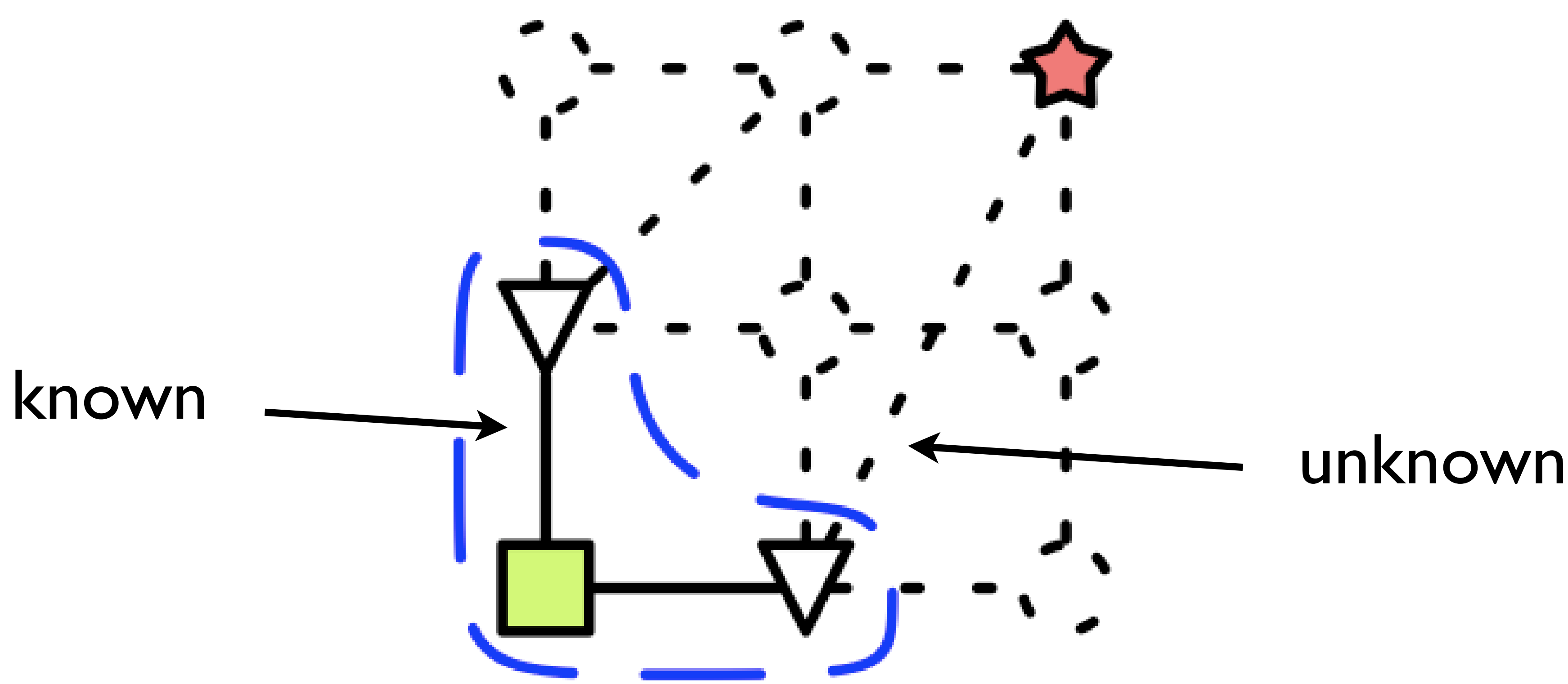
Advantage: the estimation function is calculated by using properties of the spatial embedding alone, and does not involve the knowledge of the graph as a whole.

Finding the best path: A decentralized algorithm

Under these assumptions, we build the estimation function

$$f\left(i,j;t\right)=\begin{cases}\displaystyle\int_0^t dt'\,\left(\mathop{\ast}_{k=1}^{\bar{k}}\bar{p}\right)(t')\,, & \text{if } i\neq j,\\ 1\,, & \text{if } i=j\end{cases}$$

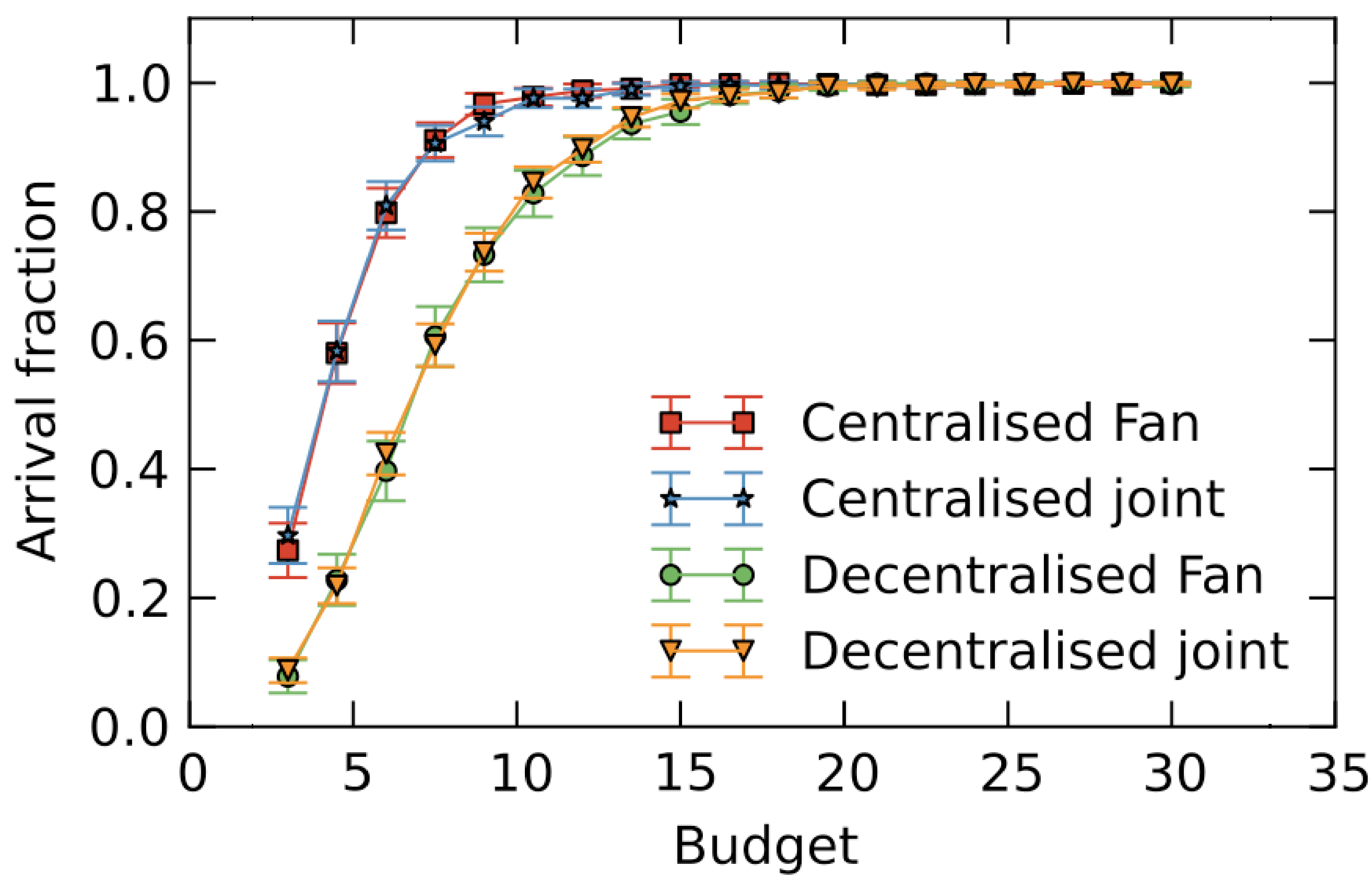
and use it to build a local algorithm, where we estimate the exact cdf for nodes where we have no information



Numerical tests

Simulations on two-dimensional lattices with short-cuts (à la Kleinberg)

To determine edge weights, we assign lognormal PDFs $p_{\ln}(\mu, \sigma; t)$ with mean μ and standard deviation σ chosen uniformly at random in the interval $[0.5, 1.5]$.

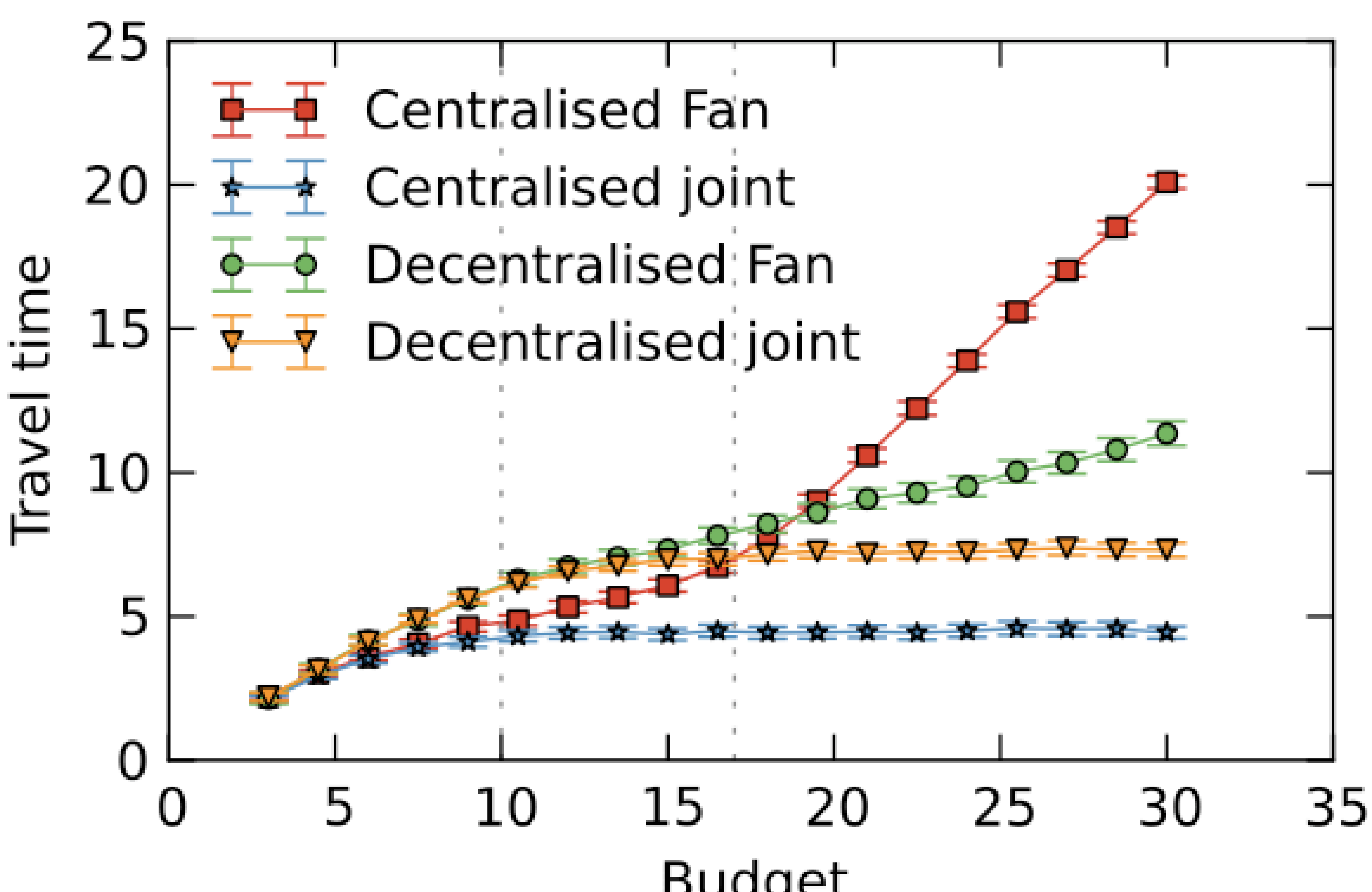


Numerical tests

Simulations on two-dimensional lattices with short-cuts (à la Kleinberg)

To determine edge weights, we assign lognormal PDFs $p_{\ln}(\mu, \sigma; t)$ with mean μ and standard deviation σ chosen uniformly at random in the interval $[0.5, 1.5]$.

Mean travel time of successful routing



When budget increases, the joint criterion outperforms Fan criterion, as expected

Numerical tests

Simulations on two-dimensional lattices with short-cuts (à la Kleinberg)

To determine edge weights, we assign lognormal PDFs $p_{\ln}(\mu, \sigma; t)$ with mean μ and standard deviation σ chosen uniformly at random in the interval $[0.5, 1.5]$.

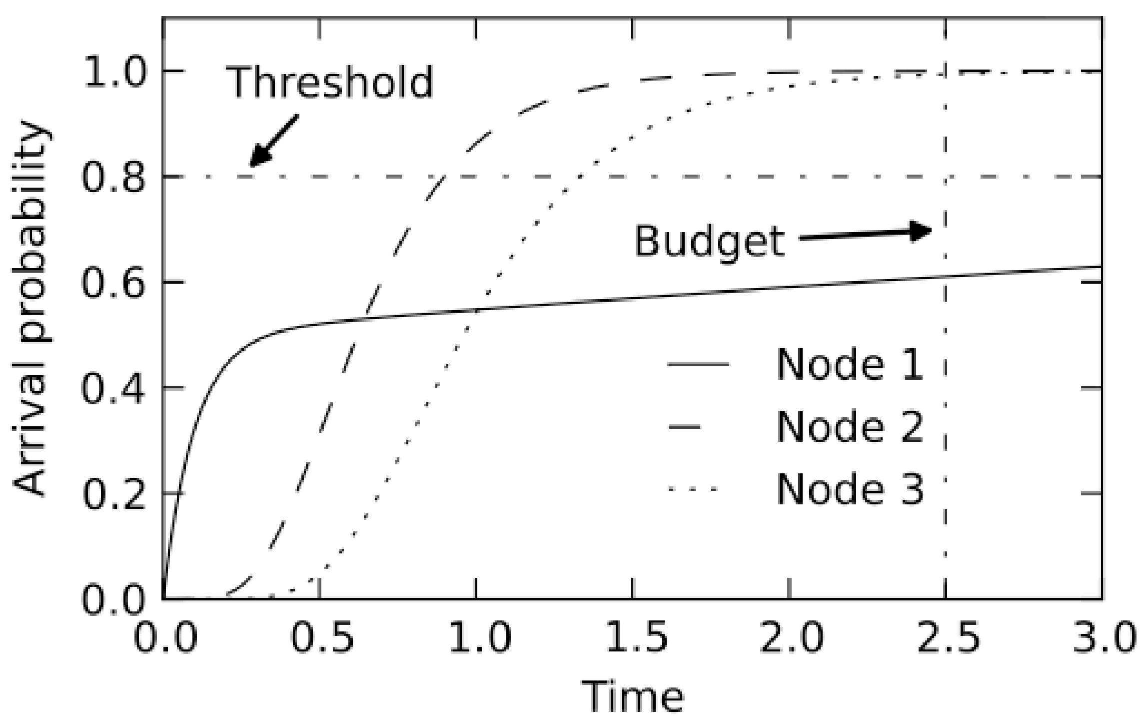


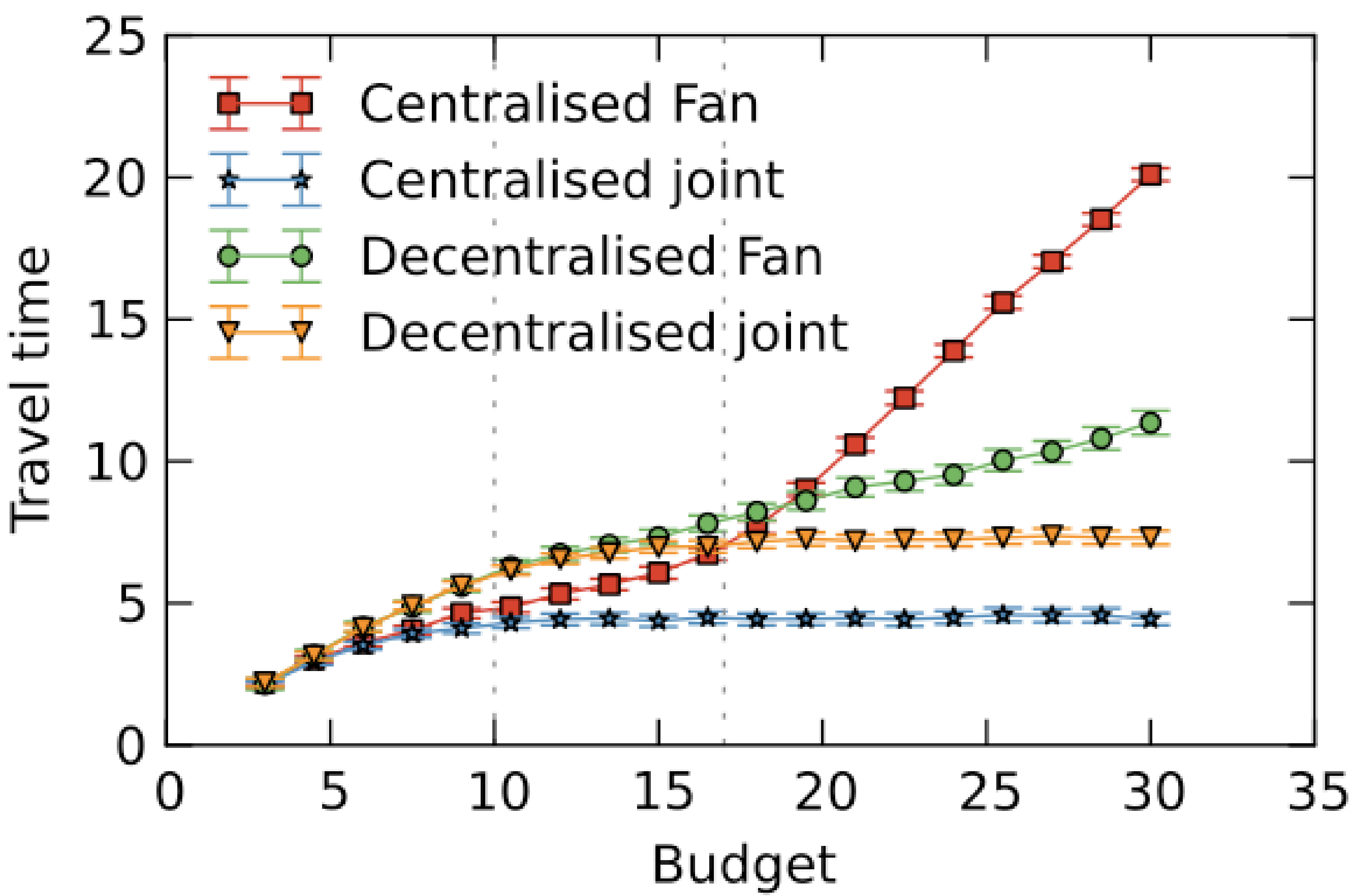
FIG. 1: Comparison of path optimality criteria using CDFs of three paths. Fan et al.'s criterion prefers paths 2 and 3 to path 1 but cannot discriminate between the CDFs of paths 2 and 3. Frank's criterion prefers path 2 to path 3 but cannot be applied to path 1. The joint criterion is applicable to all CDFs and chooses path 2 as the optimal one.

When budget increases, the joint criterion outperforms Fan criterion, as expected

Numerical tests

Simulations on two-dimensional lattices with short-cuts (à la Kleinberg)

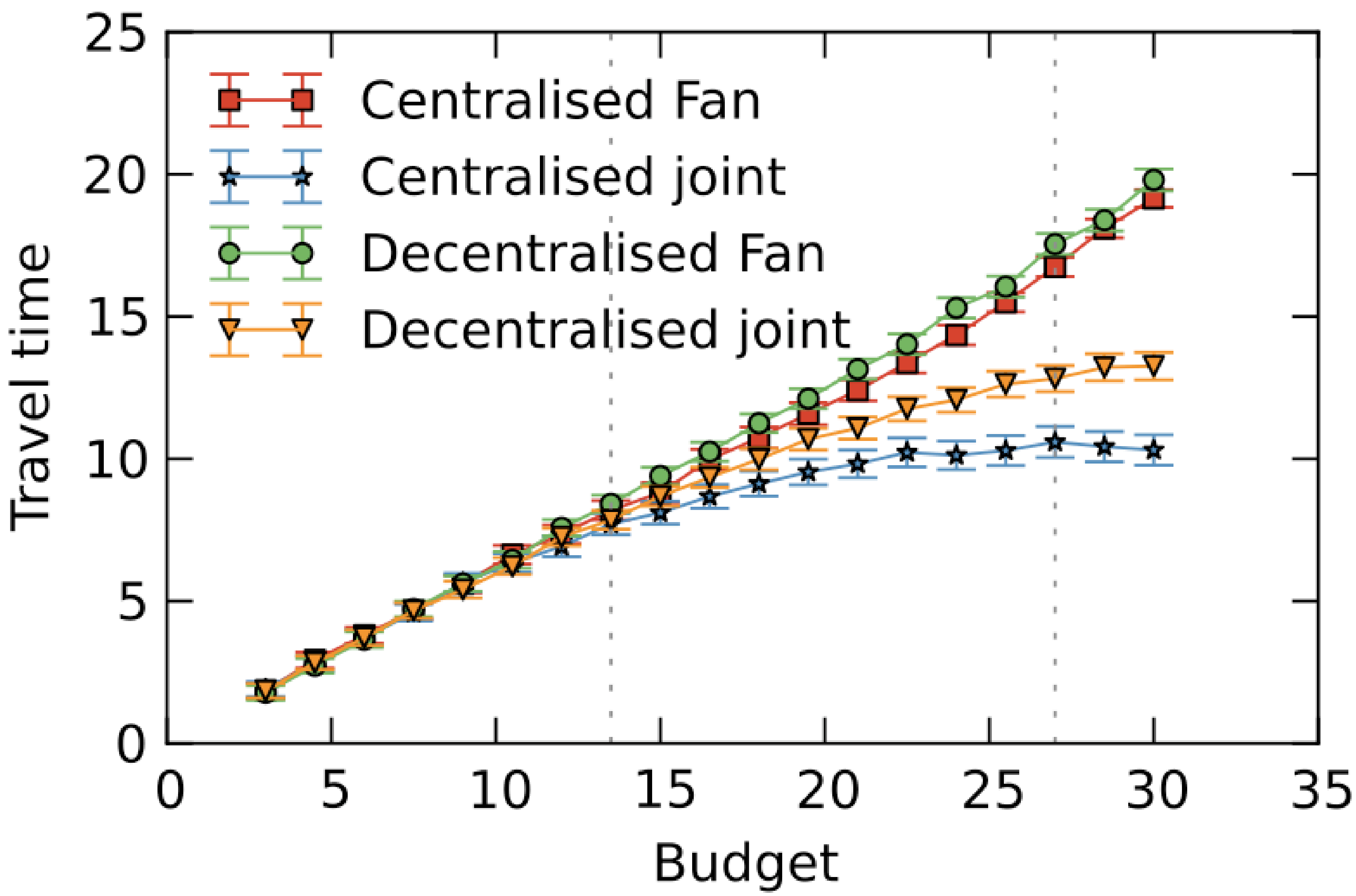
To determine edge weights, we assign lognormal PDFs $p_{\ln}(\mu, \sigma; t)$ with mean μ and standard deviation σ chosen uniformly at random in the interval $[0.5, 1.5]$.



When budget increases, the joint criterion outperforms Fan criterion, as expected

Numerical tests

Chicago road network (with 542 junctions)

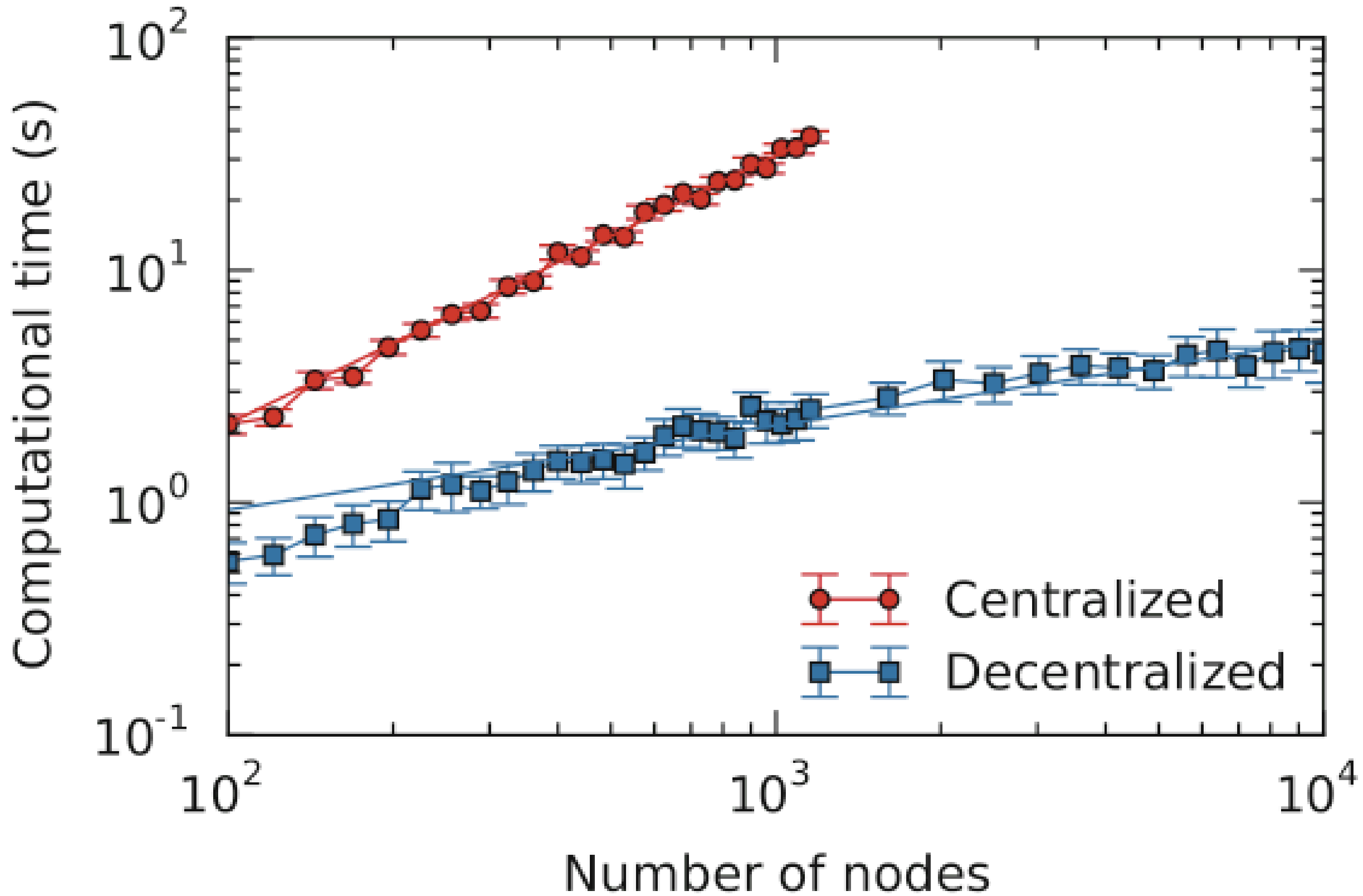


When budget increases, the joint criterion outperforms Fan criterion, as expected

Computational time

Centralized algorithm: finding the shortest path between two nodes requires operations on the full network => $t \sim N^a$, with a close to 1.

Decentralized algorithm: only local searches => sublinear scaling $t \sim N^b$, with $b<1$



a = 1.16
b = 0.36

Conclusion

Need for algorithms for temporal networks

Temporal network is modelled as a stochastic process, where each edge is assigned a probability distribution

Routing: centralized (more efficient) versus decentralized (faster) algorithms

Efficient algorithms for community detection, block modelling in such stochastic systems?
Possibility to take advantage of the observed temporal patterns to guide the development of algorithms?

Decentralized Routing on Spatial Networks with Stochastic Edge Weights, Till Hoffmann, R.L. and Mason Porter, PRE 2013

Thanks to:

Till Hoffmann and Mason Porter (Oxford): Temporal networks

Martin Rosvall (Umea) and Andrea Lancichinetti (Northwestern): Pathways and Memory

J.C. Delvenne (Louvain), Luis Rocha (Louvain->Namur/Karolinska) and Lionel Tabourier (Namur): Burstiness

Financial support of IAP DYSCO from Belspo, Région Wallonne and FNRS