Indirect estimation of shortest path distributions with small-world experiments

Antti Ukkonen

Helsinki Institute for Information Technology HIIT
Aalto University, Finland
antti.ukkonen@aalto.fi

Abstract. The distribution of shortest path lenghts is a useful characterisation of the connectivity in a network. The small-world experiment is a classical way to study the shortest path distribution in real-world social networks that cannot be directly observed. However, the data observed in these experiments are distorted by two factors: attrition and routing (in)efficiency. This leads to inaccuracies in the estimates of shortest path lenghts. In this paper we propose a model to analyse small-world experiments that corrects for both of the aforementioned sources of bias. Under suitable circumstances the model gives accurate estimates of the true underlying shortest path distribution without directly observing the network. It can also quantify the routing efficiency of the underlying population. We study the model by using simulations, and apply it to real data from previous small-world experiments.

1 Introduction

Consider the real-life social network where two people are connected if they mutually know each other "on a first name basis".

What is the distribution of shortest path lengths in this network?

In the late 1960s, decades before the social networking services of today, Jeffrey Travers and Stanley Milgram conducted an experiment to study this question [15]. Participants recruited from Omaha, Nebraska, were asked to forward messages to a target person living in Boston, Massachusetts, via chains of social acquaintances. Given only basic demographic information of the target, such as name, hometown and occupation, the starting persons were instructed to pass the message to a friend (their neighbor in the underlying social network) whom they considered as likely to forward the message further so that it eventually reaches the target. The participants were also instructed to notify the experimenters whenever they forwarded the message.

Out of the 296 messages that were initially sent, 64 reached the target person, and took only a relatively small number of hops [15]. The main hypothesis resulting from this experiment was that we live in a "small-world", meaning

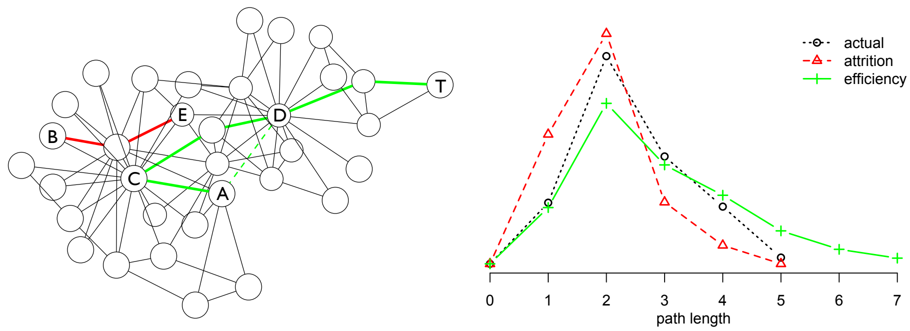

Fig. 1. *Left:* The Karate-club graph with two paths (green and red, solid lines) that are examples of observations in a small-world experiment on this network. *Right:* The shortest paht distribution of the Karate-club graph (black, dotted), together with the effects of attrition (red, dashed) and routing efficiency (green, solid) to the length distribution of observed completed paths.

that a lot of people are connected to each other by *short* paths in the underlying social network. Because of this we refer to such experiments as *small-world experiments*. Similar experiments have been repeated a number of times [12,5].

The observations in such a small-world experiment are a *set of paths* that each indicate the trajectory of one message through the social network. Some of these paths are *completed*, that is, the message reached the target individual, while others are discarded on the way. The length distribution of the completed paths can be seen as a characterization of the *social connectivity* of the population. Indeed, the popular notion of "six-degrees-of-separation" is based on the median of the length distribution of the completed paths that were observed by Travers and Milgram [15].

However, the lengths of completed paths are in practice distorted by two factors: *attrition* and *routing efficiency* of the participants. Attrition refers to messages being discarded for one reason or another during the experiment, causing some paths to terminate before reaching the target. By routing efficiency we refer to the ability of the participants to pass the message to an acquaintance who is in fact closer to the target when distance is measured in terms of the true shortest path distance.

Figure 1 illustrates examples of attrition and routing efficiency with the well known Karate-club network [17]. There are two starting persons, nodes A and B, with T as the target person, as well as two paths that originate at A and B. The path that started from B was terminated early due to attrition: node E failed to forward the message further. In real small-world experiments a message is discarded at every step with an average probabiliy that varies between 0.25 and 0.7, depending on the study [15,5]. On the other hand, the path that starts from node A does reach the target T in five steps, but it does this via node C. Observe that the true shortest path distance from C to T is four steps, while the node A is in fact only three steps away from T (via node D, dashed edge). Node A made thus a "mistake" in forwarding the message to C instead of D. This will also happen in practice. Indeed, even when all participants belong to the same organization, the messages are forwarded to an acquaintance who is truly closer to the target only in a fraction of the cases [10].

Attrition and routing efficiency *will distort the lengths of paths that we observe in a small-world experiment.* In particular, attrition makes long paths less likely to appear[1], even if they might exist in the underlying social network. As a consequence the length distribution of observed completed paths is *shifted towards short paths*, giving an "optimistic" view of the social connectivity. Routing (in)efficiency, on the other hand, implies that even in the absence of attrition we would not observe the true shortest paths. Finding these requires access to the entire network topology; information that individuals most likely do not have. Instead, we observe so called *algorithmic shortest paths* (see also [11,13]), i.e., those that the participants are able to discover using only information about their immediate neighbors and the target person. These, however, are longer than the shortest paths. The lengths of observed completed paths are thus *shifted towards longer paths*, making the network seem less connected.

Indeed, attrition and routing efficiency have contradictory effects on the observed path lengths, as illustrated on the right side of Fig. 1. Attrition makes the observed paths shorter, while routing (in)efficiency makes them longer when compared to the actual shortest path distribution. In practice we observe paths that have been affected by both. Motivated by this discrepancy between the observations and the shortest path distribution, we address the following question:

> *Can we recover the true shortest path distribution given the observed paths from a small-world experiment?*

This is an interesting task for a number of reasons. First, it is another way to address the original question of Travers and Milgram [15]. Using only the completed paths gives one view of the social connectivity, by estimating the actual shortest paths distribution results in a less biased outcome. Second, the question is related to recent work on reconstructing networks and properties thereof from observed traces of activity in the network [8,9,14,4]. The small-world experiment can be seen as another type of such activity. Moreover, a better understanding of the process that underlies the small-world experiment may lead to improvements in other propagation processes over networks, such as spreading of epidemics or opinions.

Our contributions: We propose a new model to analyse the observations of a small-world experiment that accounts for the bias caused by attrition and routing efficiency. The main difference to previous approaches [10,7] is the use a well-defined probabilistic model that can estimate the true shortest path distribution. Our technical contribution is that of devising an intuitive parametrization of the process that underlies a small-world experiment, as well as a means to

[1] If a path terminates with probability r at every step, a path of length l appears with probability $(1-r)^l$. This decreases rapidly as l increases. See also [7].

express the likelihood of the observed paths in terms of the process parameters. By fitting the proposed model to data from previous small-world experiments [15,5] we compute *estimates for the shortest path distribution* of the underlying social network, as well as *quantify the routing efficiency* of the population that participated in the experiment. To the best of our knowledge this has not been done before.

2 A model for small-world experiments

The input to our model consists of the set D of both *completed* and *failed* paths that are observed in a small-world experiment. A path is completed if it reaches the target, and failed if it terminates due to attrition. Of every path in D, we know thus its *outcome*, and its *length*. These are the only two characteristics of a path that our model is based upon.

We first discuss the parameters of the *message-forwarding process* that underlies a small-world experiment. As with any generative model, we do not claim that this process accurately represents reality, it is merely a useful and tractable representation of it. Then, we show how the process parameters *induce the parameters of a multinomial distribution* over different types of paths. The likelihood of D is determined by this multinomial distribution in a standard manner.

Model parameters and dynamics
We assume the structure of the underlying social network to be hidden. However, at every step each of the messages must be at some *shortest path distance* from the target person. These distances are unknown, but we can model how they evolve as the messages are forwarded. When a node forwards the message, the shortest path distance to the target can *decrease by one, stay the same, or increase by one.* Note that the distance can *not* increase or decrease by any other amount in a single step.

We assume that each node chooses the recipient so that the distance decreases with probability q_-, increases with probability q_+, and remains the same with probability $1 - q_- - q_+$. We also enforce the constraint $q_+ < q_-$ (the participants are assumed to be in some sense "benevolent"), as well as $q_- + q_+ \leq 1$. The probabilities q_- and q_+ are the first and second parameters of our model, and they capture routing (in)efficiency. Moreover, a node might not forward the message in the first place. We assume that a message can be discarded at every step with the constant probability r, the attrition rate, which is the third parameter of our model.

The parameters q_-, q_+, and r capture how the shortest path distance from the current holder of the message to the target person evolves, but the process must start from somewhere. The distances from the starting nodes to the target node are unknown. We assume that *the initial distances from the starting nodes to the target* follow some predefined distribution $\tau(\cdot)$. In particular, we take these to be Weibull distributed with parameters k and λ, because it has been argued [3] that this produces a good fit for shortest path distributions that are observed

Table 1. Summary of model parameters

q_-, q_+	probabilities for message to approach and move away from target
r	probability of message to be discarded
k, λ	shape and scale parameters of the initial distance distribution τ

in different types of random networks. The k and λ parameters of τ are the fourth and fifth parameters of our model. All five parameters are summarised in Table 1.

In summary, we model the small-world experiment using the following two-stage process. Initially the messages are distributed at distances from the target that is given by $\tau(\cdot \mid k, \lambda)$. Upon receiving the message a node 1) decides if it is going to discard the message, and if not, 2) the node forwards the message to a neighbor. An independent instance of this process launches from every starting node, and continues until the message reaches the target or is discrarded.

Path probabilities

Recall that of every path in the input D we know its outcome o and length l. A path is of *type* (o, l) if it has outcome o and length l, where o is either $\top$ (a completed path) or $\bot$ (a failed path). For example, the red and green paths in Fig. 1 are of types $(\bot, 2)$ and $(\top, 5)$, respectively.

Next we derive the probability to observe a path of type (o, l), assuming that the forwarding process adheres to the parameters discussed above. To this end we express the message-forwarding process as a discrete-time Markov chain. A graphical representation of this chain together with the transition probabilities is shown in Fig. 2. The states of the chain are all possible distances to the target up to some maximal distance m, denoted $\mathsf{0}$, $\mathsf{1}$, ..., m, a special attrition state, denoted A, and a special terminal state, denoted E. At every step the chain is in one of the states. The message reaches the target when the chain enters the state $\mathsf{0}$. Likewise, the message is discarded when the chain enters the state A. The state E is needed for technical reasons.

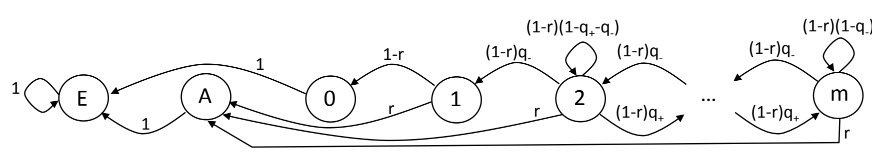

Fig. 2. The Markov chain that represents the message-forwarding process. The message is discarded with probability r independently of its distance to the target. The distance decreases with probability $(1-r)q_-$, increases with $(1-r)q_+$, and remains the same with probability $(1-r)(1-q_- - q_+)$. A neighbor of the target node can discard the message or passe it to the target. In the state m the distance can no longer increase, so we introduce a self-loop with probability $(1-r)(1-q_-)$.

Let $\mathbf{Q}$ denote the transition probability matrix associated with the Markov chain of Fig. 2 so that $\mathbf{Q}_{\mathsf{xy}}$ is the probability of taking a transition from state x to state y. From the basic theory of Markov chains we know that given any initial probability distribution π_0 over the states, the distribution after running the chain for l steps is equal to $\pi_l = \pi_0^\top(\mathbf{Q})^l$. Recall that $\tau(\cdot \mid k, \lambda)$ is our guess of the distribution of distances from the starters to the target. We initialize π_0 so that $\pi_0(\mathsf{A}) = \pi_0(\mathsf{E}) = \pi_0(\mathsf{0}) = 0$, and for the states $\mathsf{1}, \ldots, \mathsf{m}$ we set $\pi_0(\mathsf{x}) = \tau(\mathsf{x} \mid k, \lambda)$. The probabilities to observe a path of type (o, l) given the model parameters is now

$$\Pr((o,l) \mid q_-, q_+, r, k, \lambda) = \begin{cases} \pi_l(\mathsf{0}), \text{ if } o = \top, (\text{completed path}), \\ \\ \pi_l(\mathsf{A}), \text{ if } o = \bot, (\text{failed path}), \end{cases} \tag{1}$$

where $\pi_l(\mathsf{0})$ and $\pi_l(\mathsf{A})$ are the probability masses at states $\mathsf{0}$ and A after running the chain for l steps, respectively. The above equation is derived simply by noticing that the states $\mathsf{0}$ and A correspond to endpoints of completed and failed chains, and the probability of a path of length l to land in either of these is obtained by running the chain for l steps. (The special terminal state E is needed to guarantee that π_l indeed has the desired value at states $\mathsf{0}$ and A.)

Finally, to compute the probabilities in practice, the maximum distance m must be set to a large enough value. In practice we did not observe m to have a strong effect. In the experiments we use $m = 30$.

Likelihood of the input D

As in [7], we view a small-world experiment as a simple sampling procedure, where $|D|$ paths are *drawn independently*[2] *from a categorical distribution*, where every category corresponds to a path type, and the probabilities of individual categories are induced by the model parameters (Table 1) as discussed above. Let $\theta = (q_-, q_+, r, k, \lambda)$ denote a vector with all parameters, and define the probability of a type $p_\theta(o, l) = \Pr((o, l) \mid \theta)$ as in Eq. 1.

In theory the number of distinct types is infinite, because our model does not impose an upper limit on the length of a path. However, we only consider types up to some length, because the probability of long paths decreases rapidly. Given θ, we choose $l_{\max}$ to be the smallest integer so that the inequality $\sum_{l=1}^{l_{\max}} (p_\theta(\top, l) + p_\theta(\bot, l)) \geq 1 - \epsilon$ holds for some small ϵ (in the experiments we let $\epsilon = 10^{-8}$). That is, we assume that the paths in D are an i.i.d. sample from a discrete distribution with categories $C(\theta) = \{(\top, 1), (\bot, 1), \ldots, (\top, l_{\max}), (\bot, l_{\max})\}$. Note that depending on θ, $l_{\max}$ can be larger than the longest path we observe in D.

Furthermore, Let $c_D(o, l)$ denote the *number of paths* of type (o, l) in D. Of course $c(o, l) = 0$ for any type (o, l) that does not occur in D. The numbers $c_D(o, l)$ are a sample from a multinomial distribution with parameters $|D|$ and

[2] This may not fully hold in real small-world experiments, as lengths and outcome of paths having e.g. the same source or target may be correlated [5]. However, we consider this to be a reasonable simplification to make the model tractable.

$p_\theta(o,l)$ for every $(o,l) \in C$. Therefore, the likelihood of D can be expressed as

$$\Pr(D \mid \theta) = \frac{|D|!}{\prod_{(o,l)\in C(\theta)} c_D(o,l)!} \prod_{(o,l)\in C(\theta)} p_\theta(o,l)^{c_D(o,l)}. \qquad (2)$$

The likelihood depends on the model parameters θ thus through the probabilities $p_\theta(o,l)$. To compute the likelihood, we first find $p_\theta(o,l)$ for every $(o,l) \in C(\theta)$ using the Markov chain of Fig. 2, and then apply Equation 2.

Parameter estimation
For parameter estimation we can use any available optimisation technique. After trying out several alternatives, including different numerical optimisation algorithms as well as MCMC techniques, the Nelder-Mead method was chosen[3]. It does not require partial derivatives, finds a local optimum of the likelihood function, and is reasonably efficient for our purposes. It also allows to use fixed values for some of the paramters, and solve the model only for a subset of them to obtain *conditional estimates*. For instance, we can fix q_- and q_+ and solve the model only for r, k and λ.

It is worth pointing out that the model can be non-identifiable. (Meaning its true parameters are hard to find even given an infinite amount of data.) Some of the parameters have opposite effects. For instance, in terms of the observed paths in D, it might not matter much if a) the initial distances to the target are long and the routing efficiency is high, or b) the initial distances are short but the routing efficiency is low. This means that a simultaneous increase in both e.g. the median of $\tau(k,\lambda)$ as well as the routing efficiency parameter q_- can result in only a very small change in $\Pr(D \mid \theta)$.

3 Experiments

Estimation accuracy
As there in general is no ground truth available in a small-world experiment, it is important that the estimates obtained from the model are at least somewhat stable. The estimates are affected by size of the input data, as well as indentifiability issues of the model itself. We start by quantifying these effects by using paths that are obtained from simulations of the message forwarding process with known parameters.

As expected, estimating all model parameters simultaneously is a hard problem. The top half of Table 2 shows both the true parameter values as well as the median of their estimates from 250 inputs of 5000 paths each. (The estimate of r is always very accurate, and is omitted from the table.) We can see that while the median is often fairly close to the true value, the quantiles indicate a high variance in the estimates, especially for q_-.

[3] We use the implementation provided by the `optim` function of GNU R, `http://www.r-project.org`.

Table 2. Estimating parameters from simulated data (5000 paths)

true values					simultaneous estimates (with 5% and 95% quantiles)			
q_-	q_+	r	k	λ	q_-	q_+	k	λ
0.5	0.0	0.25	4	5	0.60 (0.44,0.82)	0.03 (0.00,0.10)	3.71 (3.11,4.46)	5.53 (4.56,6.76)
0.5	0.1	0.25	4	5	0.47 (0.29,0.69)	0.07 (0.00,0.21)	4.10 (3.53,5.08)	4.86 (3.83,6.01)
0.3	0.0	0.25	4	5	0.43 (0.30,0.90)	0.05 (0.00,0.13)	3.45 (2.34,4.10)	6.13 (4.87,11.93)
0.3	0.1	0.25	4	5	0.33 (0.15,0.70)	0.12 (0.00,0.31)	3.86 (2.79,5.50)	5.27 (3.60,8.51)
0.5	0.0	0.50	4	5	0.65 (0.35,1.00)	0.04 (0.00,0.17)	3.65 (2.79,5.05)	5.80 (3.92,8.38)
0.3	0.0	0.50	4	5	0.39 (0.18,0.88)	0.05 (0.00,0.17)	3.67 (2.44,5.41)	5.64 (3.71,11.95)
true values					conditional estimates (with 5% and 95% quantiles)			
q_-	q_+	r	k	λ	$q_- \mid k, \lambda$	$q_+ \mid k, \lambda$	$k \mid q_- q_+$	$\lambda \mid q_-, q_+$
0.5	0.0	0.25	4	5	0.50 (0.49,0.53)	0.00 (0.00,0.03)	4.01 (3.83,4.21)	4.99 (4.88,5.10)
0.5	0.1	0.25	4	5	0.50 (0.46,0.54)	0.10 (0.05,0.15)	4.01 (3.79,4.24)	4.98 (4.87,5.14)
0.3	0.0	0.25	4	5	0.31 (0.29,0.33)	0.00 (0.00,0.05)	4.02 (3.75,4.24)	5.00 (4.84,5.18)
0.3	0.1	0.25	4	5	0.30 (0.27,0.33)	0.10 (0.04,0.16)	4.00 (3.75,4.30)	4.99 (4.80,5.22)
0.5	0.0	0.50	4	5	0.51 (0.46,0.59)	0.00 (0.00,0.18)	4.00 (3.64,4.42)	5.00 (4.71,5.31)
0.3	0.0	0.50	4	5	0.31 (0.27,0.37)	0.00 (0.00,0.22)	3.99 (3.55,4.58)	5.01 (4.61,5.53)

However, by fixing some parameters it is possible to improve the quality of the resulting estimates. The bottom half of Table 2 shows conditional estimates for q_- and q_+ *given* k and λ, and vice versa, the estimates for k and λ given q_- and q_+. Now the estimates are very accurate in all of the cases. This suggests that the model is most useful when we have some prior knowledge either about the routing efficiency, or the shortest path distribution. For example, [10] suggests to use $q_- = 0.3$, and [1] provides the exact shortest path distribution of the entire Facebook social network, which could be used in place of $\tau(\cdot \mid k, \lambda)$.

The next question is how sensitive are the estimates of k and λ to errors in our assumptions of q_- and q_+? That is, suppose the true q_-, denoted q_-^{True}, is 0.5, but we solve the model with fixed a $q_- = 0.35$ for example. Let $\hat{\mu}$ denote the median of the estimated shortest path distribution, while μ^{True} is the median of the true distribution. (This is a more intuitive quantity than k and λ when interpreting the shortest path distribution.) The left panel in Figure 3 shows $\Delta(\mu) = \hat{\mu} - \mu^{\text{True}}$ as a function of $\Delta(q_-) = q_- - q_-^{\text{True}}$ when k and λ are estimated given different q_- from 1000 paths generated by the model. (The plot shows distributions over 100 runs for every q_-. Input was generated with $q_- = 0.5$, $q_+ = 0.05$, $r = 0.25$, $k = 4$, $\lambda = 5$.) We observe that in this range of q_- an under- or overestimate of 0.15 will make the conditional estimate of $\hat{\mu}$ about one step too low/high.

The variance of the estimates will also depend on the size of the input D. The right panel in Figure 3 shows effect of $|D|$ on the variance of $\hat{\mu}$ when k and λ are estimated with q_- and q_+ fixed to their correct values. (The paths were generated using the same parameter values as above.) We find that the conditional estimate of $\hat{\mu}$ is both *unbiased* and *consistent*. In practice a high accuracy requires > 1000 paths, but even only 300 paths seem to give reasonable results.

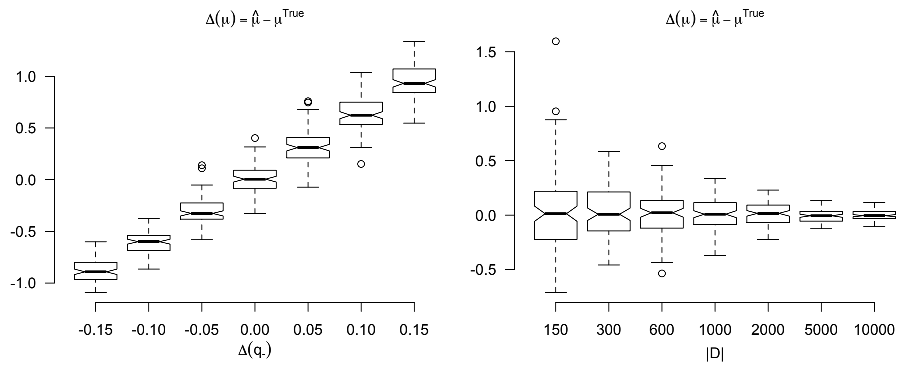

Fig. 3. *Left:* Effect of the error when fixing q_- to the conditional estimate of $\hat{\mu}$. *Right:* Effect of $|D|$ to the conditional estimate of $\hat{\mu}$. (Boxes show 1st and 3rd quartiles.)

Estimating shortest path distributions

We continue by estimating the shortest path distribution[4] in random networks. The top row of Fig. 4 shows conditional estimates for an Erdős-Rényi [6] as well as a Barabási-Albert [2] graph. In both cases we considered three values for q_-: the correct one, and ones that under- and overestimate the true value by 0.1. We find that in both cases the $\Delta(q_-) = 0$ estimate (solid line) captures the qualitative properties of the true distribution (bars). And like above, an under- or overestimate in q_- leads to a slight under- or overestimate in the path lenghts.

Finally, we apply the model to observed data from previous small-world experiments. Travers & Milgram (TM) provide the numbers of completed and failed paths that we need for our model (in Table 1 of [15]), while for Dodds et al. (DMW) we obtain the numbers through visual inspection of Fig. 1 in [5]. We fit both the full model, as well as conditional estimates given fixed q_- and q_+. Resulting shortest path distributions are shown in the bottom row of Fig. 4.

While the full estimation is known to be unstable, it is interesting to see that it produces reasonable estimates. For TM the estimates for q_-, q_+, and r are 0.92, 0.08 and 0.21, respectively, while for DMW we obtain $q_- = 0.52$, $q_+ = 0.00$, $r = 0.71$. The attrition rate estimates are very close to the ones reported in literature [15,5], while the routing efficiency parameters tell an interesting story. It appears that in the TM experiment the participants almost always chose the "correct" recipient, while in the DMW experiment they did this only half of the time. This is not inconceivable, as the TM study was carried out by regular mail, while DMW used email as the means of communication. Subjects in the TM study might indeed have been more careful when choosing the recipient as participating took more effort also otherwise.

[4] Notice that here the resulting shortest path distributions reflect the initial distances to a single target from the starters. To obtain the all-pairs shortest path distribution, all paths should have independently sampled starters and targets.

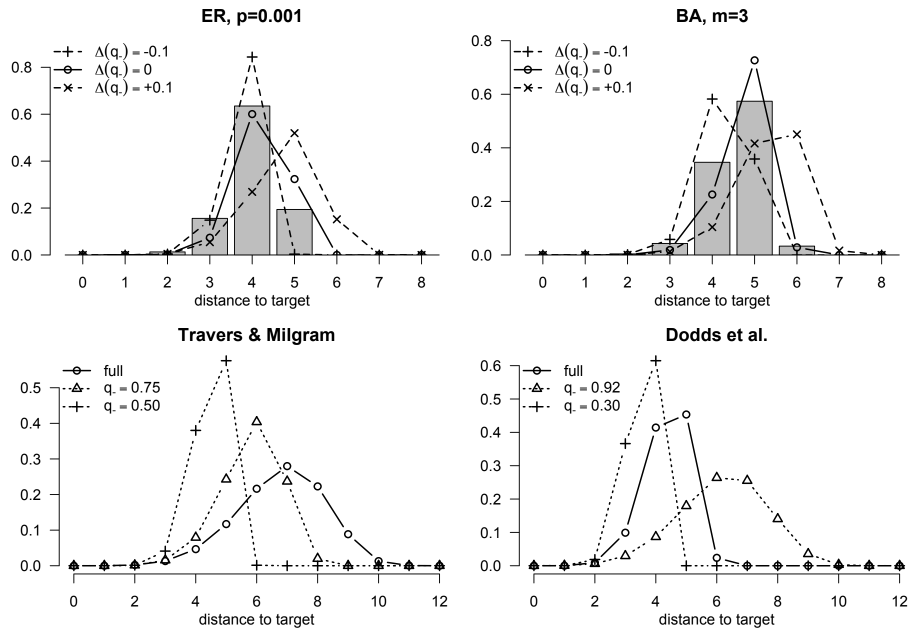

Fig. 4. *Top row:* True shortest path distribution (bars) together with three estimates for different values of $\Delta(q_-)$ computed from 1000 paths in an Erdos-Renyi (left) and a Barabasi-Albert (right) graph. *Bottom row:* Estimated shortest path distributions given data from the Travers & Milgram [15] as well as the Dodds et. al. [5] experiments.

The conditional estimates suggest that the shape of the estimated shortest path distributions is more or less the same in both experiments if we assume an identical q_-. The conditional estimate for $q_- = 0.5$ in TM is qualitatively somewhat similar to the full DMW estimate. This applies also to the conditional estimate for $q_- = 0.92$ in DMW and the full TM estimate.

4 Discussion

Our model has common aspects with the method devised in [10] as they also use a similar Markov chain to infer frequencies of observed completed paths given the q_- parameter. But there are some important differences. We allow the message also to move away from the target, and our model fully separates attrition from routing efficiency. Moreover, we propose to compute maximum-likelihood parameter estimates from observed paths. Finally, compensating for attrition has received attention in previous literature as well [5,7], but using very a different technique (importance sampling).

Small-world experiments show also that humans can find short paths in a decentralized manner. It is not obvious why this happens. There are two factors that play a role in the process: *structure* of the underlying social network and the *strategy* (or algorithm) used to forward the messages. In real small-world experiments it has been observed that participants tend to pass the message to an aqcuaintance who has some common attributes with the target. Especially geographical location and occupation have been reported as important criteria [5]. For such a “greedy” routing strategy to find short paths, the network must have certain structural properties that reflect the similarity of the nodes in terms of social attributes [11,16,13]. However, our model is independent of both routing strategy and network structure, meaning that we do not have to make assumptions about either.

Estimating the shortest path distribution from a set of very short (biased) random walks over an unobserved network is a hard problem, and the lack of a ground truth makes the results difficult to evaluate. We claim, however, that the proposed model can be a useful tool when analysing small-world experiments. Extending the model to deal with other types of data, such as information cascades [8,9,14,4] is an interesting open question, and so is improving the stability of the unconditional parameter estimates.

Acknowledgements
I would like to thank Aristides Gionis for his comments to a very early draft of this paper, and Hannes Heikinheimo for the valuable suggestions that helped to substantially improve the final version.

References

1. L. Backstrom, P. Boldi, M. Rosa, J. Ugander, and S. Vigna. Four degrees of separation. In *WebSci*, pages 33–42, 2012.
2. A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.
3. C. Bauckhage, K. Kersting, and B. Rastegarpanah. The Weibull as a model of shortest path distributions in random networks. In *MLG*, 2013.
4. F. Bonchi, G. De Francisci Morales, A. Gionis, and A. Ukkonen. Activity preserving graph simplification. *Data Min. Knowl. Discov.*, 27(3):321–343, 2013.
5. P.S. Dodds, R. Muhamad, and D.J. Watts. An experimental study of search in global social networks. *Science*, 301(5634):827–829, 2003.
6. P. Erdős and A. Rényi. On random graphs. *Publicationes Mathematicae Debrecen*, 6:290–297, 1959.
7. S. Goel, R. Muhamad, and D. J. Watts. Social search in ”small-world” experiments. In *WWW*, pages 701–710, 2009.
8. M. Gomez-Rodriguez, J. Leskovec, and A. Krause. Inferring networks of diffusion and influence. *Transactions on Knowledge Discovery from Data*, 5(4):21, 2012.
9. M. Gomez-Rodriguez and B. Schölkopf. Submodular inference of diffusion networks from multiple trees. In *ICML*, 2012.
10. P. D. Killworth, C. McCarty, H. R. Bernard, and M. House. The accuracy of small world chains in social networks. *Social Networks*, 28(1):85–96, 2006.
11. J.M. Kleinberg. Navigation in a small world. *Nature*, 406(6798):845, 2000.
12. C. Korte and S. Milgram. Acquaintance links between white and negro populations: Application of the small world method. *Journal of personality and social psychology*, 15(2):101–108, 1970.

13. D. Liben-Nowell, J. Novak, R. Kumar, P. Raghavan, and A. Tomkins. Geographic routing in social networks. *Proceedings of the National Academy of Sciences*, 102(33):11623–11628, 2005.
14. M. Mathioudakis, F. Bonchi, C. Castillo, A. Gionis, and A. Ukkonen. Sparsification of influence networks. In *KDD*, pages 529–537, 2011.
15. J. Travers and S. Milgram. An experimental study of the small world problem. *Sociometry*, 32(4):425–443, 1969.
16. D.J. Watts, P.S. Dodds, and M.E.J. Newman. Identity and search in social networks. *Science*, 296(5571):1302–1305, 2002.
17. W. Zachary. An information flow model for conflict and fission in small groups. *Journal of Anthropological Research*, 33(4):452–473, 1977.