Graph Theoretical Approaches to the Small-World Phenomenon

Felix Stahlberg

Supervisor: Björn-Oliver Hartmann
Seminar: Analysis of Social Networks
Institute for Program Structures and Data Organization (IPD)
Karlsruhe Institute of Technology, Karlsruhe, Germany

Abstract. Surely the academic area of social psychology is not considered having a large degree of overlap with mathematics. Nevertheless an experiment conducted by the social psychologist Stanley Milgram in 1967 – known as the *small-world experiment* [1] – is still in focus of current mathematic and computer scientific research for various reasons. The experiment suggests, that the most pairs of strangers in the world are connected via short *chains of acquaintance*. This paper presents several attempts at modeling this phenomenon as graph and investigating structural properties of these graphs, i.e. a low expected minimum path length of two arbitrary nodes. However, we will observe that such evidences of existence of short paths do not satisfy claims and we have to take algorithmic aspects into account as well. Therefore the paper focuses on local routing in small-world networks and presents two algorithms enabling nodes provided only with local knowledge to find short paths.

Key words: small-world experiment, small-world networks, routing, characteristic path length, clustering coefficient

1 Motivation

Most people probably came in contact with the small-world phenomenon already – maybe through own experiences like meeting a stranger and point out a common friend, maybe through applications on social network platforms calculating the usually short connection path via friends and friends of friends from you to the owner of the profile you are looking at (like [2]), or maybe through a radio morning show where they called some professor saying that you are likely to know Barack Obama personally via a very little number of intermediate acquaintances. All these notices refer to a thesis known as the *small-world phenomenon*, which is not only interesting for social psychology but for mathematic and computer scientific research as well.

The thesis was introduced by the American social psychologist Stanley Milgram in his manuscript *The small-world problem* [1], in which he conducted following experiment. A *source* person S (for instance in Boston) obtained a letter with the task of delivering it to a *target* person T (far away from the *source* person – for example in Omaha) identified by basic personal information. Therefore the source should forward the letter either directly to the target person (in the case of acquaintance) or to someone he or she knows on a first-name basis who is more likely to know the target. The result of the experiment was among other things, that the average path length of thus developed *chains of acquaintance* fell around six, which is a surprisingly small value in respect of the number of people living in USA. Furthermore these chains often passed the same dedicated person, namely Mr. G., who delivered the letter over a long distance. Mr. G. was somehow predestinated for that task, for instance due to his profession or because he lives in the same city as S but has relatives in T's occupation.

This experiment was adapted to other social networks (like [3], [4], [2]) with similar results. Even in technological (see [5]) or biological (for example the neural network of *Caenorhabditis elegans* [6]) networks, the small-world phenomenon is pointed out as an essential factor for the high efficiency of networking. We see that understanding this kind of network is important to many scientific areas, especially to the field of computer science. Routing in computer networks and designing efficient routing tables for such networks is only one example of use.

We will consider a historical overview over various attempts at an mathematical explanation of the small-world phenomenon. Furthermore we will discuss latest research results on this topic focusing on algorithmic approaches. Both aspects are highly nontrivial problems. Human social life and interpersonal communication are way too complex to represent it in a reasonable clear mathematical model. Human behavior is highly nondeterministic and complex as well, so it is questionable if the concept of algorithms is capable to provide a proper framework at all. However when we move our focus away from social psychology and consider other scientific fields of interest, the results help us understanding real-world networks and improving man-made networks as mentioned above.

In the second section, this paper presents different efforts of modeling small-world networks as graph. Therefore the central theme in this section can be summarized in following question. *How a network should be structured in order to be a small-world network?*. The third section focuses on algorithmic aspects and *how vertices provided only with local information are able to find such short chains of acquaintance.*

2 Modeling Approaches

2.1 Basic Paradigms

2.1.1 Modeling Principles All of the modeling approaches discussed in this paper follow the basic design idea suggested in [7]. According to this, small-world networks can be modeled by structured graphs plus a certain amount of randomness. Therefore the edges of such graphs can be divided into *local* and *long-range* connections. An arbitrary node is rich on local connections, but also has few long-range connections with endpoints somehow randomly distributed over the network.

This concept often has a "geographic" interpretation. A node n in a small-world network usually has a neighborhood consisting of nodes not far away from n, but is also associated with few nodes distributed more broadly across the network. An example is given by acquaintance relationships in social networks, where the nodes are individuals which are communicating with their neighbors as well as with friends all over the world. Another example are national, continental, and intercontinental air connections in the network of the world's air traffic where the nodes represent airports [8].

Because of technical reasons, we normalize most of the parameters defined in the following section in order to assure that they lay between 0 and 1. Thus we define a helper function

$$\forall A \subseteq V : c(A) := |A|(|A| - 1)\frac{1}{2} \tag{1}$$

where V is an arbitrary vertex set. Note that for a $A \subseteq V$, $c(A)$ gives the maximal number of edges which could exist between nodes in A.

2.1.2 Crucial Graph Properties Thinking of necessary structural properties for small-world networks forces the investigation of the *diameter* of the graph. When we take a closer look to the experiment setting, the *average length* of the shortest path between two arbitrary nodes should be more meaningful. Consequently [7] defines the *characteristic path length* L for an undirected connected graph $G = (V, E)$ as shown in (2).

$$L := \frac{\sum_{v_1,v_2 \in V} d_s(v_1, v_2)}{2 * c(V)} \tag{2}$$

where $d_s : V^2 \to \mathbb{N}$ gives the length of the shortest path between two nodes. Figure 1 gives two examples how to calculate this value. In figure 1a, all the nodes are directly connected with each other, so the average shortest path length between two arbitrary nodes is $L = 1$. In figure

(a) Complete graph structure with $L = 1$, $C = 1$.

(b) Star graph structure with $L = 72/42 \approx 1.714$, $C = (6 + 2/7)/7 \approx 0.898$.

Fig. 1: Calculating characteristic path length and clustering coefficient.

1b, the denominator is $2 * c(V) = 7 * (7-1) = 42$ since there are 7 nodes. For the same reason there are 42 positive summands in the numerator. $d_s(A, v) = d_s(v, A) = 1$ for all nodes $v \neq A$. Consequently 12 of the 42 summands are 1. All the other summands are 2 because if A is not an end- or start node of a shortest path, its length is 2 (having A as the "middle" node). This leads to $L = (12 + (42-12) * 2)/42 = 72/42 \approx 1.714$.

Unlike L, which describes the long-range connections, the *clustering coefficient* C (defined in (4)) is used to examine the structure of the local neighborhood of the nodes.

$$\forall v \in V : C_v := \frac{|E \cap N_v^2|}{c(N_v)}, N_v := \{n \in V | \{n, v\} \in E\} \tag{3}$$

$$C := \frac{\sum_{v \in V} C_v}{|V|} \tag{4}$$

Thus C_v is the fraction number of connections which exists in the neighborhood of v over the maximum number of edges which could possibly exists. This value is often called *local clustering coefficient* and describes how close the neighborhood of v is to being a clique. C then denotes the average over these C_vs. Consequently a high C implies a highly clustered graph. Figure 1a shows a graph with $C = 1$, because the neighborhood of each vertex is the whole (complete) graph. In figure 1b $C_v = 1$ for all $v \neq A$ since the neighborhood consists only of the directly connected v and A. Since nodes in N_A are connected with 6 edges and $c(N_A) = \frac{1}{2} * 7 * 6$, we can write $C_A = (2 * 6)/(7 * 6) = 2/7$. Consequently $C = (6 * 1 + 2/7)/7 \approx 0.898$.

[7] suggests, that small-world networks are networks with a high clustering coefficient C and a low characteristic path length L. However, the quantification via clustering coefficient is controversial, since it does not respect social roles of vertices in the social network application and does not allow "urban" sub-networks typically for human interaction behavior.

[6] extended these definitions for generic weighted graphs $G = (V, E, d : E \to \mathbb{R}_+)$ (therefore possibly nonmetric and nonconnected) by introducing the terms *local* (which describes the *fault tolerance* of a system) and *global efficiency* (which is an analogon to L) of a network. When we calculate the efficiency of unweighted graphs, we set $d \equiv 1$. Let now $\epsilon : V^2 \to \mathbb{R}$ and for $v_1, v_2 \in V$ let $\epsilon(v_1, v_2)$ be the inverse minimum weight of all paths in G between v_1 and v_2 (set $\epsilon(v_1, v_2) = 0$ if there is no such path in G or $v_1 = v_2$). For a subgraph $H \subseteq G$ we define its *efficiency* in (5).

$$E(H) := 2 * \frac{\sum_{v_1, v_2 \in V(H)} \epsilon(v_1, v_2)}{|V|(|V|-1)} \tag{5}$$

We set the *global efficiency* $E_{glob} := E(G)$ and the *local efficiency* $E_{loc} := \frac{\sum_{v \in V} E(G_v)}{|V|}$ where $G_v \subseteq G$ is the subgraph inducted by v and its neighbors. $\frac{1}{L}$ may be a reasonable approximation for E_{glob} as long as we are aware of the fact that $\frac{1}{L}$ measures the efficiency on *sequential* systems while E_{glob} handles *parallel* systems in the communication network scenario. C is a good approximation for E_{loc} on nonsparse graphs. Consequently small-world networks are characterized by high local as well as high global efficiency.

In order to normalize the measurement, [6] suggests to divide $E(G)$ with a suitable value to guarantee $0 \leq E(G) \leq 1$. In general this value depends on the specific graph. However, in this paper we will consider only unweighted graphs ($d \equiv 1$), so we chose the constant 2 in (5). This is correct, because the most efficient unweighted graphs are the complete graphs K_n, since for

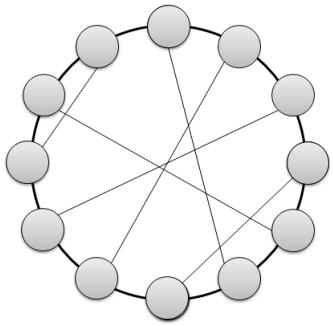


Fig. 2: Graph formed by adding a random matching to a 12-cycle.

all nodes $v_1, v_2 \in V(K_n), v_1 \neq v_2$ is $\epsilon(v_1, v_2) = 1$ (inverse of the shortest path, which has length 1). Thus for all unweighted graphs $G = (V, E)$ we can estimate $E(G)$ as follows.

$$0 \leq E(G) \leq E(K_{|V|}) = 2*\frac{\sum_{v_1,v_2 \in V(K_{|V|})} \epsilon(v_1, v_2)}{|V|(|V|-1)} = 2*\frac{\sum_{v_1,v_2 \in V(K_{|V|}), v_1 \neq v_2} 1}{|V|(|V|-1)} = \frac{|V|(|V|-1)}{|V|(|V|-1)} = 1 \tag{6}$$

2.2 Interpolating Between Random and Structured Graphs

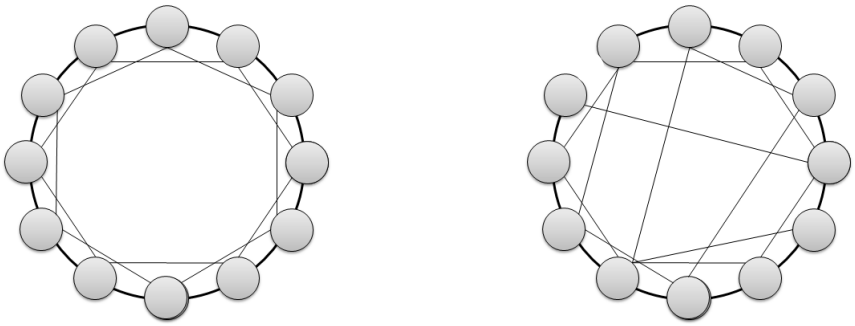
2.2.1 Diameter of Structured Graphs Plus Random Matchings One of the first mathematic approaches to the small-world phenomenon considered the diameter of random graphs, which usually is in $O(\log n)$ and thereby very small [9]. Therefore it is very likely that a short path between two generic vertices exists in a random graph. However, modeling small-world networks with pure random graphs was given up quickly because it does not represent a fundamental aspect of small-world networks: locality. For instance in social networks, when A and B know C, it is very likely that A and B know each other. In order to overcome this defect, the model has to somehow interpolate between a structured and a completely random graph. In [10], B. Bollobás and F. R. K. Chung considered the diameter of such interpolating graph structures. They generalized their theorem and showed that inequalities similar to (7) hold true basically for all bounded degree graphs plus random matchings. In this paper we will only cover the basic case, namely that if "a graph $G = (V, E)$ [is] formed by adding a random matching to an n-cycle [as shown in figure 2], then with probability tending to 1 as n goes to infinity, G has diameter $D(G)$ satisfying

$$\log_2 n - c \leq D(G) \leq \log_2 n + \log_2 \log n + c \tag{7}$$

where c is a small constant (at most 10)" [10]. Consequently, the characteristic path length L lays in $O(\log n)$. Therefore for $v_1, v_2 \in V, v_1 \neq v_2$ is $\epsilon(v_1, v_2) \in \Omega(\frac{1}{\log n})$ because G is connected. Thus, analog to (6) we write

$$E_{glob} \in \frac{|V|(|V|-1) * \Omega(\frac{1}{\log n})}{|V|(|V|-1)} = \Omega(\frac{1}{\log n}). \tag{8}$$

L and E_{glob} assure a good global efficiency of G. We can see that $\mathbb{E}(C) = \frac{2}{3}$, since every vertex has two neighbors, so the maximum number of edges in such neighborhood is three (K_3



(a) Rewired lattice with $n = 12$, $k = 4$, $p = 0$. (b) Rewired lattice with $n = 12$, $k = 4$, $p = 1/8$.

Fig. 3: Random rewiring procedure of Watts and Strogatz.

has three edges), but actually there exist only two edges (assuming there random matchings point to vertices outside this neighborhood which is the case as n goes to infinity). Analog consideration leads to

$$\mathbb{E}(E_{loc}) = 2 * \frac{1 + 1 + \frac{1}{2}}{3 * (3 - 1)} = \frac{5}{6}. \tag{9}$$

The analysis done above shows that graphs formed by adding a random matching to a cycle meet our paradigms constituted in section 2.1. Nonetheless we will generalize this modeling approach in the following and show that even the generalized model is inappropriate for explaining the small-world phenomenon in section 3.

2.2.2 Randomly Rewired Lattices In [7], Watts and Strogatz suggest a modeling procedure which allows to interpolate between random and structured graphs in a more detailed way. Therefore they started with a ring lattice with n vertices of the degree k. With a probability of

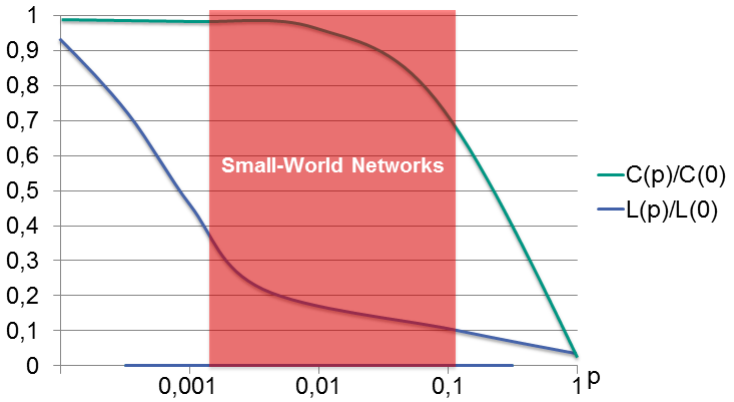


Fig. 4: Small-world networks in randomly rewired lattices, characterized by a low characteristic path length L and a high clustering coefficient C. The diagram sketches their distribution depending on the rewiring probability p (logarithmic scaled). Diagram adapted from [7].

$0 \leq p \leq 1$ each edge is then rewired to a randomly but uniformly chosen new target. Note that $p = 0$ leads to a regular lattice while setting $p = 1$ produces completely random graphs. Two examples of this modeling procedure are shown in figure 3. For the clustering coefficient C and the characteristic path length L, they detected that with increasing p, $C(p)$ stays on a high level for a long time while $L(p)$ quickly drops to a low level. So there are values of p with $L(p)$ is close to L_{random} but $C(p) \gg C_{random}$, which identifies small-world networks (as outlined in figure 4). They test this conclusion successfully with three real world small-world networks (collaboration network of actors, electrical power grid of the western United States and the neuronal network of the worm *C. elegans*).

2.2.3 Kleinberg's Model The models presented so far were based on the assumption that the long-range contacts are uniformly distributed over the network. [11] showed that this assumption may prevent local algorithms from finding short paths effectively (see section 3), so he used a more generalized graph model $G = (V, E)$, introducing the parameter $p \geq 1$, $q \geq 0$ and $r \geq 0$. The vertices in Kleinberg's model are identified by tuples $(i, j) \in \{n_0 \in \mathbb{N} | n_0 \leq n\}^2$ and therefore represent lattice points in a $n \times n$ square ($n \in \mathbb{N}$ fixed). We define the metric $d((i, j), (k, l)) := |k - i| + |l - j|$, which would have the meaning of "number of lattice steps between (i, j) and (k, l)" in our $n \times n$ square analogon. This metric is known as the *Manhattan distance*. A vertex v is connected with all vertices u for which $d(v, u) \leq p$ holds true (local contacts) and additionally to q long-range contacts, whereby the endpoints of the long-range contacts are chosen by independent random trials using a probability distribution called *inverse r^{th} power distribution*. Thus there is a long-range edge from v to u with probability

$$P((v, u) \in E) = \frac{[d(v, u)]^{-r}}{\sum_{x \in V} [(d(v, x)]^{-r}}. \tag{10}$$

When we set $r = 0$ we get the rectangular distribution for long-range contacts. Note that when we additionally set $q = p = 1$ and use a one dimensional lattice, our resulting graphs are isomorph to graphs formed by adding a random matching to a cycle discussed in section 2.2.1. Consider figure 5 to become more familiar with the model.

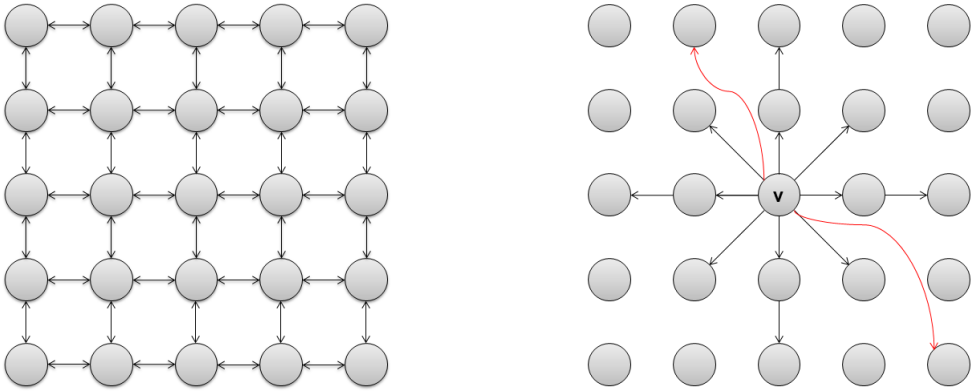
[11] provides no further analysis of its network model. [12] suggests that the expected diameter is $\Theta(\log n)$ when we set $r = 2$. Evaluating the clustering coefficient seems to be nontrivial and would go beyond the scope of this paper.

In the following paragraph we will refer to the parameters from [7] with p_{sw} and k_{sw}.

Counterintuitively Kleinberg's model is *not* an exact generalization of the approach from [7] which we considered in section 2.2.2, since we cannot simulate this rewiring procedure for all feasible values of p_{sw}. For instance let G be such a rewired lattice with $0 < p_{sw} < 1$. In order to transform such a structure to Kleinberg's regime, we have to set $p = 0$ because otherwise each vertex would have a guaranteed number of local contacts which is not the case in G (each outgoing edge of a vertex may be rewired). Thus all contacts have to be long-range contacts ($q = k_{sw}$). The rewiring procedure in [7] implies that for a vertex v the probability of having an edge to vertex u is

$$P((v, u) \in E) = \begin{cases} c_1 & \text{if } v \text{ was connected with } u \text{ before rewiring} \\ c_2 & \text{otherwise} \end{cases} \tag{11}$$

for some constants c_1 and c_2 depending on p and n. Such a case-by-case analysis could obviously not be simulated by an inverse power distribution unless one of the cases would be impossible. However since both cases are possible (because of $p_{sw} \notin \{0, 1\}$), a model of this kind has no equivalent in Kleinberg's regime.



(a) Regular two-dimensional 5×5 lattice produced by setting $n = 5$, $p = 1$, $q = 0$.

(b) Outgoing edges of a vertex v in a model with $n = 5$, $p = 2$, $q = 2$. The red edges (long-range connections) point to nodes distributed more broadly over the lattice.

Fig. 5: Modeling approach in [11].

2.2.4 Further Classifications of Small-World Networks The question of vertex connectivity distributions of small-world networks was discussed in [8] as well. For this purpose they studied a number of real-world small-world networks (such as social, biological, physical, economical and technical networks) and devided them into three classes characterized by their vertex distribution of the long-range connections:

1. *scale-free networks* with a vertex distribution decaying as a power law similar to Kleinbergs model from section 2.2.3,
2. *single-scale networks* with a vertex distribution with a fast decaying tail, for instance exponential or Gaussian decay and
3. *broad-scale networks* with a vertex distribution similar to scale-free networks up to a dedicated value, but similar to *single-scale networks* for distances greater than this value.

They pointed out several factors responsible for the variant vertex distributions. *Scale-free networks* arise if new vertices added to the network link themself to a number of existing vertices, more likely to those with high degree (also known as *rich get richer* model). This growing behavior causes the characteristic inverse power vertex distribution [13]. There are several factors which could prevent a network from being scale-free. One of these factors is called *aging*, which we will explain using the example of the collaboration network between actors. In time, an actor stops acting and the degree of his representing vertex will remain static form that moment on. However, the vertex is still part of the network. Besides that a vertex can have a maximum degree which could be handled. For example considering the network of world's air traffic, where vertices represent airports, the size of an airport limits the maximum number of links possible. Aside from that, adding new links to this network involves costs, which tamper with the growing behavior of scale-free networks and lead to vertex distributions different from the basic case with power law like vertex distributions.

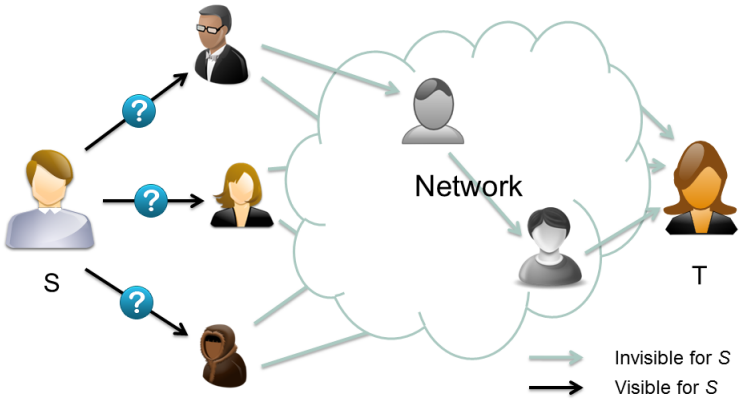


Fig. 6: Local algorithms in the small-world setting.

3 Algorithmic Aspects

3.1 Formal Framework

[11] pursues the question *how vertices provided only with local knowledge are able to find short paths in a small-world network* as shown in figure 6. This is one of the most fundamental questions of scientific analysis on small world networks, because it is combined with possible applications for example for routing problems in real world networks. Analyzing this problem requires a proper framework. Namely we have to define

1. network model we will use,
2. how the *local knowledge* of each node is defined and
3. what a *short path* is.

We will use the network model presented in section 2.2.3 and define a message holder's v *local knowledge* as

1. knowledge of the underlying grid structure,
2. the lattice coordinates of the target t and
3. the lattice coordinates of all former message holders and their long-range contacts,

whereby (3.) will not be used in any of the presented algorithms, because (1.) and (2.) already might fulfill locality in a regular sense while (3.) would overfulfill it. A usage in impossibility statements however will only strengthen their meaning, which will be a statement about locality.

A *short path* is defined as a path from vertex s to t whose length is polynomial in its expected minimum path length – or in other words a polynom in $\log n$.

3.2 Algorithmic Results

3.2.1 Impossibility Statements [11] obtains two major results. Firstly becomes apparent that there is no way for a local algorithm to find short paths in a Kleinberg network with $r \neq 2$ (even if they exist). This has an extensive impact on former research. Recall the network model discussed in section 2.2.1. As shown in section 2.2.3 graphs produced by adding a random matching to an n-cycle form a subset of Kleinberg's model which we use here (generated by

setting $p = q = 1$, $r = 0$ and using a 1-dimensional lattice). Therefore local algorithms are unable to find short paths in this model and consequently it is inappropriate to explain the small-world phenomenon (the vertices [individuals] in Milgram's small-world experiment indeed have only local knowledge). Probably the Strogatz-Watts-Model (section 2.2.2) also proves to be unsuitable, even though it is no subset of Kleinberg's model, as shown in section 2.2.3. Their differences however turned out to be slight and on top of this the uniform probability distribution used by Strogatz and Watts is likely to prevent local algorithms from finding short paths.

3.2.2 Kleinberg's Algorithm As second major result, [11] suggests a simple local greedy algorithm $\mathcal{A}$, which is able to find short paths in Kleinberg networks with $r = 2$ and $p = q = 1$ from a source node s to a target node t. When we combine both results, we see that using the inverse-square distribution for generating long-range contacts is a necessary condition for local algorithms being able to find short paths.

$\mathcal{A}$ is probably the first algorithm which comes in mind when thinking about the problem. First of all, it declares s as the *current message holder* u and set the path $P =< s >$. Then, until $u = t$, $\mathcal{A}$ chooses a vertex, which is adjacent to u and as close as possible to t. This vertex is added to P and serves as the new *current message holder* u. At the end, P is returned. $\mathcal{A}$ is shown in algorithm 1.

Figure 7 gives an example of a possible output of this algorithm. The red colored edges mark the returned path. In its first step, the algorithm chooses a local contact since the long-range contact of s is far away from t. In the second step, $\mathcal{A}$ uses a long-range connection and thus shortens the remaining distance to t greatly. Note that this is actually a suboptimal decision because if $\mathcal{A}$ had chosen the subjacent local contact and afterwards long-range contacts twice, the length of the returned path would have been 4 instead of 6. Nevertheless keep in mind that we do not search for shortest paths but only for *short paths* (polynomial in the expected minimum path length).

In order to see that the algorithm is truly capable to find short paths, we will discuss the basic proof idea for that claim. Consider the green color gradation in figure 7. We divide the course of action into phases. We say $\mathcal{A}$ is in phase $i \in \mathbb{N}_+$, iff $2^i < d(v,t) \leq 2^{i+1}$, and in phase 0, iff $d(v,t) \leq 2$, where v denotes the current message holder. It is obvious that the number of phases $\mathcal{A}$ walks through is at most $1 + \log n$. Kleinberg shows, that the expected number of iterations the algorithm performs in each phase lays in $O(\log n)$. Altogether, the cumulative number of iterations lays in $O(\log^2 n)$. Since $\mathcal{A}$ increments the length of the returned path by one in each iteration, the total length of the returned path lays in $O(\log^2 n)$ as well and thus polynomial in the expected minimum path length (which is in $O(\log n)$ as mentioned in section

Algorithm 1 Greedy algorithm $\mathcal{A}$ for finding short paths for $r = 2$. Note that this implementation is supposed to encourage better understanding. In fact, $\mathcal{A}$ has no knowledge of the entire network structure $G = (V, E)$ like it is passed to the function.

```
1: function GREEDY(G = (V, E), s, t)        ▷ Find a short path in G from s ∈ V to t ∈ V
2:     P ←< s >
3:     u ← s
4:     while u ≠ t do
5:         u ← arg min_{x∈{v∈V|(u,v)∈E}} d(x, t)
6:         P.append(u)
7:     end while
8:     return P
9: end function
```

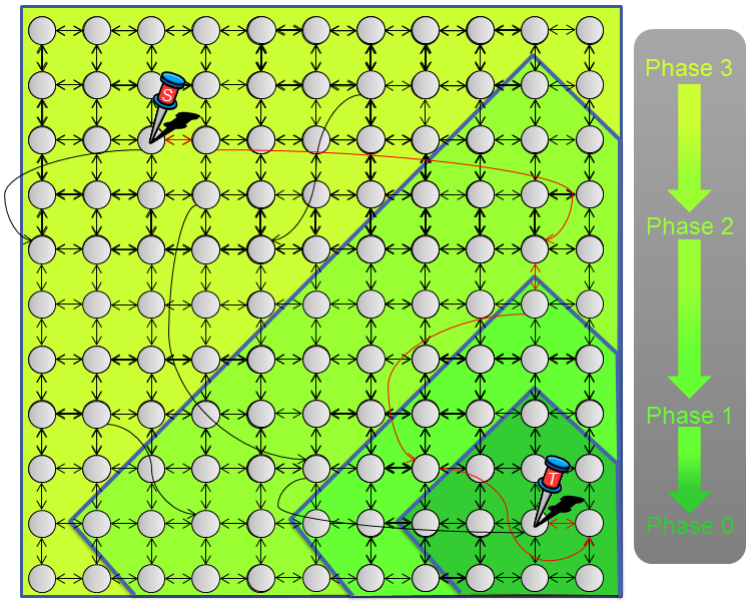


Fig. 7: A network of Kleinberg's modeling framework with $n = 11$, $p = q = 1$ and $r = 2$. Most of the long-range connections are dropped to simplify matters. The red edges identify an example output of algorithm 1. The green color gradation visualizes the phases of the algorithm which are used for analysis.

2.2.3). Consequently $\mathcal{A}$ does return short paths. The correctness follows when we recall that $\mathcal{A}$ starts with s and consider that the remaining distance between t and the current message holder is decreasing strictly monotonic in each iteration.

3.2.3 Indirect-Greedy Algorithm After Kleinberg's initial results concerning algorithmic aspects of small-world networks, a number of papers followed up his ideas. For example [12] showed that Kleinberg's analysis of $\mathcal{A}$ is tight and therefore $\mathcal{A}$ runs even in $\Theta(\log^2 n)$. It also suggests an algorithm which achieves path lengths in $O(\log^{3/2})$, but needs additional local knowledge (each node have to know the long-range connections of its $\log n$ closest neighbor nodes).

[14] provides each node with $O(\log^2 n)$ bits of topological awareness as well and proofs thereby a general tight $O(\log^{1+1/k} n)$ upper bound for greedy algorithms on k-dimensional lattices. This means that short path finding algorithms are getting more effective with increasing number of dimensions of the underlying grid structure. In the case of social networks, a two-dimensional lattice will be turn out to be inappropriate to represent the human nature, because it would leave out many intrinsic dimensions of our (world) society like religion, culture, social standings or different professions. [14] also specifies an *oblivious* algorithm (*indirect-greedy*) which achieves the upper bound mentioned above. In this context *oblivious* means that the algorithm does not need to store data about former message holder at any step. Thus decisions made by the algorithm only depend on the target t and local information about the current message holder, which reduces protocol overhead in network routing scenarios, enhances fault-tolerance and is more similar to the way people behaved in Milgram's original experiment [1].

In the following, let $G = (V, E)$ be an instance of Kleinberg's network model (with $r = 2$, $q \geq 1$ and $p \geq 1$), $\mathcal{B}$ denote the *indirect-greedy* algorithm operating on G in order to find a

Algorithm 2 Indirect-greedy algorithm $\mathcal{B}$ on $G = (V, E)$. $A_v \subseteq E$ denotes the topological awareness and $A'_v := A_v \cup (E \cap \{v\} \times V)$ the edges known by v. Note that this implementation is supposed to encourage better understanding. In fact, $\mathcal{B}$ has no knowledge of the entire network structure G like it is passed to the function.

function IndirectGreedy($G = (V, E), s, t$) ▷ Find a short path in G from $s \in V$ to $t \in V$
  $P \leftarrow < s >$
  $u \leftarrow s$
  while $u \neq t$ **do**
    $a \leftarrow \arg\min_{x \in \{v \in V \mid \exists y \in V : (y,v) \in A'_u\}} d(x, t)$ ▷ Find target of intermediate destination
    $b \leftarrow \arg\min_{x \in A'_u \cap V \times \{a\}} d(u, x)$ ▷ Find intermediate destination
    $u \leftarrow \arg\min_{x \in E \cap \{u\} \times V} d(x, b)$ ▷ Route to intermediate destination
    $P.append(u)$
  end while
  return P
end function

short path from $s \in V$ to $t \in V$ and $A_v \subseteq E$ the topological awareness of node $v \in V$. A_v is a list of long-range connections in G, for instance the long-range connections of the $\log n$ closest neighbors of v. $\mathcal{B}$ is shown in algorithm 2. Each step of $\mathcal{B}$ can be divided into two phases. In the first phase, $\mathcal{B}$ chooses the edge $(x, y) \in E$ that is known by the current message holder u and minimizes the distances $d(y, t)$ (on first priority) and $d(u, x)$ (on second priority). In the second phase, it finds a node which is adjacent to u and as close as possible to x. This node is appended to the output path and declared as the new current message holder. Then $\mathcal{B}$ starts again with the first phase applied to the new current message holder until it reaches t.

Reconsidering figure 7 we can now observe that when assuming A_v contains the long-range connections of the respective subadjacent node, $\mathcal{B}$ would not choose the suboptimal long-range contact in the second step like $\mathcal{A}$ did (recall section 3.2.2) but would see the advantageous long-range connection of the current message holder's neighbor and choose this node as intermediate destination. Thus $\mathcal{B}$ would output a path with length 4 instead of 6.

4 Conclusion

The small-world phenomenon states, that the most pairs of strangers in the world are connected via short chains of acquaintance. This thesis does not only spark interest of social psychologists but of scientists from various other academic areas as well.

We have discussed several attempts at a mathematical explanation of the small-world phenomenon. Small-world networks can be modeled by mixtures of structured and random graphs combining locality and low minimum path lengths. In order to measure these both polarities we introduced the terms *characteristic path length* and *clustering coefficient* and alternatively *global* and *local efficiency* of networks, based on which we evaluate different modeling approaches. Cycles plus random matchings score well in our measurement framework, but do not allow to interpolate between the polarities mentioned above. Randomly rewired lattices overcome this defect introducing the parameter p. Nevertheless we detected the modeling approach of [11] as a more generalized framework for small-world networks. Moreover [11] gives fresh impetus to the discussion raising the question how vertices in such graphs are able to find short paths. It suggests that the network has to provide some *clues* to the vertices how to find a suitable vertex in order to deliver the message. These clues disappear for example if the long-range connections are generated by random trials using the rectangular distribution like it was often assumed in former modeling approaches. [11] and follow-up papers like [14] and [12] presented simple local greedy algorithms capable of producing short paths in dedicated model instances using a inverse-square vertex connectivity distribution for long-range contacts.

However, the modeling frameworks are still very clean and skip important factors of small-world networks observed in the real-world. The discussed models for instance assume uniform or power law connectivity distributions of the long-range connections. In section 2.2.4 we detected that other distributions (like exponential, Gaussian or mixtures) are possible. When we consider the world community, even these distributions may be inappropriate and a gradual distribution would be a better approximation. Reducing to the "geographic" dimensions, it is probably more likely to know people from our home city (state, country, continent...) then from our neighbor city (state, country, continent...), even if the geographic distance to second ones might be shorter. Furthermore the models ignore the wide variety of connectivity characteristics of the individual nodes, for instance the prominent position of Mr. G. in Milgram's experiment (mentioned in the introduction section 1).

Research on small-world networks has many fields of application. In computer science one of its many applications is helping to evaluate and optimize the trade-off between size and efficiency of routing tables in the computer network scenario. Popular public key infrastructures (systems providing secure user identification of network participants) like PGP are based on a *Web of Trust* [15], which has to show characteristics of small-world networks for being useable. Besides that, very large social networks turned out to be small-world networks (conversation graph of MSN [16] and the friendship network of Facebook [17]) recently. Research results might offer valuable clues to information flow in such communication systems.

To sum up, small-world networks will stay in focus of research and have not been fully explored yet. We can hope that future research will enable us to combine theory with practice even more strongly.

Acknowledgments. Thanks to Ernst-Victor Prohl for his valuable review.

Bibliography

[1] Milgram, S.: The small world problem. Psychology Today (1967)

[2] Quinn, B., et al.: Facebook app six degrees. `http://apps.facebook.com/sixdegreesearch/` (December 2010)

[3] Liljeros, F., et al.: The web of human sexual contacts. Nature **411** (June 2001)

[4] Adamic, L.A.: The small world web. In: Research and advanced technology for digital libraries : third European Conference, ECDL 99, Paris, France, September 22 - 24, 1999; proceedings. Lecture notes in computer science ; 1696. Springer, Berlin (1999)

[5] Bernardo A. Huberman, L.A.A.: Diameter of the world-wide web. Nature **401** (September 1999)

[6] Latora, V., Marchiori, M.: Efficient behavior of small-world networks. Phys. Rev. Lett. **87**(19) (Oct 2001) 198701

[7] Duncan Watts, S.S.: Collective dynamics of small-world networks. Nature **393** (June 1998)

[8] Amaral, L.A.N., Scala, A., Barthelemy, M., Stanley, H.E.: Classes of small-world networks. Proceedings of the National Academy of Sciences of the United States of America **97**(21) (2000)

[9] Bollobás, B.: Random graphs. 2. ed., repr. edn. Cambridge studies in advanced mathematics ; 73. Cambridge Univ. Press, Cambridge (2004) Previous ed.: London : Academic Press, 1985.

[10] Bollobás, B., Chung, F.R.K.: The diameter of a cycle plus a random matching. SIAM Journal on Discrete Mathematics **1**(3) (1988) 328–333

[11] Kleinberg, J.: The small-world phenomenon: an algorithm perspective. In: STOC '00: Proceedings of the thirty-second annual ACM symposium on Theory of computing, New York, NY, USA, ACM (2000) 163–170

[12] Martel, C., Nguyen, V.: Analyzing kleinberg's (and other) small-world models. In: Proceedings of the twenty-third annual ACM symposium on Principles of distributed computing. PODC '04, New York, NY, USA, ACM (2004) 179–188

[13] Ijiri, Y., Simon, H.A.: Skew distributions and the sizes of business firms. Studies in mathematical and managerial economics ; 24. North-Holland Comp., Amsterdam [u.a.] (1977)

[14] Fraigniaud, P., Gavoille, C., Paul, C.: Eclecticism shrinks even small worlds. Distributed Computing **18** (2006) 279–291

[15] Zimmermann, P.R.: The official PGP user's guide. MIT Press, Cambridge, MA, USA (1995)

[16] Leskovec, J., Horvitz, E.: Planetary-scale views on a large instant-messaging network. In: Proceeding of the 17th international conference on World Wide Web. WWW '08, New York, NY, USA, ACM (2008) 915–924

[17] Wilson, C., Boe, B., Sala, A., Puttaswamy, K., Zhao, B.: User interactions in social networks and their implications. In: Proceedings of the 4th ACM European conference on Computer systems. EuroSys '09, New York, NY, USA, ACM (2009) 205–218