# The Beacon Number Problem in a Fully Distributed Topology Discovery Service

D. Ficara[‡], F. Paolucci[*], L. Valcarenghi[*], F. Cugini[†], P. Castoldi[*], S. Giordano[‡]
[*]Scuola Superiore Sant'Anna, Pisa, Italy
Email: {fr.paolucci, luca.valcarenghi, piero.castoldi}@sssup.it
[†]CNIT, Pisa, Italy
Email: filippo.cugini@cnit.it
[‡] Università degli Studi di Pisa, Italy
Email: d.ficara@netserv.iet.unipi.it, s.giordano@iet.unipi.it

***Abstract*—In grid computing the need for collecting information about both distributed computational resources and network topology and performance is constantly growing. Several tools are currently under development to provide such information. They are based either on a centralized or a distributed architecture.**

**Distributed tools are commonly based on IP-level application-oriented network metrology. The measurements are done by means of beacons running in some network nodes (e.g., grid hosts) and collecting the required information. However, the number of utilized beacons might heavily impact the final result, i.e. the discovered topology and the collected performance information.**

**In this study a model is developed to estimate the percentage of discovered links provided that a specific number of beacons is placed in the network. The model is developed for Erdös-Rényi (ER) graph network models but it can be applied also to other networks. Numerical evaluation shows that the model closely approximate the percentage of discovered links obtained through simulation for ER networks. For other theoretical and real networks the model overestimates the percentage of discovered links. However experimental results show that for networks with realistic average nodal degree the overestimate is less than 20%.**

## I. Introduction

Emerging distributed applications (e.g., grid-enabled applications) rely on computational resource information to optimize their elaboration time. However, when the elaborations involve also a heavy utilization of network resources, such as a large file transfer or intense interprocess communication, information about network resource status (e.g., delay or bandwidth capacity) becomes of paramount importance.

Currently, several tools have been developed for network topology discovery and performance monitoring. They may be classified into two main categories: the first one is related to topology builder tools and the second one includes resource monitoring and brokering tools for specific applications. Tools such as HynetD [1], Topomon [2], Skitter [3] and NetInventory [4] belong to the first group: their main goal is to derive a complete network graph by exploiting Simple Network Management Protocol (SNMP) and Internet Control Message Protocol (ICMP). Tools like MonALISA [5], based on Simple Network Management Protocol (SNMP) queries, and Network Characterization Service (NCS) [6], based on the packet train technique, belong to the second group: they provide statistical analysis of performance parameters, such as bandwidth, round-trip-time, server congestion, of a known network.

However, the aforementioned tools do not deal with the beacon number problem, that is the problem of computing the number of beacons (i.e., the number of measurement points) that provides a complete network topology and performance information. Indeed they are based on a fixed monitoring infrastructure where the number of beacons is decided a-priori and it coincides, usually, with the equipments (e.g., routers) that connect the involved grid hosts to the network.

In the recent years, several studies confirmed that poor coverage (i.e., undersampling due to an insufficient number of beacons) of the underlying network can lead to some bias in the reconstructed graph. The studies mainly focused on the exploration of scale-free (SF) graphs, i.e. graphs whose metrics of interest follow a power-law distribution. For example, in SF graphs the node degree distribution function can be expressed as $P(k) \propto k^{-\gamma}$, where $P(k)$ is the node degree distribution function, $k$ is the node degree, and $\gamma$ is a constant [7]. In terms of statistical properties of graphs, Petermann et al. [8] and Lakhina et al. [9] found that because low-degree vertices are undersampled with respect to high-degree ones, the observed value of $\gamma$ is lower than the correct exponent of the sampled SF graphs. The bias can also lead to bigger ambiguities such as making power law appear where it does not exist [10].

In particular the number of beacons might heavily impact the graph undersampling in traceroute-based sampling. In [11], the authors show that the number of beacons, building traceroute spanning trees, that yields an unbiased reconstruction of an SF network increases linearly with the average nodal degree. In [12] the authors look for qualitative results and show that a non-random choice of sources and destinations for traceroutes can reduce the number of explorations for a complete sampling. In particular, the policy of choosing source nodes, where to place beacons, with increasing degrees proves to be the best. Although it is only an empirical result, it has a justification in the difficulty to reach low-degree nodes. As for quantitative results, Barford et al. [13] define the marginal utility of topology discovery through an information-theoretic approach and show that it is more useful to increase the number of sources (i.e., beacons) for traceroute-

1930-529X/07/$25.00 © 2007 IEEE
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2007 proceedings.

like explorations than to add more destinations to the same source. Horton et al. [14] show that minimizing the number of beacons that provides a complete knowledge of an arbitrary network is NP-hard. The authors also find that the required number of beacons is at least $(n-1)/3$ and at most $(n+1)/3$, where $n$ is the number of network nodes. Therefore, under the request for a complete coverage of the underlying graph, a topology discovery scheme would require a very large number of beacons, with a corresponding huge amount of message exchanged among beacons.

In [15], Kumar et al. define the concept of *Deterministic Monitorable Edge Set* (DMES) of a beacon as the edge set that can be monitored and thus explored by the beacon. Then, they propose an approximate placement algorithm that yields beacon sets of sizes within $1+ln(|E|)$ of the optimal solution, where $|E|$ is the number of network bidirectional links. In [16] the authors present a formal analysis of the exploration process assuming that traceroute yields partially homogeneus trees with deterministic degrees. Thus all the nodes at height $j$ (i.e, with a distance of $j$ links from from the root) have fixed degree $d_j$. This is a strong simplification but [16] proposes an interesting analogy of the exploration process with the well-known "coupon collector problem" (see section III) that is used in the model developed in this study. Finally [17] shows an interesting analysis of reproducible graphs properties called $dK$*-series*. In particular the authors focus on distribution of *"d"-order* degree correlation probability in order to build a set of $dK$*-graphs* that capture the $dK$*-distribution* of an original given graph. Hence producing a $0K$*-graph* requires only the original average degree and ends up with the creation of an Erdös-Rényi (ER) graph. A $1K$*-graph* requires and reproduces the original degree distribution, the $2K$ order requires and reproduces the joint degree distribution and so on.

In this study a fully distributed topology discovery service, namely the FD-TDS, is proposed to provide network clients, i.e. applications or end-users, with network topology and performance information. The FD-TDS is based on distributed services embedding beacons that, activated in some grid hosts, collect a subset of the overall network topology and performance information that is then distributedly elaborated. Thus the FD-TDS must deal with the beacon number problem, because the number of utilized services, i.e. beacons, might impact the collected information. Indeed, the beacon number problem is technology-independent and it might arise in any distributed topology discovery and performance information tool. In the following sections the fully distributed topology discovery service (FD-TDS) architecture and implementation are described. Then the analytical model for the beacon number problem in Erdös-Rényi (ER) graphs with traceroute-based exploration is proposed and detailed. Furthermore simulative results show the performance, in terms of percentage of discovered network links, obtained by applying the developed model to different theoretical graphs. In the end, experimental results show the performance of the FD-TDS implementation in discovering a real network topology as a function of the number of utilized services (i.e., beacons).

## II. Fully Distributed Topology Discovery Service Architecture and Implementation

The Fully Distributed Topology Discovery Service (FD-TDS) is intended to be deployed in a set of grid hosts running Globus Toolkit and belonging to the same Virtual Organization (VO). Each host runs an FD-TDS instance which registers itself into a Monitoring and Discovery Service (MDS) located in some node. This way the information gathered by each service becomes available to grid clients.

Such information, called resources in MDS, is the following:

- FDTDSLinkStatus is the database of all links known by the service. Each link is described by a number of characteristics (e.g., available bandwidth capacity);
- FDTDSPeerList is the database of all known peers (i.e., hosts in which the service is running);
- FDTDSNodeList is the set of all explored nodes.

The FD-TDS, whose architecture is depicted in Fig. 1, is fully distributed. Every service is a multithreaded application where each thread is responsible for a pathchar [18] exploration toward some set of nodes hosting peer services. At the end of each exploration, the FDTDSLinkStatus is updated with the new links discovered, if any. If the discovery process finds some already known links, the corresponding structure inside FDTDSLinkStatus is updated with the new data. The data structure for each link reports pathchar data (measured bandwidth, Round Trip Time, packet drop probability), an identifier of the link, and the source and destination of the path that includes the link.

Each service keeps a list of peers (FDTDSPeerList), that is the list of hosts running the same service. This is needed to distribute the information with no need for a centralized component or a broker.

An exploration can be requested by a client or triggered by an event such as the addition of a new host running the service. In the exploration process, the requested service triggers the execution of a pathchar from each host in the FDTDSPeerList toward all the other hosts in the list, obtaining trees rooted at the service hosts. Each tree is then sent to all the other peers that independently merge all the received subtrees and obtain the topology.

When an exploration ends, the service broadcasts its new data to the members of its FDTDSPeerList. Even changes in the FDTDSPeerList are broadcast: for instance, after adding a new node to the FDTDSPeerList of a host, all other nodes are notified about the new node.

Finally one of the host of the FDTDSPeerList runs also an MDS service. All information including link status, inferred node topology, and peer database are then handed to a MDS service which is in charge for the presentation of data in a simple format, such as a webpage.

Since any change in the resources of a node is broadcast, all the information shown by the MDS service is always updated, no matter which node actually runs the service.

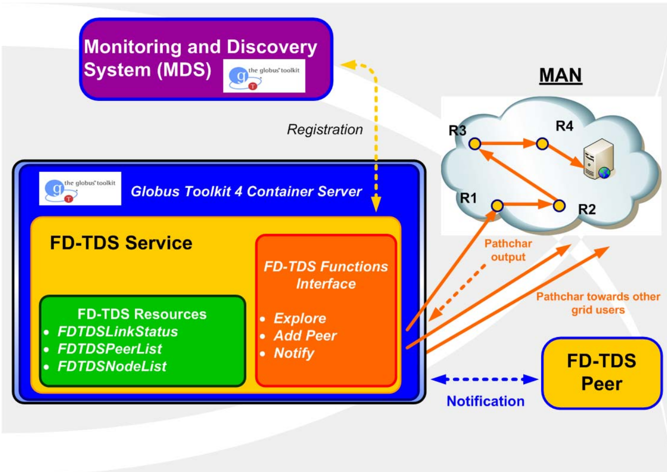


Fig. 1. FD-TDS overall Architecture.

## III. Modeling the beacon discovery

In this section a model for the beacon number problem is proposed. The model assumes that beacons explore a graph $G(N,E)$ through uniform (i.e., randomly rooted) shortest path spanning trees (USTs). Indeed a single beacon is assumed to utilize a traceroute-like exploration toward all the possible destination nodes. Thus the IP-layer topology exploration conducted by a single beacon running in a node can be modeled as a spanning tree built by a breadth-first-search algorithm [10], [16], [8]. Furthermore, in the model, the "coupon collector problem" analogy is applied in a more general case than [16]: there are $|E|$ different coupons (links) available in the boxes of a certain product (in the network), what is the probability that after buying $m$ such boxes (after $m$ explorations), one has collected $i$ different coupons (has discovered $i$ links) ?

The model is developed for Erdös-Rényi (ER) network graph models. In such a graph model, each node is connected to all other nodes with a given probability $p$. A connected graph $G(N,E)$ is considered. If this is not the case, the following discussion is valid for the giant component, i.e. a connected subgraph, of an ER graph. Another assumption is that for each link in $E$ there is at least one UST covering it. It is shown that, if the constraint of a complete topology knowledge is relaxed, a very good coverage (more than 95%) is obtained with $O(\log(|E|))$ beacons.

A tree is a graph with $n=|N|$ nodes and $n-1$ links. Therefore for each UST, $n-1$ links out the set of $|E|$ links of the ER graph are explored. The exploration process is modeled as a sequence of steps where at each step a new beacon (and thus its UST) is added to the beacon set. At the $i$-th step, $G_i'(N,E_i')$ is the reconstructed graph and $y_i=|E_i'|$ is the number of explored links. As the coupon collector, by buying $n-1$ coupons at each attempt, will eventually collect all $|E|$ coupons, the whole process stops when $y_i=|E|$.

The exploration is assumed to follow a Bernoulli process, where each link the UST explores can be undiscovered (with probability $q$) or already discovered (with probability $1-q$).

The probability to discover $k$ new links is in the $i$-th exploration can be written as:

$$P_i(k) = \binom{n-1}{k} q_i^k (1-q_i)^{n-k-1} \qquad (1)$$

Note that Eq.(1) is an approximation because $q$ changes for each explored link while in Eq.(1) it changes only after one exploration.

Since each node in an ER graph has an equal probability $p$ to have a link toward any other node and because of the utilization of the "coupon collector problem" analogy in the model, $q_i$ can be expressed as $q_i = 1-\frac{y_{i-1}}{|E|}$, that is the probability of discovering a new link when $y_{i-1}$ are already discovered until the $i-1$-th step. Now, $y_i$ is a random variable that expresses the number of known links at $i$-th step with $y_0=0$. The expected value of $y_i$ can be written as:

$$\overline{y}_i = \overline{y}_{i-1} + \overline{k}_i, \qquad (2)$$

where $\overline{k}_i = \sum_{j=0}^{j=n-1} jP_i(j) = (n-1)q_i$ represents the average number of nodes discovered at the $i$-th step.

Eq.(2) is a difference equation whose solution is:

$$\overline{y_i} = |E|\left[1-\left(1-\frac{2\alpha}{\overline{d}}\right)^i\right], \qquad (3)$$

where $\alpha=(n-1)/n \simeq 1$ and $\overline{d}$ is the average degree of $G$.

The above result can be generalized to the case of a tree with a generic number of links $n_t$ (as in the FD-TDS implementation). In this case $\overline{k}_i = n_t q_i$ and Eq. (3) requires $\alpha = n_t/n$. With Eq. (3) the number of steps that yields a certain ratio $w=\overline{y_i}/|E|$ of total coverage of the graph can be computed as $w=\left[1-\left(1-\frac{2\alpha}{\overline{d}}\right)^i\right]$.

According to the developed model, $\overline{y_i}=|E|$ is obtained only after a very large number of steps (i.e., $i\to\infty$). Thus the approximate number of explorations $i_f$ to have an "almost" complete knowledge of the graph can be computed by setting $\overline{y}_i = |E|-1$:

$$i_f = -\left\lceil \frac{log(|E|)}{log\left(1-\frac{2\alpha}{\overline{d}}\right)} \right\rceil. \qquad (4)$$

The value of $i_f$ provides also the number of beacons to be utilized to obtain an almost complete coverage of the network links.

## IV. Results

### A. Simulated explorations in network graph models

In this section the beacon discovery model developed in section III and the accuracy of Eq. (4) are evaluated in different ER topologies and other network graph models. ER topologies have been obtained through an ad-hoc Matlab generator. BRITE [19] has been used to generate Waxman topologies (a general model where also geometrical distance between

1930-529X/07/$25.00 © 2007 IEEE
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2007 proceedings.

TABLE I
SIMULATED EXPLORATION OF ER GRAPHS.

| $\overline{d}$ | $\|E\|$ | $n_f$ | $i_f$ | $\|E_{i_f}\|/\|E\|$ |
|---|---|---|---|---|
| 4 | 2020 | 15 | 12 | 99.7% |
| 6 | 3048 | 24 | 21 | 99.9% |
| 8 | 4067 | 34 | 30 | 99.95% |
| 10 | 5129 | 45 | 40 | 99.97% |
| 12 | 6086 | 56 | 49 | 99.97% |

TABLE II
SIMULATED EXPLORATION OF SF GRAPHS.

| $\overline{d}$ | $\|E\|$ | $i_f$ | $\|E_{i_f}\|/\|E\|$ |
|---|---|---|---|
| 4 | 1997 | 11 | 95.7% |
| 6 | 2994 | 20 | 94.8% |
| 8 | 3990 | 29 | 94.3% |
| 10 | 4985 | 39 | 96.3% |
| 12 | 5979 | 48 | 96.2% |
| 14 | 6972 | 58 | 97.1% |
| 16 | 7964 | 67 | 97.2% |

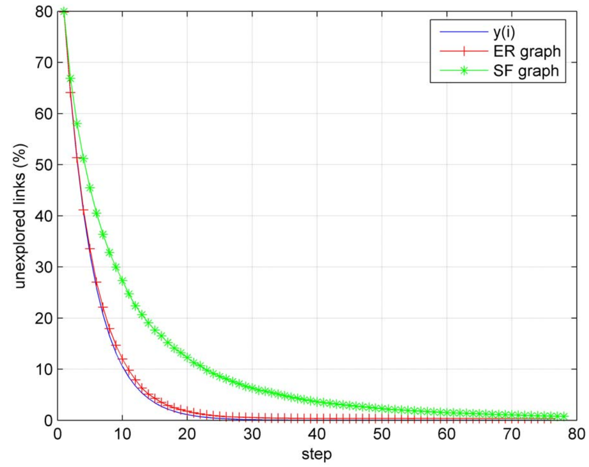


Fig. 2. Percentage of unexplored (i.e., undiscovered) links as a function of the number of steps: $y_i$ from Eq. (3) is compared with simulated results obtained on ER (red line) and SF (green) graphs.

nodes matters) and Scale-Free (SF) graphs with Barabasi-Albert model.

For each considered network $i_f$ explorations have been conducted (i.e., $i_f$ beacons have been utilized) and repeated twenty times with a random uniform choice of the beacon placement to obtain an average behaviour. For each set of simulations for a specific $i_f$ value, the percentage of discovered links, obtained from the ratio between the number of links discovered by the exploration $|E_{i_f}|$ and the number of network links $|E|$, has been computed. Moreover, in order to have a measure of the accuracy of $i_f$, the average number $n_f$ of steps needed for a complete coverage has been computed.

As reported in Tab. I and Fig. 2 the percentage of discovered/undiscovered links obtained through simulation on ER graphs and the one obtained by means of Eq. (3) fit very well. In addition the number of steps $n_f$ that yield a complete knowledge is close to $i_f$. As stated in [12], the percentage of links that remains unexplored after $i$ explorations has a heavy-tailed yet regular distribution. It is noteworthy that, as the degree grows, the knowledge of $G$ (i.e., the topology $G'_{i_f}$), given by $i_f$ steps, becomes more accurate even if the difference between $n_f$ and $i_f$ (in Tab. I) increases. Similar results have been obtained for Waxman topologies.

Tab. II and Fig. 2 show that the beacon discovery model does not directly apply to SF graphs. This is mainly due to the assumption used to evaluate $q_i$. Indeed, in power-law graphs, the probability that a link between node $i$ and node $j$ exists is proportional to the product of the nodes degrees ($p_{i,j} \propto d_i d_j$).

Nonetheless some hints on how to modify the beacon discovery model for SF graphs can be obtained from the simulations. Matlab *fit* function has been applied to the curve of unexplored link percentage as a function of the number steps. For ER and Waxman graphs the same exponential function is obtained as Eq. (3), as a further confirmation of the good approximation.

On the other hand, if the *fit* of the SF graph curve is performed, a dual exponential function (with parameters $a$, $b$, $c$ and $d$) is obtained, such as:

$$ae^{bi} + ce^{di}, \tag{5}$$

where $i$ is the discovery step and the first exponential ($ae^{bi}$) is again the same as Eq. (3). Since this behaviour is replicated on all generated SF graphs, the aforementioned result shows that the proposed beacon discovery model has some general utility.

Indeed, if the *fit* function is considered as the solution of a difference equation, as Eq. (3) is the solution of Eq. (2), a noteworthy interpretation arises. Eq. (5) can be seen as the sum of the homogeneous and the particular solution of a difference equation, where the first is given by the solution of Eq. (2) for ER graphs and the latter depends on the particular explored graph model.

According to this interpretation and to the observations in [17], Eq. (3) describes the exploration of a $0K$*-graph* (that is an uncorrelated version) of the network. Then the additional exponential terms (i.e., the particular solution of the main difference equation) provide the refinement needed to describe exploration of $dK$*-graphs* with higher "d".

*B. Simulated and experimental explorations in real networks*

In this section the evaluation of the number of links discovered by the FD-TDS in real networks is presented as a function of the number of utilized beacons. It must be noticed that in FD-TDS each beacon does not obtain a UST but it discovers a tree whose links belong to the shortest paths between the beacon and the other active beacons. Thus $\alpha = n_t/n$.

The considered performance evaluation parameter is the ratio $w$ between the number of discovered links and the number of network links:

$$w = \frac{discovered\ links}{total\ links}. \tag{6}$$

The parameter $w$ is evaluated for several networks with different average nodal degree.

A first set of networks consists of all the possible 5-node topologies in which each node has at least degree 2. The initial topology is the full mesh 5-node network; all the other considered topologies are obtained from the full mesh 5-node network by removing 1 link (10 topologies), 2 links (45 topologies), 3 links (100 topologies), 4 links (175 topologies) and 5 links at a time.

For each considered network 4 beacons have been utilized. The beacons are placed so that the average distance among them is maximized; in case of multiple equivalent placements, beacons are placed in the nodes with minimal nodal degree.

The results depicted in Fig. 3 have been obtained from four different network exploration methods.

The *spanning tree (UST) model* and the *beacon-beacon (BB) model* results have been obtained by applying the model developed in section III with $w = \left[1 - \left(1 - \frac{2\alpha}{d}\right)^i\right]$, where i=4 and $\alpha$ is $\alpha = (n-1)/n$ and $\alpha = \overline{n_t}/n$, respectively.

The *average* $w$ results have been obtained by simulating the FD-TDS exploration in each network. The simulations assume bidirectional links and that the shortest paths between source and destination span the same links in both directions. If, in a network, alternative shortest paths are available between beacons all the possible coverages are evaluated and results are averaged. Results are then averaged among all the networks with the same average nodal degree.

The *experimental traceroute* $w$ results have been obtained by building the networks through the interconnection of 5 commercial routers equipped with Fast Ethernet interfaces and running OSPF-TE as routing protocol. Beacause of the invariance on the number of discovered links for some networks with the same average nodal degree just a subset of all the generated topologies has been considered. In each experiment, each beacon runs a network probe (i.e., pathchar) towards all the other beacons. In this case the assumption that the same shortest path is used between two beacons in both the directions is not true anymore. Indeed, due to the implementation of OSPF, the choice of alternative shortest paths is random in both directions.

Results depicted in Fig. 3 show that the value of $w$ obtained by the beacon-beacon model and by the spanning tree model are close. However they overestimate $w$ with respect to the average $w$ method, especially when the nodal degree increases. This is due to the fact that the model is developed for ER graphs, while the average $w$ method is based on an exhaustive search. The non-monotonic behaviour of the experimental curve is because link coverage is function of the random choice between alternative shortest paths made by the routing protocol. Moreover Fig. 3 shows that, if the nodal degree increases, the probability to have different shortest paths forward and backward increases. In these conditions the beacon placement might additionally impact the obtained results. However the trend of $w$ obtained by the experimental traceroute method is similar to the one obtained by the other methods.

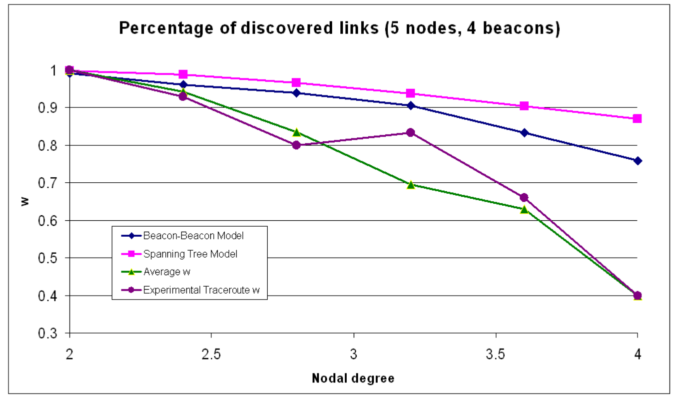


Fig. 3. Coverage experimental results.

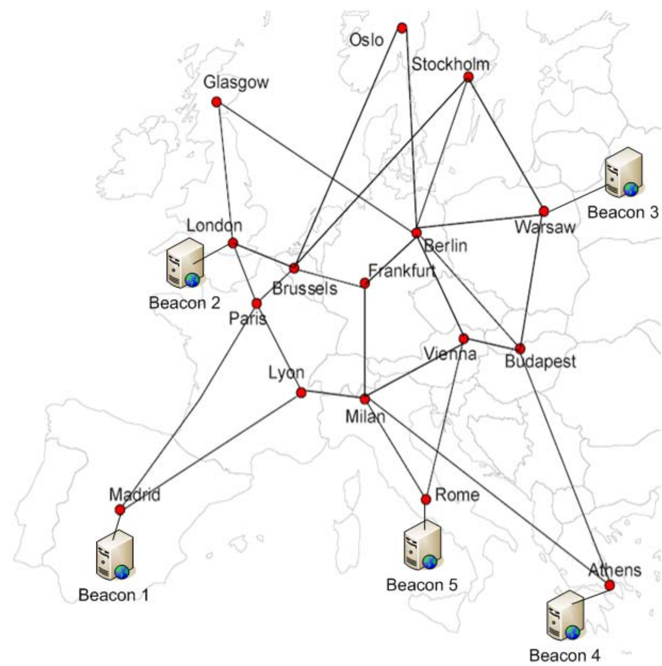


Fig. 4. Pan-European Network.

Finally, Tab. III shows the probability of not discovering one node as a function of the average nodal degree. This represents a particular event, which occurs when all the bidirectional links attached to one node are not discovered. Results show that if the average nodal degree of the network is not too high (realistic scenario) the probability of not discovering one node is quite low (under 30 %).

In the end, the UST, the BB, and the average $w$ models have been applied to the Pan-European network depicted in Fig. 4, whose average nodel degree is 3.375. The results have been obtained by placing four and five beacons as in the 5-node networks. Tab. IV shows that the BB model is capable of well predicting the undiscovered link percentage (within at most 10% of the average $w$ method) while the UST model understimates it.

TABLE III
UNDISCOVERED NODE PROBABILITY

| Nodal Degree | 1 node missed prob. |
| --- | --- |
| 2 | 0 |
| 2.4 | 0.014 |
| 2.8 | 0.214 |
| 3.2 | 0.288 |
| 3.6 | 0.666 |
| 4 | 1 |

TABLE IV
UNDISCOVERED LINKS PERCENTAGE IN THE PAN EUROPEAN NETWORK

| — | UST model | BB model | average w |
| --- | --- | --- | --- |
| 4 beacons | 0.039 | 0.445 | 0.466 |
| 5 beacons | 0.017 | 0.288 | 0.326 |

## V. CONCLUSION

In this study an issue arising in implementing a fully distributed tool for network topology discovery and performance information has been considered: the beacon number problem. The beacon number problem is the problem of computing the number of beacons (i.e., the number of measurement points) that provides a complete network topology and performance information.

A model has been developed for the beacon number problem for ER graphs and it has been applied to both network graph model and real network explorations. Numerical results show a good agreement between model and simulations for ER graph while for SF graphs the model overestimates the percentage of discovered links. Experimental results show that for networks with realistic average nodal degree the overestimate is less than 20%.

## ACKNOWLEDGMENT

This work has been partially supported by Italy-Tunisia FIRB project "Software and Communication Platforms for High-Performance Collaborative Grid" (RBIN043TKY).

## REFERENCES

[1] D. Emma, A. Pescape, and G. Ventre, "Discovering topologies at router level," in *5th IEEE International Workshop on IP Operations & Management (IPOM 2005)*, Barcelona (Spain), October 2005, pp. 118–129.

[2] M. den Burger, T. Kielmann, and H. Bal, "TOPOMON: A monitoring tool for grid network topology," in *International Conference on Computational Science (ICCS 2002)*, vol. 2330, Amsterdam, April 21-24 2002, pp. 558–567, http://citeseer.ist.psu.edu/denburger02topomon.html.

[3] B. Huffaker, D. Plummer, D. Moore, and K. Claffy, "Topology discovery by active probing," in *Applications and the Internet (SAINT) Workshops, 2002. Proceedings*, January-February 2002, pp. 90–96.

[4] Y. Breitbart and M. Garofalakis, "Topology discovery in heterogeneous IP networks: the NetInventory system," in *IEEE/ACM Transactions on Networking*, vol. 12, June 2004.

[5] I. Legrand, H. Newman, R. Voicu, and C. Cirstoiu, "MonALISA: An Agent based, Dynamic Service System to Monitor, Control and Optimize Grid based Applications," *CHEP 2004, Interlaken, Switzerland*, September 2004.

[6] G. Jin, G. Yang, B. Crowley, and D. Agarwal, "Network Characterization Service (NCS)," in *High Performance Distributed Computing, 2001 Proceedings*, 2001.

[7] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the internet topology," in *SIGCOMM Comp. Comm. rev. 29, 251*, 1999.

[8] T. Petermann and P. De Los Rios, "Exploration of Scale-Free Networks," in *The European Physical Journal B*.

[9] A. Lakhina, J. Byers, M. Crovella, and P. Xie, "Sampling Biases in IP Topology Measurements," in *INFOCOM 2003*.

[10] D. Achlioptas, A. Clauset, D. Kempe, and C. Moore, "On the Bias of Traceroute Sampling: or, Power-law Degree Distributions in Regular Graph," in *Annual ACM Symposium on Theory of Computing, Baltimore, USA, 2005*.

[11] A. Clauset and C. Moore, "Accuracy and Scaling Phenomena in Internet Mapping," in *Physical Review Letter 94, January 2005, http://arxiv.org/abs/cond-mat/0410059*.

[12] J. Guillaume, L. Latapy, and M. Magoni, "Relevance of massively distributed explorations of the internet: Simulation results," in *Proceedings of IEEE Infocom 2005*, 2005.

[13] P. Barford, A. Bestavros, J. Byers, and M. Crovella, "On the Marginal Utility of Network Topology Measurements," in *ACM SIGCOMM 2001*.

[14] J. Horton and A. Lopéz-Ortiz, "On the number of distributed measurement points for network tomography," in *ACM SIGCOMM 2003*.

[15] R. Kumar and J. Kaur, "Efficient Beacon Placement for Network Tomography," in *4th ACM SIGCOMM 2004*.

[16] Y. Azzana, F. Guillemin, and P. Robert, "A stochastic model for the topology discovery of tree networks," in *ICC, 2005*.

[17] P. Mahadevan, D. Krioukov, K. Fall, and A. Vahdat, "Systematic topology analysis and generation using degree correlations," in *ACM SIGCOMM 2006*.

[18] "http://www.caida.org/tools/utilities/others/pathchar/."

[19] A. Medina, A. Lakhina, I. Matta, and J. Byers, "BRITE: An approach to Universal Topology Generation," in *International Workshop on Modeling, Analysis and Simulation of Computer and Telecommunications Systems-MASCOTS'01*, Cincinnati,Ohio, August 2001.