Distributed and Scalable Content/Service Placement

Material:

I. Stavrakakis, “Some distributed approaches to the service facility location problem in dynamic and complex networks”, Handbook of Optimization in Complex Networks: Vol. 1 (Theory and Applications) and Vol. 2 (Communication and Social Networks), edited by Profs. My T. Thai and Panos Pardalos, Springer Publisher, fall 2011.

K. Oikonomou, I. Stavrakakis, “Scalable Service Migration in Autonomic Network Environments", special issue of ***IEEE Journal on Selected Areas in Communications*** (JSAC) on “Recent Advances in Autonomic Communications”, Vol. 28, No. 1, pp. 84-94, Jan. 2010.
http://cgi.di.uoa.gr/~istavrak/publications/2010JSAC.os.pdf

G. Smaragdakis, N. Laoutaris, K. Oikonomou, I. Stavrakakis, and A. Bestavros, “Distributed Server Migration for Scalable Internet Service Deployment,” ***IEEE/ACM Transactions on Networking***, Vol. 22, No. 3, June 2014.
http://cgi.di.uoa.gr/~istavrak/publications/2010TNET.slosb.pdf

P. Pandazopoulos, M. Karaliopoulos, I. Stavrakakis, "Distributed Placement of Autonomic Internet Services", IEEE Transactions on Parallel and Distributed Systems, Vol.25, No.7, pp.1702-12, July 2014**.**
http://cgi.di.uoa.gr/~istavrak/publications/2013TPDS.pks.pdf

Main objective - Motivation

Objective:

- bring service provision points close to demand, in order to minimize communication resource consumption
- enhance the Quality of Service (QoS) of the provided service

Motivation:

- Proliferation of user-generated content/ massive content distributions / virtualized services
- Can be hosted in end-user machines / edge clouds / datacenters / etc.

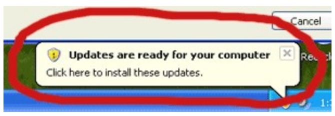
Need to move from the traditional problem of placing relatively few big services/content to one of the few (powerful) potential service provider facilities (big network elements) …..

… to a problem of placing numerous services/content to one of the numerous potential "service providers"

The environment

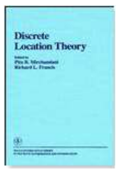
- We have:
  - a network
  - a demand
- We want:
  - the location of service facilities
- The solution:
  - [Uncapacitated *k*-median] – UKM
  - Uncapacitated Facility Location (# of facilities tbd)

a request



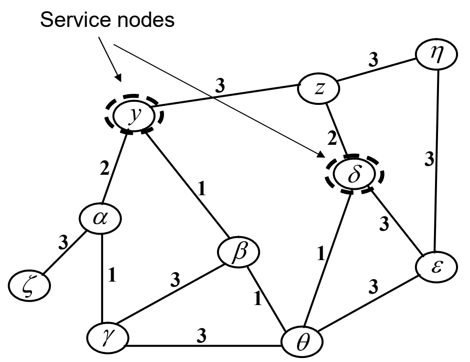
a server (software)

a really nice read



The Model

- Assume a network of the form of a graph consisting of nodes and links among them
- Nodes
  - Have service demands to be satisfied by services located at some other nodes (or even itself)
  - Are capable of "hosting" the service
- Links
  - Have been associated with a certain **weight** that corresponds to the **cost** for a certain packet to travel over this link
- Packets are traveling towards their destination over the branches of a **minimum spanning tree** defined by the employed routing protocol



- **Objective**: determine the appropriate service nodes in order to **minimize** the overall cost

The classical facility location problem

<u>INPUT</u>

V : set of nodes

w_n : demand generated by node n

d(x_j,n) : distance between nodes x_j and *n*

<u>OUTPUT</u>

F : placement

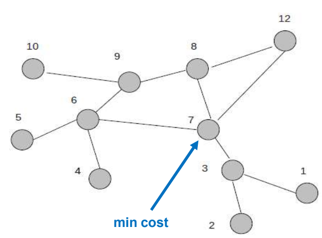
k-median problem : open up to *k* facilities so as to minimize the total service cost:

$$Cost(\mathcal{F}) = \sum_{n \in \mathcal{V}} w_n \cdot min_{x_j \in \mathcal{F}}\{d(x_j, n)\}$$

1-median: minimize the access cost of a service that is:

$$Cost(k) = \sum_{n \in \mathcal{V}} w(n) \cdot d(k, n).$$

when the service is located at node *k*

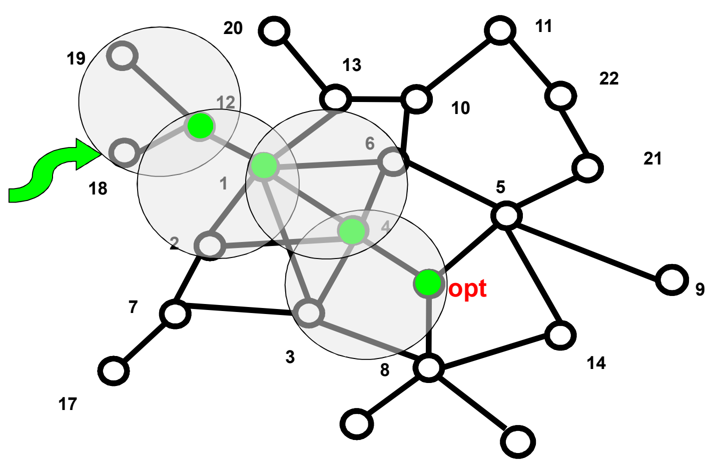


A Difficult Problem!

- The aforementioned problem corresponds to the **k-median problem** for the case that there are **k** service nodes in the net
- This is a well studied optimization problem that requires **global information** in order to be solved
- It has been shown by Kariv and Hakimi that it is **NP-hard** for the general case (directed graph)
  O. Kariv and S. L. Hakimi, "An algorithmic approach to network location problems. ii: The p-medians", SIAM Journal on Applied Mathematics, vol. 37, no. 3, pp. 539–560, 1979.
- For the case of an undirected tree an algorithm provided by Tamir allows for **O(kN²)** → $O(kN^2)$, where **N** corresponds to the network size
  A. Tamir, "An o(pn2) algorithm for the p-median and related problems on tree graphs", Operations Research Letters, vol. 19, no. 2, pp. 59–64, 1996.
- Consequently, even for the case of **k=1** and the undirected tree topology, the existing approaches offer a complexity of $O(N^2)$, that is rather high for dynamic modern environments like ad hoc networks (e.g. mobility issues)
- Several near-optimal, heuristic approaches have been proposed to reduce complexity at the expense of accuracy that can be categorized as either centralized or distributed. However, all of them require **global knowledge of the network topology and demands**
  G. Wittenburg and J. Schiller, "A survey of current directions in service placement in mobile ad-hoc networks", in Pervasive Computing and Communications, 2008. PerCom 2008. Sixth Annual IEEE International Conference on, IEEE, 2008, pp. 548–553.
- **Our objective**: to find the **optimal service location** using **local demand and topology information for large scale dynamic environments**

Distributed and scalable approaches to facility location
Heuristic local-search like approaches

The main idea: Service migrates towards the optimum host(opt) in a finite number of steps



Approach 1: Neighbor-hopping service migration (strictly local /0-hop information)

Approach 1 (Overview)
Neighbor-hopping service migration (strictly local /0-hop information)

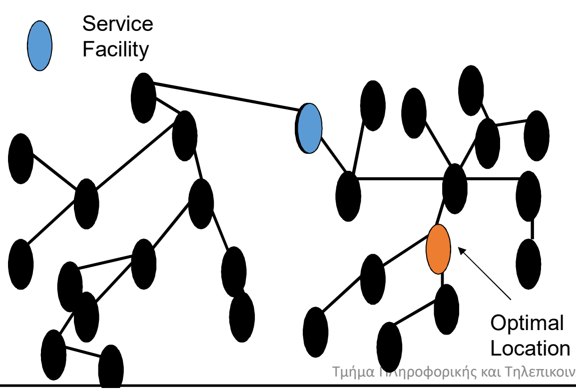
Exploits strictly local information

Bases service facility movement on the relative difference of aggregate service demands

Monotonic cost reduction for any (connected) topology
- Response to dynamic changes

Optimum is guaranteed
- for one service facility and
- in unique shortest path tree topologies (e.g., trees)

Multi-facility extension
- Monotonic cost reduction
- Optimality not guaranteed



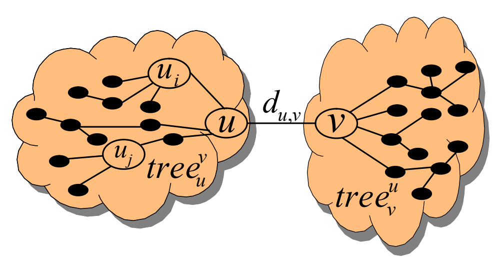
K. Oikonomou, I. Stavrakakis, "Scalable Service Migration in Autonomic Network Environments," IEEE JSAC, Vol. 28, No. 1, January 2010, pp. 84-94

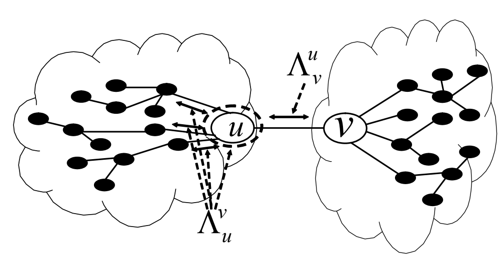
A comment on Tree Topology?

- This is the topology used frequently before considering more general network topologies
- Any network having different link weights has a unique **minimum spanning tree.** Thus, this would be the case in widely encountered networks with different link weights.
- Three different categories of (undirected) topologies may be defined
  - A tree topology
  - A general topology having one minimum spanning tree
  - A general topology with multiple minimum spanning trees
- Based on the above, the results provided later for the tree topology case **may be applied** to general topologies

Some Useful Definitions

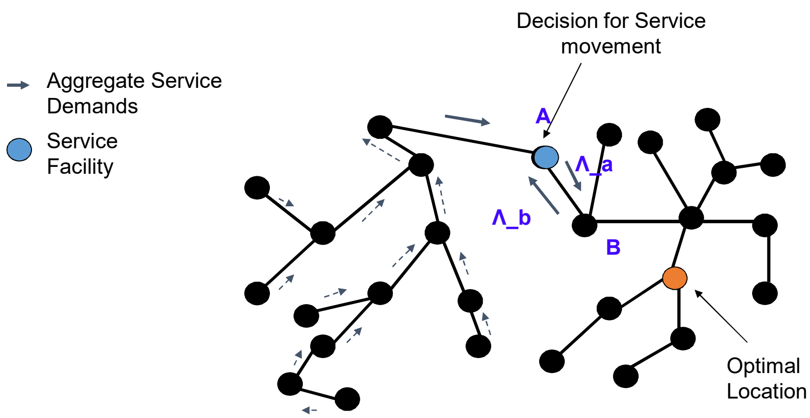
- **$tree^u_v$** : the set of nodes that belong to the **subtree** for which node **v** is the root and node **u** is an ancestor of node **v**
  - the tree considered is the minimum spanning tree when **v** is the root
- **$d_{u,v}$**: the distance (sum of the link weights over a shortest path) between node **u** and node **v**
- **λ_u**: the average service demands of node **u**
- **Λ^u_v**: the average **aggregate** service demands for the set of nodes **$tree^u_v$**
- **C_u**: the overall average cost when the service is located at node **u**
  - $C_u=\sum_{v\in V} \lambda_v d_{u,v}$



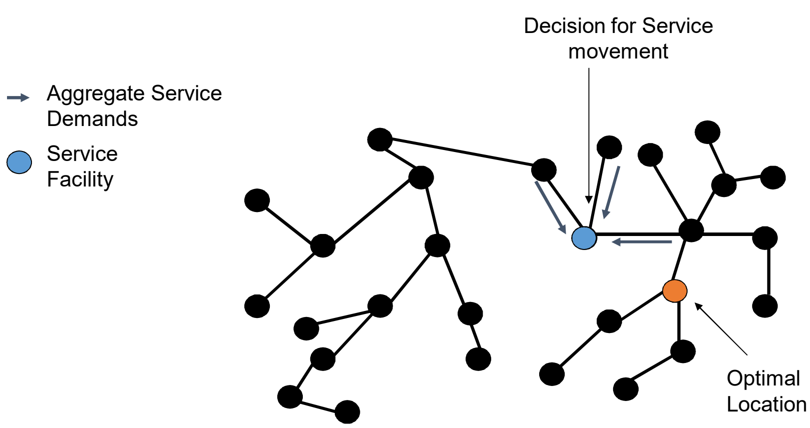


Basic Idea of Service Migration

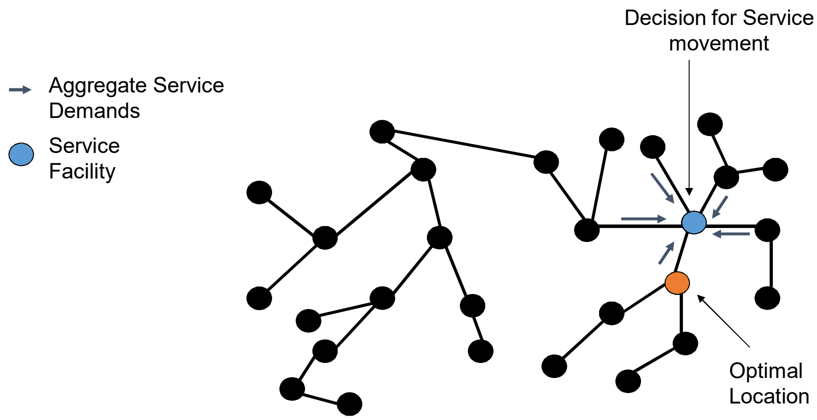
- Use strictly locally obtained topology info (links to neighbours) and service demand info (aggregate service demands over these links).
- Decide to move from A to B iff Λ_b > Λ_a



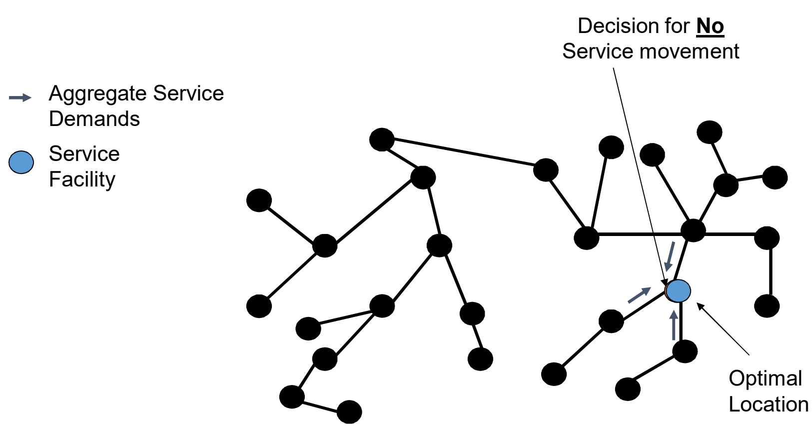
Basic Idea of Service Migration



Basic Idea of Service Migration



Basic Idea of Service Migration

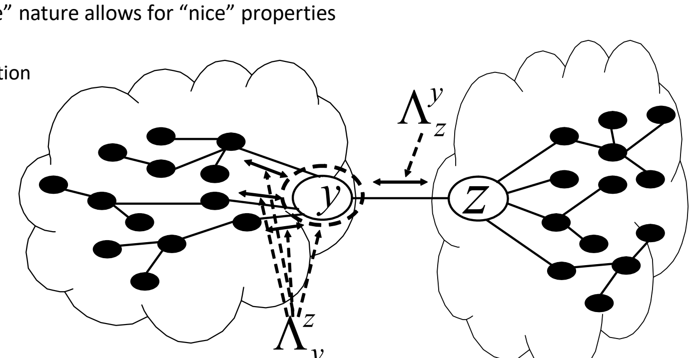


One Step Beyond

- Assume two neighbor nodes of the network **y** and **z**
- The corresponding costs are denoted as C_y and C_z respectively
  - It is interesting to identify those factors influencing difference C_z-C_y
- As it is shown : $C_z-C_y=(\Lambda^z_y-\Lambda^y_z)d_{z,y}$
  - The difference depends on the difference of the corresponding aggregate service demands $\Lambda^z_y-\Lambda^y_z$
  - The difference depends on the distance among the neighbor nodes $d_{z,y}$

A Useful Observation

- Λ^z_y and Λ^y_z can be available at node **y** (the service node)
  - e.g., using statistical observations
- It is reasonable to assume that the node that "hosts" the service (i.e., node **y**) is capable of monitoring its incoming interfaces and be aware about the aggregate service demands through them
  - Note that the "aggregate" nature allows for "nice" properties
    - Normal distribution
    - Better/faster estimation
    - …



Suitable Next Hop Migration

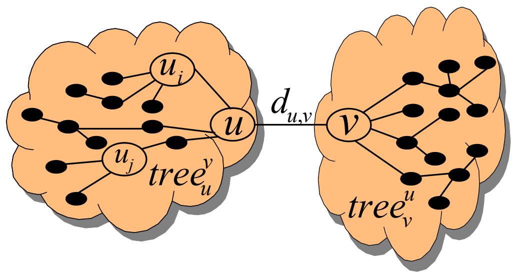
- Suppose that the service is located at node **y**
- According to equation: $C_z - C_y = (\Lambda^z_y - \Lambda^y_z) d_{z,y}$
  - $C_z < C_y$ iff $\Lambda^z_y < \Lambda^y_z$, since $d_{z,y} > 0$ **!!**
- Given the fact that Λ^z_y, Λ^y_z, are available at the service node **y**, node **y** can decide on the service migration towards its neighbor node **z** based **exclusively** on the sign of the difference $\Lambda^z_y - \Lambda^y_z$
- The previous statement corresponds to the **migration policy** proposed here
- The first immediate observation is that **cost reduction is possible** based on information that is locally available at the node
  - Note that so far there was no assumption regarding the topology. **This migration policy is capable for cost reduction for any network topology**. The tree topology will be used next in order to show that the service eventually reaches the optimal position

Some Analytical Results

- **Lemma**: If the sequence of migration steps brings the service to a location that is the optimal one, then it will stay there and will not move away in the next steps.
  - *Sketch of the Proof: A service movement takes place depending on the aggregate service demands difference which also depends on the cost difference. However, the cost is minimized at the optimal location and therefore, a service cannot be moved any longer according to the proposed migration policy*
- **Theorem**: The service eventually arrives at the optimal service location when it is moved according to the proposed migration policy
  - *Sketch of the Proof: It is proved that the service moves along a path and the corresponding cost gradually decreases. The end of this monotonically cost decreasing path is the optimal service location*

Problem with non-unique SPT topologies

- Previous migration rule (criterion A) works perfectly for topologies with a unique SPT (e.g., tree topologies) : it moves the position along and till the end of a **monotonically cost decreasing path** (**mcdp)**, reaching the optimal position
- For non-tree topologies, it can stop prematurely along a mcdp and away from the optimal location, by ignoring shorter paths that would emerge had the service moved.



Alternative criteria have been developed requiring extended local information and greater overhead.

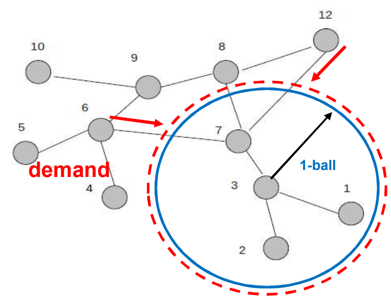
Approach2: The R-ball heuristic

Approach 2 Overview
The R-ball heuristic (local /R-hop information)

- ❑solves UKM (and UFL) through **iterative solutions of smaller** optimization problems using **limited-scope** (non-global) demand and topology info**
- ❑**within a limited neighborhood of R-hops around current facility

  Demand generated by outer nodes is *mapped* to the nodes at the outer shell of the R-hop neighborhood

$$Cost(\mathcal{F}) = \sum_{n \in \mathcal{V}} w_n \cdot min_{x_j \in \mathcal{F}}\{d(x_j, n)\}$$



G. Smaragdakis, N. Laoutaris, K. Oikonomou, I. Stavrakakis, and A. Bestavros, "Distributed Server Migration for Scalable Internet Service Deployment," ***IEEE/ACM Transactions on Networking***, Vol. 22, No. 3, June 2014

UKM and UFL

- Uncapacited K-median (UKM): Given a set of points *V* with pair-wise distance function *d* and service demands $s(v_j)$, $\forall\ v_j \in V$, select up to *k* points to act as medians (facilities) so as to minimize the service cost *C(V,s,k)*:

$$C(V,s,k) = \sum_{\forall v_j \in V} s(v_j) \cdot d(v_j, m(v_j))$$

  where $m(v_j)$ is the median that is closer to v_j.

- Uncapacited Facility Location (UFL): Given a set of points *V* with pair-wise distance function *d*, service demands $s(v_j)$, $\forall\ v_j \in V$, and facility costs $f(v_j)$, $\forall\, v_j \in V$, select a subset of points *F* to act as facilities so as to minimize:

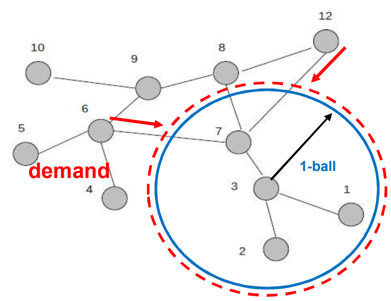
$$C(V,s,f) = \sum_{\forall v_j \in F} f(v_j) + \sum_{\forall v_j \in V} s(v_j) \cdot d(v_j, m(v_j))$$

The R-ball heuristic (local /R-hop information)

- solves UKM (and UFL) through **iterative solutions of smaller** optimization problems using **limited-scope** (non-global) demand and topology info**
- **within a limited neighborhood of R-hops around current facility

  Demand generated by outer nodes is *mapped* to the nodes at the outer shell of the R-hop neighborhood

$$Cost(\mathcal{F}) = \sum_{n \in \mathcal{V}} w_n \cdot min_{x_j \in \mathcal{F}}\{d(x_j, n)\}$$

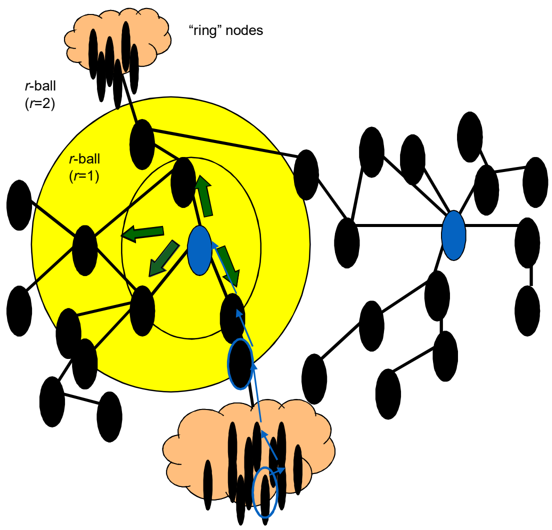


G. Smaragdakis, N. Laoutaris, K. Oikonomou, I. Stavrakakis, and A. Bestavros, "Distributed Server Migration for Scalable Internet Service Deployment," ***IEEE/ACM Transactions on Networking***, Vol. 22, No. 3, June 2014

Common framework for distributed UKM and UFL

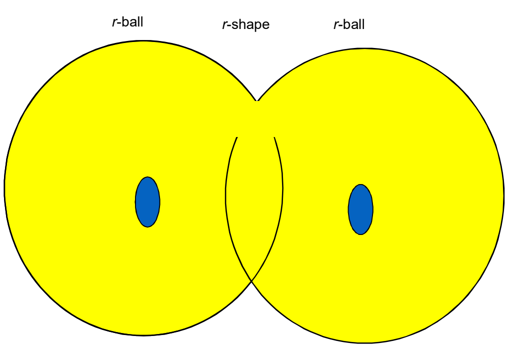
- Initialization: select an initial set of nodes to be the facilities
- Iterative improvement: select an existing facility and "process" it using <u>local information</u> only
  - change its location (in the case of UKM)
  - change its location and/or merge it with other facilities or spawn additional copies of it (in the case of UFL)
  - continue with the next facility in round-robin manner
- Stopping condition: when "processing" yields no improvement for any facility

Processing a facility



- const # facilities → **1-median in *r*-ball**
- var # facilities → **UFL in *r*-ball**
- **but there is a PROBLEM**
  - **nodes outside the *r*-ball**
  - **… are totally neglected**
- **and a SOLUTION to it**
  - **map ring demand on the "skin" of the *r*-ball**

Intersecting r-balls merge into r-shapes



- when 2 or more *r*-balls intersect we merge them
- if *J* facilities in the *r*-shape
  - *J*-median (const facilities)
  - UFL (var facilities)
- *r*-shape provides for a way to reduce the # facilities if needed
- we put a restriction on the max-size of *r*-shapes

Selecting the radius r

Small radius:

\+ limited local information for the *r*-balls (scalability)

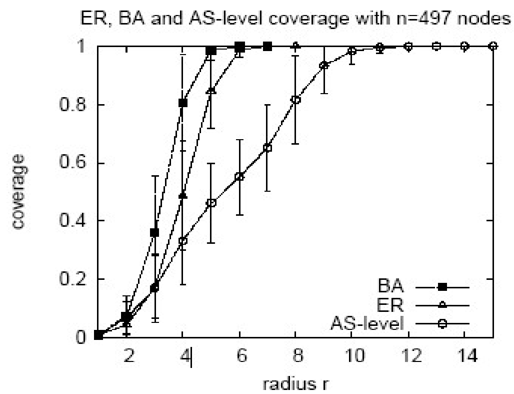
— performance penalty (easier to run into bad local minima)

Since most networks are small-worlds → we keep *r* small (1≤*r*≤3)

ER Erdos-Reni (random) graphs

BA (Barabasi) graphs (with hubs)

Radious of 1-2 is large enough, including a good number of nodes.



Case Study: The AS-level Topology

- 497 peer AS's in the core of the Internet (Subramanian et al. '02)
- load $s(v_j)$= # AS's with costumer-provider relationship to v_j
- distance $d(v_i,v_j)$= # intermediate AS's from $v_i \rightarrow v_j$

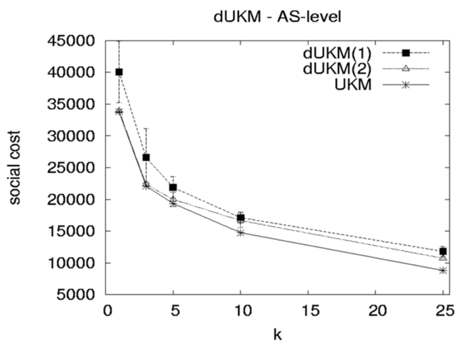
- centralized vs distributed
  - UKM vs dUKM(*r*)
  - UFL vs dUFL(*r*)

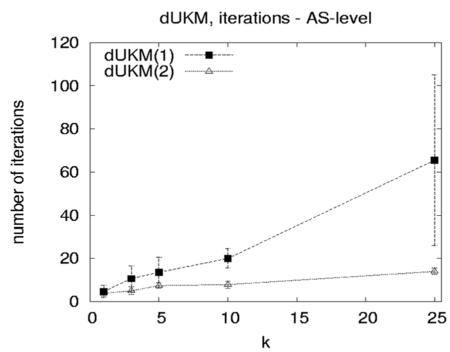
- social cost and # iterations

$$C(V,s,k) = \sum_{\forall v_j \in V} s(v_j) \cdot d(v_j, m(v_j))$$

$$C(V,s,f) = \sum_{\forall v_j \in F} f(v_j) + \sum_{\forall v_j \in V} s(v_j) \cdot d(v_j, m(v_j))$$

Placing k servers on the AS-level map





1% 3% 5%

#facilities: **% of nodes**

1% 3% 5%

#facilities: **% of nodes**

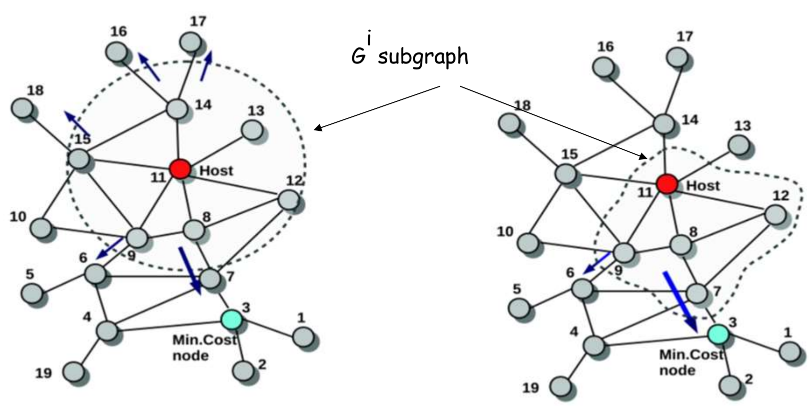
Wrap up

- Placement of service facility can be casted as a discrete location problem
- Existing centralized solutions are not practical
- Instead → multiple local re-optimizations
  - exact info for a limited neighborhood of radius *r*
  - approximate info for the surrounding "ring"
- Good approximation (experimental) even for very small radius

Approach3: centrality-driven distributed service migration

Approach 3 Overview
cDSMA (centrality-driven distributed service migration approach): a more informed search

R-ball heuristic cDSMA



P. Pandazopoulos, M. Karaliopoulos, I. Stavrakakis, "Distributed Placement of Autonomic Internet Services", ***IEEE Transactions on Parallel and Distributed Systems***, Vol.25, No.7, pp.1702-12, July 2014.

cDSMA: how it captures the topology factor

a measure of the importance of node's *u* social position : lies on paths linking others

Betweenness Centrality (*u*): sums the portions of all pairs' shortest paths in G that pass through node *u*

$$BC(u) = \sum_{s=1}^{|V|} \sum_{t=1}^{s-1} \frac{\sigma_{st}(u)}{\sigma_{st}}$$

Conditional Betweenness Centrality (*u,t)* : portion of all shortest paths **towards target node t** in G, that pass through node *u*

a measure of the relevant social standing of node *with respect to target* t

$$CBC(u;t) = \sum_{s \in V, u \neq t} \frac{\sigma_{st}(u)}{\sigma_{st}}$$

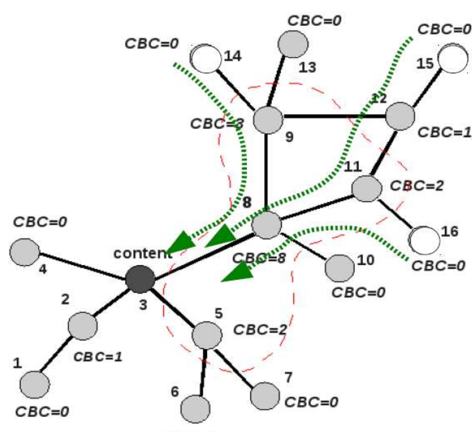
cDSMA: how it captures the demand factor

a high number of shortest paths through the node *u* (e.g. node 8) does not necessarily mean that equally high demand stems from their sources!

- *weighted* conditional BC:

$$wCBC(u;t) = \sum_{s \in V, u \neq t} w(s) \cdot \frac{\sigma_{st}(u)}{\sigma_{st}}.$$

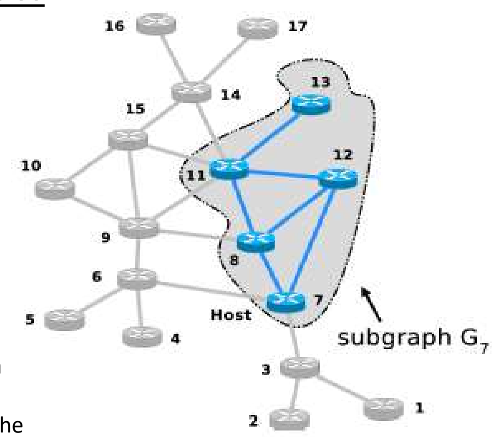
- wCBC assesses to what extent a node can serve as demand concentrator towards a given service location
  - The top a% wCBC-valued nodes are included in the 1-median subgraph



cDSMA: how it captures the demand factor

Projecting the world outside on the selected nodes

wCBC metric eases the demand mapping of the $G \backslash G^i$ nodes (world outside), on the selected G^i ones



- nodes in G^i exhibit an effective demand:

$$w_{eff}(n;Host) = w(n) + w_{map}(n;Host)$$

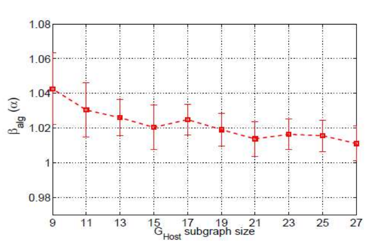
- $w_{map} \neq 0$ only for the outer nodes of the 1-median subgraph
  - demand of node z ϵ $G \backslash G^i$ is "credited" *only to the first* G^i nodes encountered on each shortest path from z towards the host

Some Results (I)

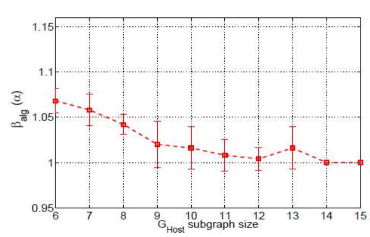
Synthetic topologies (100 nodes)

- Demand model : Zipf(s) distribution
- Performance metric : normalized access cost

$$\beta_{alg}(\alpha; G, \overline{w}) = E[\ \frac{C_{alg}(\alpha; G, \overline{w})}{C_{opt}(\alpha; G, \overline{w})}\ ]$$



a. B-A (s=0)



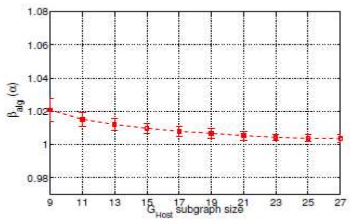
b. 10x10 Grid (s=0)

Some Results (II)

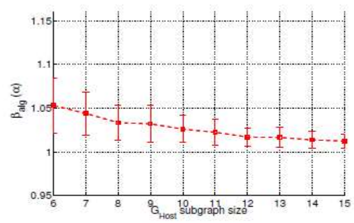
Synthetic topologies (100 nodes)

- Demand model : Zipf(s) distribution
- Performance metric : normalized access cost

$$\beta_{alg}(\alpha; G, \overline{w}) = E[\frac{C_{alg}(\alpha; G, \overline{w})}{C_{opt}(\alpha; G, \overline{w})}]$$



c. B-A (s=1)



d. 10x10 Grid (s=1)

Some Results (III)

Real world topologies

- Datasets correspond to different snapshots of 9 ISPs collected by mrinfo multicast tool*
- Less than 6% of the total network nodes suffices to yield placements of cost 2.5% close to the optimal

ISP	Dataset id/AS#	mCC nodes	Diameter	<Degree>	s=0 $< h_m >$	s=0 $\lceil \lvert G_{Host} \rvert \rceil$	s=1 $< h_m >$	s=1 $\lceil \lvert G_{Host} \rvert \rceil$	s=2 $< h_m >$	s=2 $\lceil \lvert G_{Host} \rvert \rceil$
type: **Tier-1**										
Global Crossing	36/3549	76	10	3.71	1.00±0.23	6	1.63±0.35	3	3.32±0.78	2
-//-	35/3549	100	9	3.78	1.30±0.34	7	1.26±0.16	6	1.45±0.22	4
NTTC-Gin	33/2914	180	11	3.53	1.0±0.0	18	1.11±0.13	10	1.08±0.09	8
Sprint	23/1239	184	13	3.06	1.40±0.36	8	1.22±0.15	7	1.95±0.31	4
-//-	21/1239	216	12	3.07	1.40±0.41	7	1.42±0.19	6	1.50±0.12	5
Level-3	27/3356	339	24	3.98	2.23±0.58	4	3.15±0.24	3	2.41±0.40	5
-//-	13/3356	378	25	4.49	2.27±0.59	4	2.48±0.37	4	2.22±0.34	6
Sprint	20/1239	528	16	3.13	1.40±0.62	11	1.27±0.21	21	1.09±0.11	24
type: **Transit**										
TDC	52/3292	72	9	3.28	1.20±0.29	5	1.09±0.12	5	1.46±0.28	4
DFN-IPX-Win	41/680	253	14	2.62	1.40±0.36	7	1.35±0.23	6	1.49±0.19	6
JanetUK	40/786	336	14	2.69	1.23±0.31	11	1.13±0.08	9	1.30±0.12	6
Iunet	39/1267	711	13	3.45	1.0±0.0	11	1.03±0.06	43	0.99±0.01	9

* J.-J. Pansiot, P. Mérindol, B. Donnet, and O. Bonaventure, "Extracting intra-domain topology from mrinfo probing," in Proc. Passive and Active Measurement Conference (PAM), April 2010.

Approach 1 (Overview)
Neighbor-hopping service migration (strictly local /0-hop information)

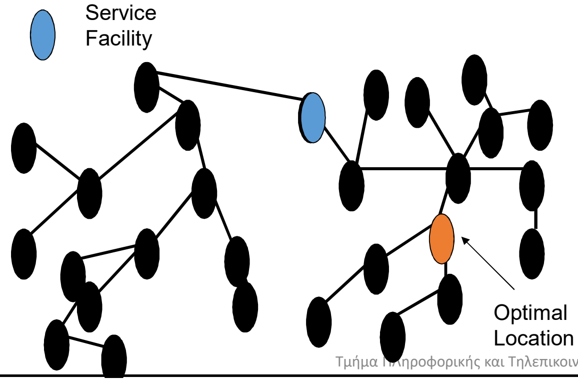
Exploits strictly local information

Bases service facility movement on the relative difference of aggregate service demands

Monotonic cost reduction for any (connected) topology
Response to dynamic changes

Optimum is guaranteed
for one service facility and
in unique shortest path tree topologies (e.g., trees)

Multi-facility extension
Monotonic cost reduction
Optimality not guaranteed

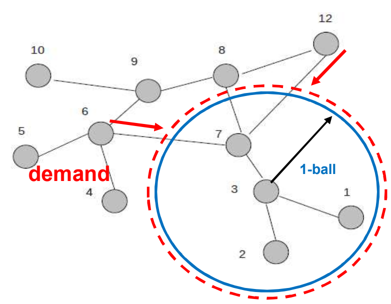


K. Oikonomou, I. Stavrakakis, "Scalable Service Migration in Autonomic Network Environments," IEEE JSAC, Vol. 28, No. 1, January 2010, pp. 84-94

Approach 2 Overview
The R-ball heuristic (local /R-hop information)

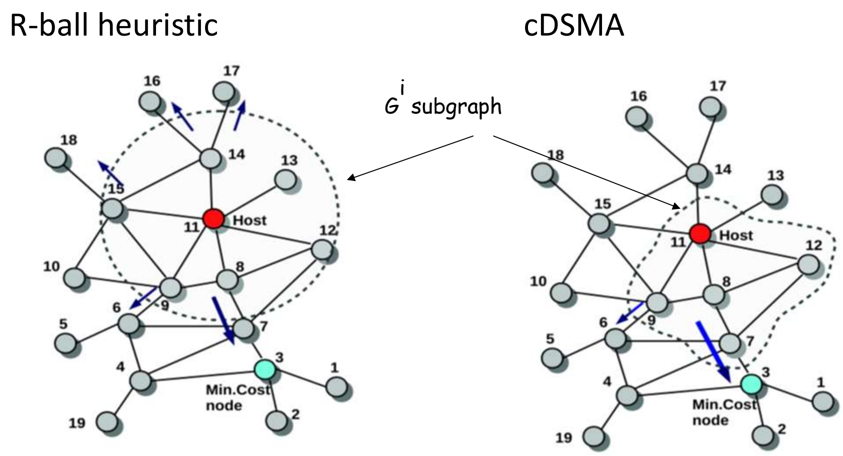
- solves UKM (and UFL) through **iterative solutions of smaller** optimization problems using **limited-scope** (non-global) demand and topology info**
- **within a limited neighborhood of R-hops around current facility
  Demand generated by outer nodes is *mapped* to the nodes at the outer shell of the R-hop neighborhood

$$Cost(\mathcal{F}) = \sum_{n \in \mathcal{V}} w_n \cdot min_{x_j \in \mathcal{F}}\{d(x_j, n)\}$$



G. Smaragdakis, N. Laoutaris, K. Oikonomou, I. Stavrakakis, and A. Bestavros, "Distributed Server Migration for Scalable Internet Service Deployment," ***IEEE/ACM Transactions on Networking***, Vol. 22, No. 3, June 2014

Approach 3: cDSMA (centrality-driven distributed service migration approach): a more informed search



P. Pandazopoulos, M. Karaliopoulos, I. Stavrakakis, "Distributed Placement of Autonomic Internet Services", ***IEEE Transactions on Parallel and Distributed Systems***, Vol.25, No.7, pp.1702-12, July 2014.

MAIN REFERENCES - READING MATERIAL

I. Stavrakakis, “Some distributed approaches to the service facility location problem in dynamic and complex networks”, Handbook of Optimization in Complex Networks: Vol. 1 (Theory and Applications) and Vol. 2 (Communication and Social Networks), edited by Profs. My T. Thai and Panos Pardalos, Springer Publisher, fall 2011.

K. Oikonomou, I. Stavrakakis, “Scalable Service Migration in Autonomic Network Environments", special issue of ***IEEE Journal on Selected Areas in Communications*** (JSAC) on “Recent Advances in Autonomic Communications”, Vol. 28, No. 1, pp. 84-94, Jan. 2010.
http://cgi.di.uoa.gr/~istavrak/publications/2010JSAC.os.pdf

G. Smaragdakis, N. Laoutaris, K. Oikonomou, I. Stavrakakis, and A. Bestavros, “Distributed Server Migration for Scalable Internet Service Deployment,” ***IEEE/ACM Transactions on Networking***, Vol. 22, No. 3, June 2014.
http://cgi.di.uoa.gr/~istavrak/publications/2010TNET.slosb.pdf

P. Pandazopoulos, M. Karaliopoulos, I. Stavrakakis, "Distributed Placement of Autonomic Internet Services", IEEE Transactions on Parallel and Distributed Systems, Vol.25, No.7, pp.1702-12, July 2014.
http://cgi.di.uoa.gr/~istavrak/publications/2013TPDS.pks.pdf