國立台灣大學資訊管理研究所
碩士論文審查

考慮智慧型惡意攻擊下之網路存活度最大化

Maximization of Network Survivability against Intelligent and Malicious Attacks

指導教授：林永松　博士
顏宏旭　博士
研究生：陳建宏　撰

中華民國九十四年七月二十二日

Outline

- Introduction & Motivation
- Problem Description & Formulation
- Solution Approach
- Computational Experiments
- Summary & Future Work

Introduction

- The critical information infrastructures protection has received more attention since the event of 911.
- How computer networks function normally under random errors or/and malicious attacks has become an even more important issue.
- Information security experts have suggested different tools and strategies that focus on different network attack modes.

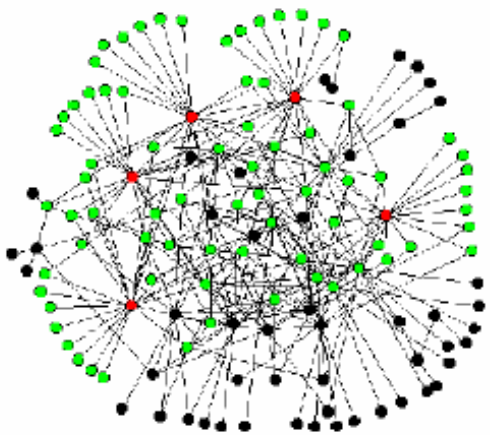
Introduction (cont.)

- We should not ask “Is the system secure?” but “How secure is the system?”
- Survivability is roughly defined as how well a network or a system sustains under random errors or malicious attacks or both.
  - Time sustainable to accidents
  - Probability to function normally
  - Availability of desired services under abnormal conditions
  - Other interesting metrics

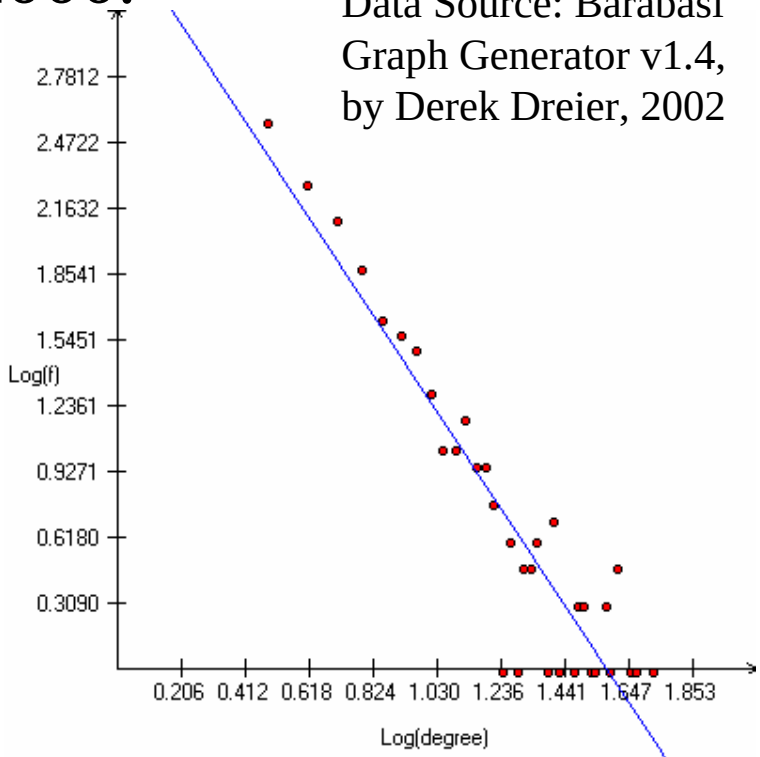
Data Source: Vickie R. Westmark, “A Definition for Information System Survivability,” *IEEE Proceedings of the 37th Hawaii International Conference on System Sciences*, 2004

Introduction (cont.)

- Faloutsos et al. obeserved that the Internet follows a power-law distribution.
- Albert-Laszlo Barabasi and Reka Albert proposed scale-free networks in year 2000.
  - Growth
  - Preferential attachment



Data Source: Reka Albert, Hawoong Jeong, and Albert-Laszlo Barabasi, “Error and Attack Tolerance of Complex Networks,” *Nature* 406, 378-381, 2000



Data Source: Barabasi Graph Generator v1.4, by Derek Dreier, 2002

Introduction (cont.)

- Some characteristics of the scale-free networks
  - Robust to random errors
  - Vulnerable to malicious attacks
- The Internet remains unaffected by the random removal of as high as 2.5% of the nodes.
- The diameter of the Internet grows more than triple if 2.5% of the nodes are maliciously attacked.

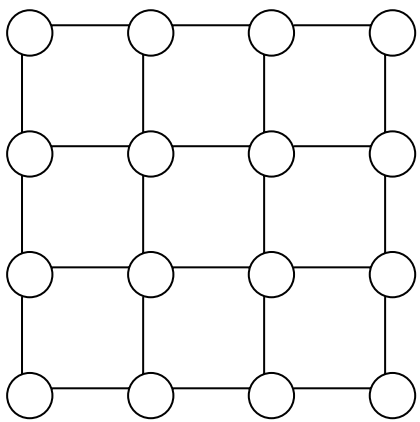
Data Source: Reka Albert, Hawoong Jeong, and Albert-Laszlo Barabasi, "Error and Attack Tolerance of Complex Networks," *Nature* 406, 378-381, 2000

Motivation

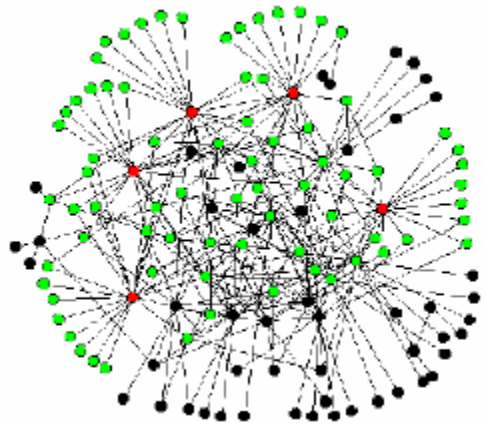
- Since the survivability issue has drawn much attention, a network operator may invest a fixed amount of budget to enhance the survivability of the existing networks.
- We therefore want to evaluate the level of robustness of a network against malicious attacks once the budget allocation policy has been determined.
- We want to know how budget allocation scenarios influence the levels of robustness.

Motivation (cont.)

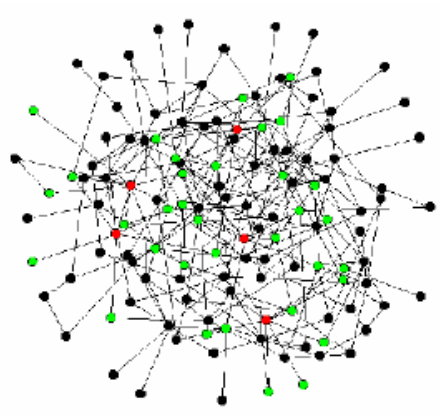
- We also want to examine the levels of robustness of different topologies.
  - Random networks
  - Grid networks
  - Scale-free networks



A grid network



A scale-free network



A random network

Motivation (cont.)

- Moreover, we want to know how a network operator should allocate a fixed amount of budget so that the survivability of a network can be maximized.
- We believe that a network operator's budget allocation policy should consider responses from an attacker, due to the fact that an attacker may change his strategies to a better one if he finds other easier ways to attain the same goal.

Problem Description
Model 1

- Assume the budget allocation policy is given, we want to know the minimal attack cost for an attacker to compromise a network.
- The system is survivable if there is at least one available path between one of the critical OD-pairs.

Problem Description
Model 1 (cont.)

Problem scenarios:

1. The survivability metric is defined as the connectivity of the given critical OD-pairs.
2. The objective of the attacker is to minimize the total attack cost of destroying all paths between given critical OD-pairs.
3. The attacker and the defender have complete information about the network topology.
4. The defender's budget allocation strategy is a given parameter.
5. We consider node attacks only. (No link attacks are considered.) If a node is attacked, its outgoing links are not functional.
6. We consider malicious attacks only. (No random errors are considered.)

Problem Description
Model 1 (cont.)

Given:

1. The network topology and the network size
2. The defender's budget allocation policy
3. A set of critical OD-pairs

Objective:

To minimize the total cost of an attack

Subject to:

1. There is no available path for each given critical OD-pair to communicate.

To determine:

1. Which nodes will be attacked

Problem Description
Model 1 (cont.)

- Problem Formulation Tips
  - To show that all paths of one critical OD-pair are destroyed, we associate a very small cost to each functional link, and a very large cost to each nonfunctional link, and require that the shortest cost path for each OD-pair involves at least one nonfunctional link.
- For simplicity, we assume that the minimal attack cost to compromise a node equals the allocated budget for it.

Problem Description
Model 1 (cont.)

Given Parameters :

Notation	Description
V	The index set of all nodes
L	The index set of all links
W	The index set of all given critical origin-destination pairs
OUT^i	The index set of outgoing links of node i, where $i \in V$
M	A large number that represents the link disconnection
ε	A small number that represents the link connectedness
P_w	The index set of all candidate paths of an OD-pair, w, where $w \in W$
δ_{pl}	An indicator function, which is 1 if link l is on path p, and 0 otherwise
b_i	Budget allocated to node i, where $i \in V$

Problem Description
Model 1 (cont.)

Decision Variables :

Notation	Description
y_i	1 if node *i* is compromised, and 0 otherwise
t_{wl}	1 if link *l* is used by OD pair, *w*, and 0 otherwise
x_p	1 if path *p* is chosen, and 0 otherwise
c_l	Cost of link *l*, which is ε if link *l* functions normally, and $M+\varepsilon$ if link *l* is broken

Problem Description
Model 1 (cont.)

Objective function: $\min_{y_i} \sum_{i \in V} y_i b_i$ (IP 2)

Subject to

$$c_l \le y_i M + \varepsilon \qquad \forall i \in V,\ l \in OUT^i \qquad \text{(IP 2.1)}$$

$$\sum_{l \in L} t_{wl} c_l \le \sum_{l \in L} \delta_{pl} c_l \qquad \forall p \in P_w,\ w \in W \qquad \text{(IP 2.2)}$$

$$\sum_{p \in P_w} x_p \delta_{pl} \le t_{wl} \qquad \forall w \in W,\ l \in L \qquad \text{(IP 2.3)}$$

$$M \le \sum_{l \in L} t_{wl} c_l \qquad \forall w \in W \qquad \text{(IP 2.4)}$$

$$\sum_{p \in P_w} x_p = 1 \qquad \forall w \in W \qquad \text{(IP 2.5)}$$

$$x_p = 0 \text{ or } 1 \qquad \forall p \in P_w,\ w \in W \qquad \text{(IP 2.6)}$$

$$y_i = 0 \text{ or } 1 \qquad \forall i \in V \qquad \text{(IP 2.7)}$$

$$t_{wl} = 0 \text{ or } 1 \qquad \forall w \in W,\ l \in L \qquad \text{(IP 2.8)}$$

$$c_l = \varepsilon \text{ or } M + \varepsilon \qquad \forall l \in L \qquad \text{(IP 2.9)}$$

$$\sum_{i \in V} y_i \ge V_{lb.} \qquad \text{(IP 2.10)}$$

Problem Description
Model 2

- A network operator decides how to distribute a fixed amount of budget so that the minimal attack cost to compromise a network can be maximized.
- The system is survivable if there is at least one available path between one of the critical OD-pairs.
- The problem objective becomes to $\max_{b_i} \min_{y_i} \sum_{i \in V} y_i b_i$, where both the sets of b_i and y_i are decision variables.

Solution Approach
Model 1

- We apply the Lagrangean Relaxation method to solve the proposed problem.
- By relaxing the difficult constraints in the original problem and associating Lagrangean multipliers, we have the Lagrangean Relaxation (LR) problem.
  - LR is easier to be solved than the original problem.
  - Lower bounds (for minimization problems) of the optimal objective function value to the original problem are gained through the process.

Solution Approach
Model 1 (cont.)

- By applying the Lagrangean relaxation method, the primal problem (IP2) can be transformed into the Lagrangean relaxation problem (LR), where constraints (IP 2.1), (IP 2.2), (IP 2.3), and (IP 2.4) are relaxed.
- Optimization Problem (LR):

$$Z_D(u_1,u_2,u_3,u_4) = \min_{y_i} \sum_{i\in V} y_i b_i + \sum_{i\in V}\sum_{l\in OUT^i} u^1{}_{il}[c_l - (y_i M + \varepsilon)] +$$

$$+\sum_{w\in W}\sum_{p\in P_w} u^2{}_{wp}\sum_{l\in L}[t_{wl}c_l - \delta_{pl}c_l] + \sum_{w\in W}\sum_{l\in L} u^3{}_{wl}[(\sum_{p\in P_w} x_p\delta_{pl}) - t_{wl}] +$$

$$\sum_{w\in W} u^4{}_w \left[M - \sum_{l\in L} t_{wl}c_l \right]$$

- We further decompose LR into three independent Subproblems.

Solution Approach
Model 1 (cont.)

- Subproblem 1 **(related to decision variable x_p)**

$$Z_{sub1}(u_3) = \min \sum_{w \in W} \sum_{l \in L} \sum_{p \in P_w} u^3{}_{wl} \delta_{pl} x_p \qquad \text{(Sub 1)}$$

subject to

$$\sum_{p \in P_w} x_p = 1 \qquad \forall w \in W \qquad \text{(Sub1.1)}$$

$$x_p = 0 \text{ or } 1 \qquad \forall p \in P_w, w \in W. \qquad \text{(Sub1.2)}$$

- Subproblem 1 can further be decomposed into $|W|$ independent problems. We apply Dijkstra's shortest cost path algorithm to optimally solve each independent problem.

Time Complexity $O(|W| \times |V|^2)$

Solution Approach
Model 1 (cont.)

- Subproblem 2 **(related to decision variable y_i)**

$$Z_{sub2}(u_1) = \min \sum_{i \in V} y_i b_i + \sum_{i \in V} \sum_{l \in OUT^i} u^1_{il}(-M) y_i \qquad \text{(Sub 2)}$$

subject to

$$y_i = 0 \text{ or } 1 \qquad \forall i \in V \qquad \text{(Sub 2.1)}$$

$$\sum_{i \in V} y_i \geq V_{lb.} \qquad \text{(Sub 2.2)}$$

- To solve Subproblem 2, we first apply a quick sort on $b_i - M \sum_{l \in OUT^i} u^1_{il}$. After satisfying (Sub 2.2), we determine the value of each y_i by examining its associated parameters.

Time Complexity $O(|V| \log |V|)$

Solution Approach
Model 1 (cont.)

- Subproblem 3 **(related to decision variable t_{wl}, c_l)**

$$Z_{sub3}(u_1,u_2,u_3,u_4) = \min \sum_{i\in V}\sum_{l\in OUT^i} u^1{}_{il}c_l + \sum_{w\in W}\sum_{p\in P_w} u^2{}_{wp}\sum_{l\in L}(t_{wl}c_l - \delta_{pl}c_l) +$$

$$\sum_{w\in W}\sum_{l\in L} u^3{}_{wl}(-t_{wl}) + \sum_{w\in W} u^4{}_{w}(-\sum_{l\in L} t_{wl}c_l) \qquad \text{(Sub 3)}$$

subject to

$t_{wl} = 0$ or 1 $\quad \forall w \in W, l \in L$ (Sub 3.1)

$c_l = \varepsilon$ or $M + \varepsilon \quad \forall l \in L.$ (Sub 3.2)

- Subproblem 3 can be further decomposed into $|L|$ independent problems.

Time Complexity $O(|W| \times |L|)$

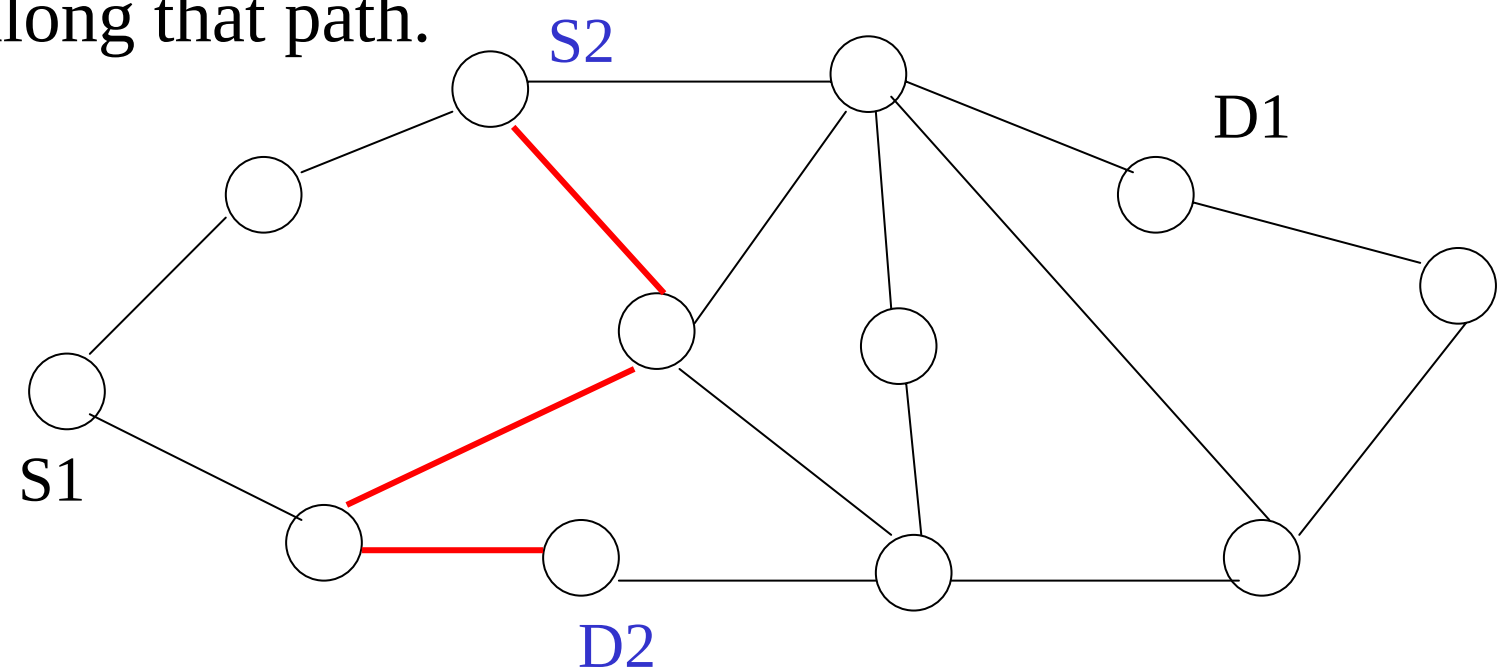
Solution Approach
Model 1 (cont.)

- Solutions to the dual problem and the Lagrangean multipliers provide good hints and a starting point to get primal feasible solutions.
- We derive a primal algorithm by using the solutions of y_i and c_l in the dual problem.

Time Complexity $O(|W| \times |V|^5)$

Solution Approach
Model 2

- Despite the complicated max-min form, we show that the optimality solution is a trivial case.
- We find the minimal hop path among all given OD-pairs, and evenly distribute the total budget, *B*, along that path.



Computational Experiments

- Simple algorithm 1: we adopt the well-known maximum flow algorithm, take the union of the minimum cuts, and “recover” some nodes.
- Simple algorithm 2: we attack the most connected node sequentially, until all possible paths between given OD-pairs are destroyed.

Computational Experiments (cont.)

Experimental Parameters

Number of Nodes	16 ~ 100
Number of Links	60 ~ 400
Number of critical OD-pairs	8 ~ 250
Testing Topology	Random, Grid, and Scale-free
Number of Iterations	2000
Non-improvement Counter	80
Initial Upper Bound	Solution of SA_1
Initial budget allocation policy	Uniform, Degree-based distribution
Test Platform	CPU: Intel Pentium-4 2.0 GHz RAM: 512 MB OS: MS Windows XP

Computational Experiments (cont.)

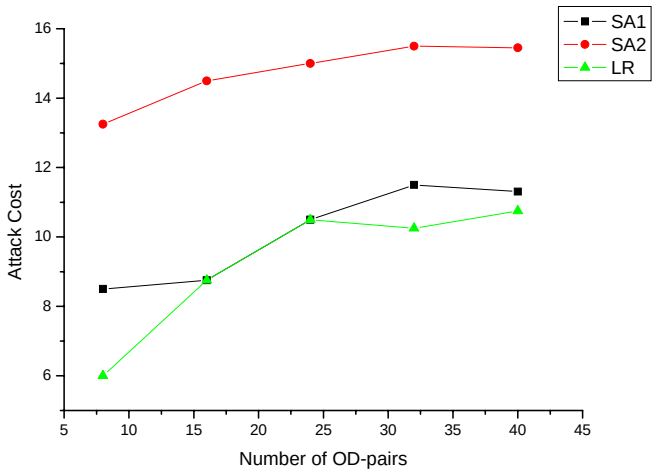


Fig 4-1

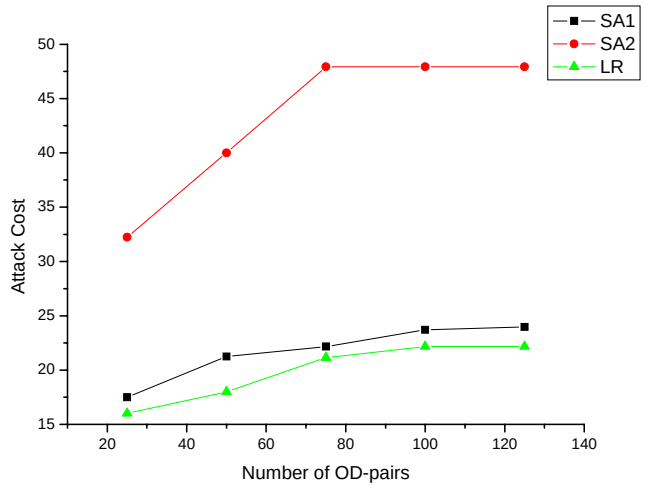


Fig 4-2

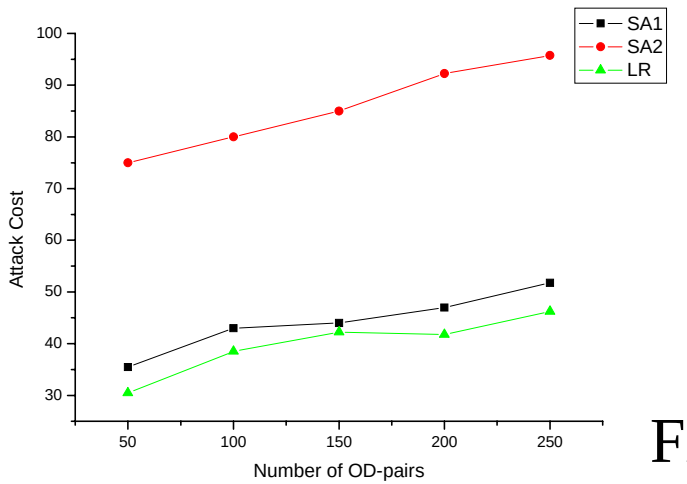


Fig 4-3

Computational Experiments (cont.)

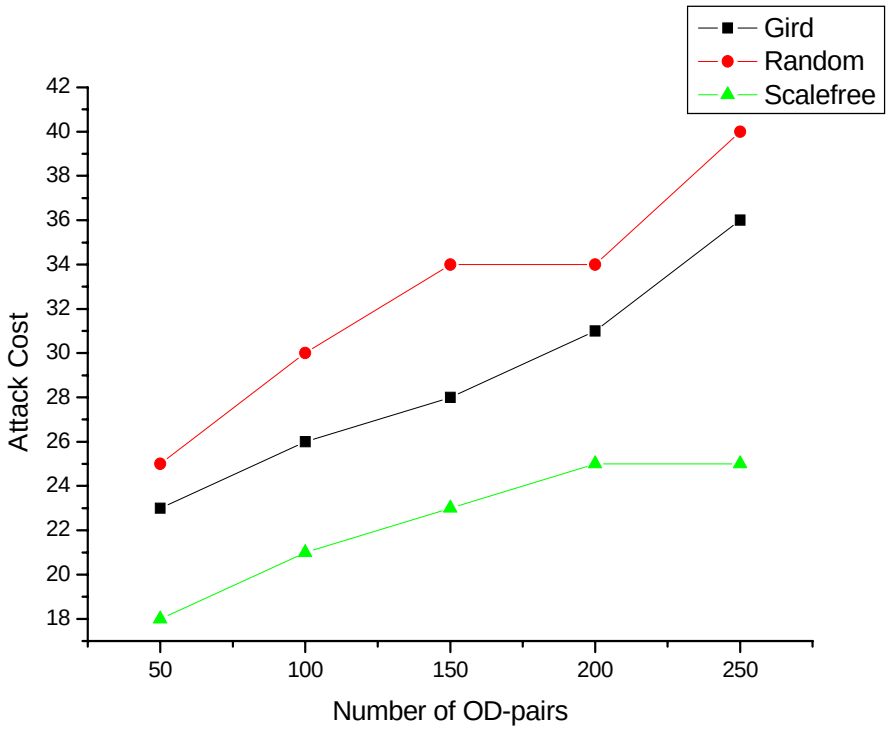


Fig 4-4

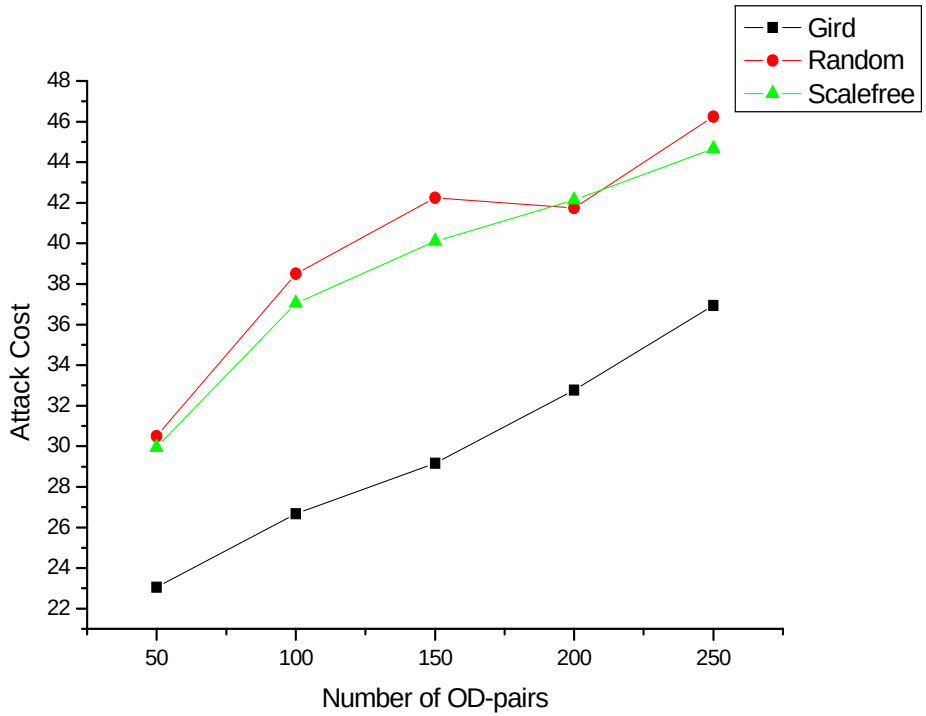


Fig 4-5

Time Complexity

Problem	Time Complexity to solve this problem
Subproblem 1	$O(\|W\|\|V\|^2)$
Subproblem 2	$O(\|V\|\log\|V\|)$
Subproblem 3	$O(\|W\|\|L\|)$
Getting Primal Feasible Heuristics	$O(\|W\|\|V\|^5)$
Simple Algorithm 1	$O(\|W\|\|V\|^3)$
Simple Algorithm 2	$O(\|W\|\|V\|^3)$

Summary

- We have proposed a mathematical model to evaluate the network survivability against malicious attacks.
- We have presented a lemma of the optimality condition for a defender under the given scenario.
- We have shown that our solution approach is effective comparing with simple algorithms.
- We have evaluated the robustness of different topologies and concluded that a proper budget allocation policy will enhance the level of robustness against attacks.

Future Work

- The best budget allocation strategy for an initial budget is a very challenging issue.
- We can further consider different definitions of survivability.
  - The connectivity of the largest fragment in a network
  - QoS constrained survivability
- We can address different attack behaviors.
  - An attacker wants to reach one or multiple core nodes through the most likely path.
- We can also research applications on different transmission media.

Contribution

- We well-formulate the problem of an attack and defense scenario.
  - To the best of our knowledge, our proposed approach is the first attempt to solve an attack and defense problem considering survivability issues in general networks via mathematical programming techniques.
- We evaluate the robustness of different topologies and conclude that, with a proper budget allocation policy, a scale-free network may achieve the same level of robustness as a random network.
- We derive a lemma of the optimality condition for the defender.

Thanks for Your Attention.

Appendix

Problem Description
Model 2 (cont.)

Given:

1. The network topology and the network size
2. A set of critical OD-pairs
3. The total budget of the defender

Objective:

To maximize the attacker's minimal total attack cost

Subject to:

1. The total budget constraint of the defender
2. No path is available for each given critical OD-pair to communicate.

To determine:

1. The budget allocated to each node
2. Which nodes the attacker has decided to target

Problem Description
Model 2 (cont.)

Problem scenarios:

1. The survivability metric is measured as the connectivity of the given critical OD-pairs.
2. The attacker and the defender have complete information about the targeted network topology.
3. The objective of the attacker is to minimize the total attack cost of destroying all paths between given critical OD-pairs.
4. The objective of the defender is to distribute the total amount of budget effectively so that the minimal total attack cost can be maximized.
5. We consider node attacks only. (No link attacks are considered). If a node is attacked, its outgoing links are not functional.
6. We consider malicious attacks only. (No random failures are considered.)

Problem Description
Model 2 (cont.)

- **Argument**: We claim that the optimality condition for the defender holds if and only if the total budget, B, is fully used.
- Note that this argument holds only when the set of decision variables b_i is continuous.

Problem Description

Given Parameters : Model 2 (cont.)

Notation	Description
V	The index set of all nodes
L	The index set of all links
W	The index set of all given critical origin-destination pairs
OUT^i	The index set of outgoing links of node i, where $i \in V$
M	A large number that represents the link disconnection
ε	A small number that represents the link connectedness
P_w	The index set of all candidate paths of an OD-pair, w, where $w \in W$
δ_{pl}	An indicator function, which is 1 if link l is on path p, and 0 otherwise
B	**Total budget of the defender**

Problem Description
Model 2 (cont.)

Decision Variables :

Notation	Description
y_i	1 if node i is compromised, and 0 otherwise
t_{wl}	1 if link l is used by OD pair, w, and 0 otherwise
x_p	1 if path p is chosen, and 0 otherwise
c_l	Cost of link l, which is ε if link l functions normally, and $M+\varepsilon$ if link l is broken
b_i	**The budget allocated to node i**

Problem Description
Model 2 (cont.)

$$\max_{b_i} \min_{y_i} \sum_{i \in V} y_i b_i$$

Subject to

$c_l \le y_i M + \varepsilon$	$\forall i \in V, l \in OUT^i$	(IP 3.1)
$\sum_{l \in L} t_{wl} c_l \le \sum_{l \in L} \delta_{pl} c_l$	$\forall p \in P_w, w \in W$	(IP 3.2)
$\sum_{p \in P_w} x_p \delta_{pl} \le t_{wl}$	$\forall w \in W, l \in L$	(IP 3.3)
$M \le \sum_{l \in L} t_{wl} c_l$	$\forall w \in W$	(IP 3.4)
$\sum_{p \in P_w} x_p = 1$	$\forall w \in W$	(IP 3.5)
$x_p = 0 \text{ or } 1$	$\forall p \in P_w, w \in W$	(IP 3.6)
$y_i = 0 \text{ or } 1$	$\forall i \in V$	(IP 3.7)
$t_{wl} = 0 \text{ or } 1$	$\forall w \in W, l \in L$	(IP 3.8)
$c_l = \varepsilon \text{ or } M + \varepsilon$	$\forall l \in L$	(IP 3.9)

Problem Description
Model 2 (cont.)

Subject to (cont.)

$\sum_{i \in V} b_i = B$		(IP 3.10)
$0 \le b_i \le B$	$\forall i \in V.$	(IP 3.11)

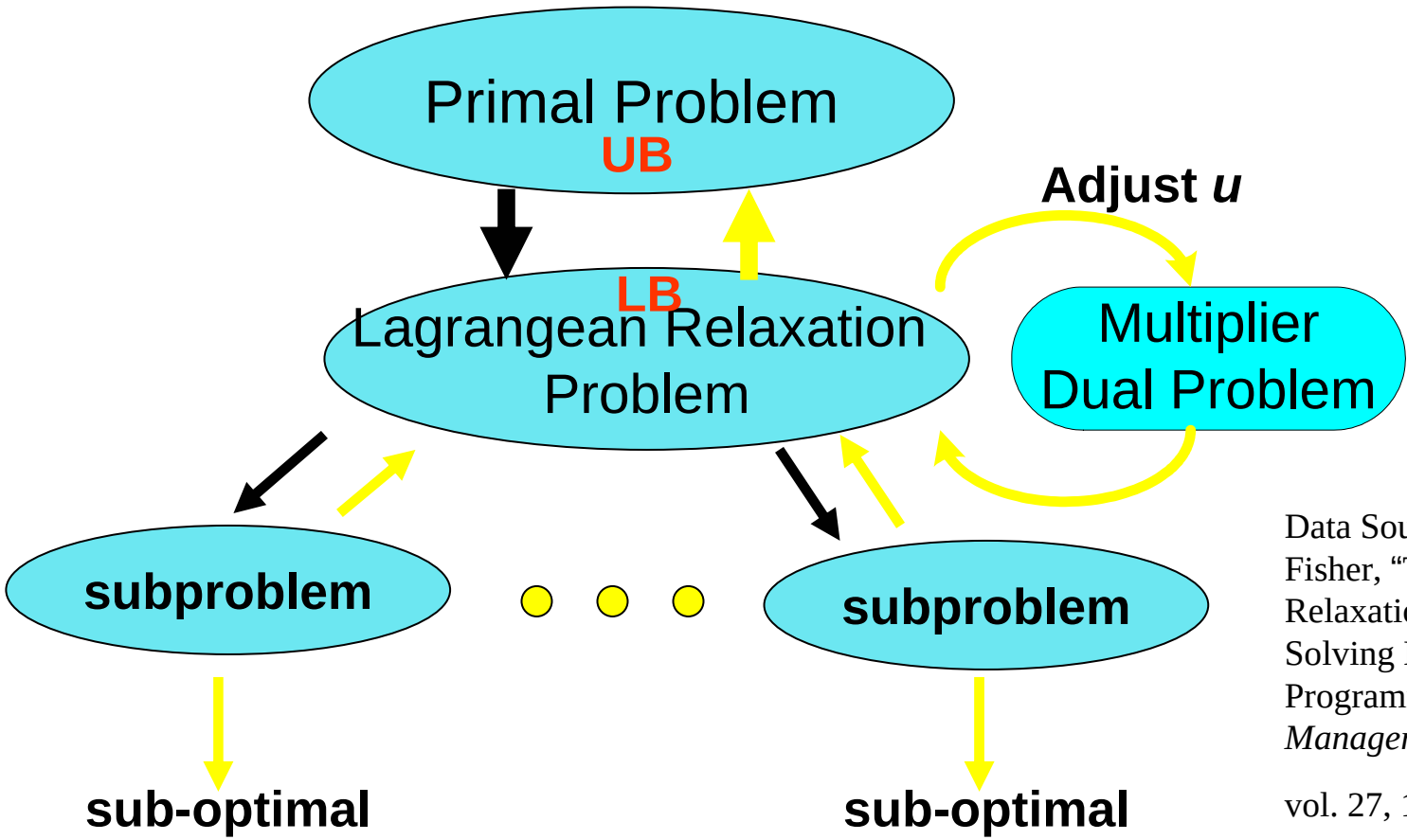
Solution Approach
Model 1 (cont.)

- The most popular method to solve the dual problem is the subgradient method.
  - In iteration *k* of the subgradient optimization procedure, the multiplier vector $\pi = (u_1, u_2, u_3, u_4)$ is updated by $\pi^{k+1} = \pi^k + t^k g^k$, where g is a subgradient of $Z_d(u_1, u_2, u_3, u_4)$ and step size t^k is determined by

$$t^k = \delta \frac{Z_{IP}^h - Z_D\left(\pi^k\right)}{\left\| g^k \right\|^2} .$$

Solution Approach
Lagrangean Relaxation

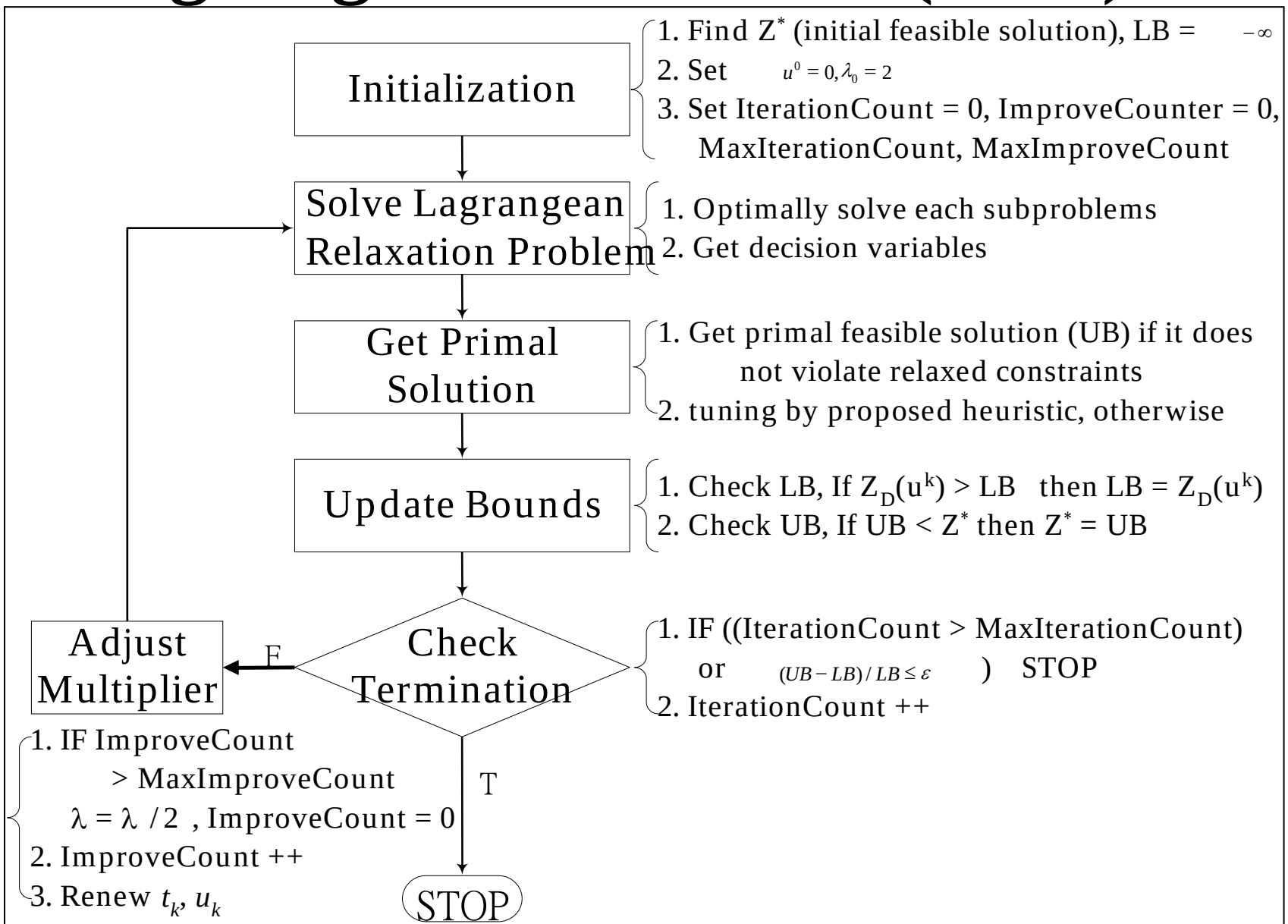
LB <= Optimal solution <= UB



Data Source: M. L. Fisher, "The Lagrangean Relaxation Method for Solving Integer Programming Problems", *Management Science*,

vol. 27, 1-18, 1981

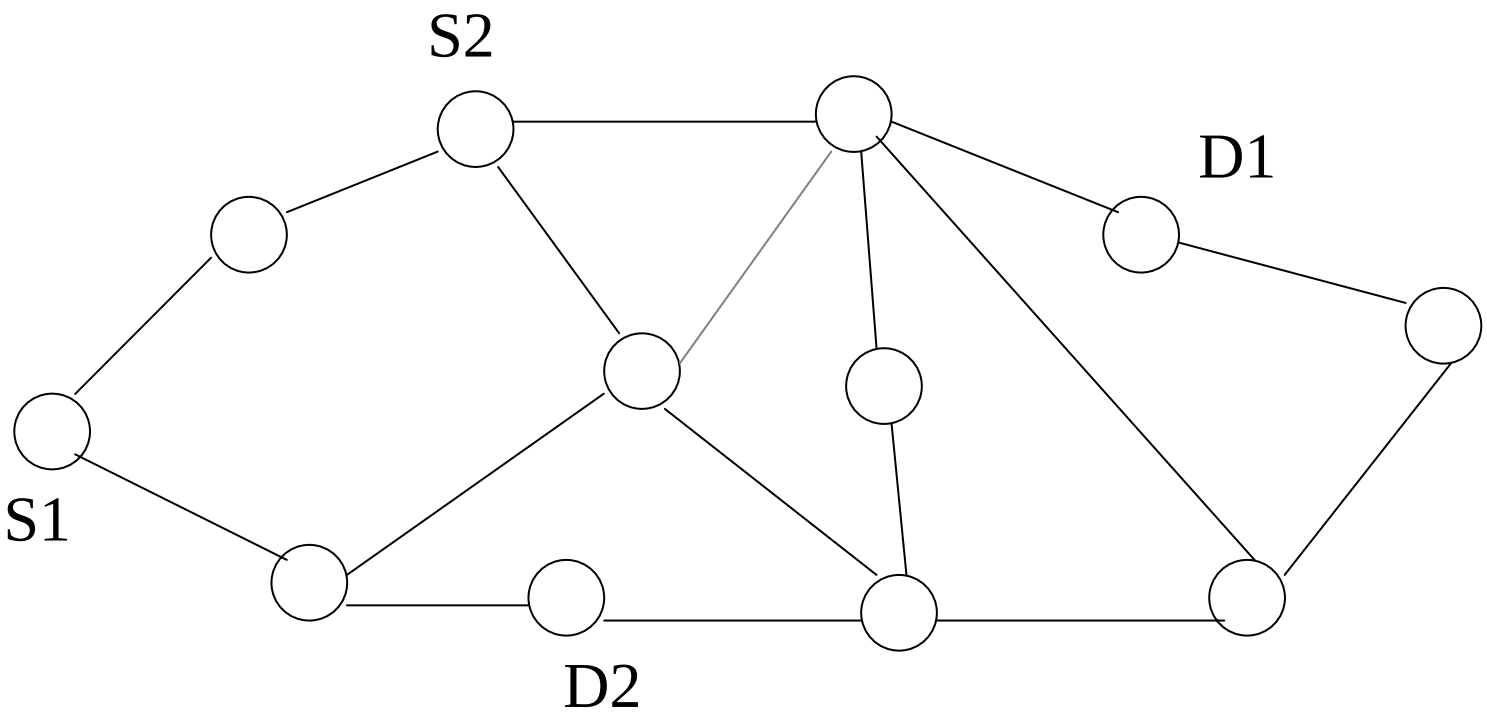
Solution Approach
Lagrangean Relaxation (cont.)



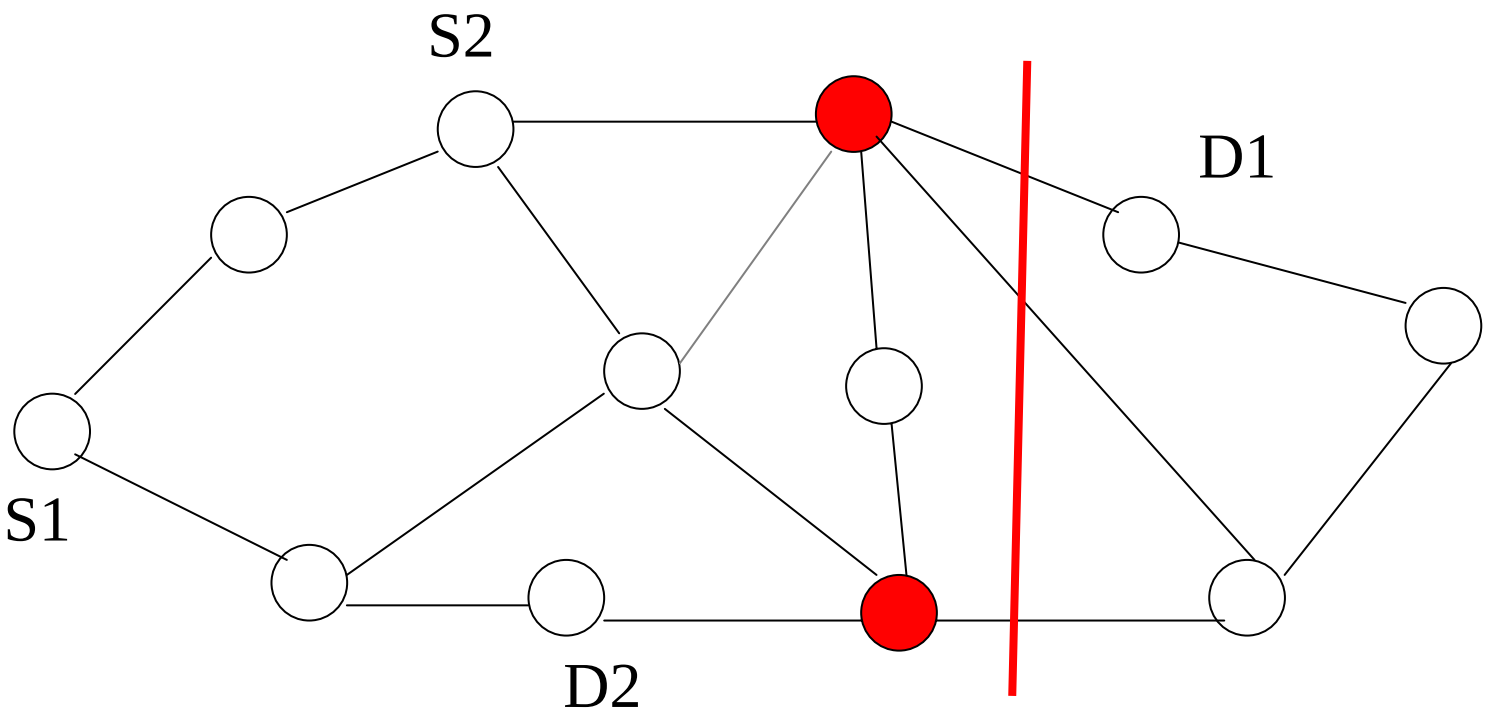
Sort the nodes in ascending order *w.r.t.* the parameters of y_i we mentioned in Subproblem2.
While (there is an available path for at least one OD-pair to communicate, and some nodes remain unexamined){
  One at a time, attack the leftmost unexamined node with a negative parameter of y_i or a large M of its outgoing link cost.
}
While (there is an available path for at least one OD-pair to communicate){
  One at a time, attack the left-most node which was not determined to be attacked yet.
}
While (some nodes remain unexamined){
  Apply a greedy algorithm; we sequentially recover the attacked node with the largest budget, b_i, and test if this recovery will lead to any available path for any OD-pair. If yes, we do not recover this node.
}
While (some nodes remain unexamined){
  Apply a greedy algorithm; we sequentially examine if a recovery of any two combinations of the attacked nodes will lead to any available path for any OD-pair. If yes, we do not recover the nodes.
}

Getting Primal Feasible Heuristics

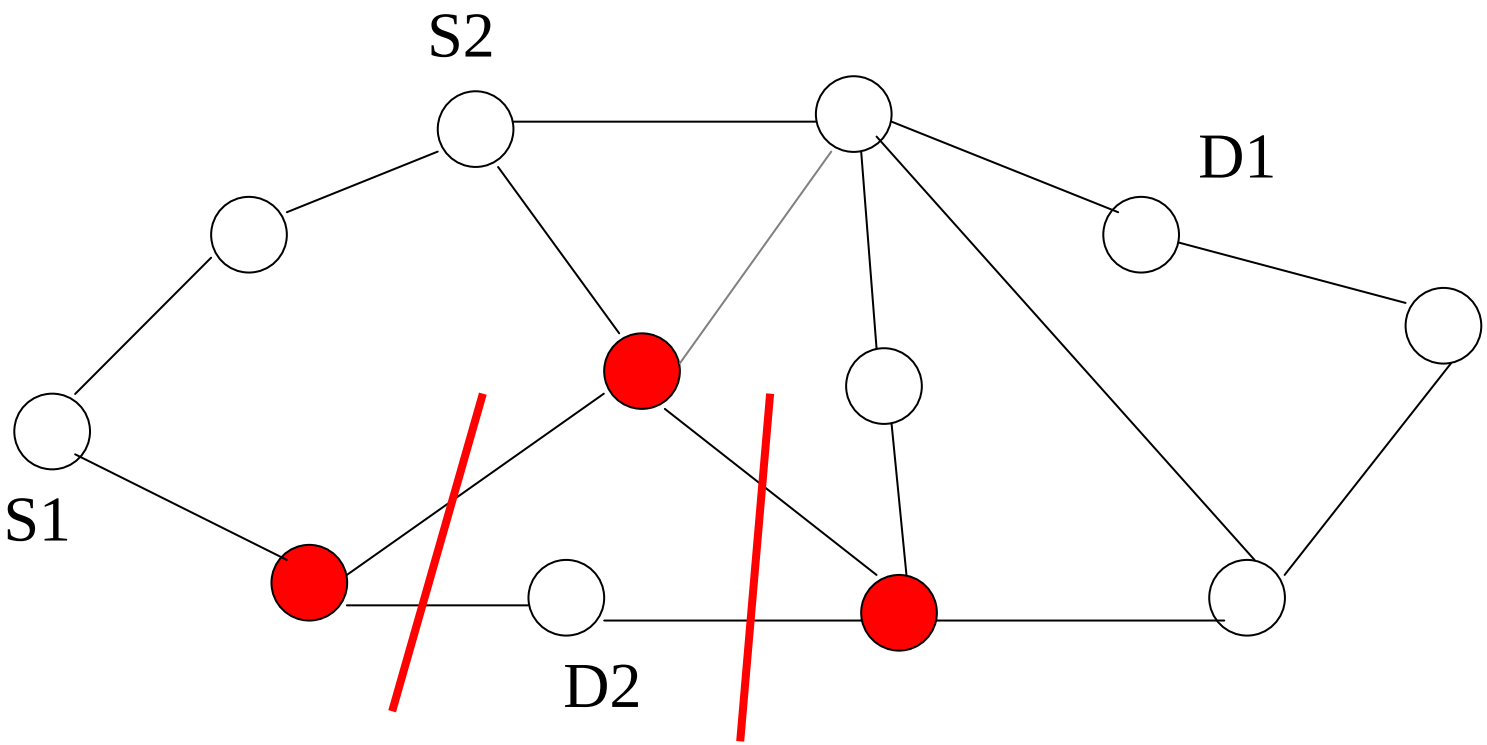
Computational Experiments
Simple Algorithm 1



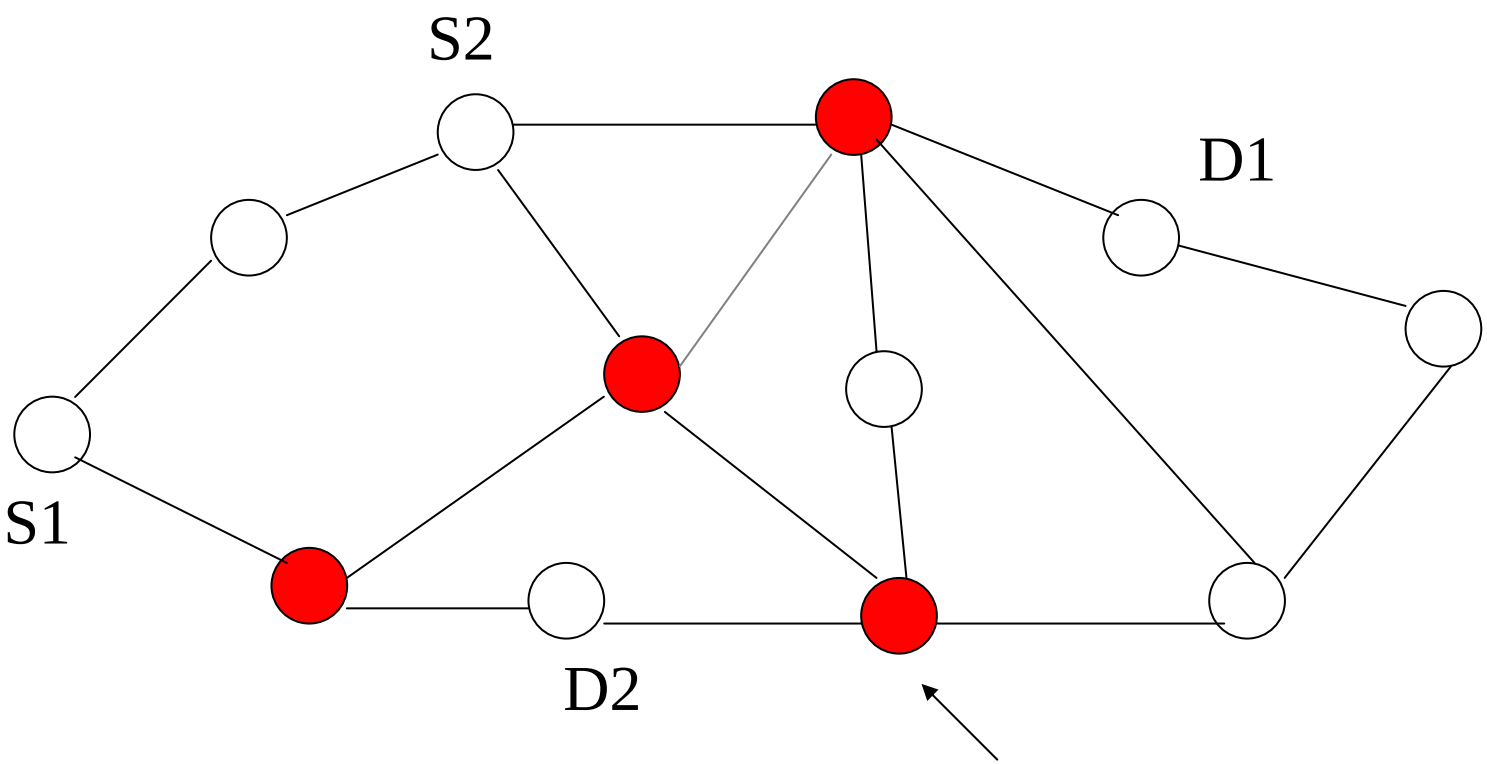
Computational Experiments
Simple Algorithm 1 (cont.)



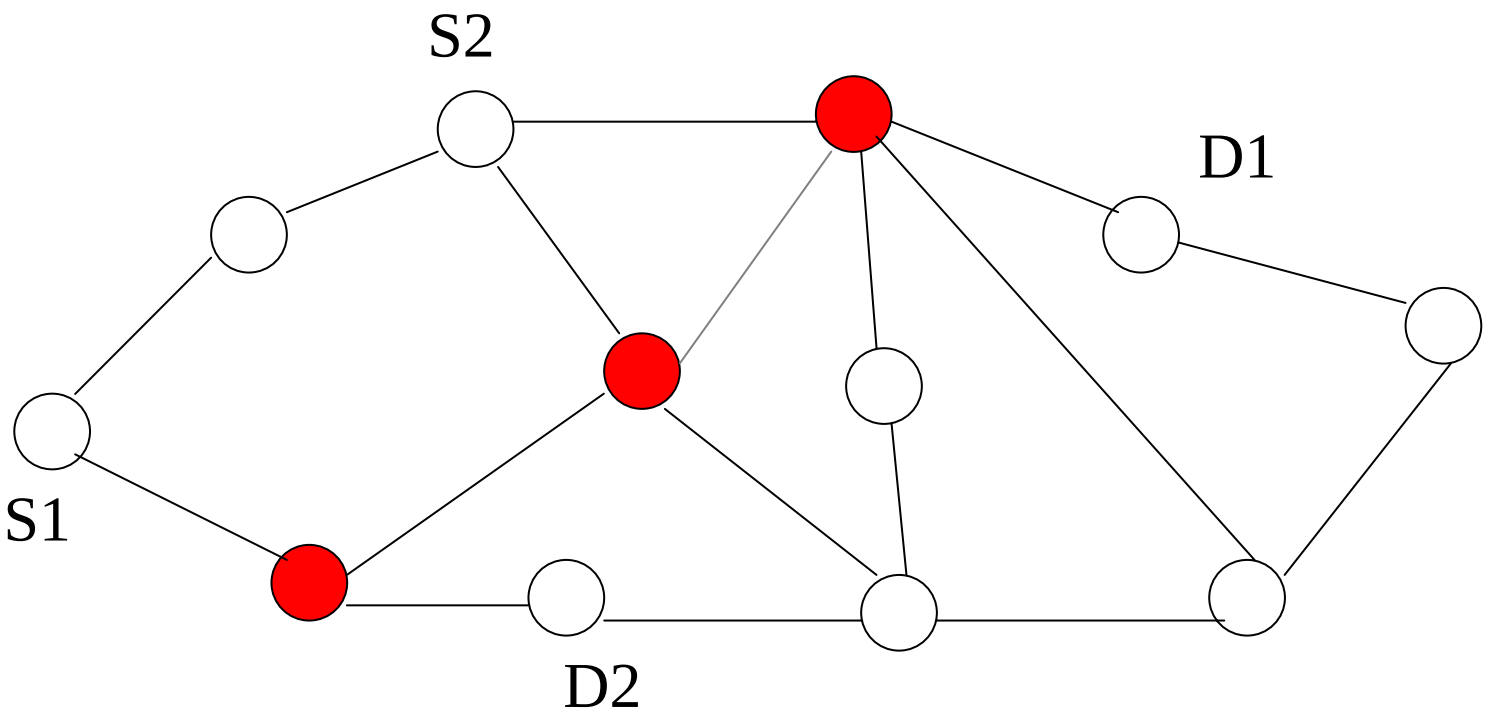
Computational Experiments
Simple Algorithm 1 (cont.)



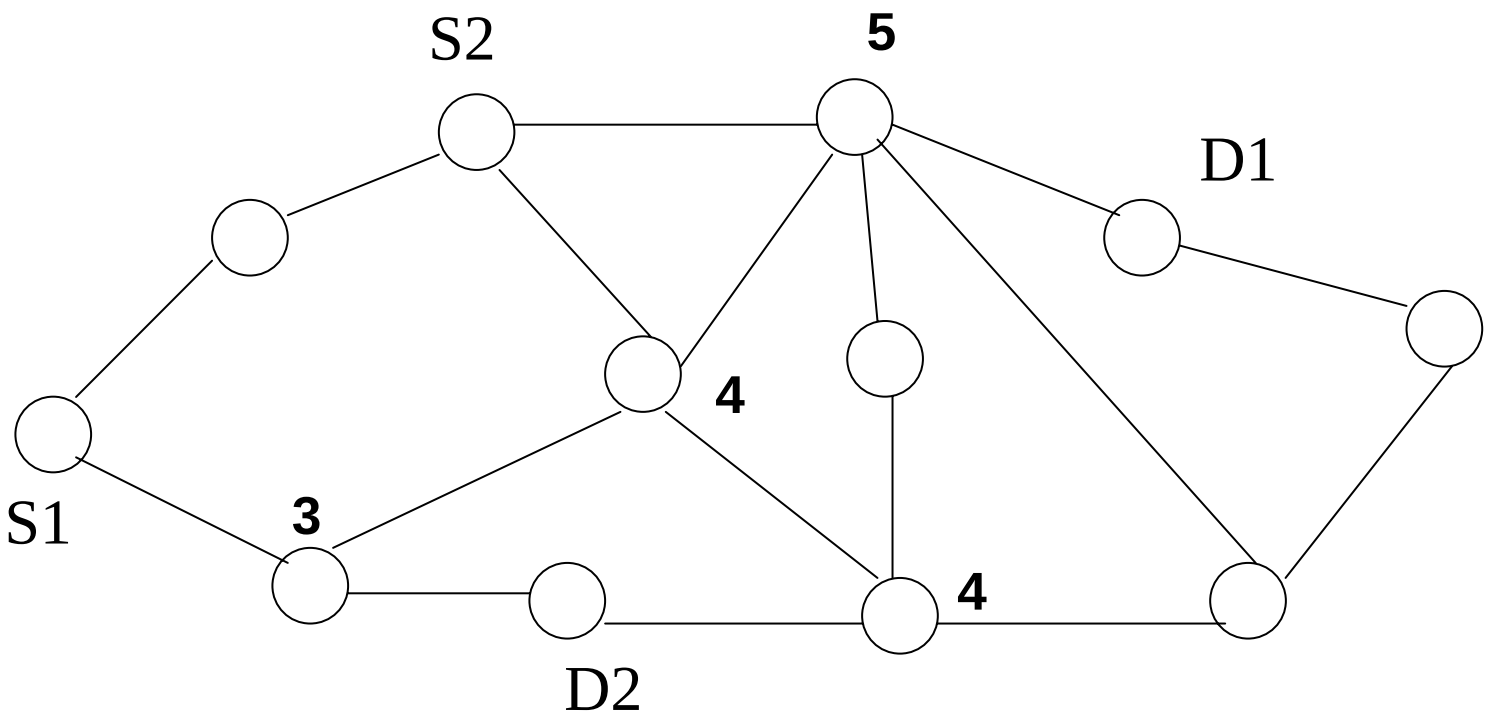
Computational Experiments
Simple Algorithm 1 (cont.)



Computational Experiments
Simple Algorithm 1 (cont.)



Computational Experiments
Simple Algorithm 2



Computational Experiments
Simple Algorithm 2 (cont.)