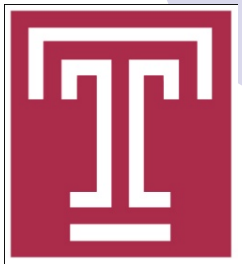
Filter Assignment Policy Against Distributed Denial-of-Service Attack



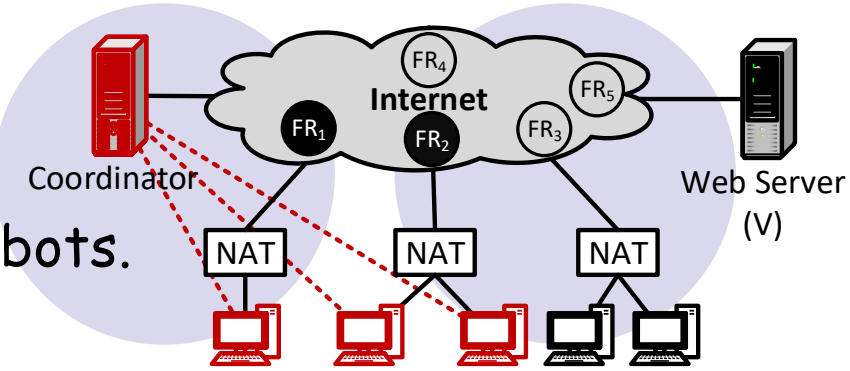
Rajorshi Biswas and Jie Wu

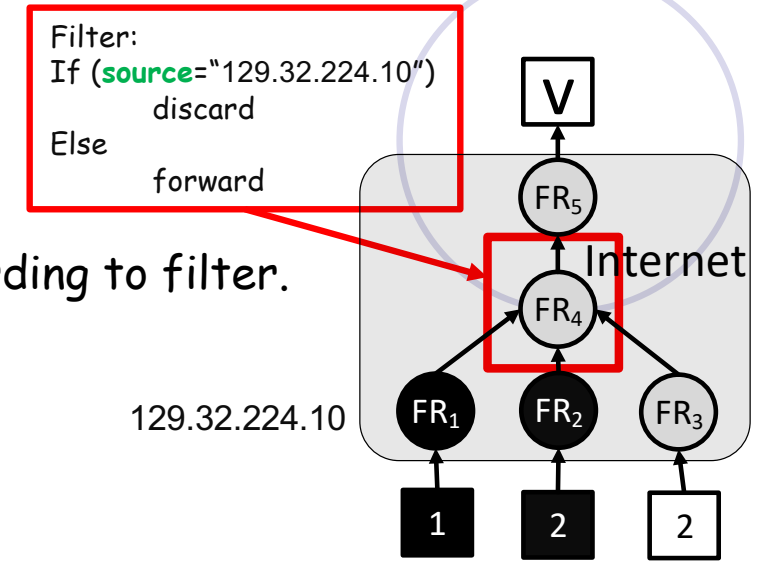
Dept. of Computer and Info. Sciences

Temple University

DDoS & Four-phase Protection System

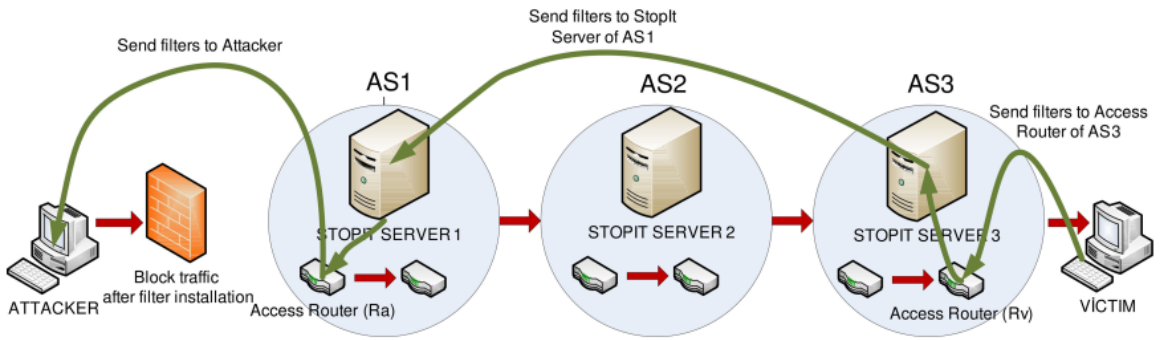
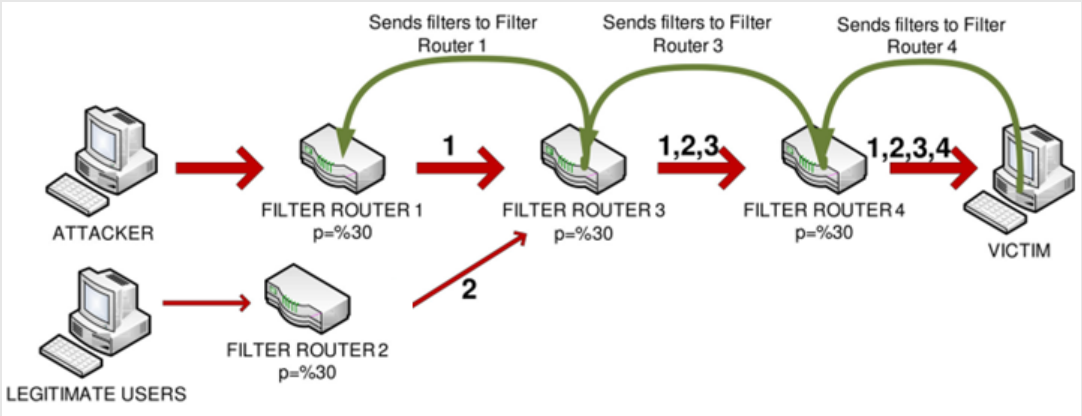
- DDoS
  - Attacker keeps the victim busy.
  - Millions of requests are fired by bots.
  - Bots are controlled by a master.
- Background
  - Filter router
    - Does packet marking.
    - Applies filter and block traffic according to filter.
  - Filter
    - Simple packet blocking rule.
    - Source-based, destination based.





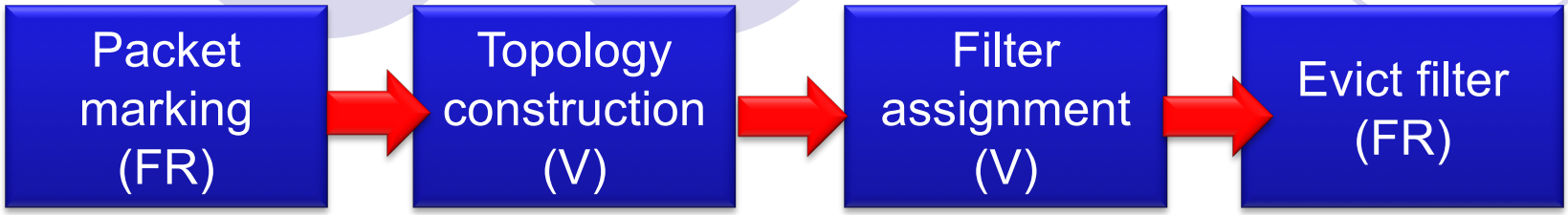
Router plugins: a software architecture for next generation routers (Dan Decasper et al. in ACM SIGCOMM '98)

Previous work

Systems	Limitations
StopIt (put filter to StopIt server)  To filter or to authorize: Network-layer DoS defense against multimillion-node botnets (X. Liu et al. at *ACM SIGCOMM Comput,* 2008)	• Needs a server to send filter to appropriate server. • Does not consider limited budget on filters.
Probabilistic Filter Scheduling (packet marking)  PFS: Probabilistic filter scheduling against distributed denial-of-service attacks (D. Seo et al. in IEEE 36th Conf. Local Comput. Netw, Oct. 2011)	• Does not consider limited budget on filters. • Filter propagation takes some time. • Hard to send huge number of filters.

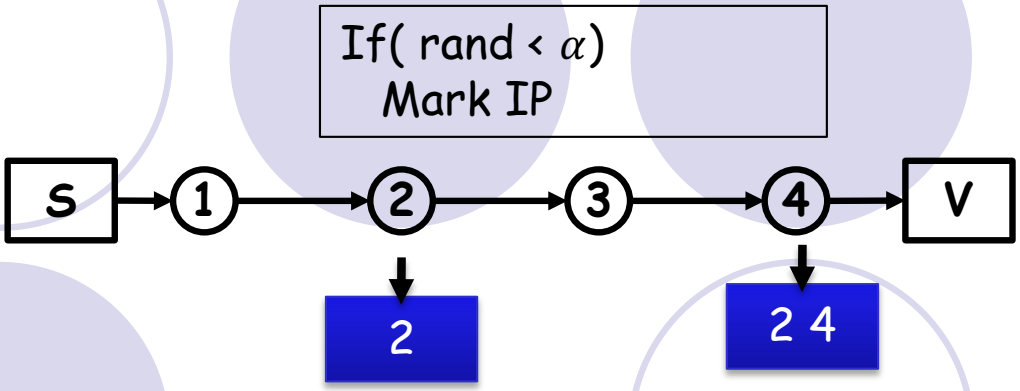
A Four-phase Protection Process

- Phase I: Packet marking by Filter Router.
- Phase II: Traffic topology and filter construction.
- Phase III: Assign filters to filter router.
- Phase IV: Evict unused filter from filter router.

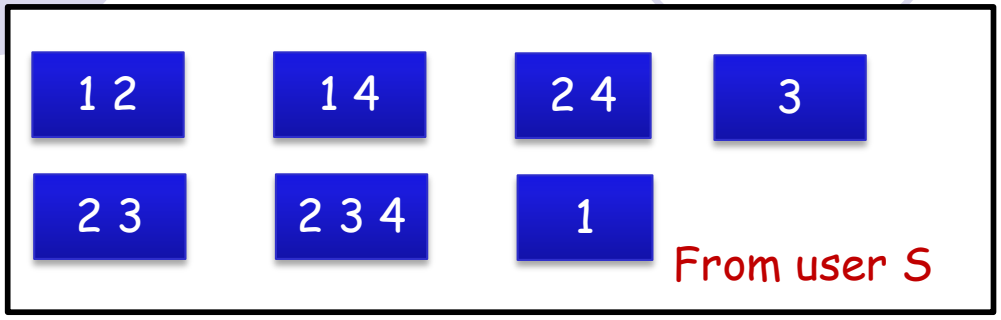


Phase I: Packet Marking by FR

- Filter router (FR) probabilistically appends its own IP address to the packet.
- α = marking probability



Example received packets, $\alpha = 0.5$

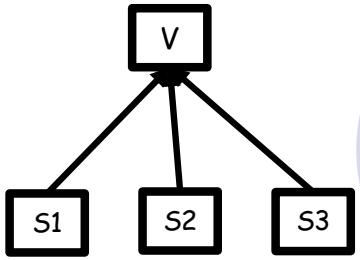


Phase II: Topology Construction

S1

S2

S3

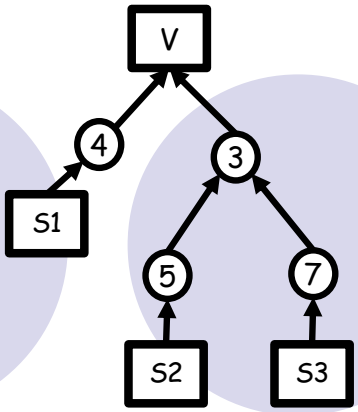


Without any marking

S1 4

S2 5 3

S3 7 3

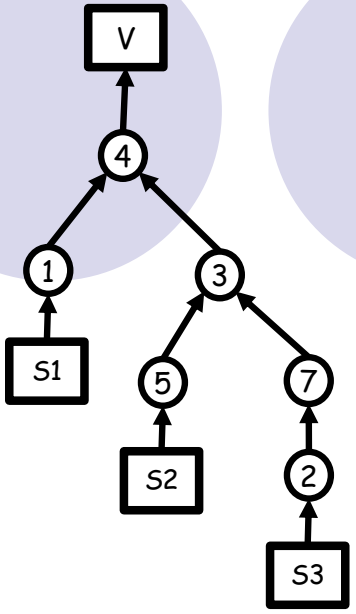


After few markings received

S1 1 4

S2 3 4

S3 2 7

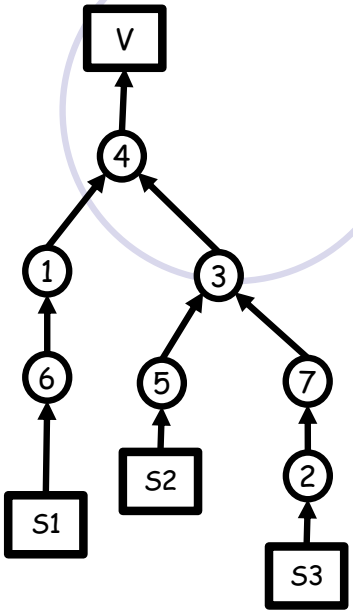


After few more markings received

S1 6 4

S1 6 1

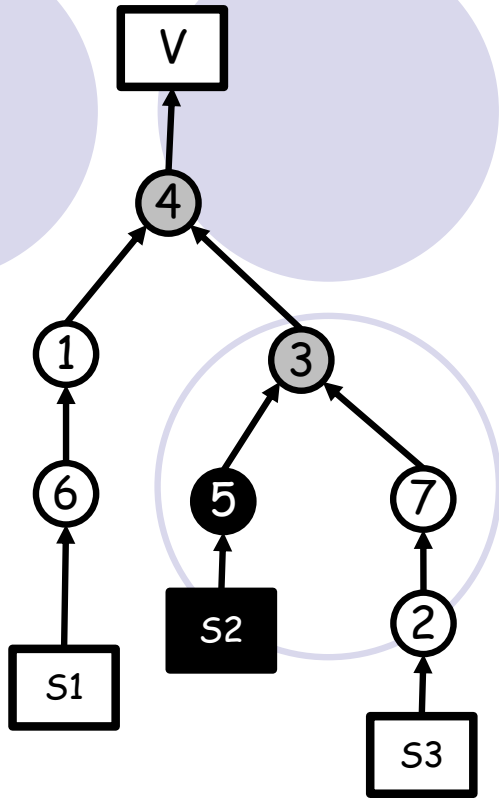
S3 3



After some more markings received

Identifying Attackers' IP

- Victim can identify attacker.
  - Statistical approaches, packet arrival time, entropy, etc.

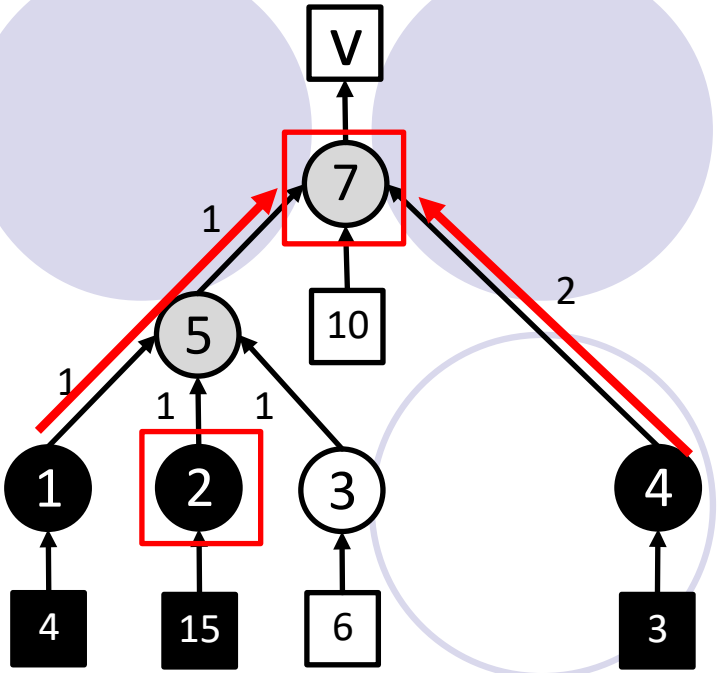


- Black=only attacker traffic
- White= only legitimate traffic
- Gray=mixed traffic

- The number of attackers is very large. Sending filters to all of them takes a lot of time.
- The capacity of filters in a FR is limited. So the hosting ISP of FR may charge money.

Problem1: Minimizing Contamination

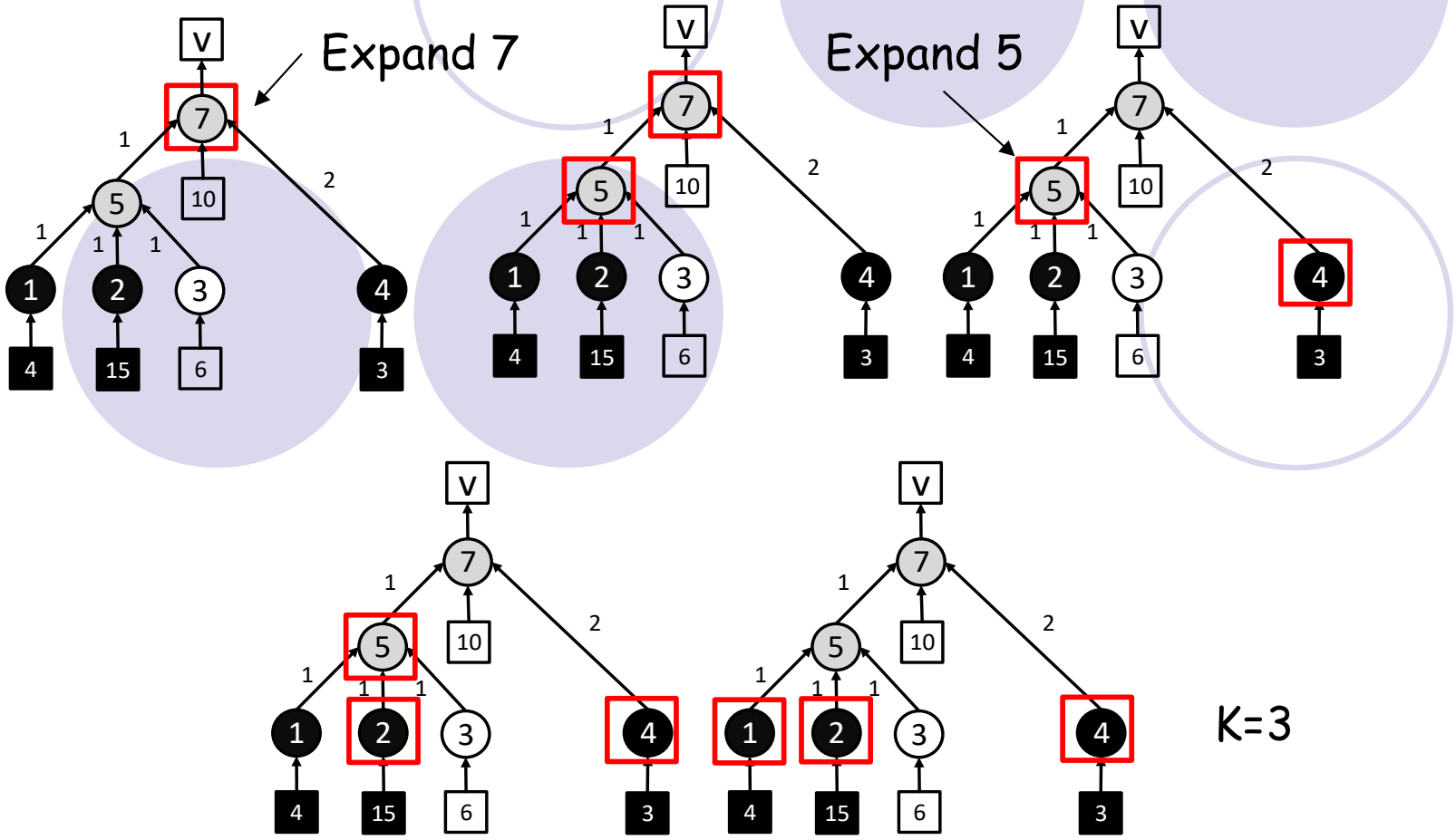
- Select K filters so that the contamination is minimum.
- Constraint: Block all attack traffic before it reaches v.
- Contamination Model
  - $C = \sum distace \times traffic\ load$
- Best assignment for k=2
  - {2,7}
  - $C = 4\times2 + 3\times2 = 14$



Problem complexity still unknown.

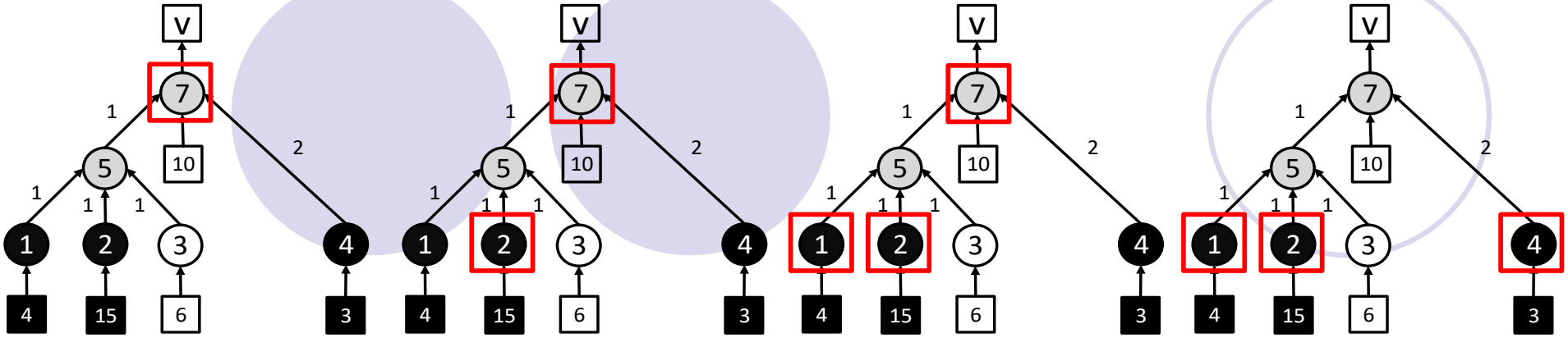
Naive Approximation (Top-down)

- Start from the root. Expand node with highest $\frac{\text{total traffic load}}{\text{number of branches}}$ until K number of filters are assigned.
- Complexity: $O(K^2)$



Greedy Approximation 1

- Start from the root. Pick the highest weighted node and recalculate weight. Continue until K nodes are picked. Remove already covered nodes.
- Weight=distance_to_the_first_filter x load
- Complexity: $O(NK)$



1	2	4	5	7
8	30	6	19	0

Select root

1	2	4	5	7
8	0	6	4	0

Select 2

1	2	4	5	7
0	0	6	0	0

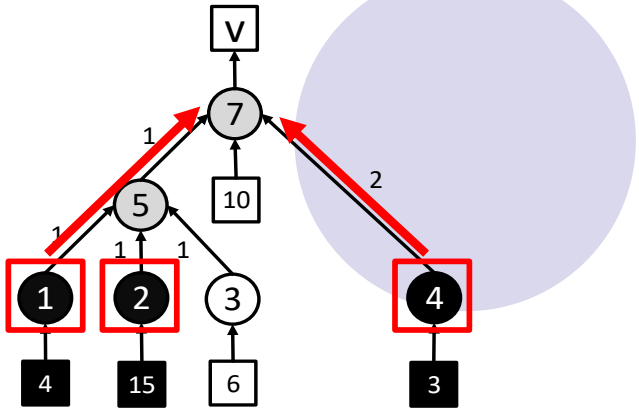
Select 1

1	2	4	5	7
0	0	0	0	0

Select 4, remove 7

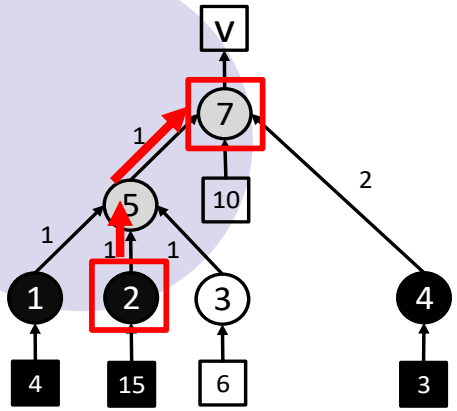
Greedy Approximation 2 (Bottom-up)

- Start by selecting all non-white entry nodes. Continue merging a pair of filters which add least penalty until the total assignment is K and put the merged filter on their least common ancestor.
- Complexity: $O(N^2(N-K))$
  - Using heap: $O((N-K)^2 \log N)$



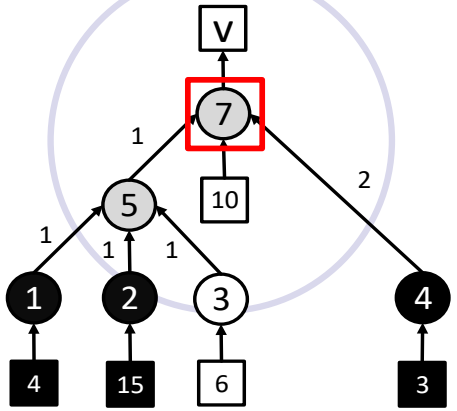
Penalty (1,2)=4+15
Penalty (2,4)=6+15x2=36
Penalty (1,4)=4x2+3x2=14

K=3



Penalty (2,7)=15x2+0=30

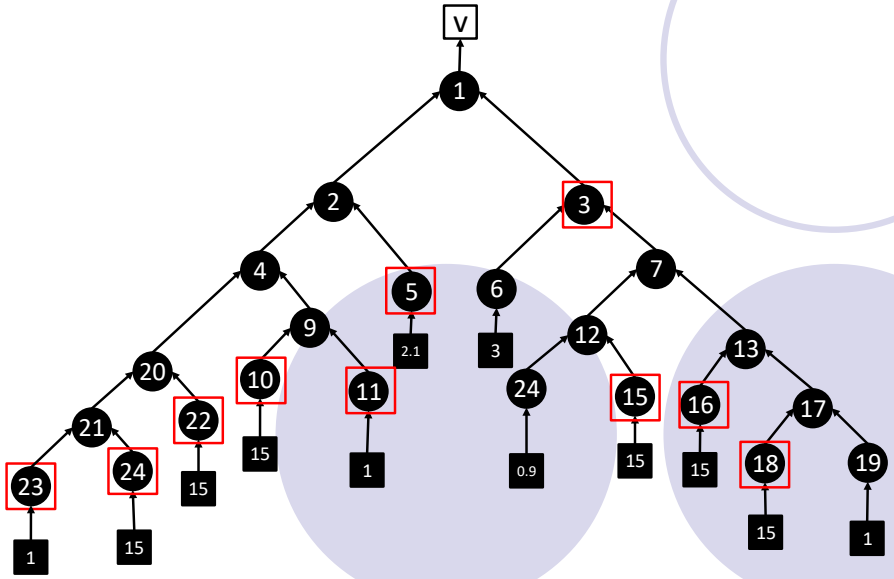
K=2



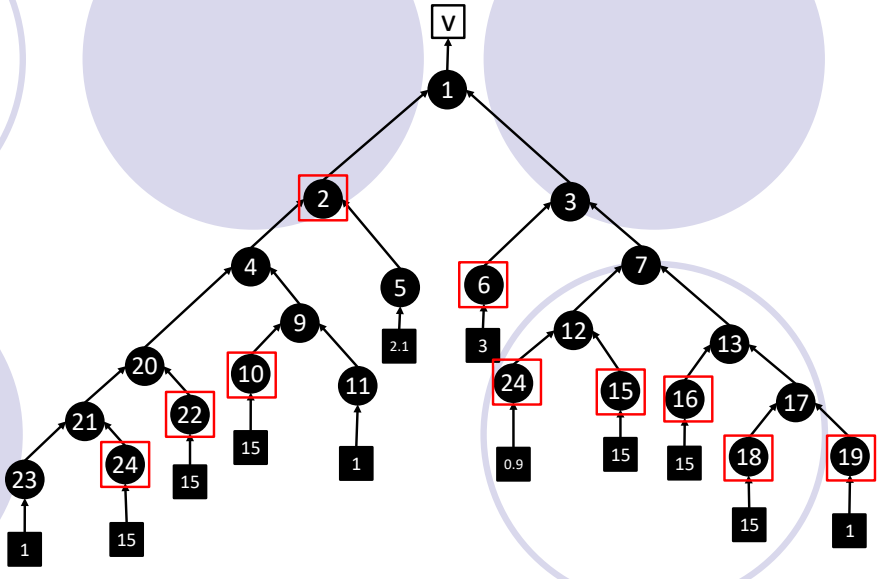
K=1

Penalty= Amount of contamination increase for a merge.

Greedy Approximation 2 is Not Optimal



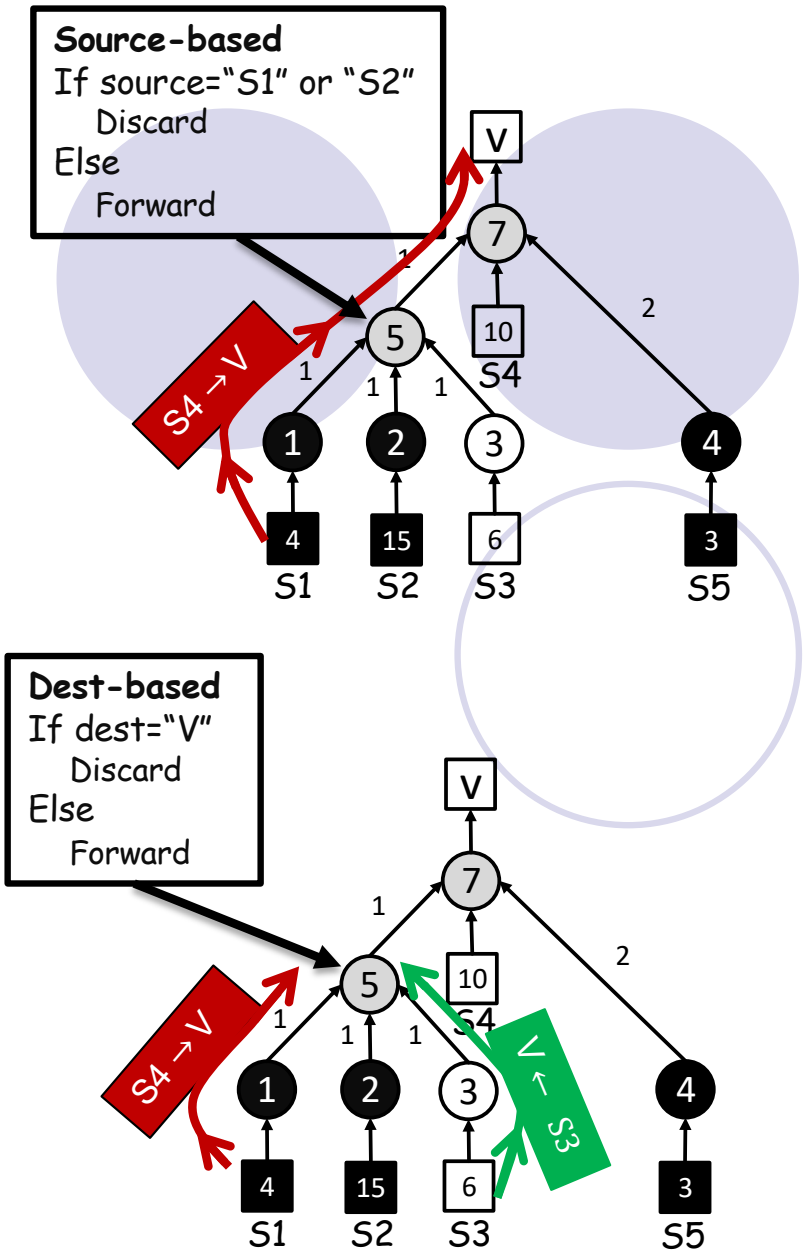
Greedy Approximation 2
K=10
Contamination=9.7



Optimal
K=10
Contamination=9.1

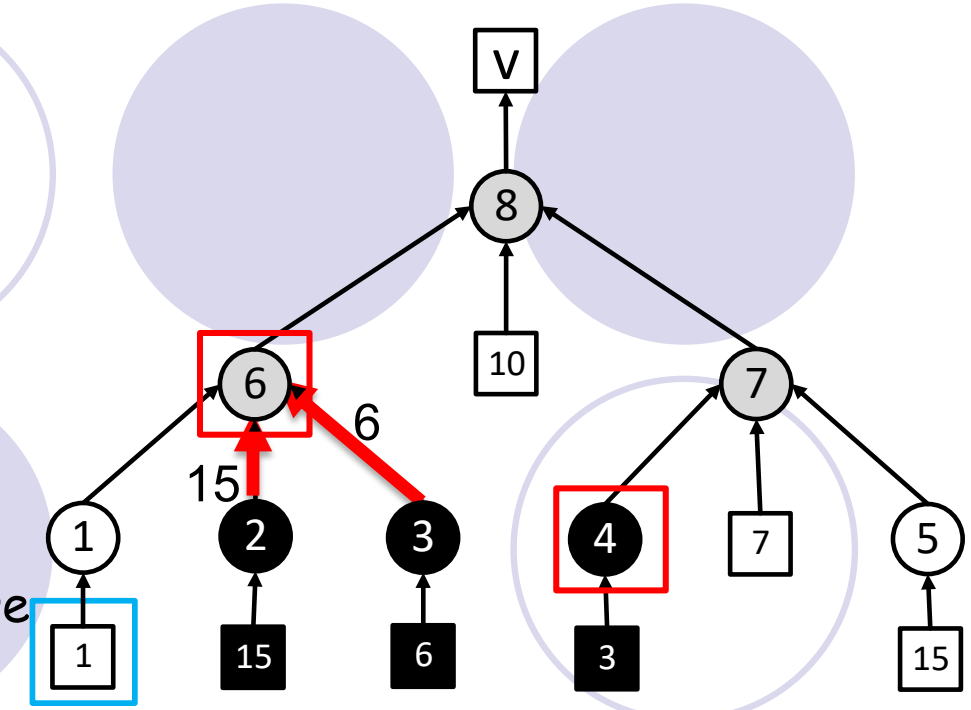
Source-based and Destination-based filters

- Source-based filter
  - Filter by source address of packet.
  - Cannot protect IP spoofing DDoS.
- Destination-based filter
  - Filter by destination address of packet.
  - Can protect against IP spoofing DDoS.
  - Blocks legitimate traffic.

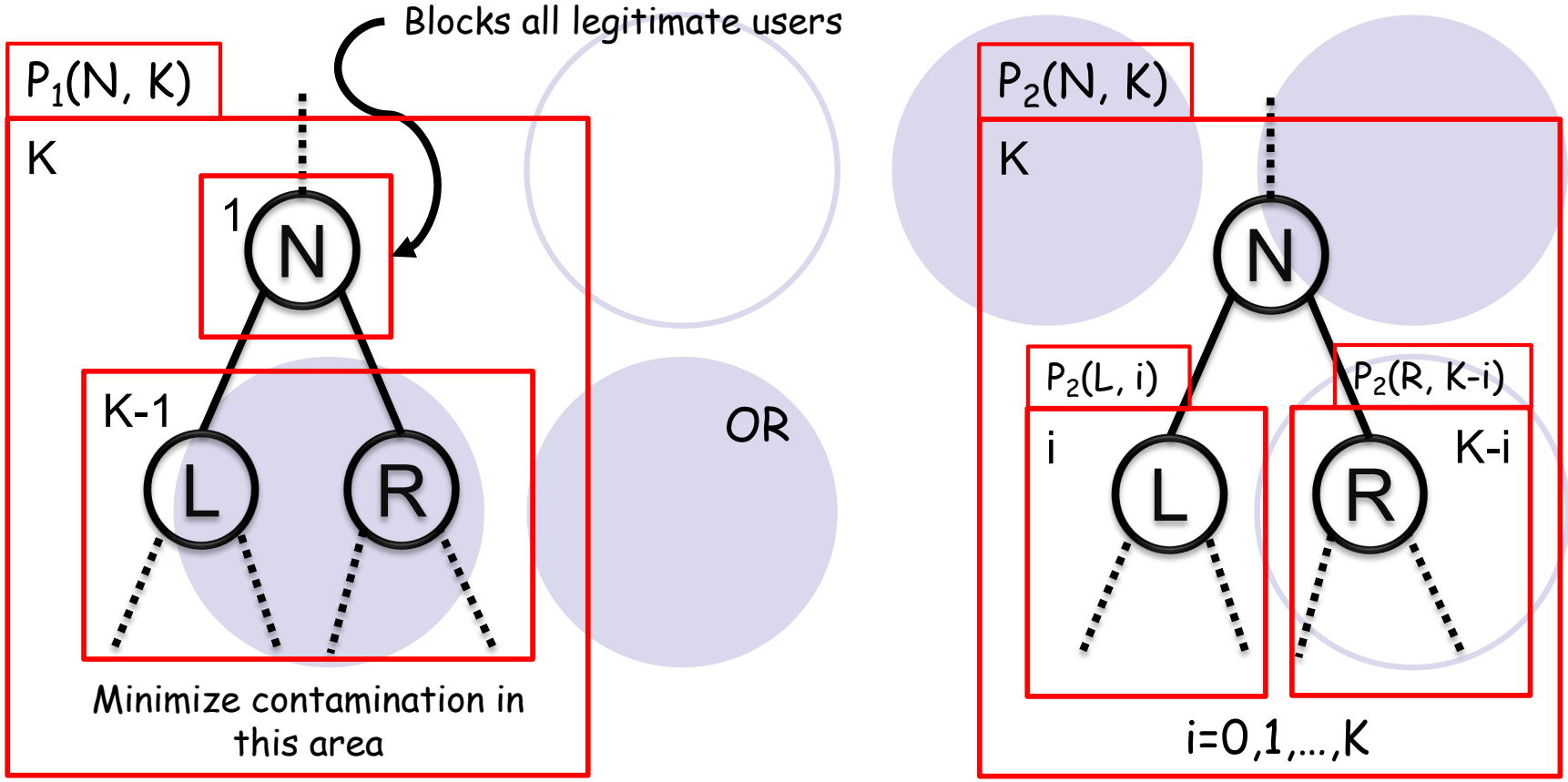


Problem 2: Minimizing Contamination and Blocked Legit Users

- Given ω and topology, select K filters so that C is minimum.
- Cost model
  - $C = \omega \times C_1 + (1-\omega)C_2$
  - C_1 = Contamination
  - C_2 = Number of blocked legit users
- Constraint
  - Block all the attack traffic before reaching v.
- Best assignment for k=2 is {6,4}
  - $\omega = 0.5, C_1 = 21,\ C_2 = 1$
  - $C = 0.5 \times 21 + (1-0.5) \times 1 = 11$



A dynamic programming solution



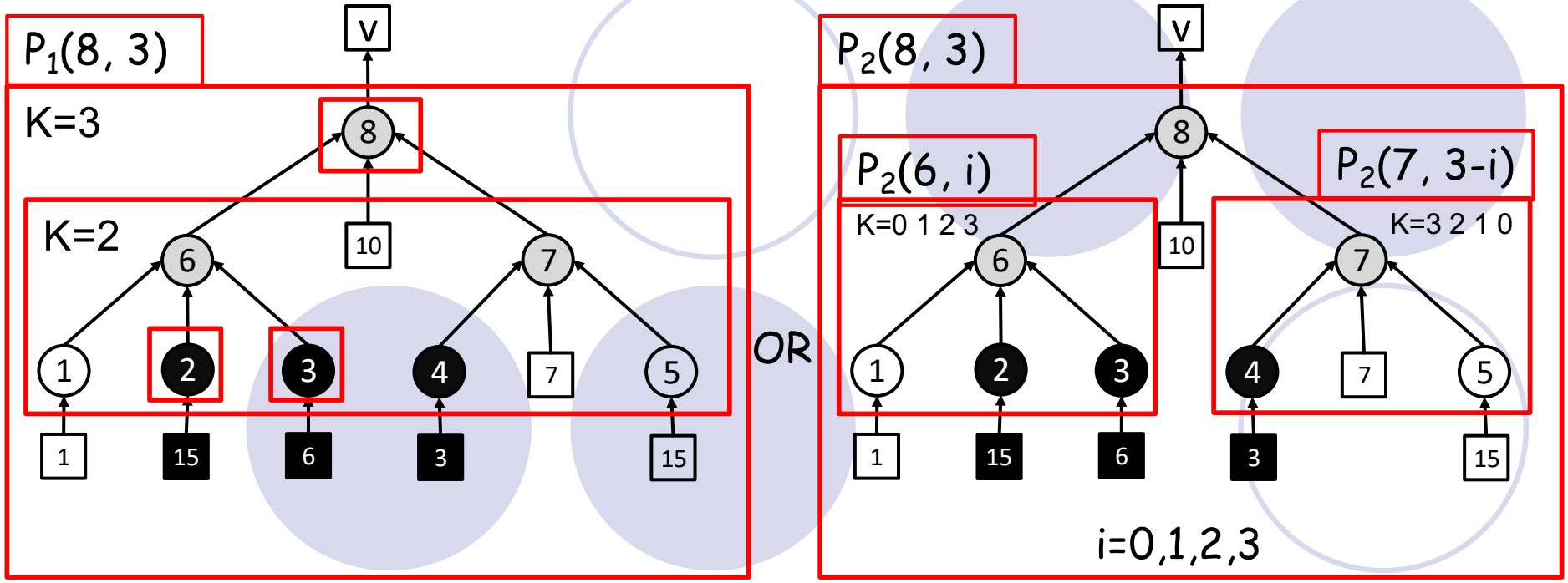
In subtree rooted by N for K filters:

$P_1(N, K)$= Minimum contamination rooted at N.

$P_2(N, K)$= Minimum cost.

Complexity: $O(NK^{D-1})$, where D: node degree.

A Dynamic Programming Solution: An Example



Greedy Approximation 2 : {2,3,8}

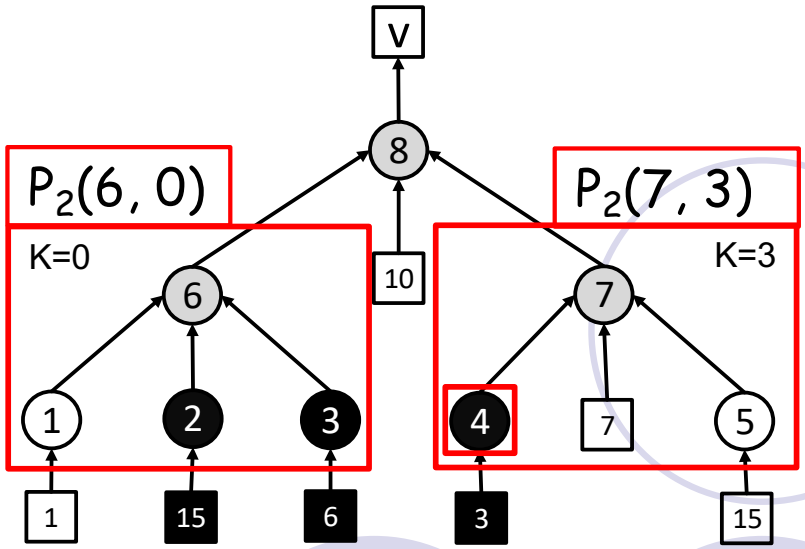
$P_1(8,3)$= 3x2=6

L(8)=1+7+15=23, L(N): number of eligt users rooted at N

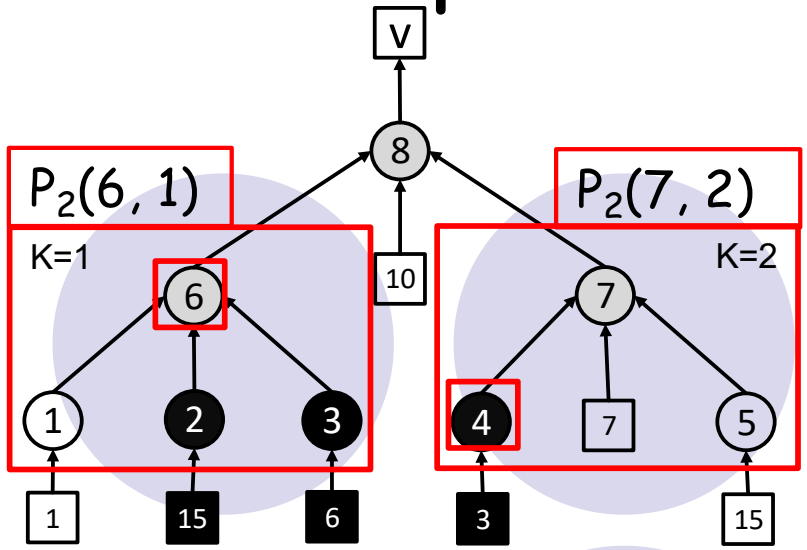
$$\text{Cost} = \frac{1}{2}23 + \left(1 - \frac{1}{2}\right)6 = \mathbf{14.5}$$

0

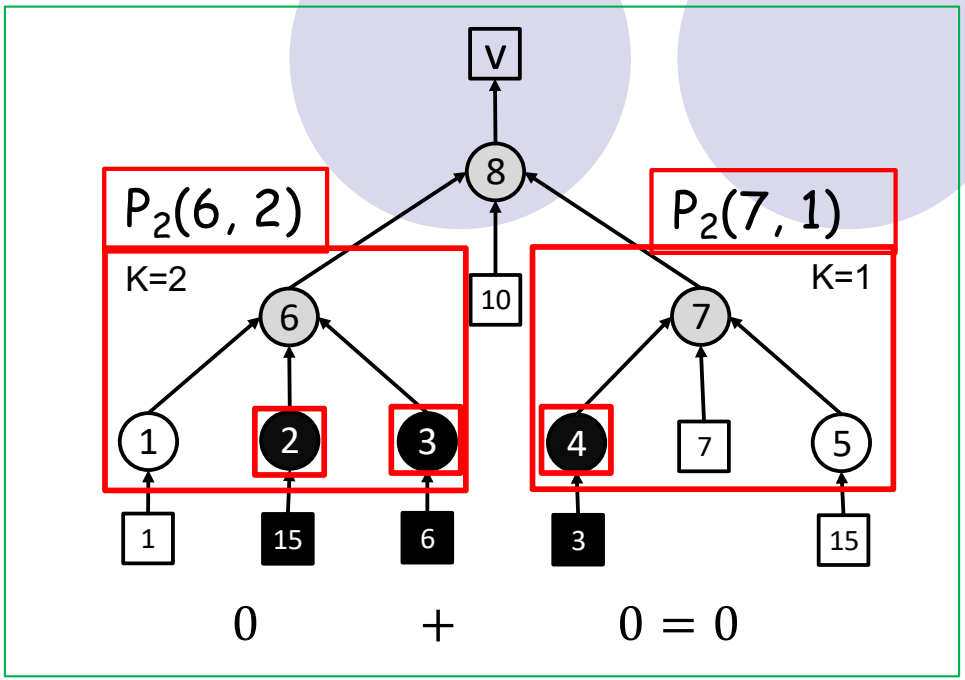
A DP Solution: An Example



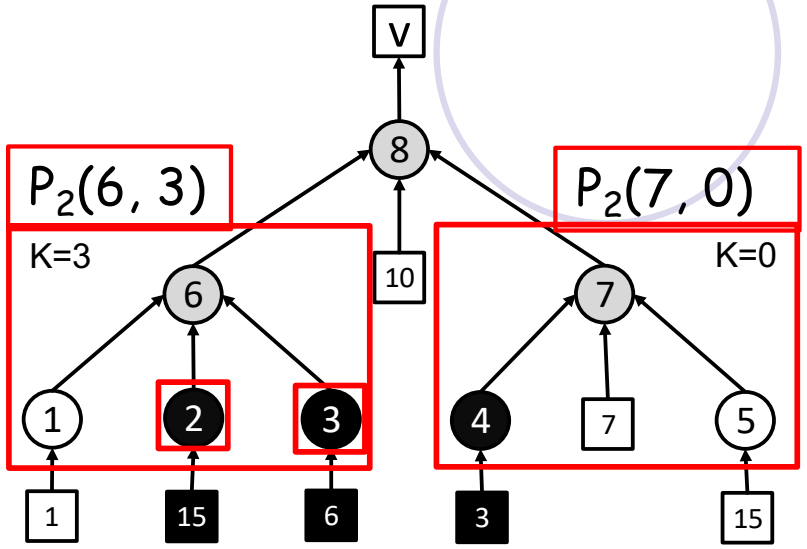
$\infty + 0 = \infty$



$11 + 0 = 11$

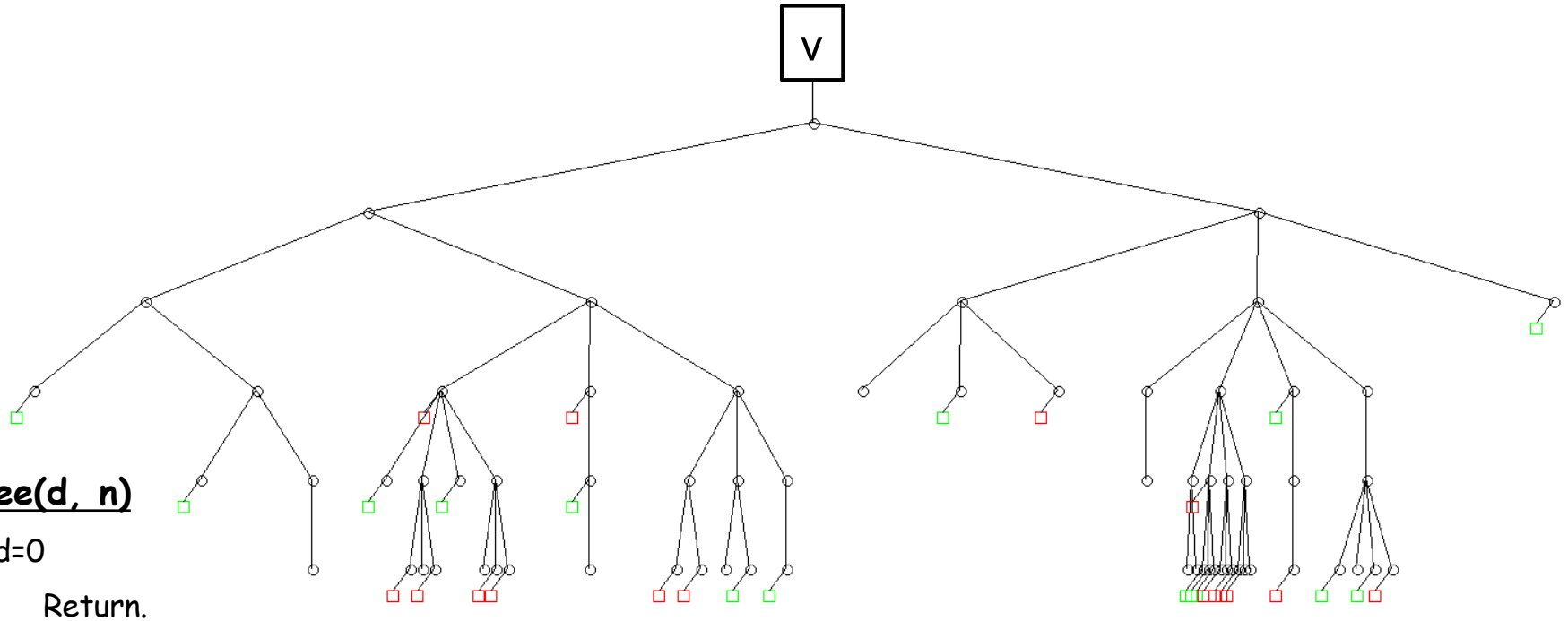


$0 + 0 = 0$



$0 + \infty = \infty$

Simulation: Random Tree Generation



Tree(d, n)

If d=0

Return.

Else

For i=0 to rand [0, Δ]

Create node c_i.

Make c_i child of n.

Tree(d-1, c_i)

Topology: 1

# of nodes : 66

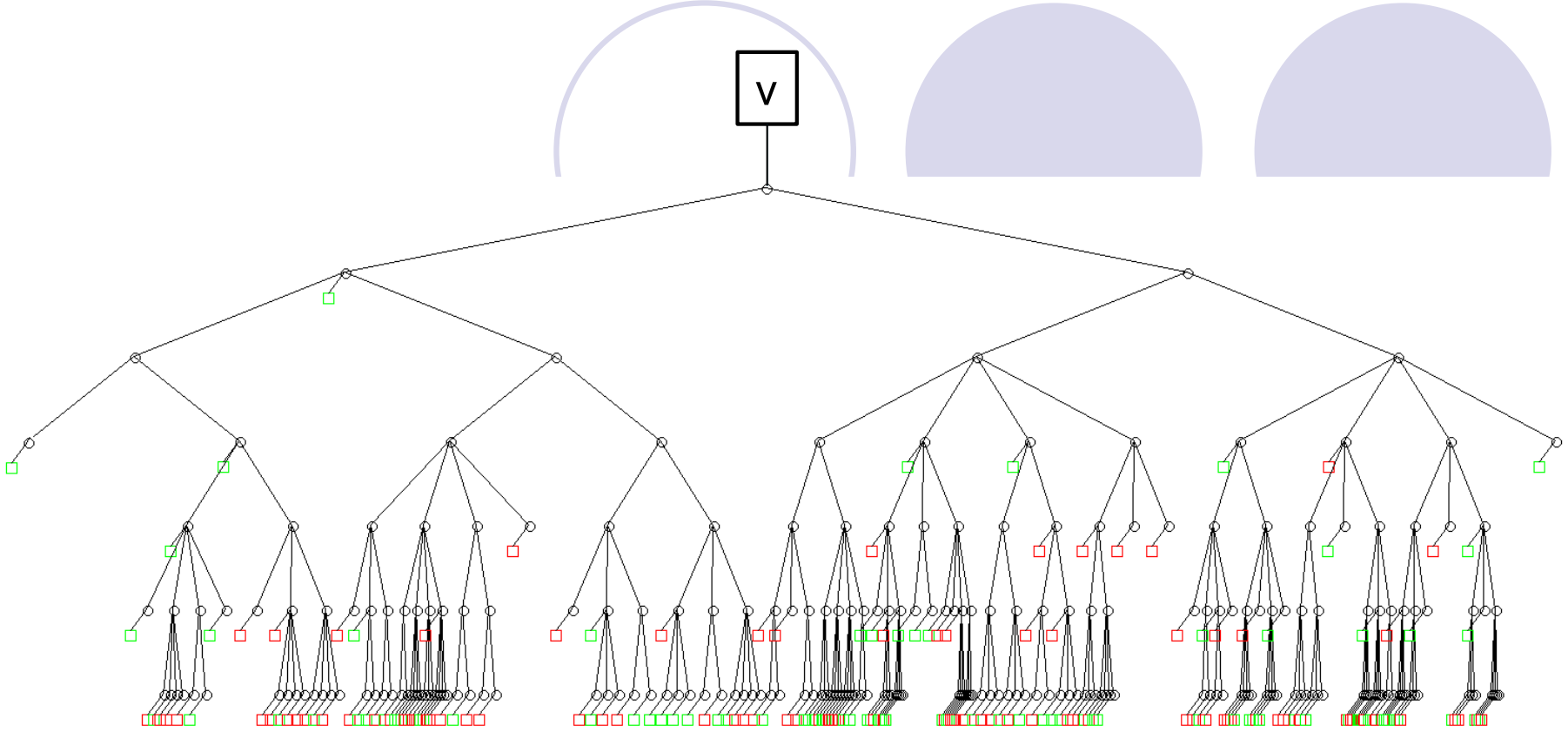
Attacker ratio: 50%

Maximum degree=4

Depth=5

Data rate= 1 to 4

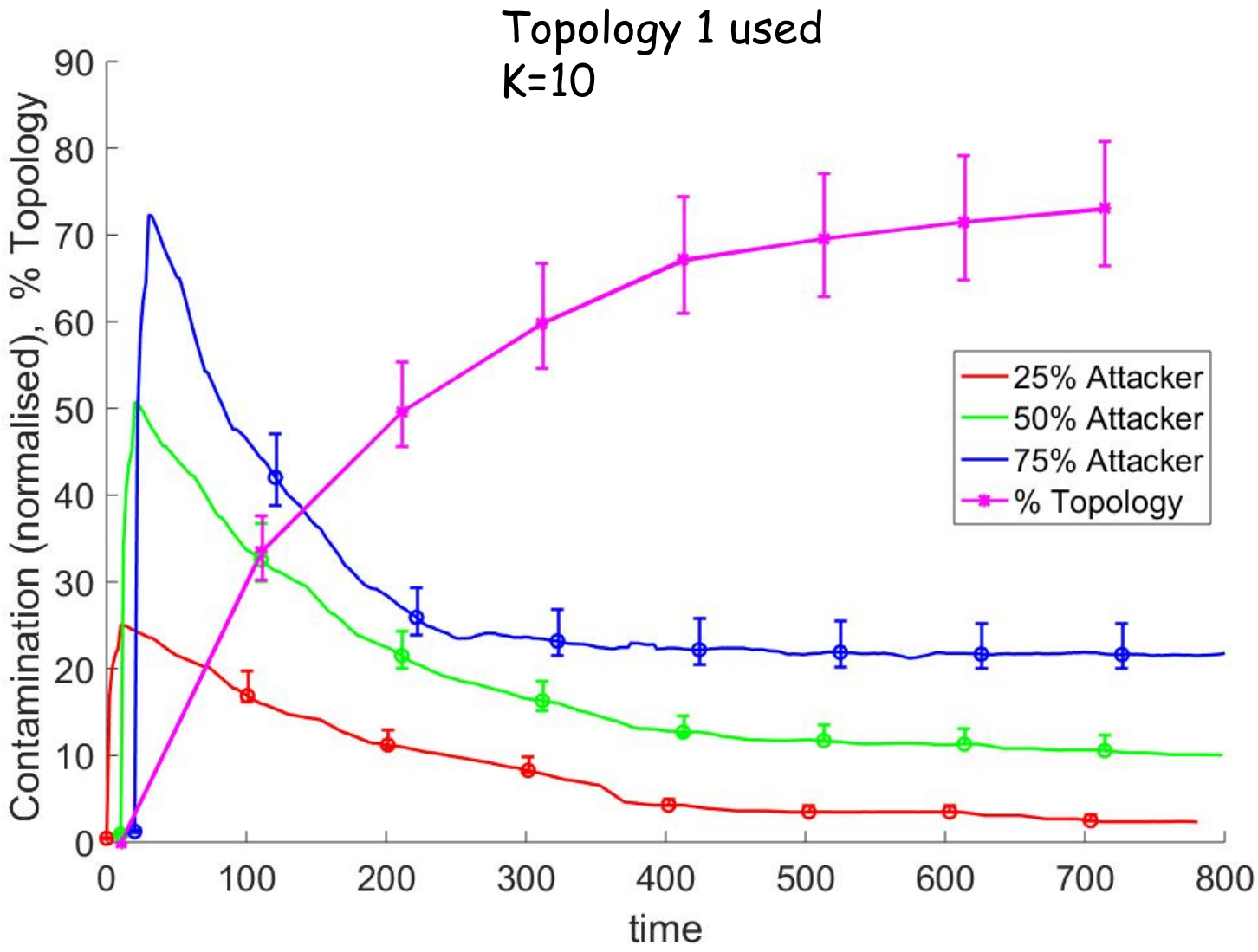
Simulation: Random Tree Generation



Topology: 2
# of nodes : 250
Attacker ratio: 60%

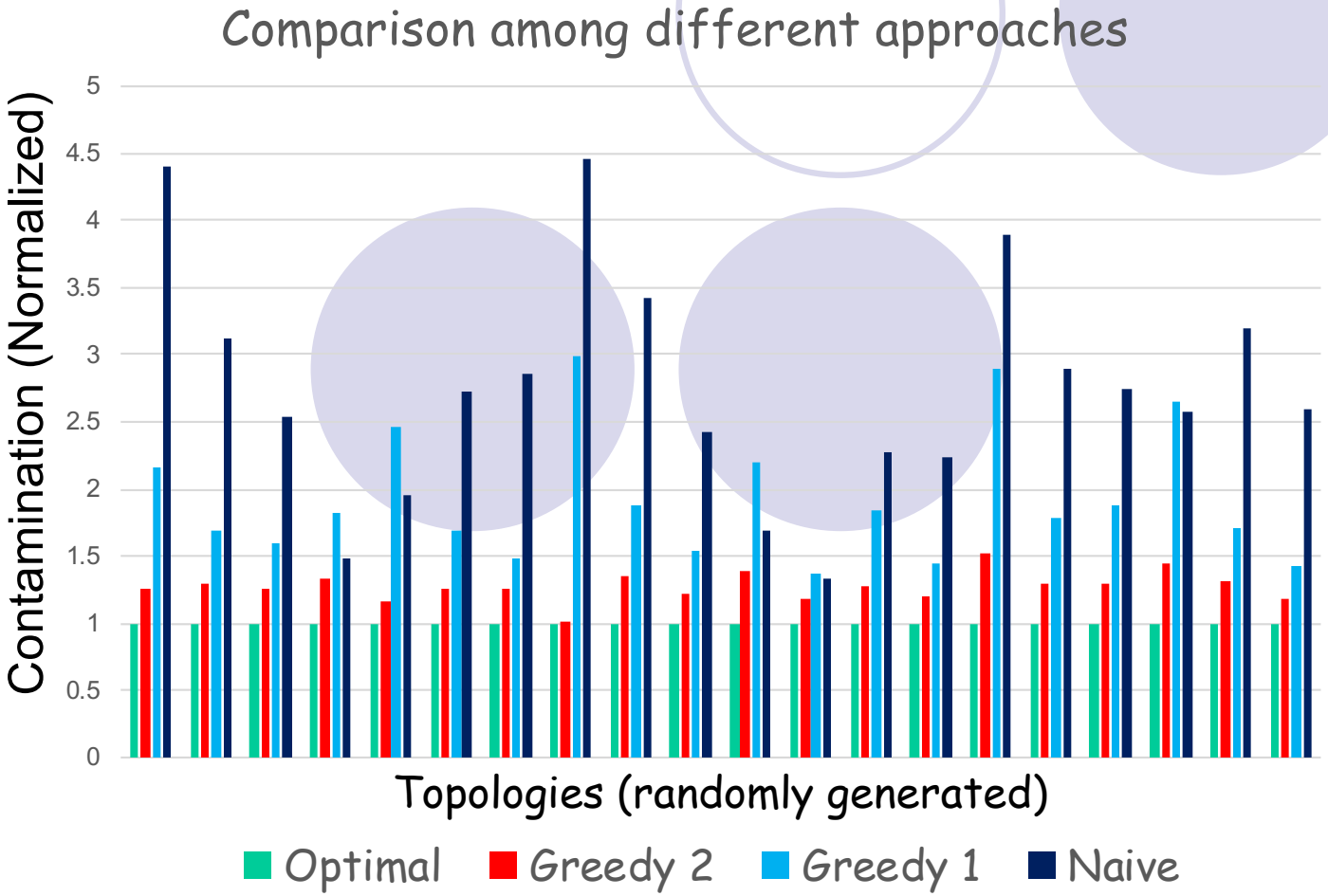
Maximum degree=4
Depth=6
Data rate= 1 to 10

Problem 1: Greedy 2 Timeline



Problem 1: Different Approaches

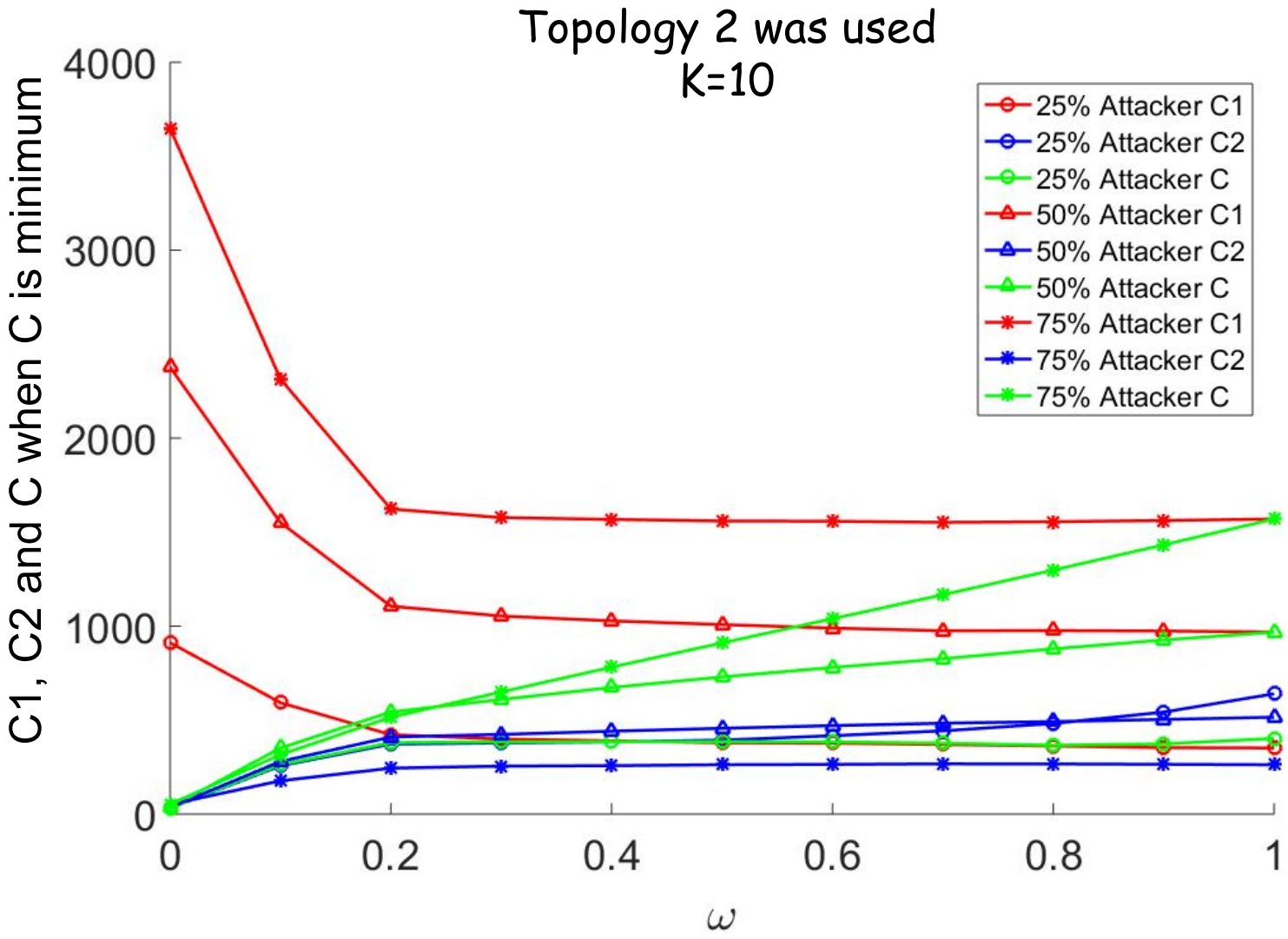
Subset of 200 Topologies are shown



Greedy 1: 43% more
Greedy 2: 26% more
Naive: 167% more

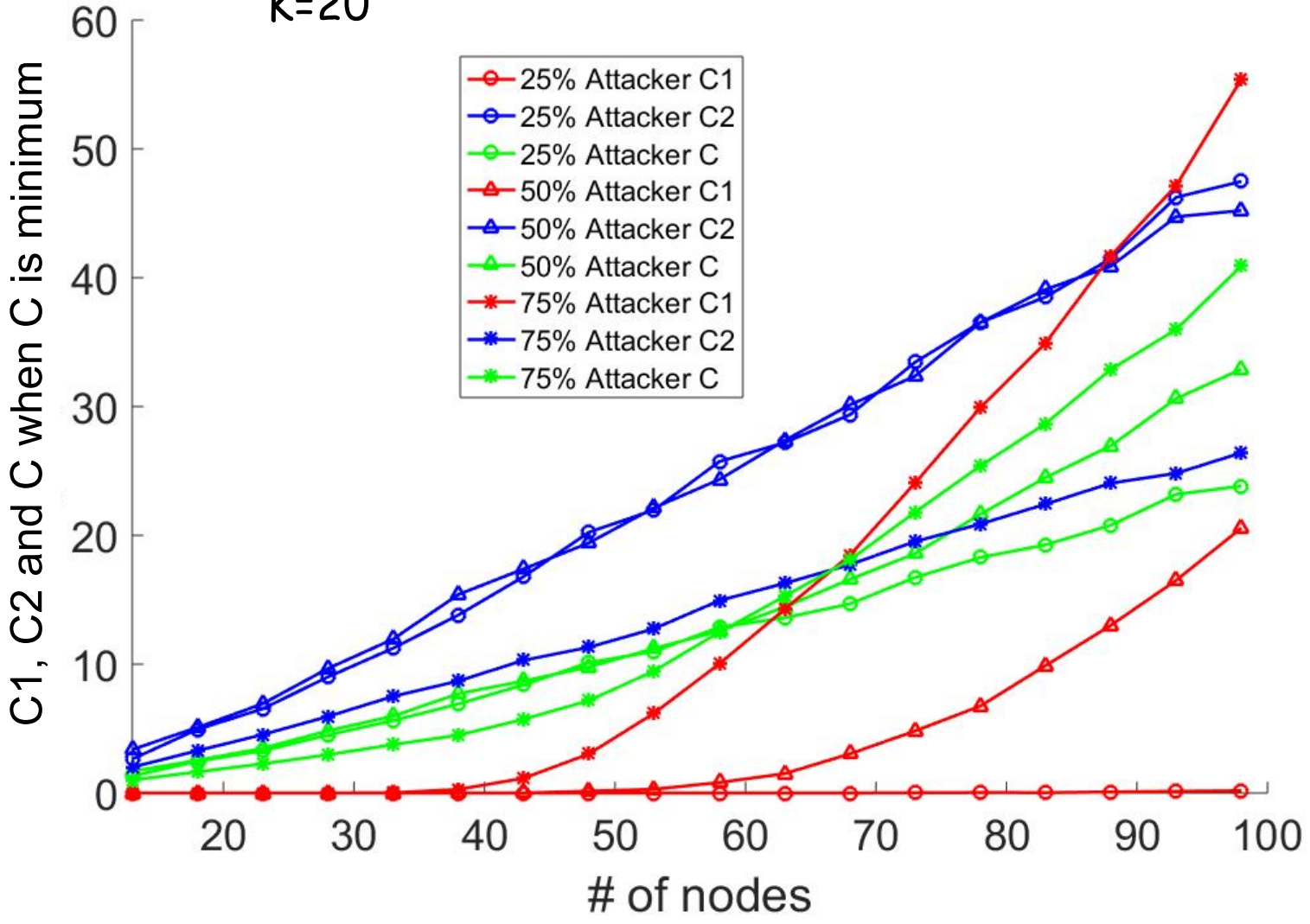
Samples=200
Nodes=25-35
Data rate=1-3
Max depth=4
Max degree=3
Attacker ratio= 50%

Problem 2: Effect of ω

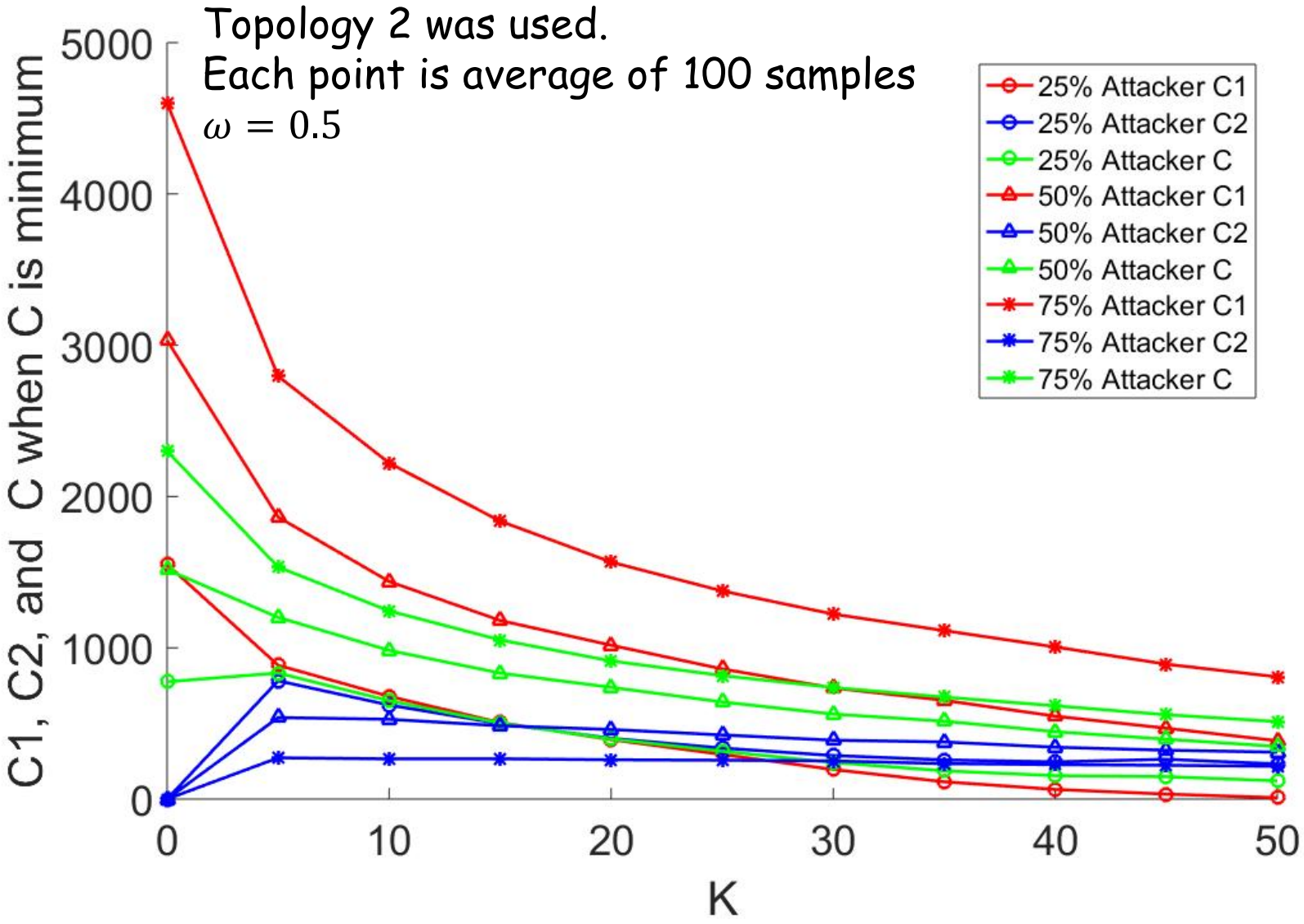


Problem 2: Effect of # of Nodes

Randomly generated topologies were used.
Each point is average of 100 samples
K=20

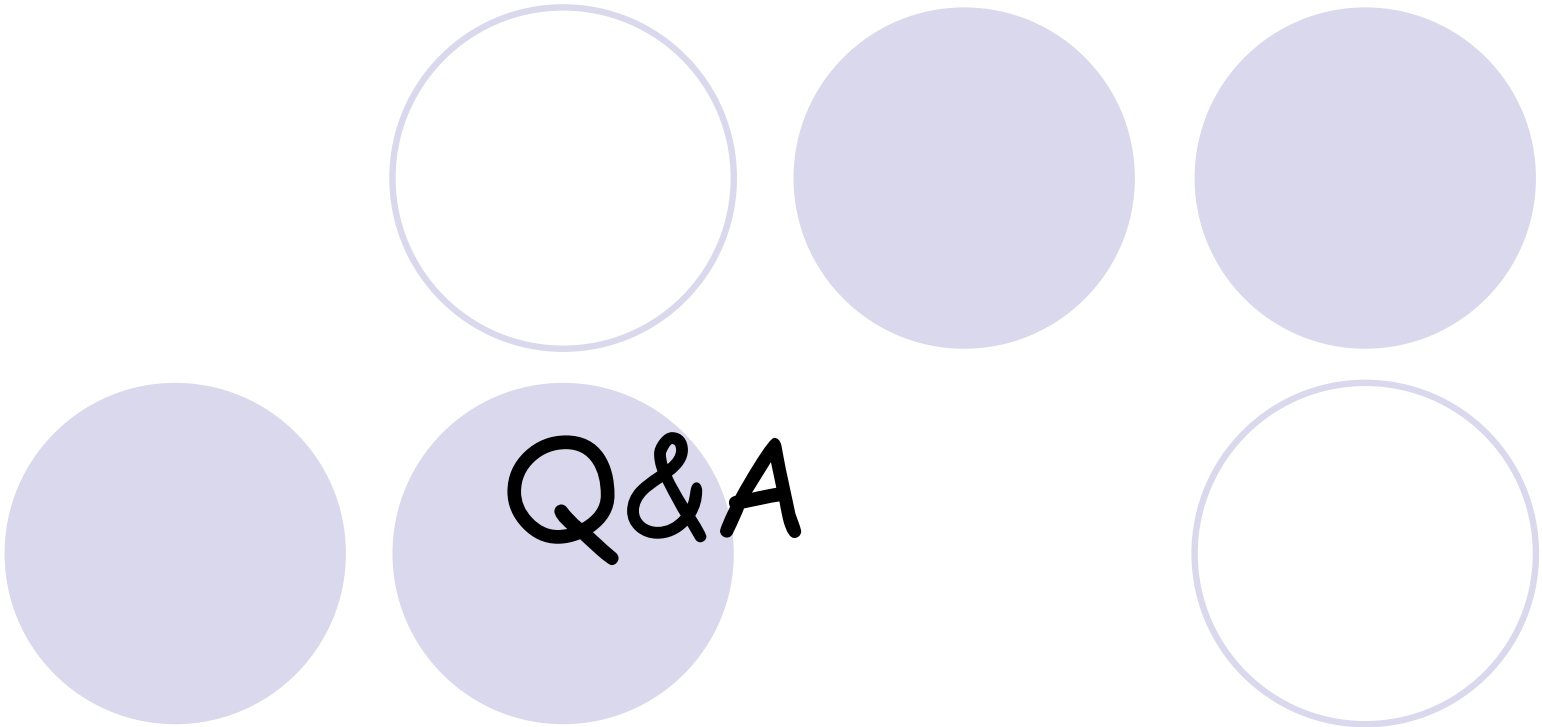


Problem 2: Effect of K



Summary and Future Work

- Two unique filter assignment problems
  - Problem 1: Source based
  - Problem 2: Destination based
- The greedy approximation 2
  - The best solution for Problem 1
- Optimality of DP solution for problem 2
  - Depends on optimality of problem 1


Q&A