On an Efficient Computational Resource Scheduling for Prioritized Multiple Target Contents Searching in Large-scale Networks

Chikara Miura, Shinji Sugawara, Yutaka Ishibashi

Nagoya Institute of Technology, Japan

2012 Annual IEEE CQR International Workshop
May 15-17,2012, San Diego, USA

Outline of Presentation

- Introduction
- Objective
- Assuming Conditions
- Proposed Scheme
- Comparison Schemes
- Simulation Conditions
- Simulation Results
- Conclusion

Introduction (1/2)

A volume of and variety of contents are exchanged and accumulated on the Internet

Search of contents is becoming difficult

Demand an effective searching for Information

Introduction (2/2)

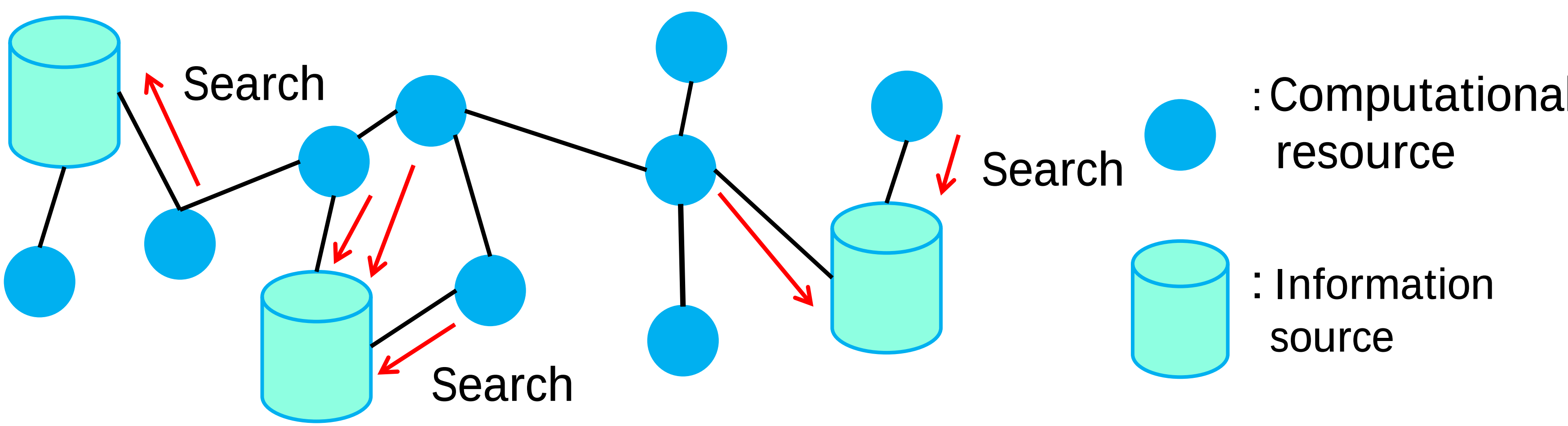
Necessity of a lot of processing capacity

A lot of computational resources scattered over the Internet (ex; Idle personal computers)

Use the resources to get a large amount of processing power for searching contents

Objective (1/2)

- Our previous Study
  - Searching multiple target contents directly
  - Using computational resources connected to the Internet
  - Computational resources are assigned and scheduled to search on the information sources



Objective (2/2)

However…

Priority of the content searching request has not been considered

- This Study

We propose an efficient scheme of computational resource scheduling for prioritized multiple target contents searching

Assuming Conditions (1/3)

- Information Sources
  - The total quantity of possessing contents varies by information source
  - The existence probability of each target content item in each information source can be estimated
  - Data processing labor for a target searching on an information source can be divided into some resources
- Computational Resources
  - Computational capacity generally varies by computational resource

Assuming Conditions (2/3)

- User
  - A user demands multiple target content simultaneously and their searching priorities are different in general.
- Switching Time
  - If a combination of computational resources and information sources is changed for the next investigation, it takes a certain switching time.

Assuming Conditions (3/3)

- Time Cost
  - The total elapsed time weighted by the user's utility from the moment the searching request of the item is issued until instant the searching content is discovered
- Communication Cost
  - The total amount of the transmitted data

Minimize the sum of time cost and communication cost

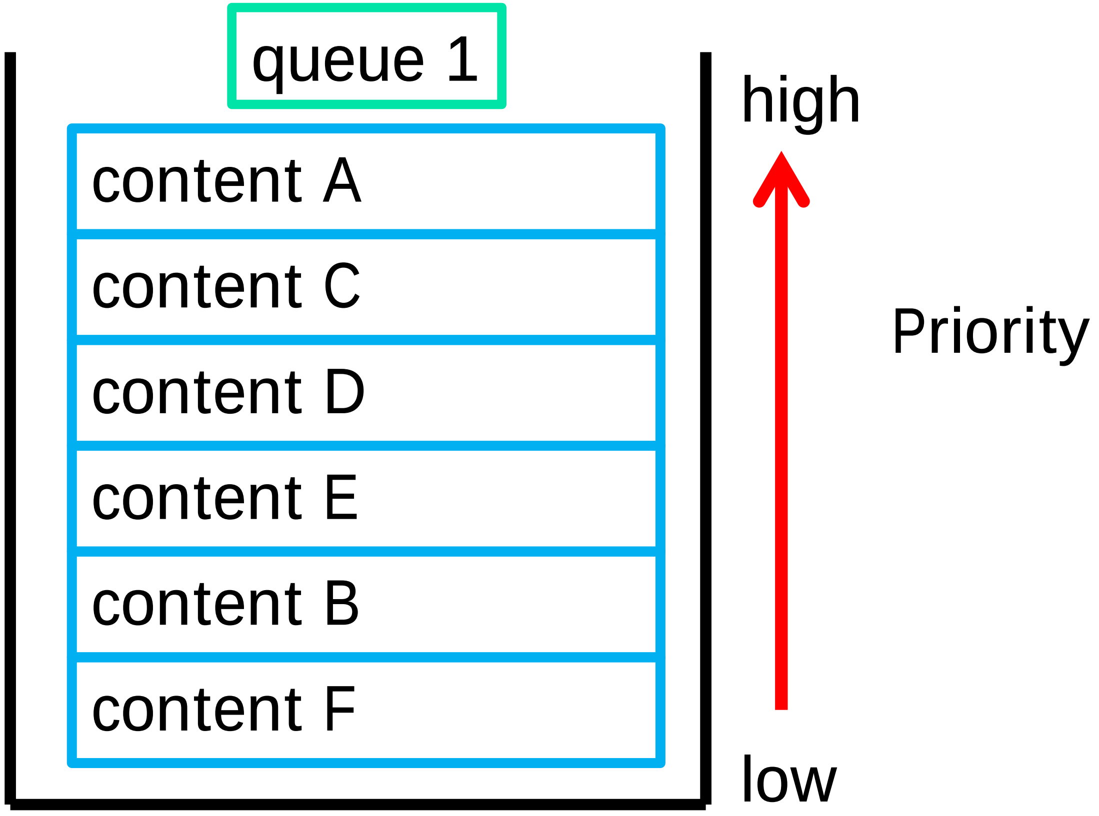
Proposed Scheme (1/3)

Proposed scheme has three scheduling parameters α, β, γ

- α : The number of target content items searched scheduled for search at a time
- β : The number of computational resources allocated to an information source
- γ : The upper limit of the total number of computational resources used simultaneously

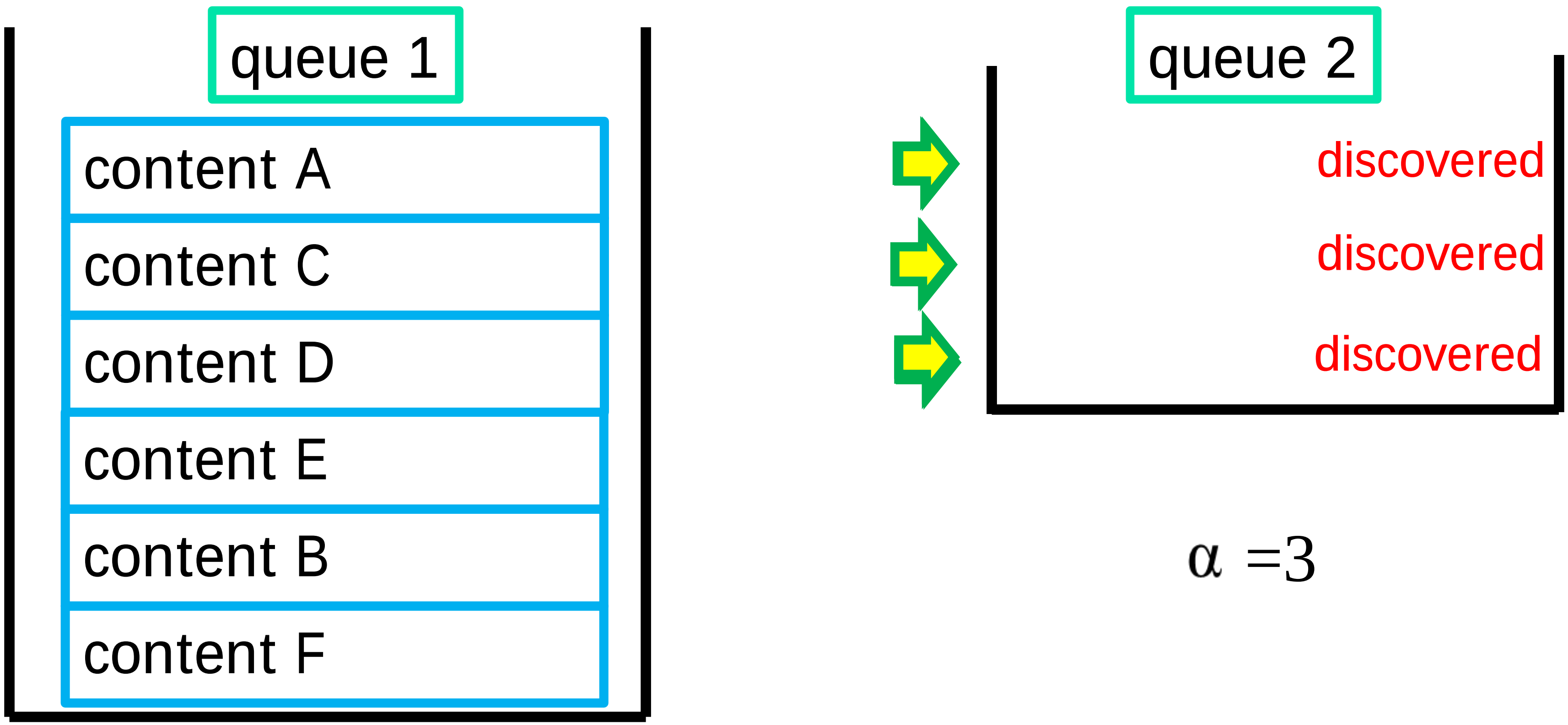
Proposed Scheme (2/3)

Target content items requested by the user are queued in queue 1 in descending order of priority.



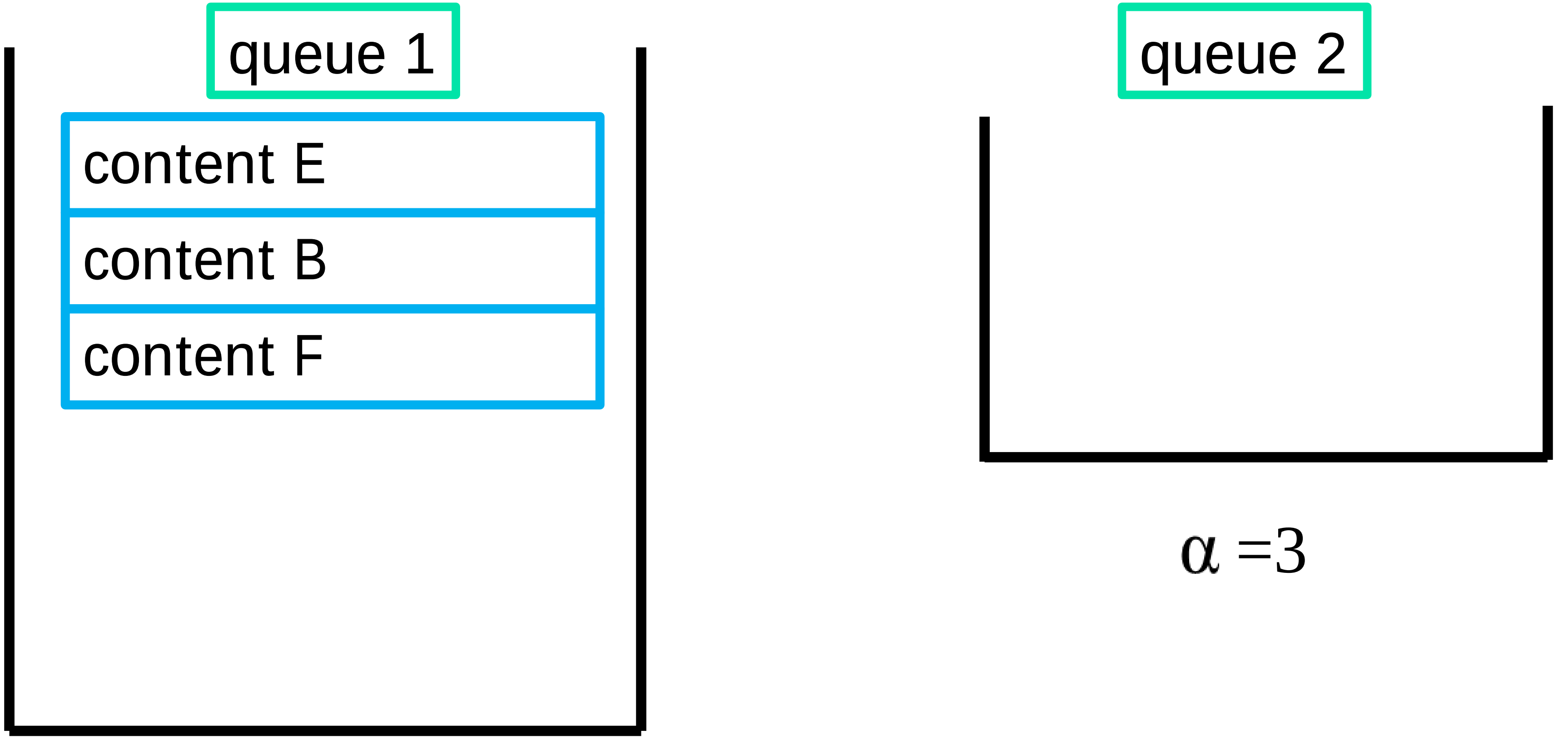
Proposed Scheme (2/3)

Pick up the top α items of queue 1 and queue them in the same order in queue 2.
Content items in queue 2 are searched by round robin manner.



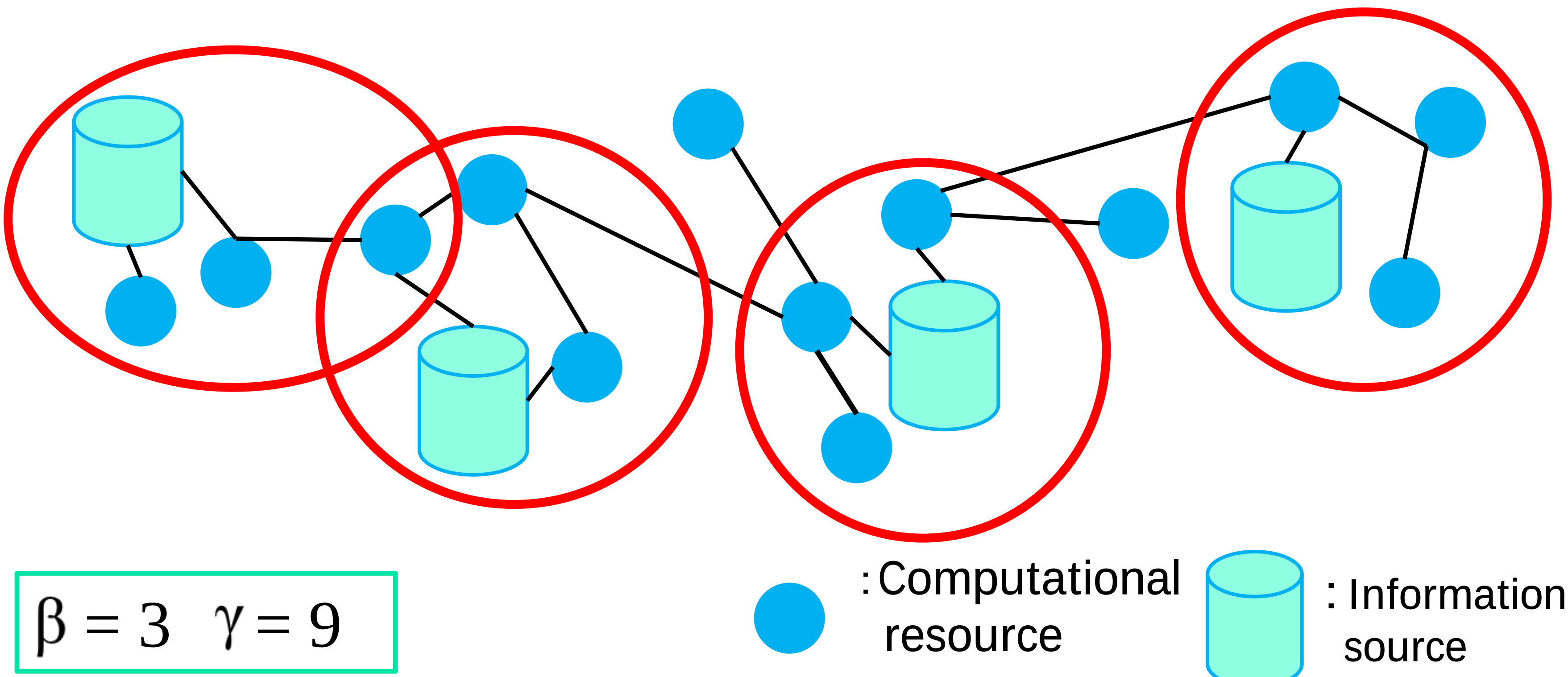
Proposed Scheme (2/3)

After discovering all the content in queue 2, pick up the top α items of queue 1 and queue them in the same order in queue 2 again.



Proposed Scheme (3/3)

Allocate neighboring computational resources to information source with the highest existence probability of the target until the number of assigned resources reaches β or γ.



$\beta = 3 \quad \gamma = 9$

Comparison Schemes (1/2)

- Cluster scheduling

  Some neighboring computational resources form a cluster, and each cluster searches for information sources.

- Random cluster scheduling

  Some computational resources picked up at random from the network form a cluster, and each cluster searches for information sources.

Comparison Schemes (2/2)

- TOS（Time cost Oriented Scheme）

Calculate the product of searching priority and existence probability for every target in every information source. And search the source in descending order of the product by using all the computational resources.

- COS（Communication cost Oriented Scheme）

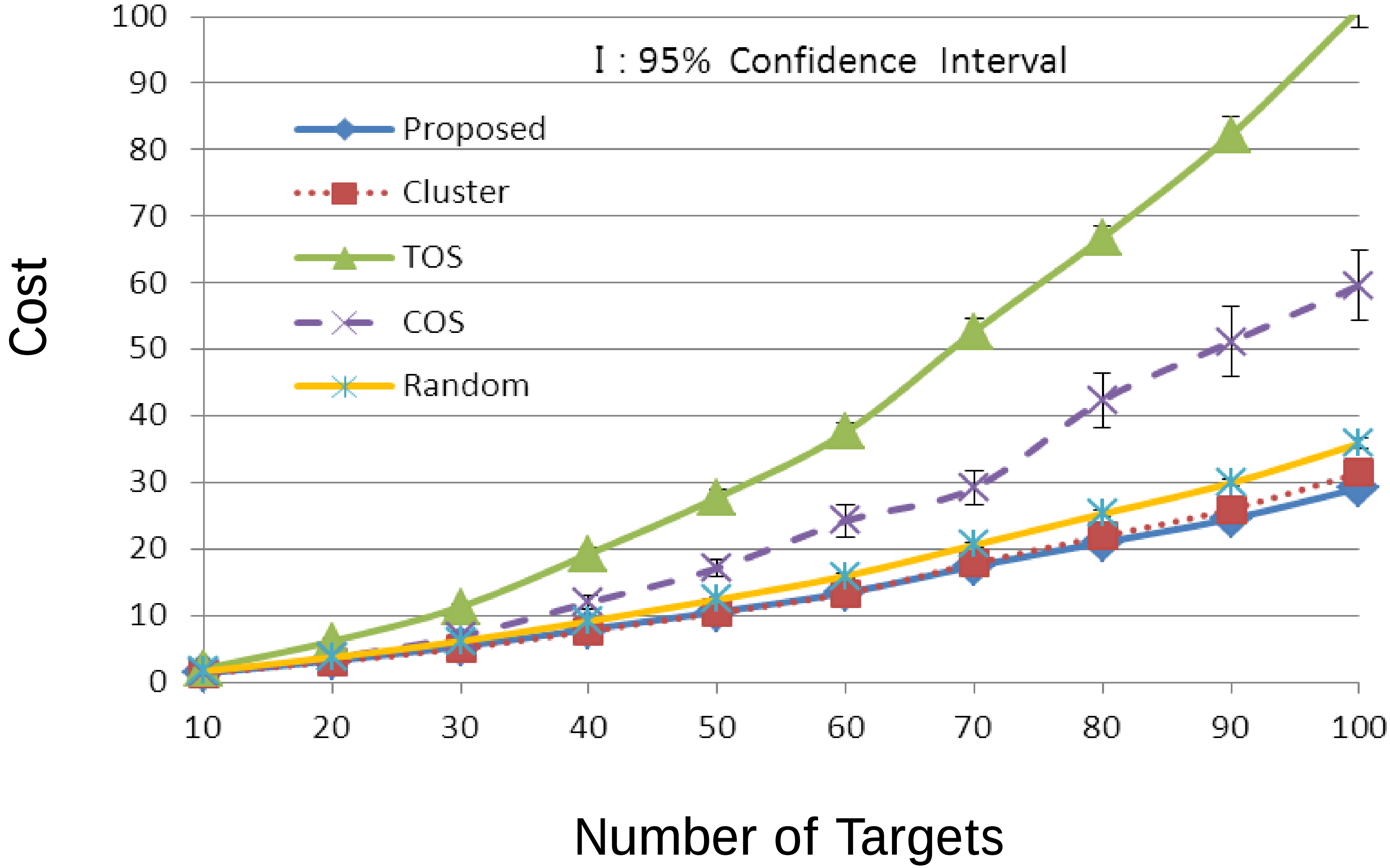
Check the highest existence probability of any target in every information source. And search the source in descending order of the probability by using the nearest single computational resource.

Simulation Conditions

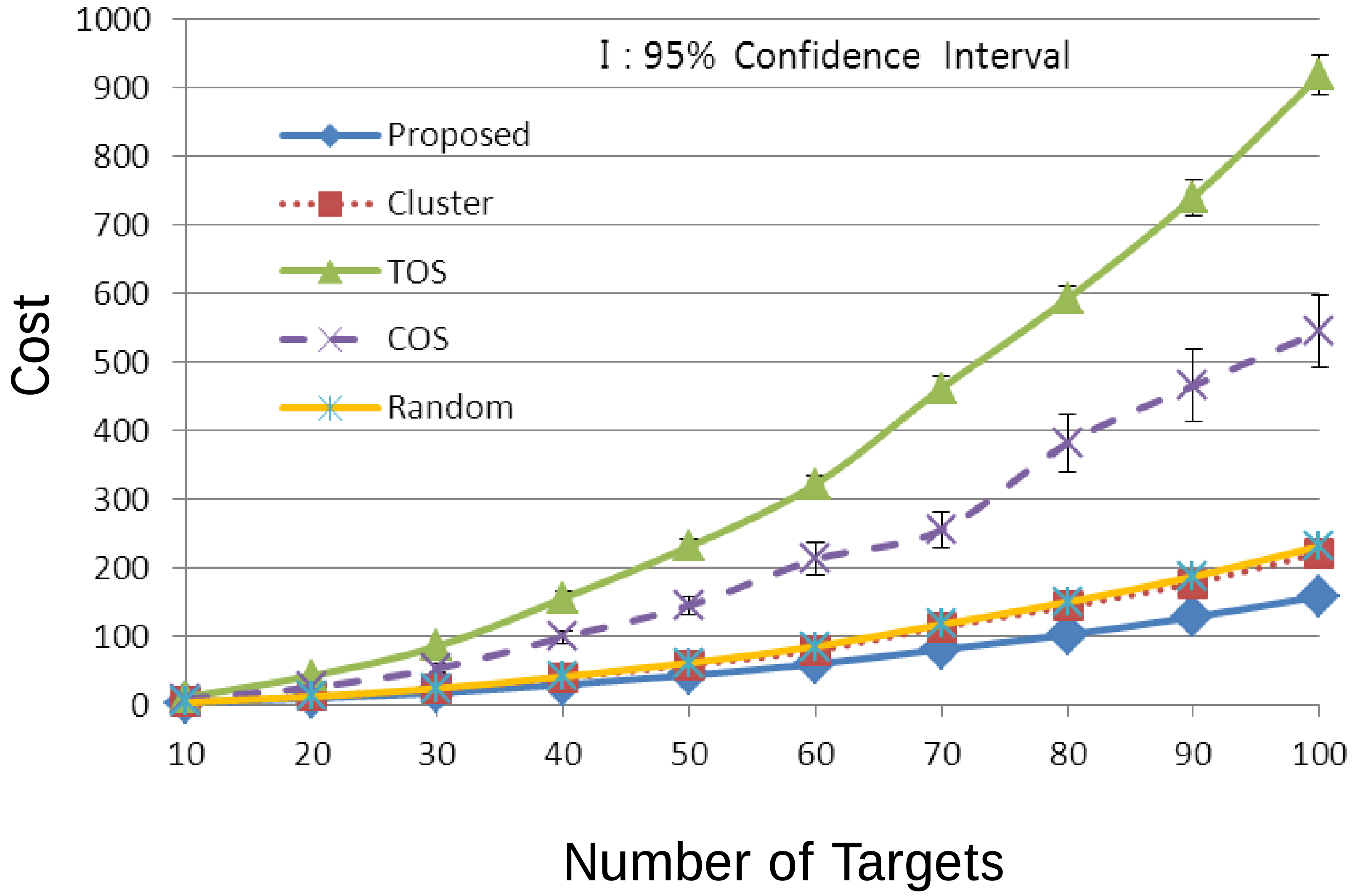
Network topology	Barabási Albert model
Number of nodes	5000
Number of information sources	900
Number of computational resources	100
Volume of information sources	min=10, max=1000
Processing capacity of computational resources	min=100, max=1100
Searching priority for targets	min=1, max=10
Probability of target existence	min=0.01, max=0.10
Switching time	5

→ Distribution of the values is based on Zipf's law

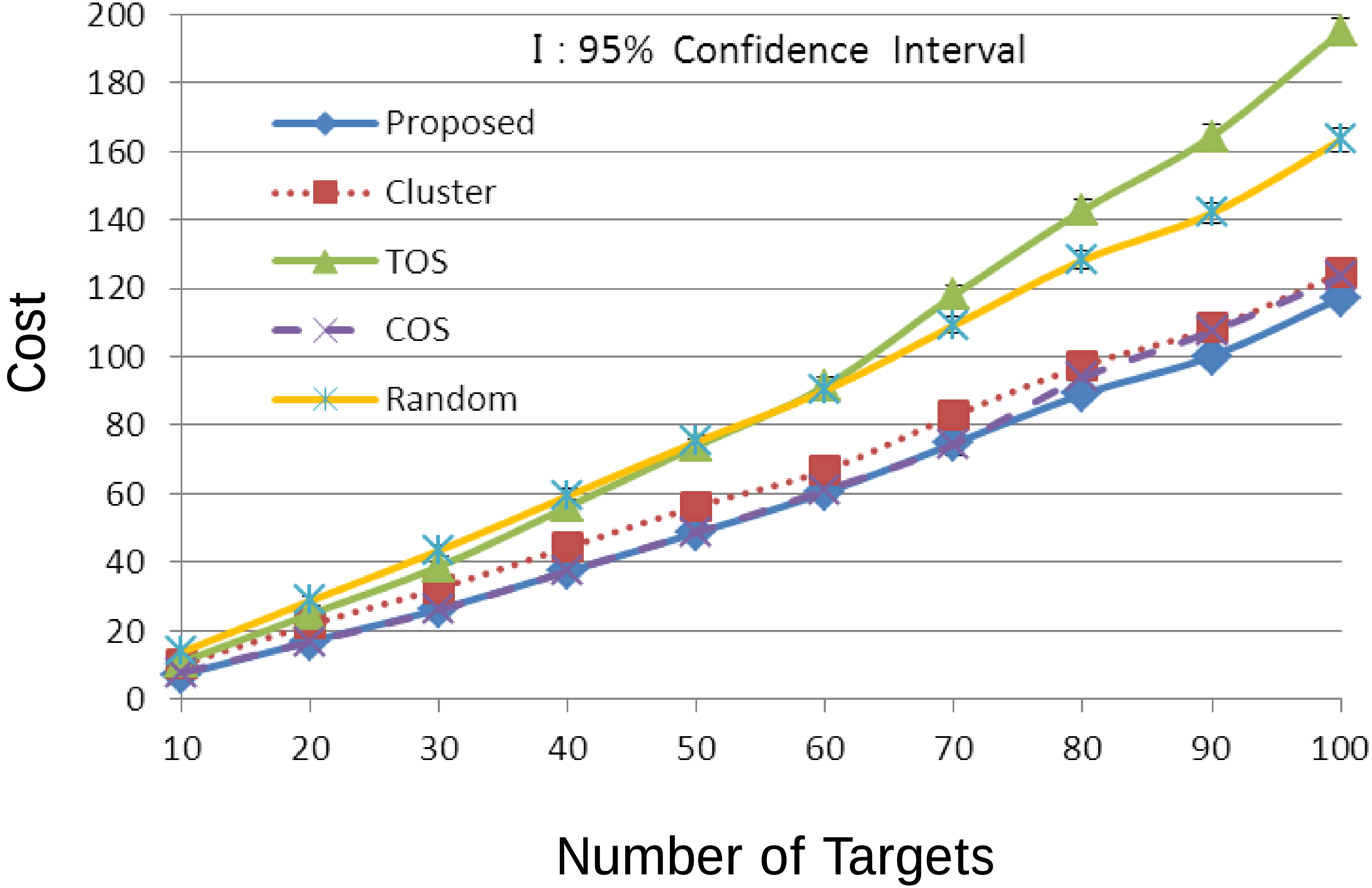
Simulation Results（1/3）
Time : Communication = 1 : 1



Simulation Results (2/3)
Time : Communication = 10 : 1



Simulation Results (3/3)
Time : Communication = 1 : 10



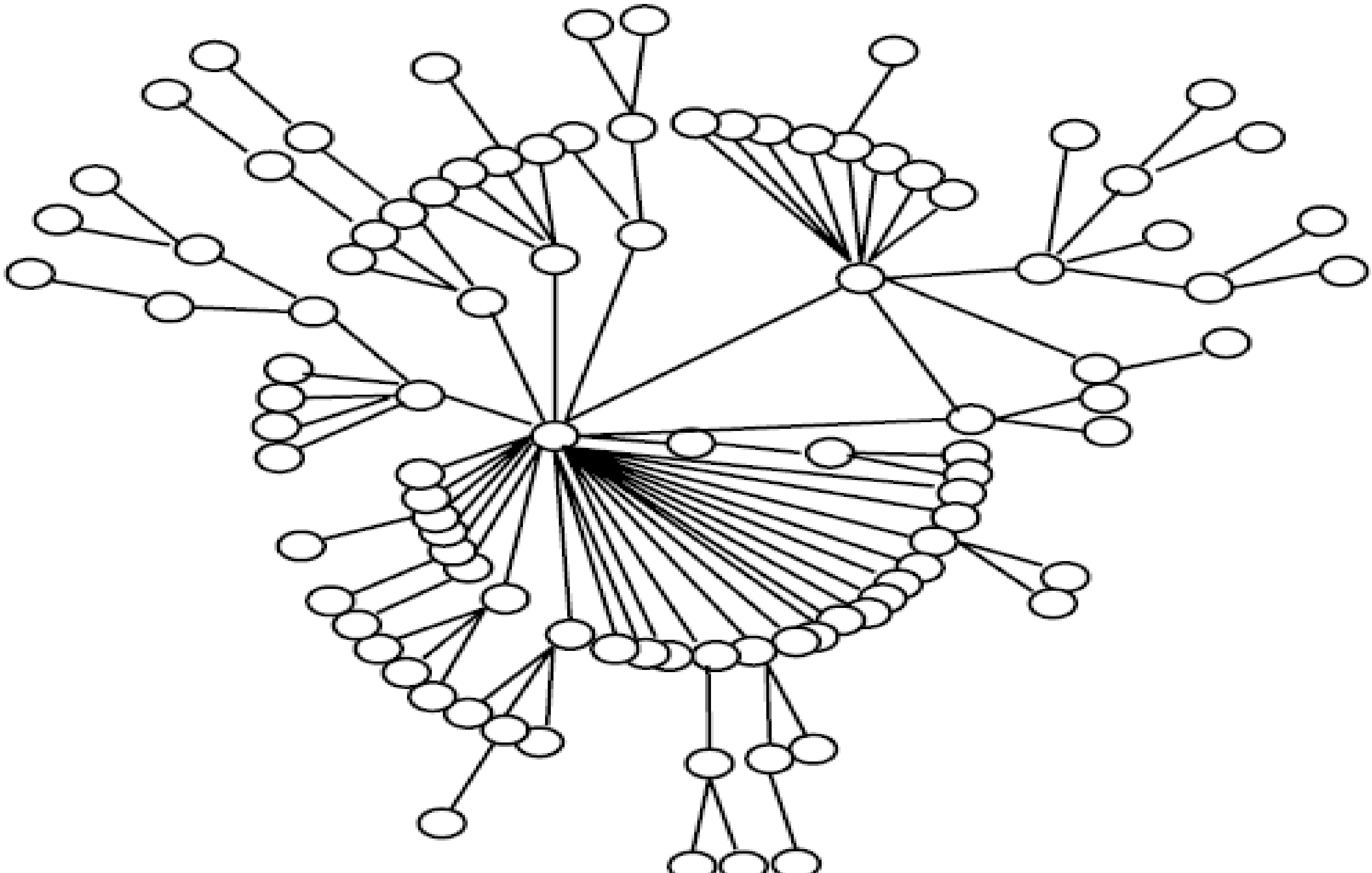
Conclusion

- We proposed a new scheme of computational resource scheduling for prioritized multiple target contents searching.

We showed the effectiveness of the scheme by computer simulations.

- Future works
  - Investigate the effectiveness of the proposed scheme in the case of using much larger network topology
  - Improve the proposed scheme

Barabasi Albert model



Simulation Results
Switching time is zero