Hybrid Periodical Flooding in Unstructured Peer-to-Peer Networks

ICPP’03

Outline

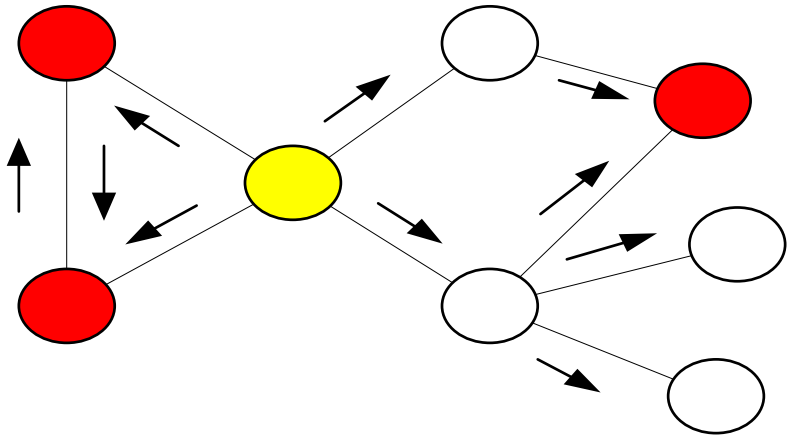
- Introduction
- Hybrid Periodical Flooding
- Performance Evaluation
- Conclusion

Introduction

- Search Mechanisms on Unstructured P2P Networks
  - Blind flooding
  - Statistics-based

Introduction

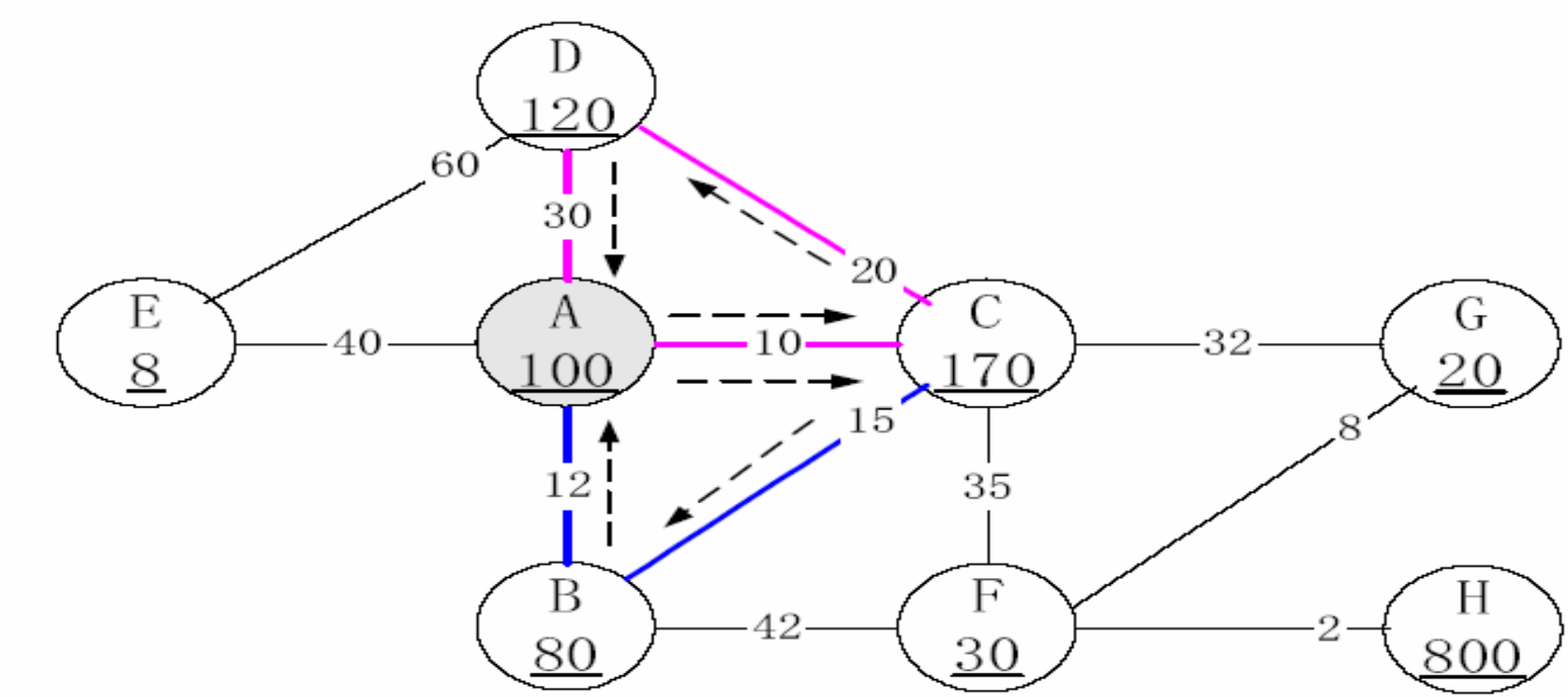
- Blind flooding
  - Blind flooding mechanism relays the query message to all its logical neighbors, except the incoming peer
  - Large volume of unnecessary traffic



Introduction

- Statistics-based
  - a peer selects a subset of its neighbors to relay the query based on some statistics information
- Partial Coverage Problem
  - Large percentage of the peers may be unreachable no matter how large the TTL value is set

Introduction



Introduction

- Uniformed selection of relay neighbors
  - Breadth-first search (BFS)
  - Depth-first search (DFS)
- Weighted selection of relay neighbors
  - Directed BFS (DBFS)

Hybrid Periodical Flooding

- Goal
  - Reducing the unnecessary traffic
  - Solving partial coverage problem
- Statistics-based + Periodical flooding

Hybrid Periodical Flooding

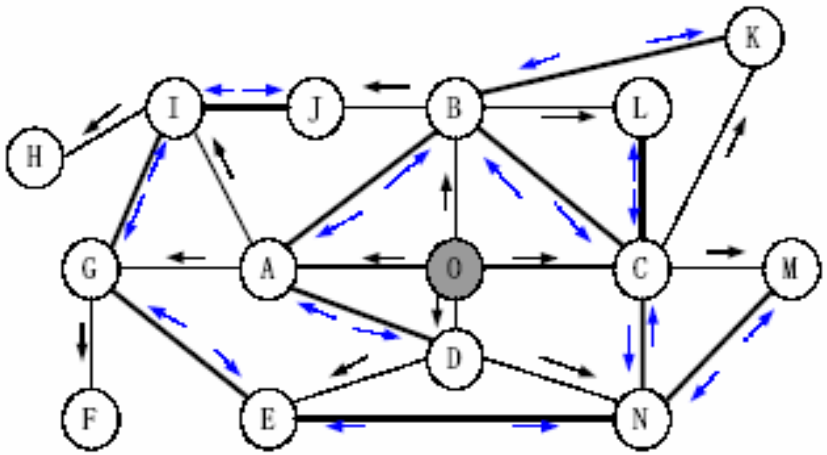
- Periodical flooding (PF)
  - Given a peer with *n* logical neighbors and the current value of TTL, the number of relay neighbors, *h*, is defined by the following function $h=f(n,TTL)$

Ex. $h=f_{BFS}(n,TTL)=n$, $h=f_{DFS}(n,TTL)=1$.

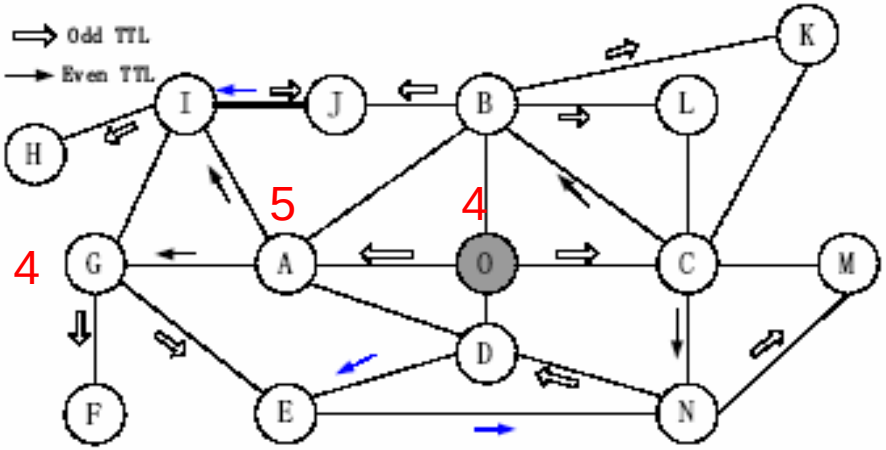
Hybrid Periodical Flooding

$$f\ (n, TTL\ ) = \begin{cases} \left\lceil \frac{1}{2} n \right\rceil, & if\ \ TTL\ \ is\ \ odd \\ \left\lceil \frac{1}{3} n \right\rceil, & if\ \ TTL\ \ is\ \ even \end{cases}$$

TTL=7



(a) BFS



(b) PF

Hybrid Periodical Flooding

Table 1. PF and Blind Flooding

	TTL	Query Msg	New Peers	Msg Per Peer
BFS	7	4	4	1.00
	6	17	8	2.12
	5	15 36	2	7.50
PF	7	2	2	1.00
	6	4	4	1.00
	5	9 15	8	1.12

Hybrid Periodical Flooding

- Hybrid Periodical Flooding (HPF)
  - The number of relay neighbors can be changed periodically based on a periodical function and the relay neighbors are selected based on multiple metrics in a hybrid way

$h = h_1 + h_2 + \ldots + h_t$ t : # of metrics

$$h_i = \lceil h \times w_i \rceil \quad 1 \leqq i \leqq t$$

Performance Evaluation

Simulation setup

Physical topology：10000 nodes

Logical topology ：1000 to 5000 nodes

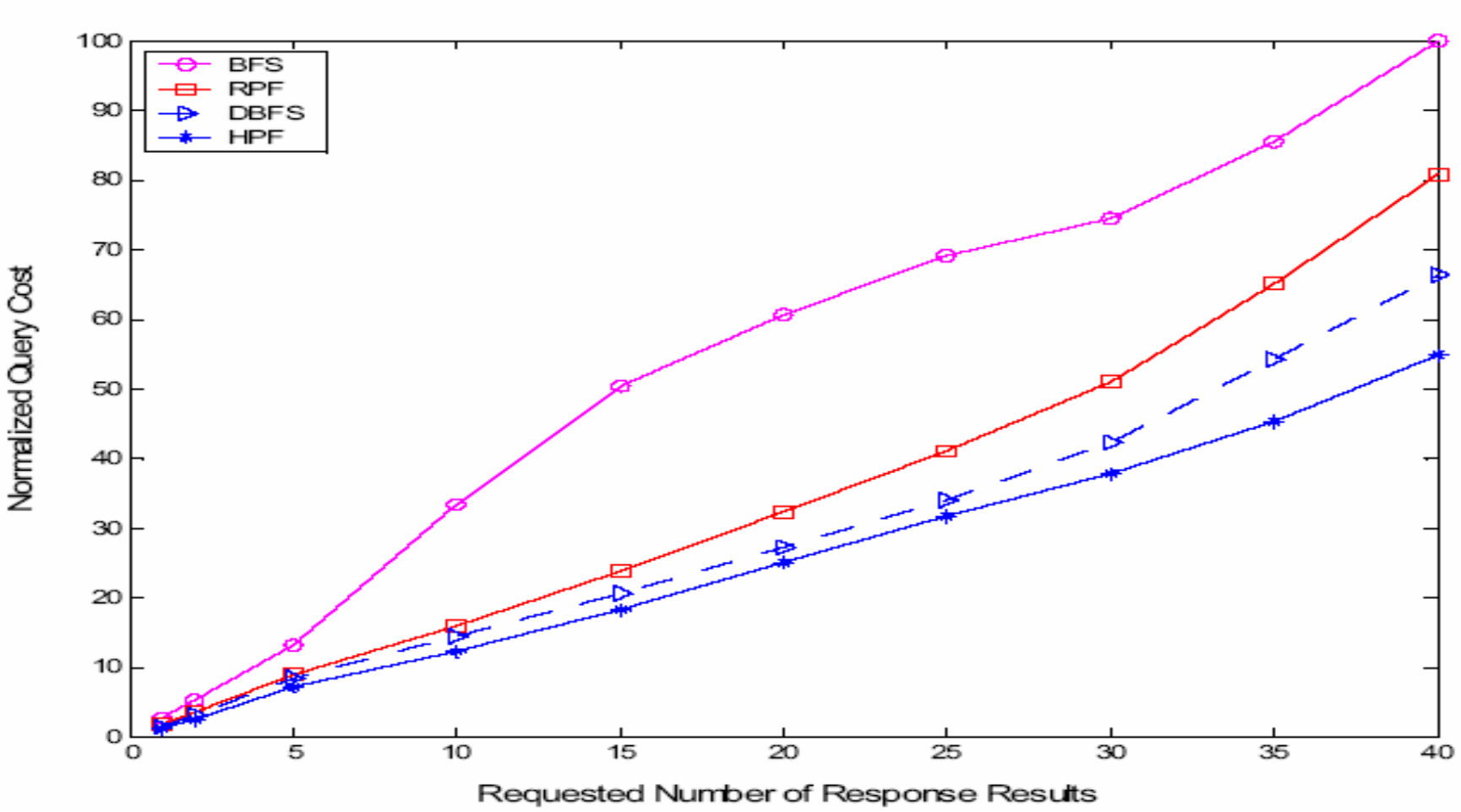
Metric：communication cost(0.6) , shared # of files(0.4)

Period function：

$$f\ (n, TTL\ ) = \begin{cases} \left\lceil \frac{1}{2} n \right\rceil, & if\ \ TTL\ \ \ is\ \ \ odd \\ \left\lceil \frac{1}{4} n \right\rceil, & if\ \ TTL\ \ \ is\ \ \ even \end{cases}$$

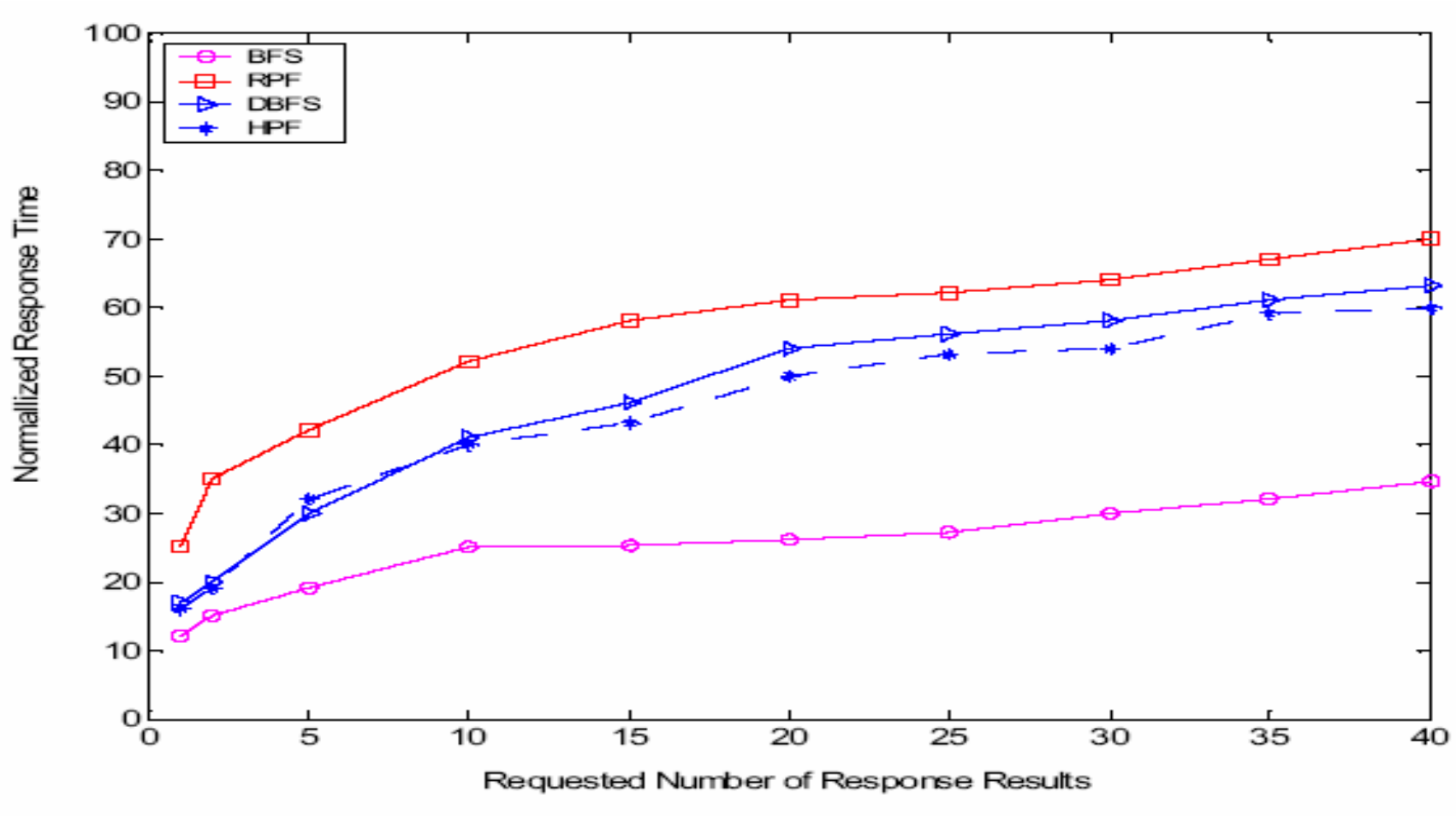
Performance Evaluation

Traffic comparison



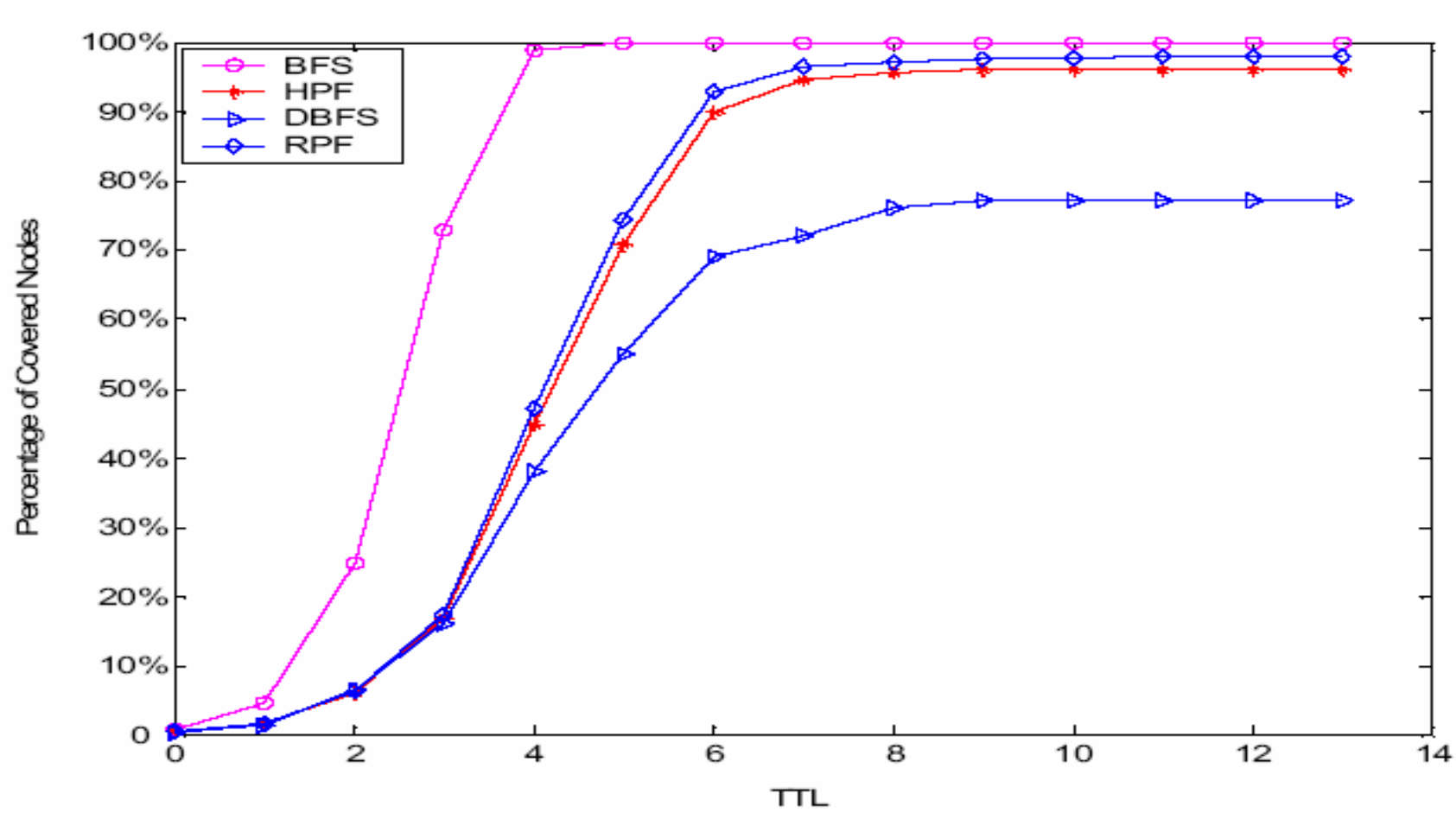
Performance Evaluation

Response time comparison



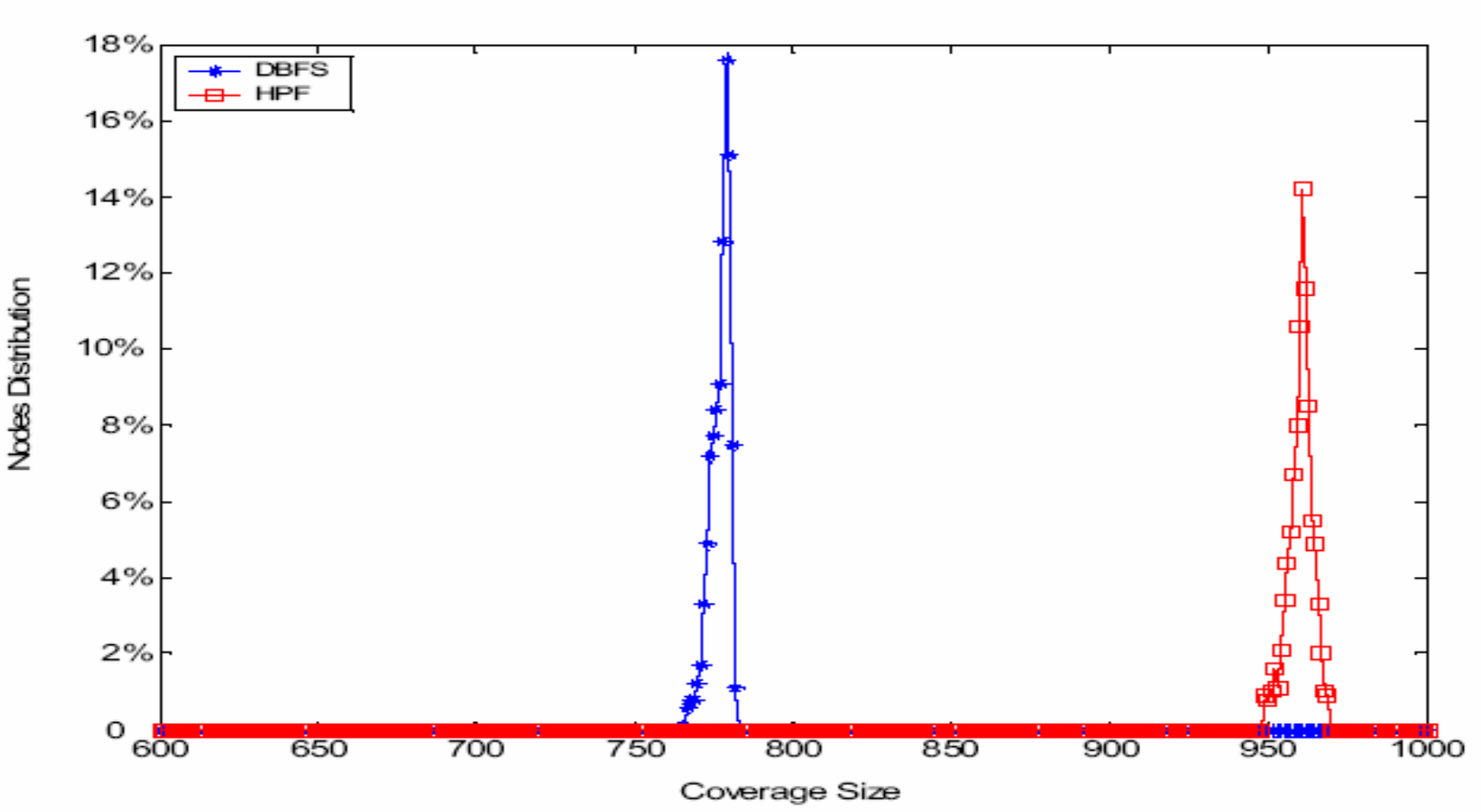
Performance Evaluation

Coverage percentage comparison



Performance Evaluation

Coverage percentage comparison



Conclusion

- HPF improves the efficiency of blind flooding by retaining the advantages of statistics-based search mechanisms and by alleviating the partial coverage problem
- Still have some work to do on the performance of response time