A Space Efficient Local Look-up Search Peer-to-Peer Protocol for Trustworthy Key Distribution

Graciela Perera
Department of Computer Science and Information Systems
Youngstown State University
gcperera@cis.ysu.edu

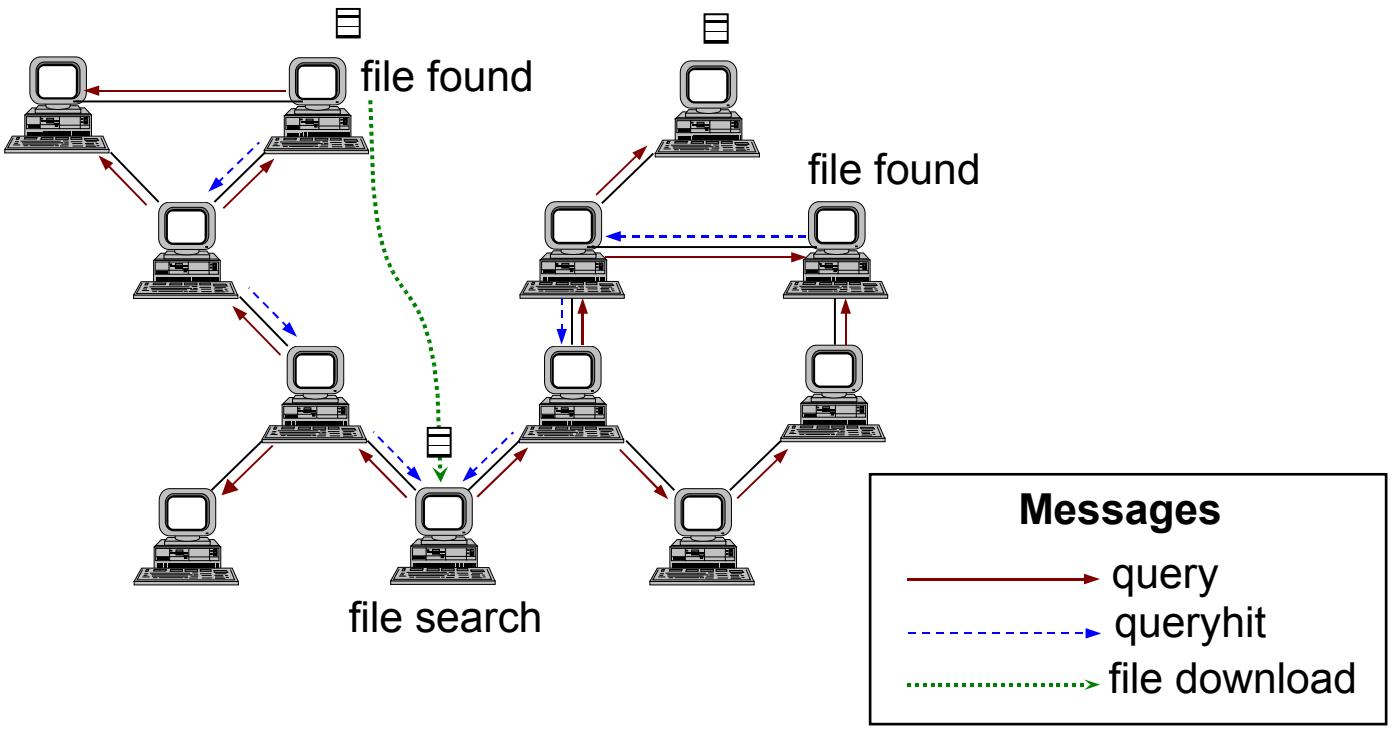
Agenda

- **Introduction**
- **Space efficient local search**
- **Summary and future work**

Graciela Perera – April 9, 2008
Department of Computer Science and Information Systems
Youngstown STATE UNIVERSITY

Introduction

- **A P2P network is an overlay on the Internet**
  - Gnutella uses a broadcast query for file search
  - Used for distributed sharing of resources



Graciela Perera – April 9, 2008
Department of Computer Science and Information Systems
Youngstown STATE UNIVERSITY

Introduction continued

- **Over 20% of the traffic on the Internet is from P2P**
- **Inefficient use of bandwidth and security limits full potential**
  - There is a large potential for improving efficiency and security
- **Can global knowledge be enabled in all hosts?**
  - Global knowledge = list of files shared by all hosts
- **Can this be done in a way so that b/w usage is not increased over query paradigm?**

Graciela Perera – April 9, 2008
Department of Computer Science and Information Systems
Youngstown STATE UNIVERSITY

BULLS

- **Hypothesis…**

By broadcasting file updates and using local search, the overhead traffic can be reduced over a broadcast query.

- **Reduce overhead by broadcasting file updates**
  - Instead of queries
  - All hosts know what files are shared by other hosts
  - Each host maintains a global data struture

→ **BULLS = Broadcast Update, Local Look-up Search**

Graciela Perera – April 9, 2008
Department of Computer Science and Information Systems
Youngstown STATE UNIVERSITY

BULLS data structure

- **Global structure stores files shared by each host**
  - Table storing basic types of data

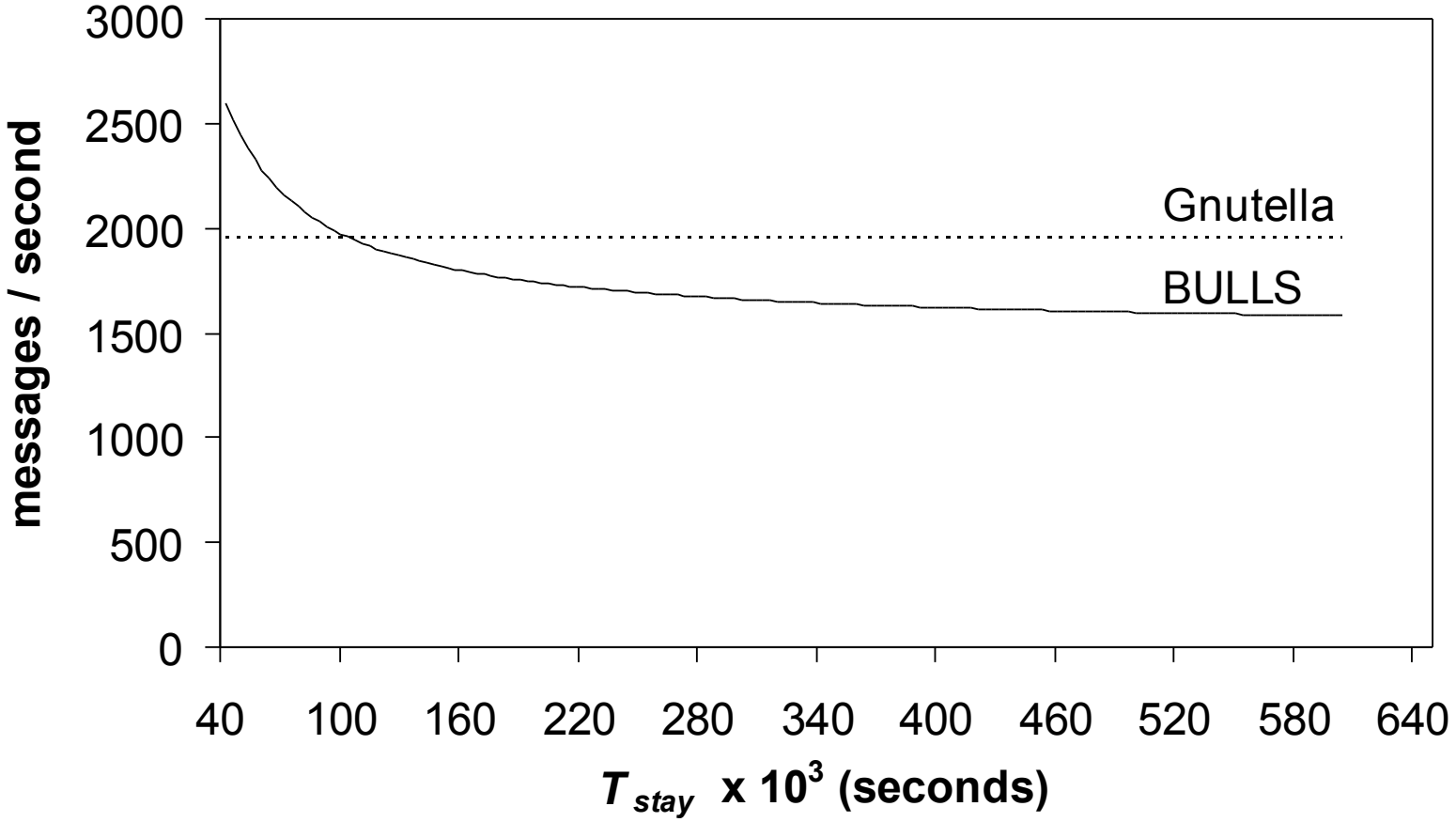
Host_name	List of file_names
host_name1	file_name1, file_name2, …
host_name2	file_name1, file_name2, …
⋮	⋮
host_nameN	file_name1, file_name2, …

- **Updates correspond to adding/deleting a file**
- **BULLS stores these updates in global structure**

Graciela Perera – April 9, 2008
Department of Computer Science and Information Systems
Youngstown STATE UNIVERSITY

BULLS overhead

- **Impact of overhead traffic (numerical results)**
  - BULLS less overhead for T_{stay} > 28 hrs



Graciela Perera – April 9, 2008
Department of Computer Science and Information Systems
Youngstown STATE UNIVERSITY

Agenda

- Introduction to P2P
- Space efficient local search
- Summary and future work

Graciela Perera – April 9, 2008
Department of Computer Science and Information Systems
Youngstown STATE UNIVERSITY

My objectives

- **My work seeks to:**

#1 Improve BULLS global data structure (SELLS)
- Use bloom filters to enhance and enable secure searches

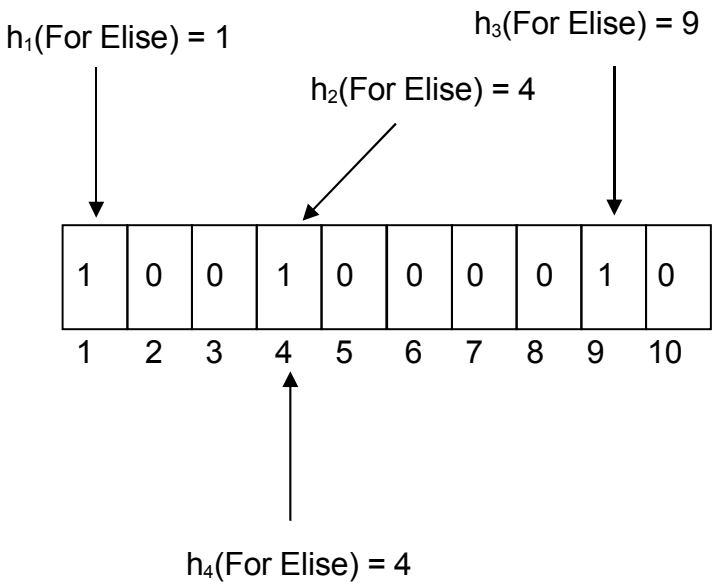
#2 Study trustworthy key distribution using SELLS
- Store secret keys generated by Diffie-Hellman protocol

#3 Improve Diffie-Hellman protocol
- Authenticate via steganography avoid man-in-the-middle attack

Graciela Perera – April 9, 2008
Department of Computer Science and Information Systems
Youngstown STATE UNIVERSITY

Objective #1 – improve BULLS

- **Bloom filter: space efficient probabilistic structure**
  - Represent a set and supports membership queries $(x \in S)$
  - Can incorrectly recognizes an element as a member of a set
  - Use hash functions (h_1, h_2, h_3, h_4) for filename ("For Elise")



Store the files shared by a host in a Bloom filter

Graciela Perera – April 9, 2008
Department of Computer Science and Information Systems
Youngstown STATE UNIVERSITY

BULLS with bloom filters

- **Impact on searching**
  - Searching in BULLS is $O(n*m)$ where n = # host, m = # files
  - Redesign BULLS data structure using Bloom filters

Inverse Bloom Filter (IBF)

Local Bloom Filters (LBF)

- There is a Bloom filter per host
- Each Bloom filter is set of shared files
- False positive yields file not found

Global Bloom Filter (GBF)

- Set of improbable (unique) files
- XOR of LBFs
- False positive yields file not found

Graciela Perera – April 9, 2008
Department of Computer Science and Information Systems
Youngstown STATE UNIVERSITY

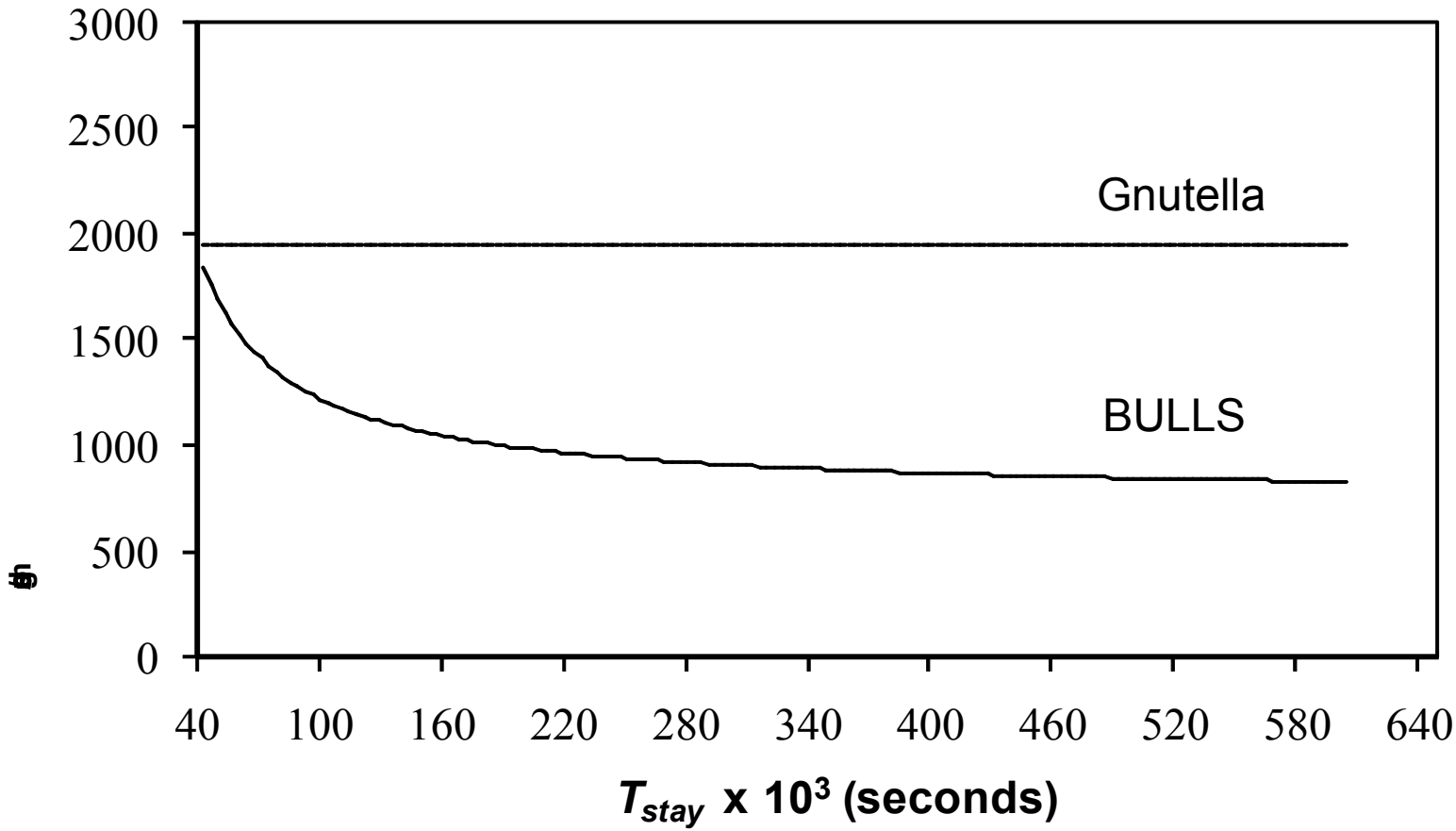
BULLS with bloom filters continued

- **Impact on searching**
  - Searching in SELLS for unique files is $O(1)$
  - Filenames are encoded by hash functions, thus confidential
- **Search cost**
  - Tradeoff between of space and false positive rate
  - Preliminary results show potential
  - Experiments will be conducted using a real trace
    - Daniel Stutzbach, University of Oregon

Graciela Perera – April 9, 2008
Department of Computer Science and Information Systems
Youngstown STATE UNIVERSITY

BULLS with bloom filters continued

- **Impact of overhead traffic (numerical results)**
  - Reduce update rate by storing updates (2) in bloom filter



Graciela Perera – April 9, 2008
Department of Computer Science and Information Systems
Youngstown STATE UNIVERSITY

Objective #2 – trustworthy distribution

- **Goal provide trust by distributing keys with SELLS**
- **Hypothesis…**

Previously shared secret key with Diffie-Hellman protocol can be distributed using SELLS

- **Can you trust a host if he has previously shared a secret key with another host?**
  - Store keys (trusted host) in LBF and distribute with SELLS

Graciela Perera – April 9, 2008
Department of Computer Science and Information Systems
Youngstown STATE UNIVERSITY

Objective #2 – trustworthy distribution

- **SELLS can be used to exchange keys**
  - Assume some set of hosts share a secret key (Diffie-Hellman)
  - Secret key between host store in LBF
  - LBF distributed with SELLS
  - Trust connection from a host if
    - Secret key found in at least one LBF (more LBF found implies more trust)
    - Secret key not found in GBF (found in GBF implies less trust)

- **If Pete trusts Gecko and Gecko trust caveman**
  - Can Pete trust Caveman based on the trust of Gecko?

Pete

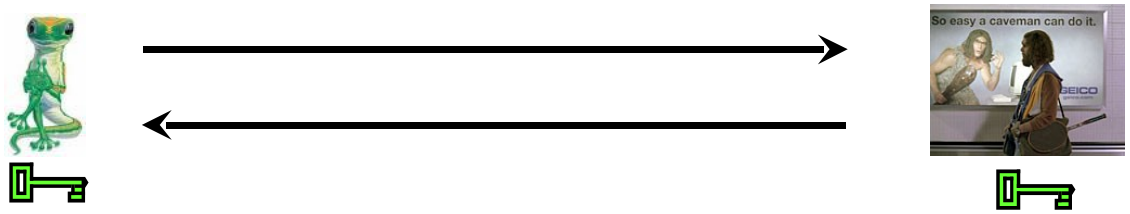
Gecko

Caveman

Graciela Perera – April 9, 2008
Department of Computer Science and Information Systems
Youngstown STATE UNIVERSITY

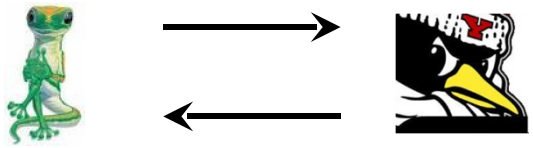
Objective #2 – trustworthy distribution

- **Gecko and Caveman share key (Diffie-Hellman)**



- **Gecko stores in his LBF**
  - host = Caveman
  - insert in Bloom filter key

- **Gecko exchanges with PETE his IBF (SELLS)**



- **Pete can then trust Caveman because of Gecko**

Graciela Perera – April 9, 2008
Department of Computer Science and Information Systems
Youngstown STATE UNIVERSITY

Objective #3 – improve Diffie-Hellman

- **Goal provide authentication using steganography**
- **Hypothesis…**

Personal data stored in hosts can be used in authentication by combining steganography and cryptography

- **Can steganography be used for authentication?**
  - Combine steganography and cryptography for session keys

PSP = Personal Secure Protocol for P2P

Graciela Perera – April 9, 2008
Department of Computer Science and Information Systems
Youngstown STATE UNIVERSITY

Agenda

- Introduction to P2P
- Space efficient local search
- Summary and future work

Graciela Perera – April 9, 2008
Department of Computer Science and Information Systems
Youngstown STATE UNIVERSITY

Summary and future work

- **BULLS reverses query search paradigm**
  - Reduce query overhead traffic in most cases
- **SELLS improves BULLS space and search efficiency**
  - Tradeoff between efficiency and precision (false positive)
- **Improve Diffie-Hellamn by using steganography for authentication (PSP)**

Graciela Perera – April 9, 2008
Department of Computer Science and Information Systems
Youngstown STATE UNIVERSITY

Summary and future work continued

Future work

- Evaluate and implement a SELLS prototype
- Evaluation of PSP to exchange session keys and distribute keys with SELLS
- Application of the knowledge learn to wireless sensor networks

Graciela Perera – April 9, 2008
Department of Computer Science and Information Systems
Youngstown STATE UNIVERSITY