

CmpE 473 Internet Programming

Pınar Yolum
pinar.yolum@boun.edu.tr

Department of
Computer Engineering
Boğaziçi University

Searching the Web using P2P Systems

Peer-to-Peer Architectures

- Richer models of interaction (compared to client-server)
 - Pull
 - Push
- Control on all peers
- Peers enter and leave; the system stays
- Difficult to shut down
 - No bottleneck

Pull

- *Pull*: (callbacks) each party can be the client of the other.
- Example:
 - File sharing
 - Each peer hosts some files (e.g., mp3s)
 - No central server, no central index
 - Query peers to find the file you need
 - Different protocols: Gnutella, Freenet

Push

- *Push*: Intelligence on the pushing side
- Example:
 - Reminders
 - Give updates

Fall 2007— Pinar Yolum

5

Gnutella

- Exchange of files (first designed for recipes, later used for mp3s)
- Simple protocol
 - Network messages: Ping, Pong
 - Data messages: Query, QueryHit, Push
 - Flooding
- Connect a network through a known peer
- Time-to-live set for message routing
- Open permanent connections with peers (usually four)
- No network structure

Fall 2007— Pinar Yolum

6

File Sharing

- No explicit modeling of others
- Trust is not considered
- Usually the protocol is based on flooding
- Quality of service for found items are not considered

Fall 2007— Pinar Yolum

7

Desired Properties

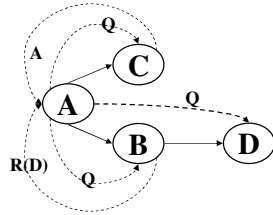
- Peer Autonomy
 - Choose who to interact with
 - Choose what to cache
- Peer Heterogeneity
 - Follow distinct policies
 - Offer distinct services and varying levels of trustworthiness
- Neighbor Choice
 - Choose and change neighbors locally as appropriate

Fall 2007— Pinar Yolum

8

Referral Systems Architecture

- An agent represents a principal who is offering a service and/or searching for services.
- An agent generates a query for a service and sends it to its *neighbors* (a small subset of its acquaintances). The neighbors can provide the service or refer other agents (based on their *referral policies*).
- Each agent models the *expertise* (quality of a service) and *sociability* (quality of the referrals) of its acquaintances (the agents it interacts with).
- Based on these models, each agent can change its set of neighbors (using its *neighbor selection policy*).



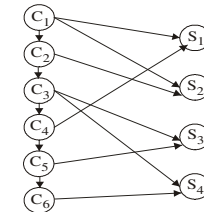
Fall 2007— Pinar Yolum

9

Application Domains

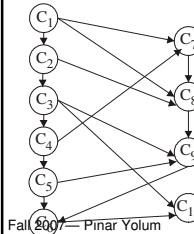
Commerce:

- Distinct service producers and consumers.
- Producers have expertise, consumers have sociability.
- Expertise of consumers does not increase.
- Answers are easy to evaluate.



Knowledge Management:

- All agents can be producers and consumers.
- Expertise of the producers can be cached by others.
- Answers are harder to evaluate.



Fall 2007— Pinar Yolum

10

Representation

- Interests used to generate queries
- Query, answer, interest, and expertise are vectors from Vector Space Model where each dimension corresponds to a domain
- Example Query: [0.3,0.8,0.3]
- Dimension of the vectors varies with domain

Fall 2007— Pinar Yolum

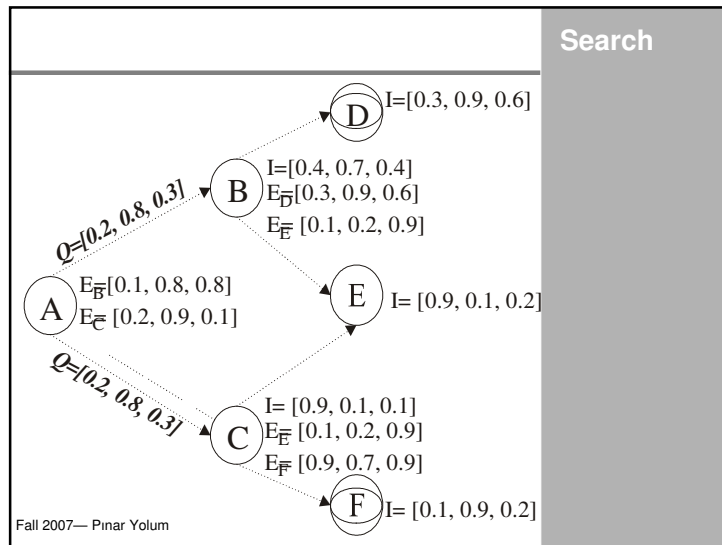
11

Referral Policies

- Refer all neighbors:**
 - Does not consider which neighbors would be more likely to answer.
- Refer all matching neighbors:**
 - Calculate how capable each neighbor is in answering a particular query.

$$Q \otimes E = \frac{\sum_{t=1}^n (q_t e_t)}{\sqrt{n \sum_{t=1}^n q_t^2}}$$
 - Refer those neighbors with “sufficient” expertise (those greater than a given threshold **T**).
- Refer best neighbor:**
 - Refer the most capable neighbor.
 - Guarantees that at least one neighbor is referred.

Fall 2007— Pinar Yolum



- Neighbor Selection Policies**
- Consider weighted sociability (W) and expertise.
- **Weighted Average ($W=0.25, 0.50, 0.75$)**
 - Choose the best m based on weighing both the expertise and the sociability of the acquaintances.
 - **Providers ($W=0$)**
 - Choose the best m agents whose expertise matches the agent's interests.
 - **Sociables ($W=0.90, 1$)**
 - Choose the most sociable m acquaintances.
- Fall 2007— Pinar Yolum 14

- Recommending Products vs. Services**
- **Products (by a product vendor)**
 - The recommender is the provider
 - Votes are known to recommender
 - Votes are given prior to usage (buying)
 - Repetition is less likely (buy the same book)
 - **Services (by a service registry)**
 - The recommender is not the provider
 - Votes are not necessarily known to recommender
 - Votes are given after usage
 - Repetition can occur but not known to registry
- Fall 2007— Pinar Yolum 15