A Control Theoretic Approach to Analyzing Peer-to-Peer Searching

Gang Ding
Qualcomm Research

Outline

- Background
- Modelling
- Analysis
- Synthesis
- Summary

Background

- Peer-to-Peer (P2P) searching algorithms
  - Unstructured
    - Napster: central directory service
    - Gnutella: distributed flooding
    - KaZaA: hierarchical super-peer
    - BitTorrent: central index server
  - Structured - Distributed Hash Table (DHT)
    - Chord, Pastry, Tapestry, CAN, Kademlia, etc.
  - Hybrid

Problem – Analyzing P2P Search

- Research challenges
  - Limited communication bandwidth
  - Limited capacity for each peer to process incoming queries
  - Network topology
  - Routing policy

- Research methods
  - Measurement-based analysis
  - Static analysis
  - Stochastic analysis

Modeling P2P Search

$$\dot{x}_i(t) = \sum_{j=i}^{n} a_{ij} x_j(t) + \sum_{k=i}^{n} b_{ik} u_k(t) \qquad i = 1, \cdots, n$$

- $x_i(t)$: state of peer i represents the aggregated query data processed by this peer
- $a_{ii}(t)$: processing rate of peer i (negative)
- $a_{ij}(t)$ ($j \neq$ i): weight of query data from peer j to i (non-negative)
- $u_i(t)$: original query data rate of peer i
- $b_{ik}(t)$: original query from peer k to i

Outline

- Background
- Modelling
- Analysis
- Synthesis
- Summary

Analysis – Stability

Will a peer's state converge asymptotically?

Yes. System will converge because $a_{ii}(t)$'s are all negative

$$\dot{\hat{x}}(t) = \hat{A}(t)\hat{x}(t) + \hat{B}(t)u_1$$

$$\hat{A}(t) = \begin{bmatrix} \hat{a}_{11} & 0 & \cdots & 0 \\ \hat{a}_{21} & \hat{a}_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ \hat{a}_{n1} & \hat{a}_{n2} & \cdots & \hat{a}_{nn} \end{bmatrix}_{n\times n}, \qquad \hat{B} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{n\times 1}$$

Analysis – Reachability

Can a query reach all peers in a given network topology?

No always, but it will reach as far as following condition holds

$$\hat{a}_{i(i-1)} \neq 0 \qquad i = 2, \ldots, n$$

Analysis – Solution

Is there a closed form solution to all states?

Yes

$$\hat{x}_1(t) = (u_1 / -\hat{a}_{11})(1 - e^{\hat{a}_{11}t})$$

$$\hat{x}_k(t) = L^{-1}\{\sum_{i=1}^{k-1} \frac{\hat{a}_{ki}\hat{x}_i(s)}{(s - \hat{a}_{kk})}\}, \qquad (k = 2, \ldots, n)$$

- When peers are of different processing power, the slowest peer will dominate the query propagation rate

Analysis Example – Super-peer Broadcast Search

For every super-peer s_i, $a_{s_i j} = 0$ and $b_{ij} = 1$ for any j in $Leaf(s_i)$

$$\dot{x}_s(t) = A_s(t)x_s(t) + B_s(t)u_s(t)$$

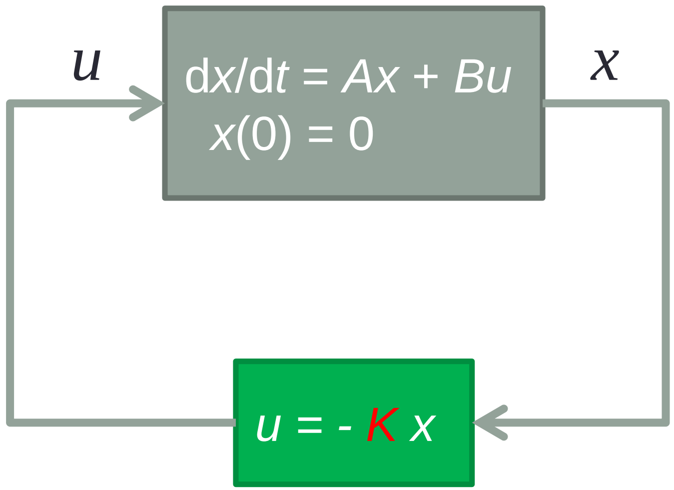
$$u_{si} = \sum_{j \in Leaf(i)} u_j$$

- The successful search rate is dependent on the lowest processing rate of super-peers
- The bandwidth capacity of a super-peer should be much larger than a regular peer

Outline

- Background
- Modelling
- Analysis
- Synthesis
- Summary

Adding Feedback Control



$$K = R^{-1} B^T P^0$$

where P^0 is the solution of algebraic Riccati equation:

$$A^T P + PA + Q - PBR^{-1} B^T P = 0$$

Distributed Synthesis

- Every peer maintains a small local system model involving the states of itself and its neighboring peers
- The peer communicates with its neighbors periodically in order to update its model with A, B, and $x(t)$
- Numerical computing method is employed to find P^0
- K is further calculated from P^0
- Control input $u(t)$ is calculated from K and state $x(t)$

Summary

- A general mathematical model for the dynamic behaviors of query during P2P searching
- Use modern control theory to analyze system stability, controllability, and state dynamics
- Synthesize the model by optimal feedback control theory
- Future studies
  - Apply feedback control theory to model, analyze, and synthesize P2P data downloading
  - Study performance of P2P networks in mobile wireless environment

Thank you!

Questions?