Network-based Matchmaking in Distributed Publish/Subscribe Systems

Jim Kurose
Department of Computer Science
University of Massachusetts at Amherst
http://www.cs.umass.edu/~kurose

Joint work with Ge Zihui (PhD student), Micah Adler, Don Towsley

Theme: *interest-based* content dissemination
**Theme: interest-based content dissemination**

**Interest-based content dissemination:**

- publish/subscribe architecture: not just moving data, but moving data based on *content*
  - publishers push content
  - subscribers specify interests
  - content delivered based on match of interest
- two approaches
  - multicasting
  - application-level content filtering within “network”
Outline

- Channelization problem
  - interest-based content dissemination via multicast groups
- Distributed matchmaker
  - interest-based content dissemination via content filters
- Summary and future work
Channelization problem: motivation

- Large scale interest-based content dissemination
  - large number of information sources, receivers
    - e.g. event notification, distributed interactive simulation
  - receivers interested only in subset of information flows
    - flooding is not desirable
  - receivers share common interests
    - unicasting is not efficient
- Multicast to scope content delivery [Calvin95, Ammar01, Levin00]
  - receivers join multicast group(s) containing flow(s) of interest
  - limited number of multicast group addresses
  - keep multicast routing tables small

Internet Multicast Service Model

multicast group concept: use of indirection
- hosts address IP datagrams to multicast group
- routers forward multicast datagrams to hosts that have “joined” that multicast group
Channelization and subscription problem

Information to multicast group mapping

limited # mcast grps

Channelization problem

Subscription problem

Requirement:

No false exclusion
Minimum false inclusion

A set of flows of interest (sources): \( S, |S| = N \)

Flow \( i \) has rate \( \lambda_i, i \in S \)

A set of multicast groups: \( G, |G| = K \)

A set of independent users: \( U, |U| = M \)

Interest matrix: \( W = (w_{i,j}), i \in S, j \in U \)

Flow-group mapping matrix: \( X = (x_{i,m}), i \in S, m \in G \)

Subscription matrix: \( Y = (y_{j,m}), j \in U, m \in G \)

\[ y_{j,m} = \begin{cases} 
1 & \text{user } j \text{ subscribes to multicast group } m \\
0 & \text{otherwise}
\end{cases} \]
Cost function

Cost function associated with mapping $X, Y$

$$C(X, Y) = w_1 \sum_{\text{flow, } i} \sum_{\text{rcvr, } j} \sum_{\text{grp, } m} x_{i, m} y_{j, m} c_{j, i} \lambda_i + w_2 \sum_{\text{grp, } m} \sum_{\text{flow, } i} x_{i, m} \lambda_i$$

$$w_1 + w_2 = 1$$

$c_{j, i}$: system cost for user $j$ to receive one copy of flow $i$

Constraint:

$$\forall i \in S \left( \forall j \in U \right) \sum_{\text{grp, } m} x_{i, m} y_{j, m} \geq w_{j, i}$$

Objective: to minimize $C(X, Y)$

Channelization problem is NP-Complete!!!

Constrained and unconstrained channelization and subscription problem

- Unconstrained version: each flow can be assigned to multiple (one or more) multicast groups

- Constrained version: a flow can be assigned to only one multicast group
  - sacrifice efficiency
  - reduce complexity ???
**Complexity study**

<table>
<thead>
<tr>
<th></th>
<th>Channelization Problem</th>
<th>Subscription Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unconstrained</strong></td>
<td>NP-Complete</td>
<td>NP-Complete</td>
</tr>
<tr>
<td><strong>Group-mapping</strong></td>
<td>🙁</td>
<td>🙁</td>
</tr>
<tr>
<td><strong>Constrained</strong></td>
<td>NP-Complete</td>
<td>Linear Time</td>
</tr>
<tr>
<td><strong>Group-mapping</strong></td>
<td>🙁</td>
<td>🐌</td>
</tr>
</tbody>
</table>

**Approximation algorithms**

- **Random assignment (RAN)**
- **Simple heuristics:**
  - sequentially assign flows to multicast groups
    - RSE: balance number of flows/group
    - RRE: balance data rate/group
- **Greedy algorithms:**
  - recursively merge two multicast groups with smallest cost increase resulting from merge
    - Flow Based Merge (FBM): initially assign each flow to a different multicast group
    - User Based Merge (UBM): initially create one multicast group for each user
Evaluating Channelization approximation

- high-rate versus low-rate flows
- popular versus unpopular flows
- baseline: cost when infinite number of multicast groups available:

\[
Cost_\infty = w_1 \sum_{r,c.r} \sum_{f,o,flow,j} w_{j,i} \lambda_i + w_2 \sum_{f,ow,l} \lambda_i
\]

\[
w_1 = w_2 = 0.5, \quad c_{j,i} = 1
\]

Evaluation results

- Simple heuristics do not perform much better than random, especially when K small.
- UBM performs well only when K small.
- FBM finds best solutions.

- Non-trivial amount of unwanted data delivered – reveals limitation of using limited multicast groups only mechanism to scope content dissemination.
Channelization problem: summary

- Channelization/subscription problem formalization
  - constrained & unconstrained version
- Complexity analysis
- Approximation algorithms, evaluation via simulation
  - FBM finds good solutions over a range of prob. configurations
  - limitation of interest-based dissemination via multicasting

Large # of flows
Large # of users \( \text{V.S.} \) Limited # of multicast groups w/ differing interests

From multicasting to content filtering

- Active networks, application-layer networks
  - \( \text{😊} \) multicast address space not a constraint
  - \( \text{😊} \) flexible filtering semantics
  - \( \text{😊} \) everyone receives exactly the needed information

? signaling for determining and installing content filters
Outline

- Channelization problem
- Distributed matchmaker
- Summary and future research

Signaling in distributed pub/sub architecture

- Content-based information dissemination

Publishers (info. provider)
- provide data flows
- provide advertisements (meta data)

Matchmaker(s)
- compare advertisements with subscriptions
- set up data path if match

Subscribers (info. receiver)
- request data flows
- provide subscriptions (meta data of its interests)
Pub/sub applications - sensor networks

- Publishers: sensors
- Subscribers: users
- Advertisement: meta data: sensor observation
- Subscription: query sensor data of interest

Pub/sub applications - distributed virtual world, multiplayer games

- Publishers & subscribers: players
- Advertisements, subscriptions contain information of virtual-geographical region
Goal: set up content filters so that content is delivered only to interested subscriber(s)

- matchmaking: detecting matches between publication advertisements, subscription requests
- install filters along the data path from content provider to content subscriber

Previous work [Wolf00, Muhl01] assumes either Broadcast publication or Broadcast subscription

Goal: set up content filters so content is delivered only to interested subscriber(s)

- matchmaking: detecting matches between publication advertisements, subscription requests
- install filters along the data path from publisher to subscriber
Previous work [Wolf00, Muhl01] assumes either
Broadcast publication
or Broadcast subscription

Goal: set up content filters so content is delivered only to
interested subscriber(s)

- matchmaking: detecting matches between publication
  advertisement, subscription requests
- install filters along the data path from publisher to subscriber

Min-cost Matchmaker example

- Goal: minimize overall signaling
  msgs propagated to match pub/sub

Examples:
- 3 hosts
- 1 active router
- Pub/Sub CDs signaled

- Method: link marking scheme -
  publication msg only go through
  P-link; subscription msg only go
  through S-link.

A: publishes two regions
B: subscribes 1 region
C: subscribes 3 regions

region: described by content descriptor(CD)
Min-cost Matchmaker example

Broadcast Subscription

Broadcast Publication

Optimal Marking

8 CDs signaled

6 CDs signaled

5 CDs signaled

Resulting filtering identical (correctness)
but signaling overhead differs (performance)

Min-cost Matchmaker problem formalization

- Pub/sub network: $G = (V,E)$
- Publishers: $P \subseteq V$, subscribers: $S \subseteq V$
- Multicast trees: $T_x \subseteq E$, $x \in P$
- Goal: find link marking $M : E \rightarrow \{p,s\}$
  - satisfying validity constraint:
    - link $t$ marked $p$ ⇒ upstream links marked $p$
    - link $t$ marked $s$ ⇒ downstream links marked $s$
  - minimizing
    $\text{Cost}(M) = \sum_{t \in E} C_t(M)$
    - $C_t(M)$: pub/sub signaling cost on link $t$. 

28
Solving Min-cost Matchmaker problem

- Shared multicast tree:
  - **Optimal valid matchmaking**: cut through tree, links one side forward publications (only), links on other side forward subscriptions
    - compare number of p, s messages, mark link p or s according to whichever smaller.
  - matching, where publications/subscriptions meet (or at end points)
  - **simple distributed algorithm guarantees valid, optimal configuration**

- source based multicast tree: more difficult
Design issues: from theory to practice

- In theory
  - polynomial time algorithm
  - utilizing optimal link-marking potentially reduce signaling cost
- In practice
  - distributed computation
  - dynamic system
    - publishing/subscribing behaviors
    - membership, topology
  - overhead to obtain optimal link-marking

Pub/sub network architecture

- Filter Mgr.: locally install content filters on data forwarding path
- AMSP: Disseminate & match publication/subscription msg efficiently; Trigger filter setup in Filter Mgr.
- ATDP: Build, maintain (reliable) signaling overlay
- AMSP: • estimate & disseminate pub/sub cost
  • perform link-marking
  • disseminate PCD/SCD
  • perform content match & disseminate MCD
Evaluating a dynamic scenario

- Transit/stub network topology:
  - 4 transit (1 domain), 96 stub (12 domains)
- Publishers/subscribers randomly chosen from stub nodes
- App model: distributed simulation
  - initial # of entities/host: \( N_e \in [50, 150] \)
  - new entities added by Poisson process \( \lambda = 1/60 \)
  - entity lifetime exponentially distributed w/ mean \( N_e / \lambda \)
  - random way-point mobility model
    - 3 types of entities: high-speed (aircraft), medium-speed (ground vehicle), low-speed (person)
    - Update publication/subscription when entities cross grid

Simulation results

![Simulation results graph](image-url)
Simulation results

Per-source based multicast trees
Number of publishers = 20

- Broadcast Publication
- Broadcast Subscription
- AMSP

Simulation demonstrates significant reduction in signaling overhead under various network and application configurations.

Distributed matchmaker: summary

- Formalized & solved Min-cost Matchmaker Prob.
- Designed simple, fully-distributed matchmaker architecture
  - AMSP dynamically adapts to application's publishing/subscribing behavior
- Simulation demonstrates significant reduction in signaling overhead under various network and application configurations
Outline

- Channelization problem
- Distributed matchmaker
- Summary and future research

Summary

- Interest-based content dissemination
  - via multicast groups
    - data efficiency (i.e., reduce unwanted information being delivered to receivers)
  - via content filters
    - signaling efficiency (i.e., reduce signaling overhead to determine and install content filters)
Future Research

- publish/subscribe as fundamental structuring mechanism for data-driven distributed systems
  - sensor network applications
  - generalized subscriptions: boolean operators over publications from multiple sites

For more information:


http://gaia.cs.umass.edu
The end

Thanks!

slides available: