Remote Priority Queue: an abstraction for application-level flow control and reliability in push-based communication

Dmitrii Zagorodnov* · University of Tromsø, Norway† · EuroSys 2006

**Setting:** items of information pushed to the user

Authors create notifications of new information and insert them into the queue.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Network Routers and Services</th>
<th>Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create</td>
<td>Store and/or Forward</td>
<td>Receive (asynchronously)</td>
</tr>
<tr>
<td>(Email messages, Usenet posts, IM messages, recommendations, pub/sub events, ...)</td>
<td>(via HTTP servers, IM servers, NNTP servers, TCP/UDP/HIP servers, pub/sub-message brokers, ...)</td>
<td>(via mail clients, newsreaders, chat clients, RSS aggregators, mobile phones, PDA's, ...)</td>
</tr>
</tbody>
</table>

**Goal:** as many “good” items as desired (high recall) and no “bad” items (high precision)

- Spam flagging allows filtering of similar email messages. Sharing this feedback improves effectiveness of spam filters for a group of people.
- Usenet postings can be delayed until a moderator approves them.
- Comments and links to RSS postings are feedback that aggregators and crawlers use to judge popularity.
- Publish/subscribe systems do not consider feedback.

**Examples of existing approaches to feedback, filtering, and recovery**

- Email is filtered on the destination server by: refusing malformed messages or messages from blacklisted sources; then, rating based on textual analysis.
- Usenet is opt-in, so the amount of data downloaded is controlled by subscribing and unsubscribing.
- Same for RSS, although centralized aggregators can filter based on popularity.
- Pub/sub subscriptions are filters on all published items.
- Email sender periodically retries sending the message up to an internally chosen time limit.
- Usenet postings persist (i.e., available to a recovering client or relay) until the server chooses to expire them.
- RSS postings persist in the feed either until they expire or until newer postings bump them off.
- Publish/subscribe message brokers do not buffer messages for long-term availability.

**Proposal:** generalizing the techniques above, connect authors and users via a “priority queue”

**Problem 1:** User feedback takes time
- User feedback on an item is useful for rating, but it is not available until at least some users see the item.
- Without feedback, initial rating can rely on automatic analysis of the content or on the reputation of its author.
- Harness the trade-off between timeliness and precision.
- Filtering close to the user is inefficient.

**Problem 2:** User may get too much
- Mismatch in production rate of all sources and the consumption rate of the user.
- Even with only opt-in sources, the user or the proxy can be overwhelmed as the rates are unpredictable.
- Need a filter based on user’s consumption.
- Need a scheme for recovering missed info.

**Problem 3:** User may miss something
- User’s proxy cannot be 100% available to receive the pushed info.
- Two options: retrying to send or saving at a location that user’s proxy can query.
- For how long to retry or to store?

**Advanatges:** general mechanism for sharing feedback, efficiency due to low control, predictable reliability

**Examples:**

- Spam flagging allows filtering of similar email messages. Sharing this feedback improves effectiveness of spam filters for a group of people.
- Usenet postings can be delayed until a moderator approves them.
- Comments and links to RSS postings are feedback that aggregators and crawlers use to judge popularity.
- Publish/subscribe systems do not consider feedback.

**Proposal:** generalizing the techniques above, connect authors and users via a “priority queue”

**Authors create notifications of new information and insert them into the queue.**

**Queues store items for some period and present them ordered by some priority.**

**Users receive items as their proxies query the queue, tracking the consumption rate.**

**Examples:**

- Email is filtered on the destination server by: refusing malformed messages or messages from blacklisted sources; then, rating based on textual analysis.
- Usenet is opt-in, so the amount of data downloaded is controlled by subscribing and unsubscribing.
- Same for RSS, although centralized aggregators can filter based on popularity.
- Pub/sub subscriptions are filters on all published items.
- Email sender periodically retries sending the message up to an internally chosen time limit.
- Usenet postings persist (i.e., available to a recovering client or relay) until the server chooses to expire them.
- RSS postings persist in the feed either until they expire or until newer postings bump them off.
- Publish/subscribe message brokers do not buffer messages for long-term availability.

**Advantages:** general mechanism for sharing feedback, efficiency due to low control, predictable reliability

- With “a thousand thanks” to Inger M. Amstein, Lars Brenna, Dag Johannessen, Harald D. Johannessen, and Cathal Gurrin
- Supported by the Research Council of Norway, IKT-2010 program