Fighting Spam in Publish/Subscribe Networks Using Information Ranking

Nikos Fotiou, Giannis F. Marias and George C. Polyzos

Athens University of Economics and Business

Mobile Multimedia Laboratory

{fotiou,marias,polyzos}@aueb.gr

6th Euro-NF Conference on Next Generation Internet
Paris, France, June 2-4, 2010

Outline

- Publish/Subscribe Architectures

- Problem Area

- Inforanking

- Approach

- Evaluation

- Conclusions, Future Work

6th NGI Conference,

Fighting Spam in Publish/Subscribe Networks Using Information Ranking

Publish/Subscribe Architectures

- 3 Basic Components
  - Publishers : Information providers that advertise information using "publications"
  - Subscribers : Information consumers that express their interest in particular pieces of information using "subscriptions"
  - A network of brokers
- Publication/Subscription matching takes place in the Rendezvous Point
- Information oriented, publisher/subscriber decouple
  - Mobility, multicast, multihoming can be easily achieved

Problem Area

- A publish/subscribe architecture in which
  - Subscribers, subscribe using keywords
  - Publishers may provide misleading publications' description and therefore lure RP ➡ Spam
- Solutions for this problem
  - Block misbehaving publishers
    - Already Done (Tarkoma S.:Preventing Spam in Publish/Subscribe)
  - Block spam publications ➡ Inforanking

Inforanking (1/2)

- A mechanism that ranks information items within a certain context

- Why do we need an information ranking mechanism?
  - Users change behavior, information no!
  - Information identification is easier
    - Hash Function VS Chain of Certificates
  - Its an information –oriented world after all!

- Ranks is based on users' votes
  - Users may vote only positively
    - Easier to be implemented
  - The weight of a user's vote is inversely proportional the number of user's vote

Inforanking (2/2)

- Inforanking example where vote weight:

1/(sum(votes of user in this set))

Information Set 1		
Information ID	**Users that voted**	**Rank**
INFO1	U1, U2, U3, U4	1.75
INFO2	U1, U2, U3, U4	1.75
INFO3	U1, U5, U6, U7	3.25
INFO4	U1	0.25

Approach (1/2)

- Create a spam rank, the bigger the rank of an information item is the less likely is a RP to select it

- Allow subscribers to vote for publications that they consider spam by applying Inforanking ➡ SuR
  - Malicious publishers will start publish even more spam publications therefore …

- Apply Inforanking in publications, considering as a vote the fact that a publisher publishes a piece of information ➡ PuR
  - PuR does not add any overhead to the architecture
  - The bigger the PuR is the better a publication is

- Spam rank = 1 – Normalized(PuR) + Normalized(SuR)

Approach (2/2)

| Usage example					
Inf. Ident.	**Sub. Votes**	**SuR (NSuR)**	**Publishers**	**PuR (NPuR)**	**Spam Rank**
INF1	S2, S5	1.5 (0.30)	P1, P2	1 (0.25)	1-0.25 +0.3 = 1.05
INF2	S2, S4	1.5 (0.30)	P1, P2	1 (0.25)	1-0.25 +0.3 = 1.05
INF3	S1, S3	1 (0.20)	P3	1 (0.25)	1-0.25 +0.2 = 0.95
INF4	S1, S3	1 (0.20)	P4	1 (0.25)	1-0.25 +0.2 = 0.95

Results (1/4)

- 100 publishers, 50% malicious, 100 subscribers

- Each malicious publisher selects a number of publications from a pool of a variable size of publications (x-axis)

- Each good publisher publishes in avg. 5 items out of a pool of 80

6th NGI Conference,
Fighting Spam in Publish/Subscribe Networks Using Information Ranking
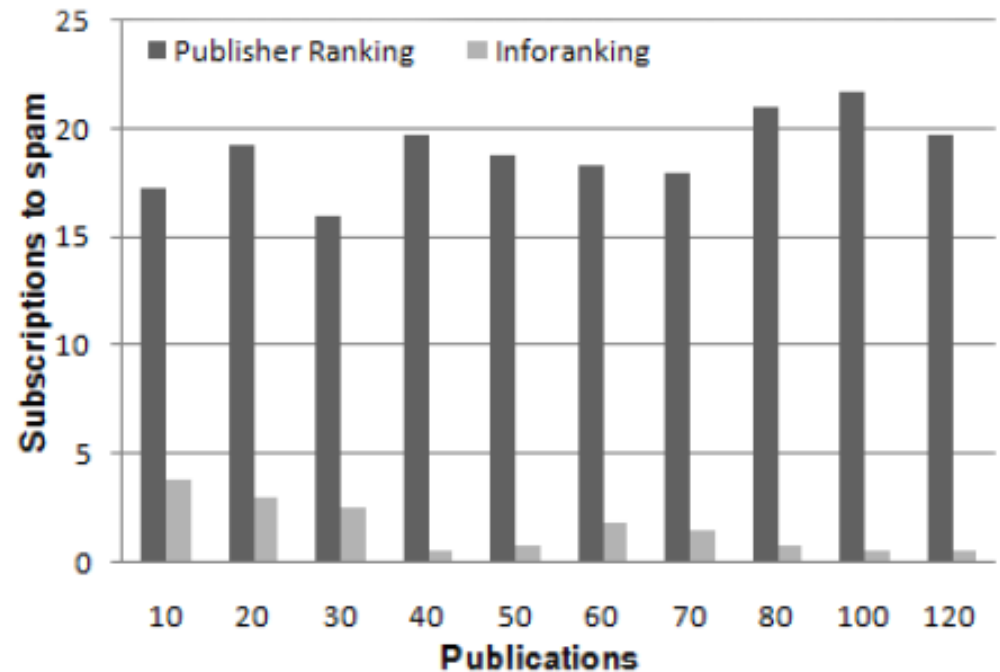
Results (2/4)

- 100 publishers, 80% malicious

- Each publisher selects a number of publications from a pool of a variable size of publications (x-axis)

6th NGI Conference,
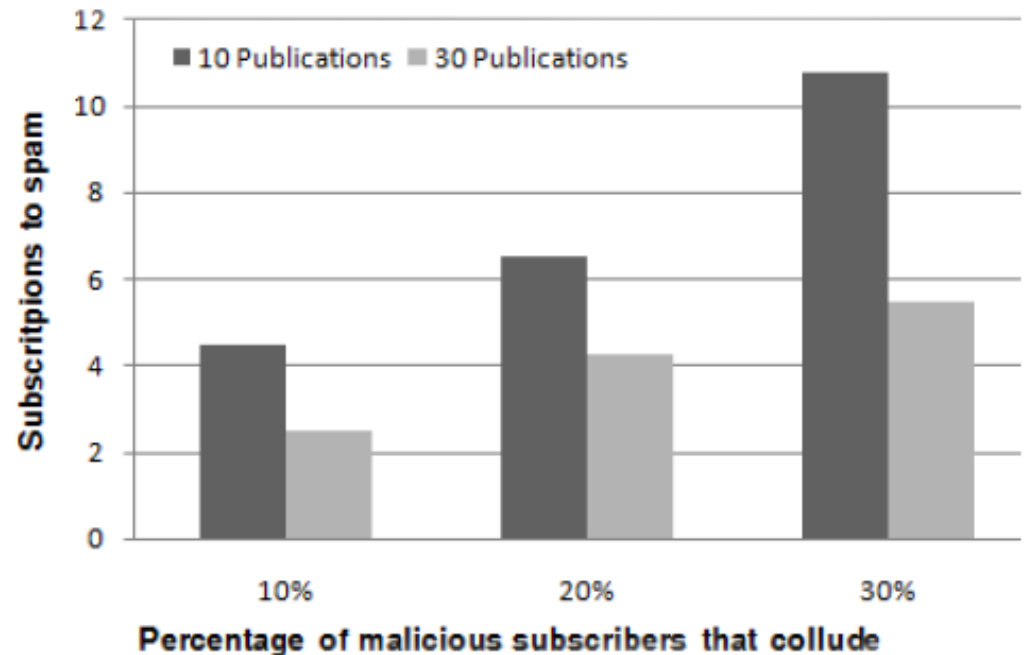Fighting Spam in Publish/Subscribe Networks Using Information Ranking
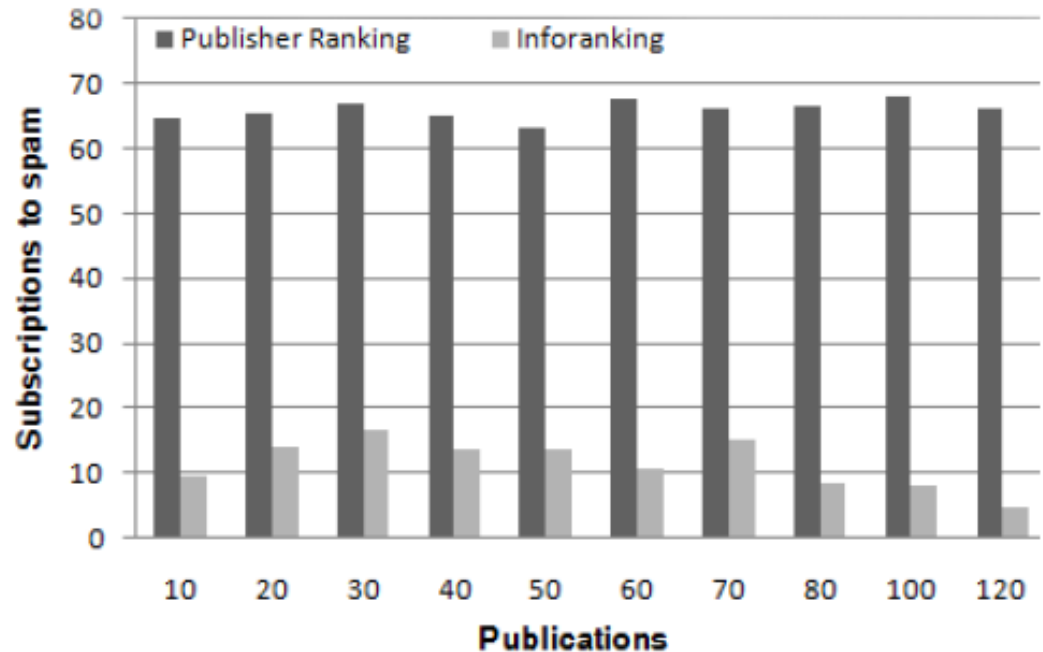
Results (3/4)

- Variable size of publishers publishers, 50% malicious, 100 subscribers

- Each publisher selects a number of publications from a pool of 30 publications

Results (4/4)

- 100 publishers, 100 subscribers, 50% of publishers and subscribers are malicious

- Each publisher selects a number of publications from a pool of 10 or 30 publications

Conclusions and Future Work

- Inforanking is a lightweight solution that can be effectively used for fighting spam in a publish/subscribe architecture
- Future works include
    - Applying inforanking in currently existing applications, such as file sharing applications, bittorrent, voting applications
    - Create a rendezvous system that integrates inforanking for future publish/subscribe architectures
    - Pre-trusted voters

6th NGI Conference, Fighting Spam in Publish/Subscribe Networks Using Information Ranking

Thank you

6th NGI Conference,

Fighting Spam in Publish/Subscribe Networks Using Information Ranking