

P2Prec: A Social-Based P2P Recommendation System*

Fady Draidi, Esther Pacitti, Didier Parigot, Guillaume Verger



Zenith team, INRIA and LIRMM, Montpellier, France

<http://www.sop.inria.fr/teams/zenith/>

Fady.Draidi,Esther.Pacitti@lirmm.fr

Didier.Parigot,Guillaume.Verger@inria.fr

Context: P2P online communities where users want to share their distributed contents (documents).

Goal: Given a query, recommend relevant documents by building a P2P recommender service.

Recommendation Model

Graph: $G = (D, U, E, T)$, where

- D is the set of shared rated documents doc_1, \dots, doc_n
- U is the set of users in the system, corresponding to autonomous peers
- E is the set of edges between the users s.t. there is an edge $e(u, v)$ if users u and v are friends
- T is the set of global topics (automatically extracted using LDA)

Topic management

- $T_u \in T$, is the set of user u 's topics of interest
- $T_u' \in T_u$, is the set of user u 's relevant topics
- q : key-word query that is mapped to topics $T_q \in T$
- A user v is considered **useful** to a user u , if v is a relevant user and has a sufficient amount of documents that are highly related to u 's topics of interest

Friendship network

- Explicit friendship (maybe combined with implicit friendship)
- Expresses trust between users
- Implemented using FOAF files (friend of friends files, in RDF/XML format)

To answer a query, we rely on **useful** and **trusted** friends wrt to T_q .

Latent Dirichlet Allocation (LDA)

We use LDA to automatically model the topics in the system, which in turn are used to extract users' relevant topics of interest. LDA processing is done in two steps: training at a global level and inference at the local level.

Computing Recommendation

Recommendation is based on similarity, rates and popularity of documents:

- $recommendation_q = rank(rec_{q_1}(doc), \dots, rec_{q_n}(doc))$
- where $rank(rec_{q_i}(doc)) = a * sim(doc, q) + b * pop(doc)$

Problem Definition

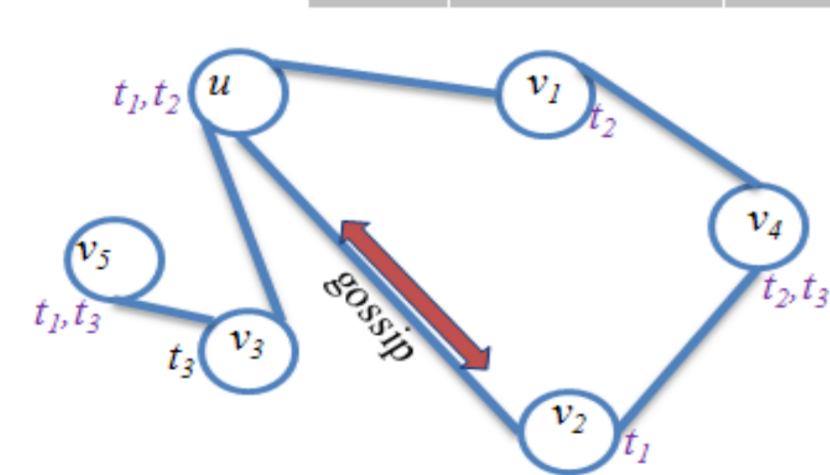
Given a key-word query q , recommend the most confident, valuable and relevant $recommendation_q$ wrt q , and selectively choose the best recommendations to provide useful documents wrt q .

General Approach

Relevant users topics of interest are periodically exchanged. These exchanges are done through **gossip protocols**, in order to explicitly establish friendship with useful users of trust. FOAF files are used as a directory to serve users' queries.

Gossiping

local-view _u before gossip			local-view _u after gossip		
user	T_{u_i}	friend(v_i)	user	T_{u_i}	friend(v_i)
v_2	$(t_1, 0)$	u, v_4	v_2	$(t_2, 1); (t_3, 0)$	v_1, v_2
			v_3	$(t_1, 1); (t_2, 1)$	v_3



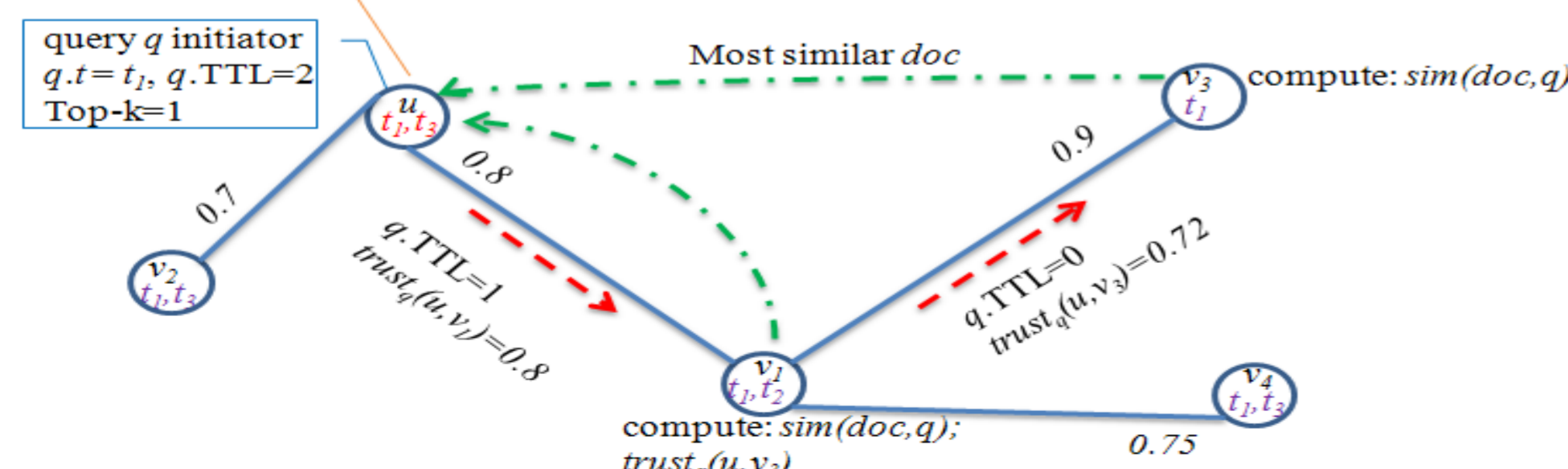
- 1) Each user u maintains a *local-view*
- 2) Each user u periodically selects
 - a random contact v to gossip with
 - a gossip message and send it to v
- 3) Each user u receives a gossip message
 - Updates its *local-view*

local-view _{v2} before gossip			local-view _{v2} after gossip		
user	T_{v_i}	friend(v_i)	user	T_{v_i}	friend(v_i)
v_2	$(t_2, 1); (t_3, 0)$	v_1, v_2	v_2	$(t_2, 1); (t_3, 0)$	v_1, v_2
			u	$(t_1, 1); (t_2, 0)$	v_1, v_2, v_3

Query Processing

A key-word query q is associated with a Time To Live (TTL) and is routed recursively in a P2P top-k manner: once a query is submitted by u , it is forwarded to u 's top-k useful and trustful friends. When a query is received at any peer, it is again redirected to its top-k useful and trustful friends, until TTL is reached. Each user v that received the query provides recommendations to u .

$$rank(doc) = a * sim(doc, q) + b * pop(doc)$$



u FOAF
T_u
Friends:
link to v_1 FOAF: T_{v_1}
$Trust(u, v_1)$
link to v_2 FOAF: T_{v_2}
$Trust(u, v_2)$

- 1) Query q is mapped to topics T_q
- 2) Select Top-k friends in the FOAF wrt to usefulness and trust
- 4) Compute: $trust_q(u, v) = \prod_{v_i, v_j \text{ path}_q(u, v)} trust(v_i, v_j)$
- 3) Redirect Query
- Do 2) and 3) Recursively until TTL

* Work partially funded by the **DataRing** project of the **ANR**

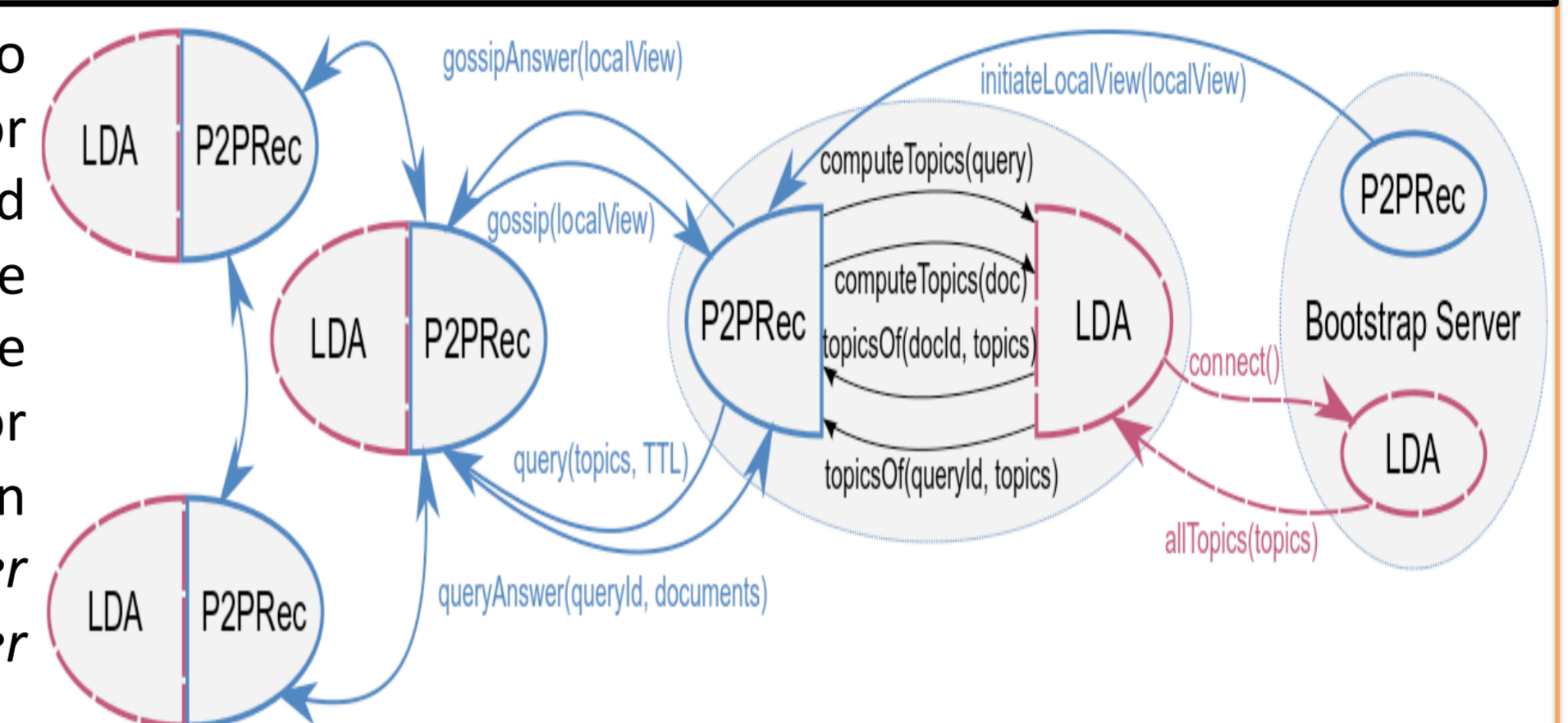
Implementation using the Shared-Data Overlay Network (SON)

SON is an open source development platform for P2P networks based on P2P and SOA concepts. SON components communicate by asynchronous message passing to provide weak coupling between system entities. To scale up and ease deployment, we rely on a Distributed Hash Table (DHT) for publishing and discovering services or data.

<http://www.sop.inria.fr/teams/zenith/SON>

P2Prec Architecture

P2Prec is a SON application with two components: the LDA component for the documents topics process and the P2Prec component for the recommendation process. The P2Prec component provides for passive and active propagation through *gossip* and *gossipAnswer* services and *query* and *queryAnswer* services.



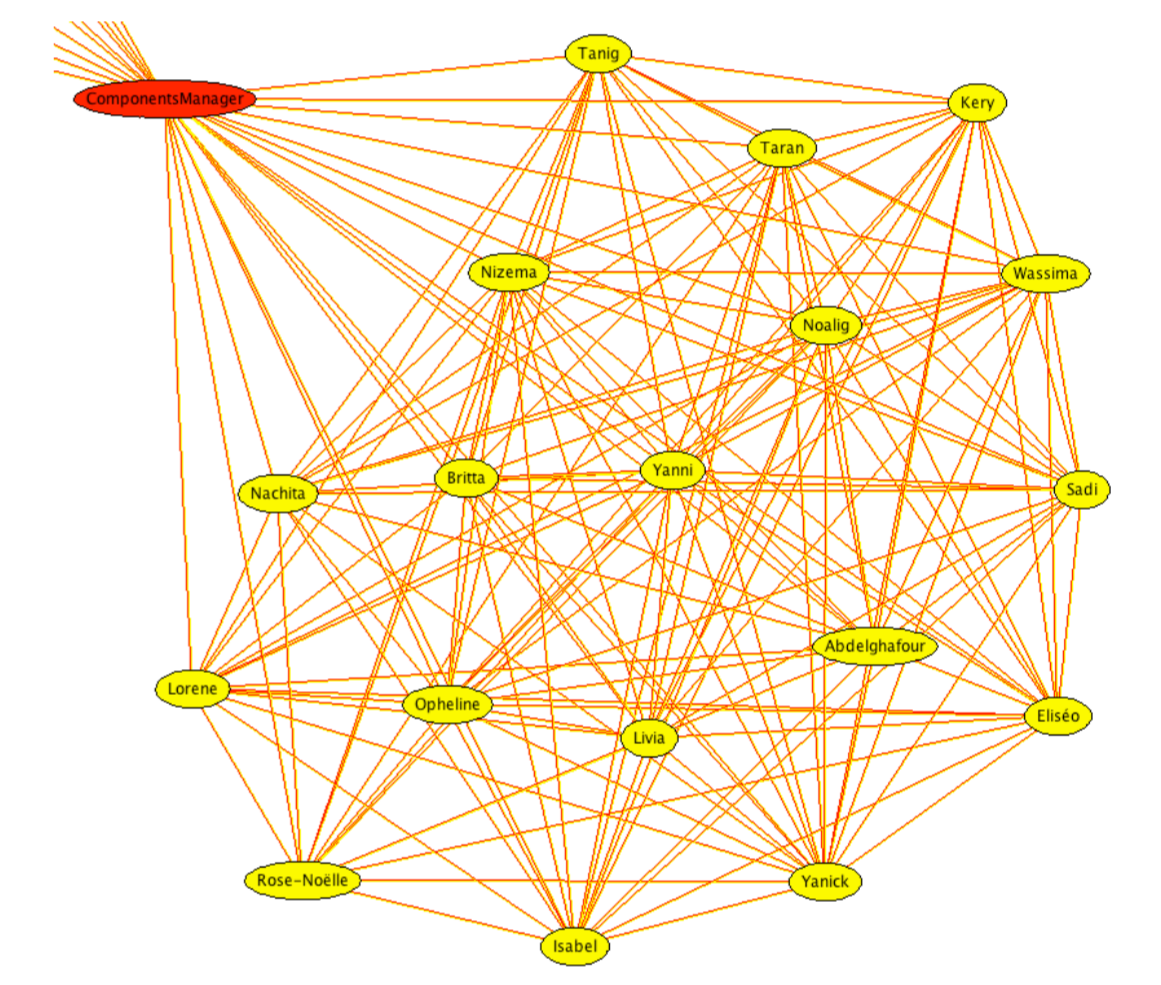
P2Prec Demonstration

The demo with MEDLINE, the on-line medical information database, can be downloaded from the P2Prec website with the complete procedure for installing, deploying and running.

<http://www.sop.inria.fr/teams/zenith/p2prec/>

Installation and Initialization

To run the P2Prec application, a bootstrap server (BS) must be started. At startup time, a new peer must first identify itself with the BS (*connect* service) which returns the current set of all topics (*allTopics* service).



Gossiping and Querying

After initialization, a peer can start performing recommendation and documents discovery without any connection with the BS. Indeed, the research of topics of a new document (*computeTopic(doc)* service) and the computing of topics of a query (*computeTopic(query)* service) can be made locally with the local peer's LDA component.

P2Prec Interface

A Web-interface shows what is internally happening during gossiping. The interface shows the current friends of the user, the gossiping messages sent and received by the peer, the gossip local-view to find friends, etc.

Publications

1. F. Draidi, E. Pacitti, B. Kemme. P2Prec: a P2P Recommendation System for Large-scale Data Sharing. *Journal of Transactions on Large-Scale Data and Knowledge-Centered Systems 3*: 87-116, Springer, 2011.
2. F. Draidi, E. Pacitti, B. M Cart : Leveraging Social and Content-based Recommendation in P2P Systems. *Int. Conf. on Advances in P2P Systems (AP2PS)*, Lisbon, Portugal, 2011.
3. F. Draidi, E. Pacitti, D. Parigot, G. Verger: Demo of P2Prec: a Social-based P2P Recommendation System. *Journées Bases de Données Avancées (BDA)*, Rabat, Morocco, 2011.
4. F. Draidi, E. Pacitti, D. Parigot, G. Verger: P2Prec: A Social-based P2P Recommendation System. *ACM Conference on Information and Knowledge Management (CIKM)*, Glasgow, Scotland, UK, 2011.