Elke Michlmayr, Sabine Graf, Wolf Siberski, Wolfgang Nejdl

Query Routing with Ants
Are ant algorithms suitable for query routing in unstructured peer-to-peer networks?

Elke Michlmayr,
Women's Postgraduate College for Internet Technologies (WIT),
Vienna University of Technology,
http://wit.tuwien.ac.at
Foraging behaviour of natural ants

- **Trail-laying**
  - Ants drop pheromones

- **Trail-following**
  - Ants sense their environment for pheromones and use existing trails

- **Pheromones evaporate over time**

- **Ant algorithms**
  - Operate on graphs
  - Rely on local knowledge only!
Ant Colony Optimization (ACO)

- Di Caro and Dorigo, 1999
- For solving graph-based optimization problems, e.g. travelling salesman problem (TSP)
  - Small number of ants (e.g., 10)
  - Large number of iterations (e.g., 10000)
- Building blocks
  - **Transition rule** derives which link to follow based on amount of pheromone and link costs
  - **Pheromone update rule** defines which amount of pheromone to drop depending on goodness of solution
  - **Evaporation rule** defines the amount of pheromones evaporating in each iteration
- Many instances of ACO: *Ant Colony System*, Ant System, MAX-MIN Ant System, ...
Ant Colony Routing (ACR)

- Routing of data packets in IP networks
- Most prominent: **AntNet** by Di Caro and Dorigo
  - Same building blocks, but no evaporation
  - Forward and backward ants
  - At regular intervals, each node $N_s$ generates a forward ant $F_{sd}$ that builds a path to a randomly chosen destination node $N_d$
    - All nodes in the network must be known!
  - When the forward ant reaches $N_d$, it creates a backward ant $B_{ds}$ that returns to $N_s$ and updates all routing table entries for $N_d$
- Differences to query routing: Each packet has
  - Only one destination node
  - Destination is known in advance
Application scenario

- Distributed search engine
  - Each peer manages a repository of documents

- Taxonomy-based peer-to-peer network
  - All documents are classified by content according to a taxonomy, e.g., ACM Computing Classification System (ACM CCS)
  - Each peer owns a copy of the taxonomy

- Peers pose queries to the network
  - Queries consist of one or more keywords
  - Keywords are connected using Boolean operator AND
  - Set of allowed keywords is limited to the concepts of the taxonomy
  - A document D is an appropriate result for query Q if D is classified to be an instance of all concepts that are keywords of Q
Proposed algorithm SemAnt

- Neither ACO nor ACR are applicable as-is
  - We combine most appropriate features from both and
  - Adapt them to peer-to-peer environment
- Queries
  - Are represented as ants
- Multiple pheromone trails
  - One for each concept in the taxonomy
- Peers
  - Each peer $P_i$ maintains pheromone trails in table $\tau$
    - One row for each neighbouring peer
    - One column for each keyword
  - $P_i$ stores link costs in a table $\eta$
    - One row for each neighbouring peer
  - Evaporation
    - Each peer locally applies evaporation rule in predefined intervals
SemAnt: If a query is issued at peer $P_Q$...

1. Check $P_Q$'s local document repository
2. Create forward ant with start time $T_{Fstart}$ and maximum lifetime $T_{max}$
3. Apply transition rule to select next peer $P_j$
4. Go to $P_j$ and check document repository
5. If results are found, create backward ant $B_Q$
   - $B_Q$ travels back hop-by-hop to $P_Q$
   - At each intermediate peer, $B_Q$
     - Updates link costs $\eta_j$
     - Drops pheromones by applying pheromone update rule
6. Add $P_j$ to list of already visited peers
7. If $T_{Fstart} + T_{max} < \text{CurrentTime}$: continue at 3
   Else: terminate
SemAnt transition rule

- Defines routing strategy
  - Adopted from Ant Colony System
  - Two strategies: weight $w_e$ decides which one to apply

- Exploiting strategy
  - Ants select the best link
  - Depending on amount of pheromones and link costs
  - Works best if paths are perfectly optimized

- Exploring strategy
  - Ants discover new paths
  - Adapted roulette wheel selection technique: for each $P_j$, compute $p_j$, and random value $q \in [0, 1]$ to decide if ant should be cloned and sent to $P_j$

$$j = \arg \max_{u \in U \land u \notin S(FQ)} \left( [\tau_{cu}] \cdot [\eta_u]^\beta \right)$$

$$p_j = \frac{[\tau_{cj}] \cdot [\eta_j]^\beta}{\sum_{u \in U \land u \notin S(FQ)} ([\tau_{cu}] \cdot [\eta_u]^\beta)}$$

<table>
<thead>
<tr>
<th>U</th>
<th>set of neighbouring peers</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S(FQ)$</td>
<td>peers already visited by F</td>
</tr>
<tr>
<td>$\eta_u$</td>
<td>link costs to neighbour $P_u$</td>
</tr>
<tr>
<td>$\tau_{cu}$</td>
<td>pheromone trail to $P_u$ for concept c</td>
</tr>
<tr>
<td>$\beta$</td>
<td>influence of link costs</td>
</tr>
</tbody>
</table>
SemAnt pheromone update rule

- Derives the amount of pheromone that must be dropped on a certain link
  - Adopted from Ant Colony System
- Amount of new pheromones \( Z \) depends on goodness of result
  - Number of documents found
  - Total link costs
- \( Z \) is derived by comparison to an (inexistent) optimal reference result
  - No problem, since always comparing to the same reference
- 50 % of \( Z \) are dropped to superconcept of \( c \)

\[
\tau_{cj} \leftarrow \tau_{cj} + Z, \text{ where}
\]

\[
Z = w_d \cdot \frac{|D|}{d^*} + (1 - w_d) \cdot \frac{T_{\text{max}}}{2 \cdot T_D}
\]
Summary

- **SemAnt**
  - Is an attempt to use ant algorithms in peer-to-peer networks
  - Combines features from Ant Colony System and AntNet
  - Optimizes trail for each keyword depending on its popularity
  - Exploits the underlying taxonomy’s knowledge by reflecting it in the pheromone trails
  - Accounts for network parameters (latency)
  - Work in progress

- **Next step: Focus on dynamic aspects**
  - Joining and leaving of peers
  - Newly added or deleted documents
  - Paper contains first ideas on that
References