Wissenschafterinnenkolleg Internettechnologien

TECHNISCHE UNIVERSITÄT WIEN
VIENNA UNIVERSITY OF TECHNOLOGY

Elke Michlmayr, Sabine Graf, Wolf Siberski, Wolfgang Nejdl

Query Routing with Ants

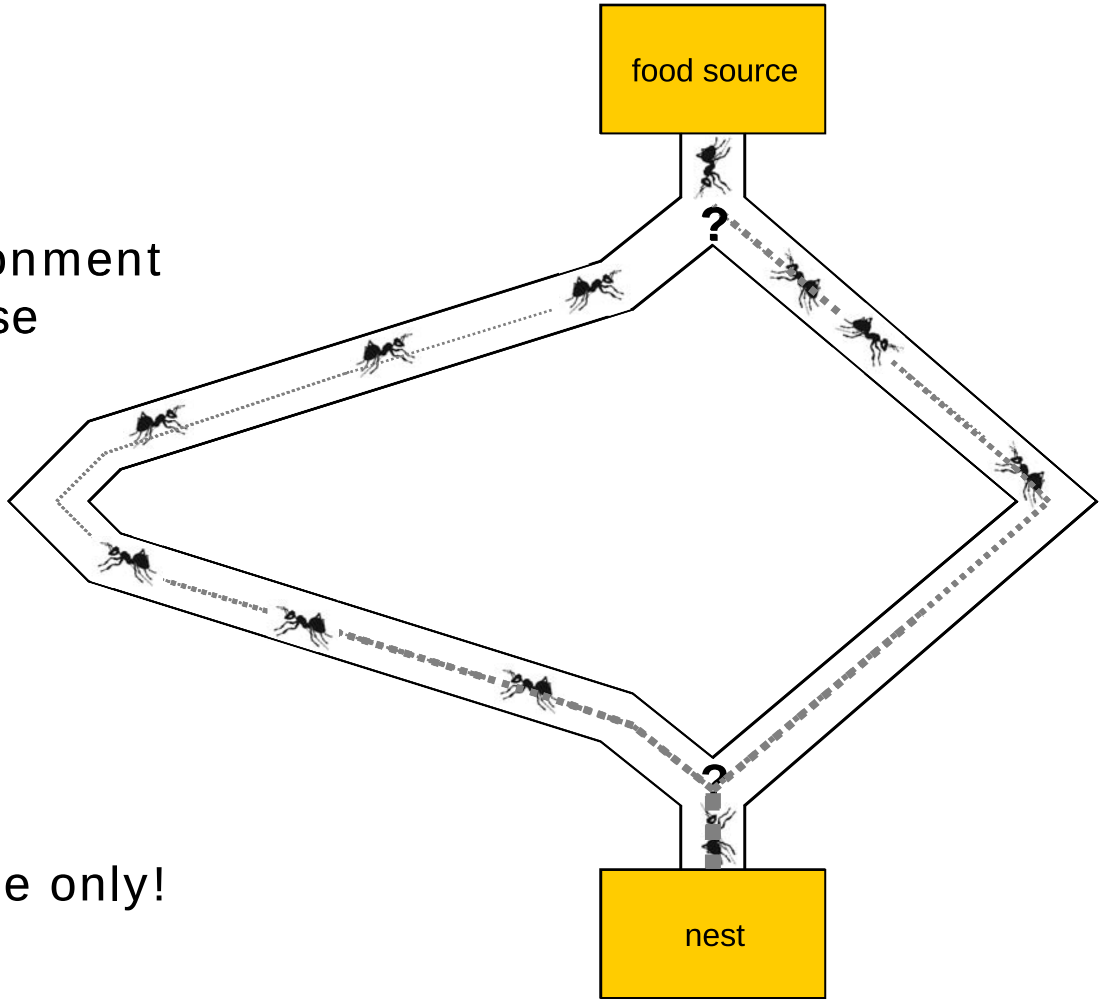
Are ant algorithms suitable for query routing in unstructured peer-to-peer networks?

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

fFORTE bm:bwk esf

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)

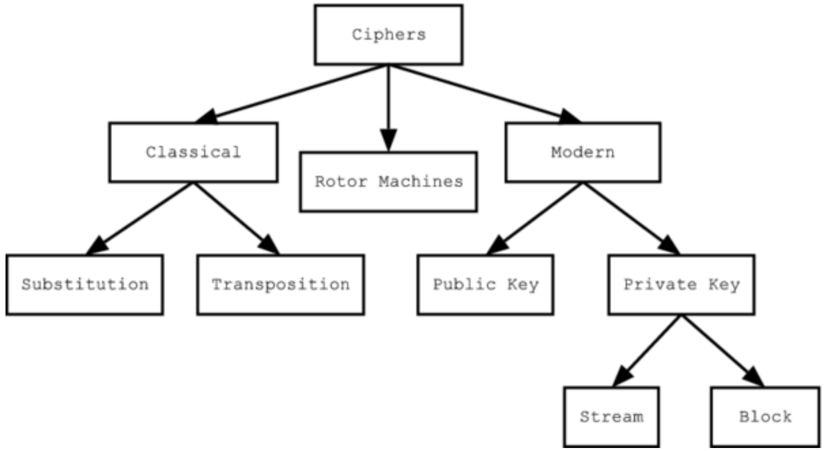
- Di Caro and Dorigo, 1999
- For solving graph-based optimization problems, e.g. travelling salesman problem (TSP)
  - Small number of ants (e.g., 10)
  - Large number of iterations (e.g., 10000)
- Building blocks
  - **Transition rule** derives which link to follow based on amount of pheromone and link 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



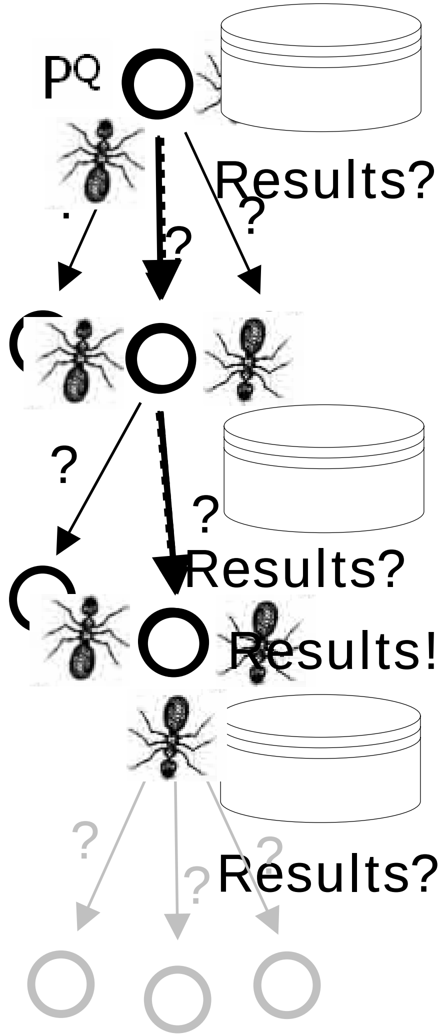
- Taxonomy-based peer-to-peer network
  - All documents are classified by content according to a taxonomy, e.g., ACM Computing Classification System (ACM CCS)
  - Each peer owns a copy of the taxonomy
- Peers pose queries to the network
  - Queries consist of one or more keywords
  - Keywords are connected 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 τ
    - One row for each neighbouring peer
    - One column for each keyword
  - P_i stores link costs in a table η
    - 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 η_j
     - Drops pheromones by applying pheromone update rule
6. Add P_j to list of already visited peers
7. If $T_{Fstart} + T_{max} < CurrentTime$: 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

U	set of neighbouring peers
$S(F^Q)$	peers already visited by F
η_u	link costs to neighbour P_u
τ_{cu}	pheromone trail to P_u for concept c
β	influence of link costs

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

$$p_j = \frac{[\tau_{cj}] \cdot [\eta_j]^{\beta}}{\sum_{u \in U \wedge u \notin S(F^Q)} \left([\tau_{cu}] \cdot [\eta_u]^{\beta} \right)}$$

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

D	documents found
d*	optimal number of docs
T_D	total link costs
T_{max}	maximum lifetime of forward ant
τ_{cj}	pheromone trail to P_j for concept c
w_d	influence of documents resp. link costs

$\tau_{cj} \leftarrow \tau_{cj} + Z$, where

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

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

- E. Bonabeau, M. Dorigo, and G. Theraulaz. Inspiration for Optimization from Social Insect Behaviour. Nature, 406:39-42, 2000.
- G. Di Caro and M. Dorigo. AntNet: Distributed Stigmergy Control for Communications Networks. Journal of Artificial Intelligence Research, 9:317-365, 1998.
- M. Dorigo and L. M. Gambardella. Ant Colony System: A Cooperative Learning Approach to the Traveling Salesman Problem. IEEE Transactions on Evolutionary Computation, 1:53-66, 1997.
- M. Dorigo and G. D. Caro. New Ideas in Optimization, Chapter The Ant Colony Optimization Meta-Heuristic, pages 11-32. McGraw-Hill, 1999.
- M. Guntsch and M. Middendorf. Pheromone Modification Strategies for Ant Algorithms Applied to Dynamic TSP. In Applications of Evolutionary Computing, Lecture Notes in Computer Science, pages 213-222. Springer, 2001.
- O. Babaoglu, H. Meling, and A. Montresor. Anthill: A Framework for the Development of Agent-Based Peer-to-Peer Systems. In Proceedings of the 22nd International Conference on Distributed Computing Systems (ICDCS 02). IEEE, July 2002.