Finding Experts and their Details in E-mail Corpora
Krisztian Balog and Maarten de Rijke

Background
Enterprise search: finding valuable information in an organization’s intranet
• Mixture of document types: web pages, news/email archives, document archives, etc.
• Searching for: services, objects, people, …

Tasks
• Finding experts: given a topic, who are experts on the topic in the organization?
• Expert details: once we’ve found an expert, how do we get in touch with them?

Approach
I. Finding Experts
• What is the probability of a candidate $ca$ being expert on the query topic $q$? -- $p(ca|q)$
• Use Bayes: rank candidates in proportion to $p(q|ca)$
• Use documents as a latent variable between the query and the candidates
• Rank documents according to the query; then determine how likely it is that a candidate is an expert by considering the set of associations
• Formally:

$$p(q|ca) \propto \sum d \cdot p(q|d) \cdot p(ca|d)$$

• Form candidate-document associations by exploiting the fielded structure of e-mail messages

$$p(ca|d) = \sum_{i=0}^{n} \alpha_i A_i (d, ca)$$

II. Mining Contact Details
• Retrieving contact details of experts is a natural component of an operational expert finder
• Signatures contain reliable information
• Signature detection using heuristics (precision oriented)
• Pattern-based extraction of details from signatures

Experimental evaluation
• Lists part of the W3C corpus (198,000 docs)
• TREC 2005 expert search data
  • list of W3C people (candidates)
  • topics, relevance assessments

Results
I. Finding Experts
• 65% of relevant experts found
• MAP = 0.24 (mean average precision)
• RR1 = 0.45 (reciprocal rank of the first relevant result)

II. Mining Contact Details
• Signatures extracted (w3c people): 15,514
• Unique signatures: 3,447
• Personal data found in signatures for 75% of (w3c) people

Example
Query: semantic web coordination

Further work
• Identify more expertise indicators
• Use external (threaded) structure of e-mail
• Extract additional personal details (eg. address, affiliation)