IDENTIFYING COMMUNITIES OF PRACTICE

Analysing Ontologies as Networks to Support Community Recognition

Kieron O'Hara, Harith Alani and Nigel Shadbolt
Intelligence, Agents, Multmedia Group, Dept. of Electronics and Computer Science, University of Southampton

Abstract: Communities of practice are seen as increasingly important for creating, sharing and applying organisational knowledge. Yet their informal nature makes them difficult to identify and manage. In this paper we set out ONTOCOPI, a system that applies ontology-based network analysis techniques to target the problem of identifying such communities.

Key words: communities of practice, ontologies, corporate knowledge management, network analysis, distributed knowledge management

1. INTRODUCTION

A *community of practice* (COP) is a relatively loose, distributed, group of people connected by a shared interest in a task, problem, job or practice. A COP is an informal self-organising group; membership goes to people who wish to improve or maintain their skills (either through pride in achievement, or through a wish to increase earning potential), and who are prepared to enter into discourse about the practice with like-minded people. Membership is not often conscious; members will typically swap war stories, insights or advice on particular problems or tasks connected with the practice (Wenger 1998). An example of a COP might be the set of people in an organisation who do the same (or overlapping) jobs. They understand each other's problems, both with the job itself and with liaison with the outside world. They are the shrewdest judges of each other's work, and often performance is aimed at impressing other COP members rather than line managers. They

---

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: 10.1007/978-0-387-35604-4_20

R. Traunmüller (ed.), *Information Systems*

generally speak a technical jargon that is impenetrable for outsiders. A *de facto* community gradually emerges from their discussions and interests.

Hence a COP is particularly interesting from the point of view of managing distributed knowledge (Prusak and Cohen 1997, pp.149-150). Members of a COP are well-placed to leverage value from the knowledge dispersed through it, to maintain knowledge repositories, and disseminate best practice. Understanding where implicit knowledge is clustered within an organisation is vital for understanding its corporate knowledge capabilities.

Furthermore, COPs can spread across organisations as well as within them, and can help to distribute knowledge beyond a single organisation. Understanding the inter-organisational aspects of a COP can help with a number of decision-making issues such as how to manage alliances and negotiations, how to locate experts, or how to achieve best practice.

However, COP membership is self-selecting and very informal (Wenger 1998); COPs are hard to manage (Wenger 1999). ONTOCOPI (ONTOlogy-based Community Of Practice Identifier - Alani et al 2002) is being developed at the University of Southampton by the Advanced Knowledge Technologies Project (AKT 2001) to help with the first stage of COP management: identifying potential COPs of interest (McDermott 1999), by analysing the strength of (informal) connections between members of an organisation.

2. COMMUNITIES OF PRACTICE

2.1 Communities of Practice: What they are and what they are not

Communities of practice are self-organising groups of people with an interest in a particular practice. The COP might be contained within one organisation, or spread across several. The members have in common a desire to develop their competence, either out of pleasure or pride in their ability, or as a way of making their jobs easier and increasing their earning power. Within an organisation, a COP will mark an autonomous system working in parallel with the (designed) organisational processes (Wenger 1998).

The COP is a focus for situated learning, as members discover how best to practise, and how best to integrate their practice with other aspects of their working and private lives. COP members spread best practice in various unofficial ways, and disseminate "war stories" about bad experiences. In this way, the COP will help its members avoid the effort of relearning what others have already learned – one of the fundamental problems of KM (cf. Smith and Farquhar 2000, for an example of corporate use of a COP).

COPs contrast with other task-based groups of people. For example, a COP is not a *functional group*, a centralised group specialising in a particular function or discipline within an organisation, such as the marketing department. Neither is it like a *team*, which is recruited from within an organisation to perform a particular task, with representatives of relevant disciplines brought in to integrate their functional knowledge to address the problem. In each of these cases, the groups are formal, determinate, and the production of new knowledge is not intended. They are not self-organising, but are set up in order to carry out some function or task concerning the organisation (Grandori 2001, Cohendet et al 2001).

2.2 Communities of Practice, Knowledge and Learning in Business

A COP adds value to an organisation in a number of ways. It has long been recognised that the value of knowledge itself is increased as its community of users grows, i.e. there are increasing returns to scale (Howitt 1996). Learning is closely associated with the understanding of a practice provided by a COP; learning involves becoming a practitioner rather than learning about practice (Brown and Duguid 1991). Hence the loose structure of COPs through an organisation will impact very strongly on the optimisation of training programmes etc. So, for example, Xerox's Integrated Customer Service project shows how understanding the COPs in a workplace can produce on-the-job learning that is more effective at transferring knowledge without removing workers to a classroom for weeks at a time (Stamps 1997).

COPs also have a vital role to play in the management of corporate memory. A large and malleable store of tacit knowledge is required to leverage value from electronic repositories of explicit knowledge (Collins 1974), and COP members understand production processes well enough to be able to identify required knowledge and integrate it into those processes (Marshall et al 1995). This also makes a COP vital for the maintenance of such repositories (Smith and Farquhar 2000).

Hence it is not surprising that COPs, suitably fostered, will be a strong source of added value. Empirical studies of, for example, the biotech industry, have shown that technology cycle time is shorter in community-focused units as opposed to hierarchically organised ones (Judge et al 1997), whereas innovation is more difficult in hierarchical bureaucracies (Dougherty and Hardy 1996).

Theoretical studies come to the same conclusion (Stewart 1996). Any approach to business that focuses on capabilities as a source of competitive advantage (Stalk et al 1992) will be inevitably drawn to analysis of COPs (Liedtka 1999), as it is the COP that supports the extraction of value from processes, as well as ongoing and situated learning (Wenger 1998).

2.3 Strategies for Dealing With Communities of Practice

COPs are difficult to manage (Liedtka 1999), and may require large investment (Stamps 1997); they are self-selecting, nebulous in form, and exist to serve the interests of their members rather than the organisations that employ them. There are, therefore, a number of difficulties with dealing with them, ranging from correct identification through to ensuring that the interests of the COP and the sponsoring organisation coincide.

Any management strategy for COPs will require a first stage of assessing the knowledge requirements of the organisation, and then identifying the COPs that could support those requirements (Wenger 1999, McDermott 1999). A COP has no formal membership, and is not a formal community in any strict sense; the issue of identification is difficult, and generally relies on interview (Wenger 1999).

AKT is currently examining the issues surrounding building a strategic capacity framework, and is attempting to develop software that can extract connections from networks, and to see what software support can be given to the subtask of identifying COPs within and across organisations.

3. ONTOCOPI

In this section, we will set out the principles underlying the ONTOlogy-based Community Of Practice Identifier ONTOCOPI. We begin by discussing the background to it, and then will move onto the system itself. In the next section, we will discuss our attempts to refine its performance.

3.1 The Background to ONTOCOPI

3.1.1 Ontology-Based Network Analysis (ONA)

By ONA we mean analysis of the network of instances and relationships in an ontology (i.e. in a classification structure and knowledge base of instantiations). Ontologies are becoming increasingly important in a number of areas as a result of their knowledge modelling capabilities, the rich infrastructure they provide for knowledge inference techniques, and the possibilities for standardised conceptualisations of domains.

Connections or relations between instances in an ontology can be measured to provide connectedness metrics. The insight behind ONTOCOPI is that if an ontology of the working domain of an organisation is created, then the links between the instances can be measured to indicate which are closely related. If certain canonical members of a COP can be isolated in advance, then ONA can be used to identify other instances related to them.

The effectiveness of ontology analysis for COP identification is dependent to a large extent on the content of the ontology and the properties of the COP. In working with a new COP or ontology, experimentation and trial-and-error must take place, for example to see which relations are important in determining the COP, and how to set relative weights. A priori constraints can be placed on this process, but there will always be domain-specific constraints too.

Note also that the more expressive the ontology (the larger the instantiations, and the wider the subject matter), the more likely that interesting and unforeseen connections will be made. However, the trade-off is that it will be more difficult to filter out noise (e.g. random connections between unrelated instances). Objects and people who exist in more than one COP will be particularly problematic.

Of course, connected groups within an ontology need not necessarily be equivalent to a COP. The essence of a COP is that it is an informal set of relations; ontologies will be wholly or largely made up of formal relations. By 'formal' here we mean relations that are determinate, fixed and cheap to establish/monitor, such as the relation of being a member of a group, being the author of a paper, having a particular telephone number. By 'informal' we mean relations that are often indeterminate and expensive to establish, such as a tendency to have a drink together after work, or a willingness to discuss each other's work. Indeterminacy and expensiveness together ensure that correctly and efficiently identifying a COP will be next to impossible. The hypothesis underlying ONTOCOPI is that (some) informal relations can be inferred from the presence of formal relations. For instance, if A and B have no formal relation, but they have both authored papers (formal relation) with C, then that indicates that they might share interests (informal relation); clearly this is not necessarily true, but is a reasonable enough assumption to support COP identification. We discuss methods of refining the output of ONTOCOPI to take account of potential mismatches between ONA-discovered groups and COPs in section 4. On the basis of this assumption, we will refer to ONA-discoverable groups as 'COPs' in the rest of this paper.

We should emphasise the advantage of using an ontology as opposed to other types of network is that the links in an ontology have semantic significance or types. So, for example, the relation *has author* is a relation between a publication and the person who wrote it. Hence, knowledge represented in an ontology can be analysed by concentrating on particular relations, or giving different weights to relations of different significance.

3.1.2 The AKT Ontology

As has already been hinted, the choice of ontology is an important step for ONA. The material for the experiment is the AKT Ontology, developed by the AKT team at Southampton University to represent the academic computer science domain, including people, projects, papers, events, research topics and so on. The ontology covers more than Southampton, and so could be used for inter- as well as intra-organisational purposes. It is implemented in Protégé 2000, a graphically-based tool for developing knowledge-based systems (Eriksson et al 1999), and was populated with information extracted automatically from departmental databases storing information about people and publications. The ontology is being used to study a number of problems, including coreference discovery and elimination, and issues relating to automatic ontology development, and currently consists of about 80 classes, 40 types of slots and 13,000 instances, mainly of university staff, postgraduates and publications.

Of course, academic ontologies are likely to have different properties from those in other domains (e.g. a focus on publications). However, the analysis techniques do not depend on this; different relations will be essential in identifying COPs, but the techniques used to analyse those relations can remain the same.

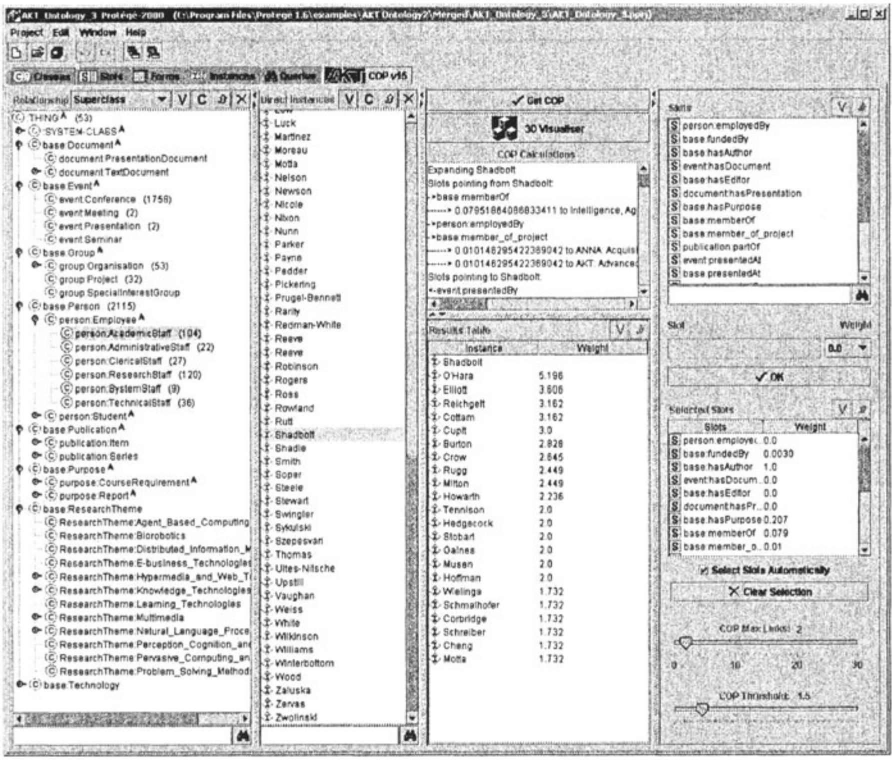
3.2 The Design of ONTOCOPI

As we have noted, the ONTOCOPI system is plugged into Protégé 2000, and uses the AKT ontology as its raw material. The user interface is shown in *Figure 2*. The two panels on the left are taken from the Instances plugin of Protégé, which allows the user to browse an ontology and select a class, for which the system will display all available instances. The user can select a specific instance from the list, and then click the *get COP* button to run a spreading activation algorithm to identify close instances in the instantiated ontology within limited expansion and weight thresholds, which can be altered from scrollbars on the right hand panel. Results are displayed on the two panels in the middle.

In this section, we will examine the design of ONTOCOPI further. See also (Alani et al 2002).

3.2.1 The System

The essence of ONTOCOPI is that relations can be selected according to their relevance to the practice to be uncovered, and different weightings can be set to reflect their relative importance. However, selecting the appropriate relations and weighting them properly is a challenging task (Alani et al 2002). Two approaches are followed: allowing the user to select and weight the relations to traverse manually; and an automatic approach to selection and weighting based on the number of times each has been used in the KB.



```
Expanding Shadbolt
Slots pointing from Shadbolt
->base:memberOf
------> 0.07951884086833411 to Intelligence, Ag
->person:employedBy
->base:member_of_project
------> 0.010148295422389042 to ANNA: Acquisi
------> 0.010148295422389042 to AKT: Advance
Slots pointing to Shadbolt
<-event:presentedBy
```

Instance	Weight
Shadbolt	
O'Hara	5.196
Elliott	3.606
Reichgelt	3.162
Cottam	3.162
Cupit	3.0
Burton	2.828
Crow	2.645
Rugg	2.449
Milton	2.449
Howarth	2.236
Tennison	2.0
Hedgecock	2.0
Stobart	2.0
Oaines	2.0
Musen	2.0
Hoffman	2.0
Wielinga	1.732
Schmalhofer	1.732
Corbridge	1.732
Schreiber	1.732
Cheng	1.732
Motta	1.732

Slots:
- person:employedBy
- base:fundedBy
- base:hasAuthor
- event:hasDocument
- base:hasEditor
- document:hasPresentation
- base:hasPurpose
- base:memberOf
- base:member_of_project
- publication:partOf
- event:presentedAt
- base:presentedAt

Slots	Weight
person:employe...	0.0
base:fundedBy	0.0030
base:hasAuthor	1.0
event:hasDocum...	0.0
base:hasEditor	0.0
document:hasPr...	0.0
base:hasPurpose	0.207
base:memberOf	0.079
base:member_o...	0.01

Figure 2: The user interface of ONTOCOPI

Manual selection is relatively straightforward, and the less weight given to a selected relation, the less its impact will be. The user must produce a set of weights that transitively records his preferences, and then can set the system going. The advantage of manual selection is that the user has full control, and can experiment to an extent with the selections and weights. The disadvantage, of course, is that finding the "right" set of relations and weights is likely to be difficult. For example, the effect of a relation's weight on the results is proportional not only to the weight of the other selected relations, but also to the number of those relations in the KB.

The user can check a box to let the system select the relations and calculate their weight depending on the frequency of use of those relations in the KB. The frequency of use of relationships in a KB can be used as an indication of the level of importance of these relationships to that ontology, or whether the ontology is good or weak in providing information related to certain relationships. The system counts the number of times each relationship is used in the KB, then normalises the results to be in the range of 0 to 1. A weight of 1 is given to the most used relationship (*has author* in the AKT ontology), and 0 to the ones that are never used. Relationships with zero weights are excluded from the analysis.

The advantage of this approach is that it is automatic, where users do not need to be aware of the ontology structure or the weighting criteria. The disadvantage of the approach is that the weighting is rough, and does not take into account the user's relationship preferences. In other words, their semantics are disregarded. This however has the advantage of making the approach more generic and can be applied to any ontology regardless of the type of relationships it represents.

An example output of ONTOCOPI can be seen in figure 1. The user selected the class *AcademicStaff*, the instance *Shadbolt*, and set the link and weight thresholds to 2 and 1.5 respectively as can be seen in the right panel. With the automatic relationship selector switched on, ONTOCOPI identified the COP of the query instance (*Shadbolt*) as displayed in the two-column table in the middle panel. The left column lists the instances and the right one lists their relative weights to the identified COP. The more an instance has in common with *Shadbolt*, the higher the rank of this instance will be in the identified COP. However, this is dependent on the type of shared information (e.g. a joint publication, a shared research interest, a common project) as some relationships can be regarded as more important than others. Selecting the relationships automatically assigns higher importance values to authorship and membership relationships, as they are often used in the KB. For this reason, *O'Hara* was found to be the most relevant person to *Shadbolt's* COP, as they are in the same department and research group, share a project together, and have co-authored more than 27 papers. Next is *Elliot* with also a shared department and project, but with fewer joint publications with the query instance, and so on. Different relationships can be selected and weighted to identify COPs that focus on other specific semantic relationships.

3.2.2 The Algorithm

After selecting the main instance and weighting the relationships, the expansion algorithm can be activated to retrieve a set of close instances and rank them accordingly. The algorithm applies a breadth first, spreading activation search where a term would be expanded by traversing relationships until an expansion threshold is reached. The algorithm takes a single instance (e.g. a specific person or project) and crawls the ontology KB one link at a time. Starting with a weight of 1, this crawler passes weights to all other instances following a set of weighted relationships. The amount of weight passed equals the weight of the primary instance multiplied by the weight of the relation being traversed. The algorithm is discussed in detail in (Alani et al 2002).

3.2.3 Visualisation

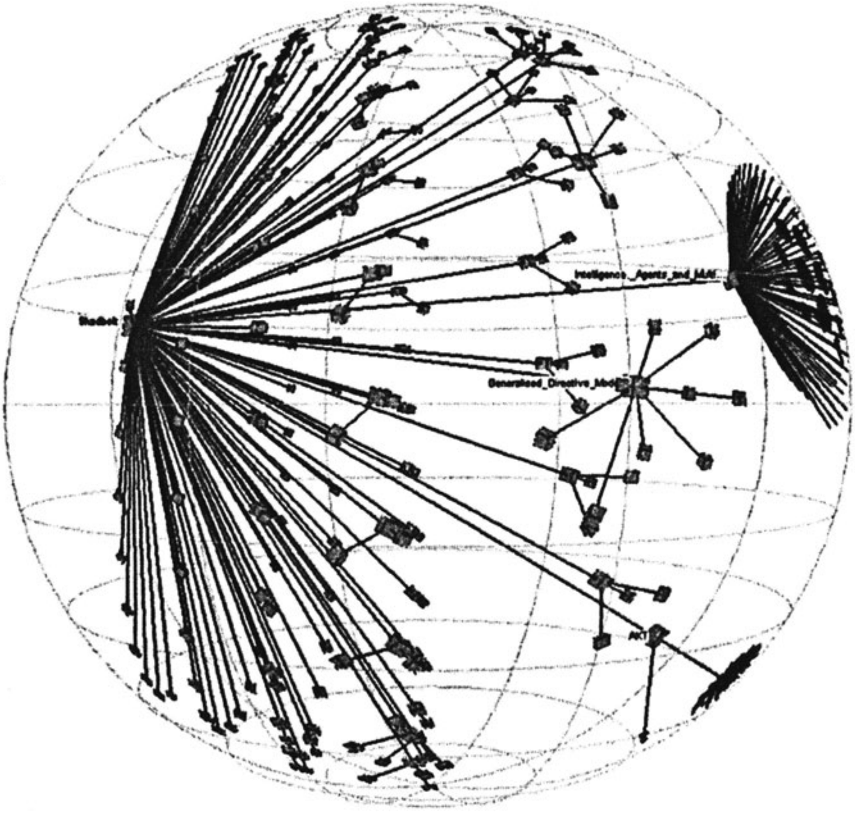


Figure 3: Visualisation of the shortest paths from *Shadbolt* to other nodes within 4 links

The visualisation uses H3View software (Munzner 1997), designed for viewing large networks, to get a 3D view of results, allowing the rotation of the result network, and bringing certain nodes to focus while dithering others. The view shown here is of the shortest paths from *Shadbolt* to other nodes four links away. Islands of nodes within the given range can be seen as dense trees. For example, the tree on the right is the IAM group, which is highly interconnected. Another tree to the bottom right follows on from the node *Glaser*, an academic at Southampton working on AKT, though not in IAM. The tree indicates a cluster of nodes for which the shortest path to them from *Shadbolt* is through *Glaser*; this is what we would expect, as Glaser works in a different group and has connections to other COPs.

4. LINK ANALYSIS AND COMMUNITIES OF PRACTICE

The system as described above is the raw version. Connections between entities can be quite arbitrary. *Shadbolt* has written papers in a number of fields, and hence instances with connections to him will not necessarily be members of the same COP; there is little connection between *Motta* and *Elliot*, both of whom have co-authored papers with *Shadbolt*. The trick is to filter out irrelevant connections where possible. Naturally the results are unlikely to be perfect, even if perfection can be defined in an informal COP. But effective filtering is essential for support for COP identification.

Research is ongoing into such filtering. Space precludes a detailed discussion of the experiments (Alani et al 2002), but we can indicate some of the degrees of freedom. First of all, the number of links that ONTOCOPI will traverse can be set. Traversing one link away from *Shadbolt* will reach his co-authors, events he has attended, supervisees, etc. Two links will reach anything connected with them, and so on. The more links chosen, the more likely interesting connections will be made, but the greater the risk of noise.

More specific types of entity can be found by manually selecting and weighting the relations of interest. For instance, by increasing the weight of the *member of* relationship, more of *Shadbolt's* colleagues are uncovered, and *their* colleagues, and *their* colleagues, and so on.

Time slots in the AKT ontology allow ONTOCOPI to look for temporal restrictions on relations, dealing with the dynamic nature of COP membership. Of course, if such temporal information is not present in the analysed ontology, ONTOCOPI will not be able to extract such information; ONTOCOPI cannot extract information that isn't there!

A research issue for ONTOCOPI is to exploit semantic information implicit in an ontology. Experiments are taking place to analyse the paper abstracts using Term Frequency Inverse Document Frequency (TFIDF) heuristics (Salton and McGill 1983). These enable key words (that appear frequently in the abstract but not in other texts) to be extracted, and used to pick out papers particular to the COP under investigation. The relevance of the key words to the COP could then determine the weight to place on the paper and its relations in the ONA. The COP could be characterised through key words selection by hand, or TFIDF analysis of selected central papers.

5. RELATED WORK

Analysis of information networks has been the focus of much research in a variety of domains. For example, there has been much research on rating the importance of web pages by identifying hubs and authoritative sites (Page et al 1999, Gibson et al 1998, Flake et al 2000, Kleinberg 1999) using the number and direction of links between pages. Even though some of these methods have proven to be successful for certain Web requirements, hyperlinks lack semantics; the reasons links are authored are lost, and therefore their significance is hard to measure.

Dunlop (2000) investigated methods to cluster people hierarchically based on their personal Web pages, to help identify employee replacements, and reconnect broken collaborations. Term frequency measures were used to compare home pages and produce a list of pages ranked according to their similarity to an employee's page. The limitation of this approach is that personal web pages aren't always representative. Furthermore, sub-pages often cause problems, as they are not included in the Web page comparison.

Kautz et al (1997) investigated building social networks automatically from web pages related to an individual as retrieved in an AltaVista search. People were considered to be related if their names co-occurred in the same web page/s. The problem here lies in the equal treatment of all inferred relations. Some pages could provide lists of people who, for example, attended a certain conference (weak social relation), or co-authored a paper together (strong social relation). Relation significance cannot be inferred.

Collaboration networks represent professional interactions within an organisation. Newman (2001) constructed networks of co-authorship and applied different methods to find the best-connected scientist, taking into account such factors as the number of papers written by each author, number of collaborators the authors have, network distance between them, etc. Strength of collaboration between pairs of scientists was measured by the number of papers they co-authored together, ignoring all other relationships.

Self-organising maps (SOMs) are neural network algorithms that can model large data sets to find central dependencies in complex data (Kohonen

2001). SOMs could detect similarity clusters in the information provided by an ontology. However, user-adjustment of feature (i.e. relation) weights must be possible to replicate the results of ONA; furthermore, SOMs would have difficulty encoding information about units further than one relation away.

6. CONCLUSIONS AND FUTURE WORK

We have set out the principles underlying ONTOCOPI, a system for supporting the identification of COPs. There are many economic reasons why COPs need to be identified, as they can be used for knowledge transmission, learning and corporate memory, recruitment and expert location. Currently COP identification is largely done by structured interviews, a resource-heavy technique. The automatic analysis of ontologies will not by itself solve this problem, but could provide useful support, particularly in widely-distributed organisations.

However, it is important to express the following caveats.

1. An ontology is likely to contain mainly or entirely formal relations, where a COP is informal. ONTOCOPI makes an explicit assumption that (some) informal relations can be inferred from formal ones.
2. The content, extent and scope of an ontology will determine the effectiveness of the COP identification at least as much as the system.
3. In any new domain, a range of trials would have to be carried out to determine the interesting link thresholds and relation weights. In particular, it is an open question as to whether the same thresholds and weights will give adequate answers for different instances.
4. Automatic relation selection/weighting relies on the reasonable but hardly watertight assumption that the most-used relations in the ontology are the most relevant to a COP.
5. Both identifying and filtering out noise will be problematic without knowledge of the domain. Note also the problem of *brokers* (people who exist in two COPs) or *boundary objects* (objects with meaning for two COPs). In cases such as these the communities identified may look more like the *union* of two or more COPs. It could be that this is a widespread problem, though Wenger (1998) does not think so. Much will depend in an individual instance on what information is represented in the ontology under investigation, and what mechanisms are available for filtering (e.g. temporal relations).

It is essential that (a) the COP can be characterised accurately enough, and (b) that an ontology is available that will contain the required information. The actual nature of the analysis will also depend heavily on the COP itself. However, by analysing the links between entities, it is possible to filter out the noise to a sufficient extent as to allow the essential connections between the knowledge resources in an organisation to be revealed, without a great deal of computational or human resource overhead.

Future research will focus on further ways of filtering out noise and making the search process more flexible. Further scenarios will also be employed to ascertain in which other knowledge management tasks such ontology-based network analysis can be exploited (for example, for the discovery of coreference in ontologies created from heterogeneous resources). Investigation will also be required into the provision of explanations to the users, both in terms of (a) the information that users will require or ask for respecting the algorithm, and (b) how to present that information to the user.

7. ACKNOWLEDGMENTS

This work is supported under the Advanced Knowledge Technologies (AKT) Interdisciplinary Research Collaboration (IRC), which is sponsored by the UK Engineering and Physical Sciences Research Council under grant number GR/N15764/01. The AKT IRC comprises the Universities of Aberdeen, Edinburgh, Sheffield, Southampton and the Open University.

8. REFERENCES

AKT (2001) *The AKT Manifesto* http://www.aktors.org/publications/Manifesto.doc

Harith Alani, Kieron O'Hara and Nigel Shadbolt (2002) 'ONTOCOPI: methods and tools for identifying communities of practice' *Proceedings of 2002 World Computer Congress, Intelligent Information Processing Stream* (Montreal)

John Seely Brown and Paul Duguid (1991) 'Organizational learning and communities of practice: toward a unified view of working, learning and innovation' in Lesser et al, 99-121

Patrick Cohendet, Frederic Creplet and Olivier Dupouet (2000) 'Interactions between epistemic communities and communities of practice as a mechanism of creation and diffusion of knowledge' in D. Delli Gatti, M. Gallegati and A. Kirman (eds.) *Interaction and Market Structure: Essays on Heterogeneity in Economics* (Springer-Verlag, Berlin)

Harry M. Collins (1974) 'The TEA set: scientific knowledge in tacit networks' *Science Studies* 4, 165-186

D. Dougherty and C. Hardy (1996) 'Sustained product innovation in large, mature organization: overcoming innovative to organization problems' *Academy of Management Journal* 39, 1120-1153

M.D. Dunlop (2000) 'Development and evaluation of clustering techniques for finding people' *Proceedings of 3rd International Conference on Practical Aspects of knowledge Management – PAKM'2000* (Basel, Switzerland)

H. Eriksson, R. Fergerson, Y. Shahr and M. Musen (1999) 'Automatic generation of ontology editors' *Proceedings of the Twelfth Workshop on Knowledge Acquisition, Modelling and Management (KAW'99)* (Banff, Canada)

Gary Flake, Steve Lawrence and C. Lee Giles (2000) 'Efficient identification of web communities' *Proceedings of the 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Boston)

D. Gibson, J. Kleinberg and P. Raghavan (1998) 'Inferring web communities from link topology' *Proceedings of the 9th ACM Conference on Hypertext and Hypermedia*

Anna Grandori (2001) *Organization and Economic Behavior* (Routledge, London)

Peter Howitt (1996) 'On some problems in measuring knowledge-based growth' in Dale Neef (ed.) *The Knowledge Economy* (Butterworth Heinemann, Boston, 1998) 97-117

William Q. Judge, Gerald E. Fryxell and Robert S. Dooley (1997) 'The new task of R&D management: creating goal-directed communities for innovation' in Lesser et al, 37-51

Henry Kautz, Bart Selman and Mehul Shah (1997) 'The hidden web' *AI Magazine* 18(2) 27-36

Jon Kleinberg (1999) 'Authoritative sources in a hyperlinked environment' *Journal of the ACM* 46

T. Kohonen (2001) *Self-Organizing Maps* 3rd Edition (Springer, Berlin)

Eric L. Lesser, Michael A. Fontaine and Jason A. Slusher, eds., (2000) *Knowledge and Communities* (Butterworth-Heinemann, Boston)

Jeanne Liedtka (1999) 'Linking competitive advantage with communities of practice' in Lesser et al, 133-150

Catherine C. Marshall, Frank M. Shipman III and Raymond J. McCall (1995) 'Making large-scale information repositories serve communities of practice' in Lesser et al, 225-247

Richard McDermott (1999) 'Why information technology inspired but cannot deliver knowledge management' in Lesser et al, 21-35

Tamara Munzner (1997) 'H3: laying out large directed graphs in 3D hyperbolic space' *Proceedings of the 1997 IEEE Symposium on Information Visualisation* (Phoenix), 2-10

Mark Newman (2001) 'Who is the best connected scientist? A study of scientific coauthorship networks' *Phys. Rev. E.* 64

Lawrence Page, Sergey Brin, Rajeev Motwani and Terry Winograd (1999) *The PageRank Citation Ranking: Bringing Order to the Web* (Stanford University working paper SIDL-WP-1999-0120)

Laurence Prusak and Don Cohen (1997) 'Knowledge buyers, sellers and brokers: the political economy of knowledge' in Dale Neef, G. Anthony Siesfeld and Jacquelyn Cefola (eds.) *The Economic Impact of Knowledge* (Butterworth Heinemann, Boston, 1998) 137-159

G. Salton and M. McGill (1983) *Introduction to Modern Information Retrieval* (McGraw-Hill, New York)

Reid G. Smith and Adam Farquhar (2000) 'The road ahead for knowledge management: an AI perspective' *AI Magazine*, Winter 2000, 17-40

G. Stalk, P. Evans and L. Shulman (1992) 'Competing on capabilities: the new rules of corporate strategy' *Harvard Business Review* 70, 57-69

David Stamps (1997) 'Communities of practice: learning is social. Training is irrelevant?' in Lesser et al, 53-64

Thomas A. Stewart (1996) 'The invisible key to success' *Fortune*, 5th August 1996, 173-176

Etienne Wenger (1998) *Communities of Practice: Learning, Meaning and Identity* (Cambridge University Press, Cambridge)

Etienne Wenger (1999) 'Communities of practice: the key to knowledge strategy' in Lesser et al, 3-20