Steffen Staab
University of Koblenz-Landau
sst@aifb.uni-karlsruhe.de

Social Networks Applied

Social networks have interesting properties. They influence our lives enormously without us being aware of the implications they raise: How does a kind of fashion become en vogue? How does a virus spread and infect people? How does a research topic become a hot topic? Why are some companies successful and others aren't? All these questions affect us, and understanding them by building and investigating computational models might give us a powerful tool to improve our health system, increase individual and general wealth, or just increase awareness about how the people around us actually influence our opinions, which we frequently believe that we shape.

Pedro Domingos investigates how to exploit our unprecedented wealth of data and how we can mine social networks for purposes such as marketing campaigns. It's fascinating how such techniques could turn current marketing strategies upside down.

Peter Mika considers a particular form of influence: the way that people agree on terminology and this phenomenon's implications for the way we build ontologies and the Semantic Web. In a nutshell, he concludes that the Semantic Web will either include social networks' influence in its architecture or wither away.

While Peter and Pedro target social networks as something we can discover from data, Jennifer Golbeck takes a constructivist view: people already provide explicit social network information in formats such as friend-of-a-friend files. If we refine this kind of information, we could offer a wealth of new applications, such as better recommendations for restaurants, trustworthy email senders, or (maybe) blind dates.

Li Ding, Tim Finin, and Anupam Joshi also take a constructivist view by investigating the richness and difficulty of harvesting FOAF information.

Andrzej Nowak and Robin R. Vallacher conclude this issue from the viewpoint of social scientists who've studied extensively how information processing is bound to social context. They point us to the intriguing ways that network topology's definition determines its outcomes.

—Steffen Staab

Mining Social Networks for Viral Marketing

Pedro Domingos, *University of Washington*

Traditionally, social-network models have been descriptive rather than predictive. They're built at a very coarse level, typically with only a few global parameters, and aren't useful for predicting the network's behavior. In the past, this was due largely to lack of data; the networks available for study were small and few and contained minimal information about each node.

Fortunately, the Internet's rise has changed this dramatically. Massive quantities of data on large social networks are available from blogs, knowledge-sharing sites, collaborative-filtering systems, online gaming, social-networking sites, newsgroups, chat rooms, and so on. These networks typically number in the tens of thousands to millions of nodes. They often contain sufficient information to build models of individual nodes, which we can then assemble into models of the networks they're part of. This gives us an unprecedented level of detail in social-network analysis, along with the potential for new understanding, useful predictions, and their productive use in decision making.

My colleagues and I have begun to build social-network models at this scale using data from the Epinions knowledge-sharing site, the EachMovie collaborative-filtering system, and others.[1,2] These models let us design viral-marketing plans that maximize positive word of mouth among customers. In our experiments, this has made it possible to achieve much higher profits than if we ignored interactions among customers and the corresponding network effects, as traditional marketing does.

Customers' network values

Customer value is usually defined as the expected profit from sales to a customer over the lifetime of his or her relationship to the company. Customer value is of critical interest to companies because it determines how much is worth spending to acquire a particular customer. However, traditional measures of customer value ignore the fact that in addition to buying products, a customer can influence others to buy them. For example, if I see a movie and persuade three friends to see it with me, my customer value with respect to that movie has effectively quadrupled, and the movie studio is justified in spending more on marketing the movie to me. Conversely, if I tend to decide what movies to see on the basis of what my friends tell me, marketing to me might be a waste of resources that would be better spent

1541-1672/05/$20.00 © 2005 IEEE
Published by the IEEE Computer Society

marketing to my friends. A customer's *network value* is the expected increase in sales to others that results from marketing to that customer.

Clearly, ignoring customers' network values, as traditional direct marketing does, can lead to suboptimal marketing decisions. But while the marketing literature has acknowledged network effects' existence, it has generally considered them unquantifiable, particularly at the individual-customer level. The data sources now available have changed this.

Our models let us measure a customer's network value. We model how likely each customer is to buy some product, as a function of the customer's and product's intrinsic properties and of the influence of the customer's neighbors in the network. By performing probabilistic inference over the joint model of all customers, we can answer questions such as, "If we market to a particular set of customers, what's the expected profit from the whole network after those customers' influence has propagated throughout?" Using this capability, we can search for the optimal set of customers to market to—that is, the set that will yield the highest return on investment. Intuitively, we can look for the customers with the highest network values, market to them, and reap the benefits of the ensuing wave of word of mouth.

Factors that influence network value

What makes for a customer with high network value? Clearly, high connectivity in the network should help, but our model identifies other factors.

Customer opinion

First, it's important that the customer like the product, preferably a lot. Customers who have high connectivity but dislike a product can have negative network value, and we should avoid marketing to them. In our experiments with EachMovie, our model took this into account, which helped it outperform a standard direct-marketing approach. The standard approach assumed that the most it had to lose by marketing to a customer who didn't like the product was the marketing's cost, which is typically small per customer. As a result, the standard approach marketed even to customers whose chances of liking the product were relatively low.

Asymmetric influence

Another key aspect is that to have high network value, a customer should influence his or her acquaintances more (ideally much more) than they influence that customer. If influence is symmetric, searching for the most influential customers has no advantages. Fortunately, asymmetric influence is widespread in practice, and our approach exploits it. While various fields have well-known opinion leaders, such as celebrities, our approach lets us identify them at the local level.

Chain of influence

Perhaps most important, a customer's network value doesn't end with his or her immediate acquaintances. Those acquaintances in turn influence other people and so on until they potentially reach the entire network. A customer who's not widely connected might, in fact, have high network value if an acquaintance is highly connected (for example, an advisor to an opinion leader). In our experiments with the Epinions Web site, the most valuable customer had a network value of over 20,000, meaning that marketing to him was as effective as marketing to over 20,000 others in the absence of network effects. However, the customer's number of direct links to others in the network—that is, people who read his reviews—was much smaller.

What makes for a customer with high network value? Clearly, high connectivity in the network should help, but our model identifies other factors.

Our model's consequences

Word-of-mouth marketing might not be effective in some markets because the requisite influence networks aren't present. Although this is known at a high level for some market types,[3] many startup companies have failed by investing heavily to unleash network effects that never materialized. Conversely, trials for some products, such as cash cards and interactive television, have failed because the companies didn't appreciate that giving the product to a small sample of isolated customers doesn't allow network effects to take hold. When the data is available, our models let us measure these effects precisely and make better decisions.

Another interesting consequence of our model is that it might pay to lose money on some customers, if they're influential enough. In traditional direct marketing, customers receive an offer only if their expected profits exceed the offer's cost. In viral marketing, giving a free product to a well-chosen customer could pay off many times in sales to other customers.

Maximizing word of mouth

Given a social-network model, we have a well-defined optimization problem: choose the set of customers to market to so as to maximize net profits—that is, profits from sales minus marketing's cost. David Kempe, Jon Kleinberg, and Éva Tardos have shown this problem to be NP-hard, but approximable within 63 percent of the optimal using a simple hill-climbing search procedure.[4] We obtained similar results with an even faster approach: we added each customer to the current marketing set as long as this improved overall profit.

With careful implementation, the potentially prohibitive cost of performing probabilistic inference over the whole network at each search step (necessary to measure the effect of adding a customer to the marketing set) also turns out not to be a problem. This is because the vast majority of customers have low network values; their influence doesn't propagate far, so the computation for them converges quickly. For the few customers with high network values, the computation can indeed take substantial time, but amortized over all search steps, it becomes quite manageable. We found the optimal marketing set for a network with tens of thousands of nodes in minutes.

No matter how much data we have, completely capturing the network of social interactions among people in the real world will never be feasible. So, the important question arises of whether our approach to maximizing word of mouth still works when our knowledge of the network is incomplete. We've tested this by randomly removing a variable number of edges from the network before passing it to the data-mining system.

We found the system to be quite robust, with 70 percent of the total increase in profit obtained when the system knew only five percent of the edges. We can also use our model to determine the most cost-effective way to gather additional knowledge. We've found that the simple heuristic of iteratively asking the customers with the highest network value who their acquaintances are is quite effective.

Prospects

Traditional marketing is in crisis because customers are increasingly inured to television commercials, direct mailings, and so on. At the same time, companies such as Amazon, Google, and Hotmail succeed with virtually no marketing, solely on the basis of word of mouth.[3] A recent study found that positive word of mouth among customers is by far the best predictor of a company's growth.[5] Word-of-mouth marketing has a key advantage: a recommendation from a friend or other trusted source has credibility that advertisements lack.[6] Because it leverages customers themselves to do the marketing, it can also produce unparalleled returns on investment. Until now, it's been somewhat of a black art. Our goal is to put it on a firmer foundation, and the results so far are promising.

Beyond marketing, word-of-mouth optimization is potentially applicable in any setting where we desire a large social outcome with limited resources. Examples include reducing the spread of HIV, combating teenage smoking, and grassroots political initiatives. Until recently, sociology lagged behind other sciences in developing a computational branch. The wealth of social data the Internet provides could change this, and we see our work as a step in this direction.

We've only begun to scratch the surface of the rich set of possibilities that building predictive social-network models opens up. Real social networks evolve and have multiple types of arcs and nodes. Multiple players' actions affect them, and we can mine them from a combination of sources. Because data points aren't independent and identically distributed, subtle statistical issues arise. We're designing a rich language for modeling these and other aspects of social networks and developing learning and inference algorithms for it. This language, called Markov logic networks, combines the probabilistic modeling of Markov random fields with first-order logic's expressiveness.[7] In preliminary experiments, it speeded development of a complex social-network model and yielded more accurate predictions than standard methods.

We're all familiar with the notion that a butterfly flapping its wings in Beijing can cause a storm in New York. At the same time, the chances that a given butterfly flapping its wings will indeed cause a storm in New York are very small. Our approach, in a nutshell, is to ask, "If we wanted to cause a storm in New York, and could make a few butterflies flap their wings, which ones would we choose?" Our experiments so far show that, at least in the marketing world, this is an effective way to unleash storms on demand.

References

1. P. Domingos and M. Richardson, "Mining the Network Value of Customers," *Proc. 7th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining*, ACM Press, 2001, pp. 57–66.
2. M. Richardson and P. Domingos, "Mining Knowledge-Sharing Sites for Viral Marketing," *Proc. 8th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining*, ACM Press, 2002, pp. 61–70.
3. R. Dye, "The Buzz on Buzz," *Harvard Business Rev.*, vol. 78, no. 6, 2000, pp. 139–146.
4. D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the Spread of Influence through a Social Network," *Proc. 9th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining*, ACM Press, 2003, pp. 137–146.
5. F. Reichheld, "The One Number You Need to Grow," *Harvard Business Rev.*, vol. 81, no. 12, 2003, pp. 47–54.
6. S. Jurvetson, "What Exactly Is Viral Marketing?" *Red Herring*, no. 78, pp. 110–112; www.redherring.com/Article.aspx?a=5485&hed=From+the+ground+floor%0a.
7. M. Richardson and P. Domingos, *Markov Logic Networks*, tech. report, Dept. Computer Science and Eng., Univ. of Washington, 2004; www.cs.washington.edu/homes/pedrod/mln.pdf.

Social Networks and the Semantic Web: The Next Challenge

Peter Mika, *Free University Amsterdam*

The 2004 International World Wide Web Conference heralded the completion of the first phase of the World Wide Web Consortium's Semantic Web Activity (www.w3.org/2001/sw/Activity): laying the Semantic Web architecture's foundations. Publications and demonstrations at past WWW events and the International Semantic Web Conference series have presented an impressive picture of standardized representation languages, from RDF Schema to variants of OWL; tools for creating, storing, querying, and reasoning with ontologies; and browsing, search, and knowledge-sharing applications, driven by the ontologies under the hood. With Semantic Web Activity Phase 2 beginning, the message is that by and large, the Semantic Web is ready for deployment.

Looking more closely, some issues seem to be left for future research—those hot potatoes Semantic Web researchers have passed around from the beginning. Two in particular stick out from the thick proceedings volumes: ontology learning and ontology mapping. Ontology learning or extraction is the attempt to recreate a conceptual model from existing knowledge sources, in particular natural text. Ontology mapping (also known as merging, alignment, and so on) refers to finding and reconciling the relations between two or more conceptual models and creating a single model that captures their intentions and the relationships between them.

The underlying reason that automating these tasks is difficult is our machines' lack of understanding when it comes to ontologies. Interestingly, this is the problem the Semantic Web set out to solve—the idea being that we would attach machine-processable descriptions to content that's otherwise unintelligible to computers. As the well-known slide about what it's like to be a computer illustrates, although the machine doesn't understand strings of human symbols, it has no problem identifying the parts in angled brackets. Many have noted that the strings in the angled brackets are just symbols as well, but at least their limits are clear. Our machines have no problems applying all kinds of rules to them, generating further symbols along the way (or pointing out to us that the way we used them is inconsistent with the rules we set out for them). If the conclusions don't seem to be intelligent or are surprising to us, we add more rules to fix it.

What the machine can't do is access what we think those symbols' interpretations are, and therein lies the problem. Unfortunately, this is the crux of ontology learning and mapping.

Semantics are us

Creating and reconciling interpretations is a human-complex problem, as Figure 1 illustrates. The picture is a puzzle, although an easier one than the computer must face, because we uncovered two of the network's three terms. The puzzle imitates the basic step of creating or recreating interpretations, namely placing or retrieving a new concept on the basis of the context. The question to the reader is, "What concept is hidden under the code?"

Unfortunately for the computer, this challenge's answer is a matter of association. What comes to your mind when you look at the combination of these terms? The answer necessarily depends on your mental schemata; you'll accept an answer as plausible based on your own response. Interestingly, we happen to know the answer that came from a set of Edinburgh college students in 1973.

Knowledge engineers collected that year's Edinburgh Associative Thesaurus by handing a list of words to students and instructing them to write as quickly as possible next to each stimulus word the first word it made them think of. The experiment's next round used those words. The engineers repeated the cycle three times; by then, the number of responses was so large that they couldn't all be reused as stimuli.

According to the EAT, the shortest path from love to money in the students' minds runs through the notions of girl and security (see Figure 2). If we consider only lines with weights less than five, the path runs as follows: love, sex, yes, please, thanks, a lot, money. So much for Edinburgh college students' love lives in the '70s.

Deeper network analysis confirms that money was central to these students' thought processes (it had the largest inward degree and betweenness centrality), even more so than sex.[1] Moreover, it appears that the students also had water and food on their mind, when not preoccupied with the ideas mentioned earlier.

The EAT is actually an ontology. The ontology-engineering method is particularly revealing, because once the initial set of words is selected, the only parameter to the process is the population chosen. In particular, the knowledge engineer has no other role than handing out questionnaires and collecting responses. Some results are likely to hold for other communities (such as the overwhelming tendency for Christians to say "Noah" when they hear "ark"), but the experiment's subjects' collective mindset drives many of the aggregated associations. The well-known dynamics of social networks create this collective mindset: interaction creates similarity and vice versa. (Caveat to the second part: in practice, physical constraints such as geographic distance often limit interaction.)

Communities and the Semantic Web

Even if we should try to avoid entering old debates about the nature of knowledge and learning, we should bow to the social sciences at this point and consider knowledge's embeddedness in the social context. (For a summary of the different views on knowledge and learning, see Figure 3.) I believe we can learn something here for the Semantic Web. Communities aren't merely sets of users, but an integral and dynamic part of the architecture. Without intending or even realizing, we've elevated them to this place by integrating the notions of ontology and semantics into the first Web's technological framework.

It's easy to predict what's going to fail if we decide to ignore social-context elements. Communities will start creating their own ontologies that reflect their identities, languages, and collective intelligence, developing them through interaction surrounding a practice or interest. With these ontologies, the communities will annotate their online or offline content. Communities, ontologies, and content make up the three layers of the Semantic Web (see Figure 4). Islands of semantics will arise unless we find a way to map these ontologies. Mappings, however, just reflect the similarities in the conceptualizations the separate communities have developed, possibly through interaction between community members.

Further, these conceptualizations change

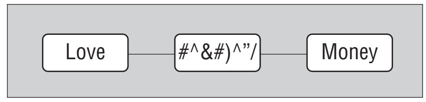


Figure 1. The ontology puzzle.

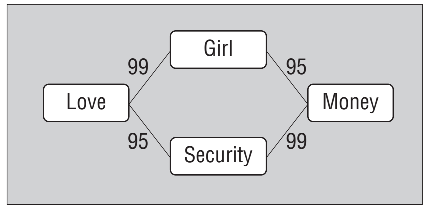


Figure 2. The shortest path between love and money according to the EAT.

as communities evolve and learn. Unless we make communities part of the system, the system will have significant difficulties catching up with changing semantics. (Ontology versioning is really an ontology mapping problem between a model's old and new versions.)[2] The more unstable knowledge is, the more difficulty we can expect in formalizing and sharing it on a large scale.[3]

If all this sounds a bit gloomy, I should quickly add that we might have already come across a way toward the solution by pure chance. Created and spread for the sole purpose of having fun (another first for the Semantic Web), the Friend of a Friend project's simple ontology lets us identify, describe, and relate users using URIs (uniform resource identifiers), RDF, and a few extra terms, just as we would with conventional, Web-accessible resources such as HTML pages. On the representation side, it's a small step from here to associating FOAF's users and groups to the Semantic Web's ontologies and metadata.

It will take time and a new, interdisciplinary mindset to find out how we could characterize a community's relationship to an ontology and its content and what this

AI and cognitive sciences		Social sciences	
Learning is ...	Knowledge is ...	Learning is ...	Knowledge is ...
Abstract	Context free	Situated	Contextual
Passive	Mostly explicit	Active	Mostly tacit
Cerebral	Model	Embodied	Practice
Individual	Resource	Social	Process
Transmission	Transfer	Reception	Experience

Figure 3. A comparison of the AI or cognitive science and social science views on knowledge and learning. Based on a presentation by Paul Duguid (Int'l PhD Seminar on Organizational Learning, Networks, and Communities, Amsterdam, 2004).

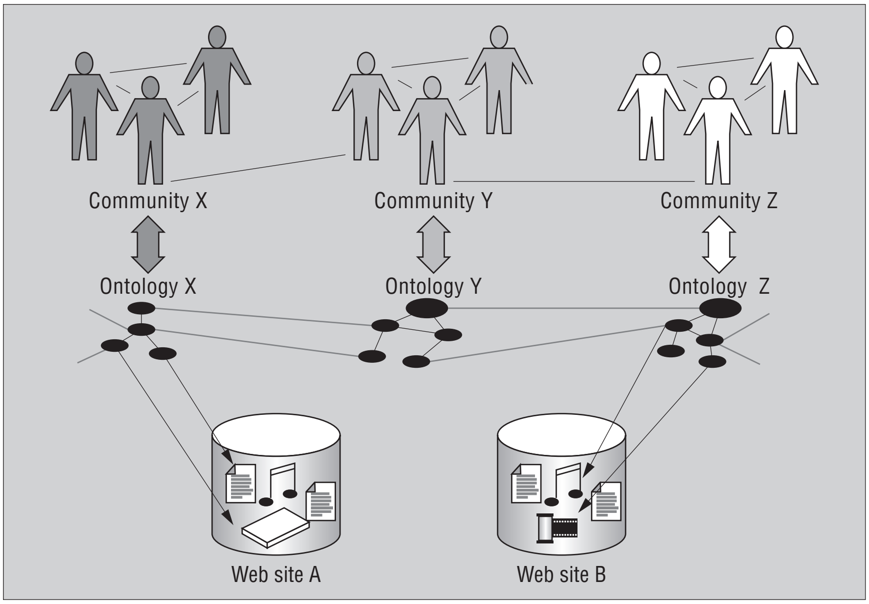


Figure 4. The Semantic Web's three layers.

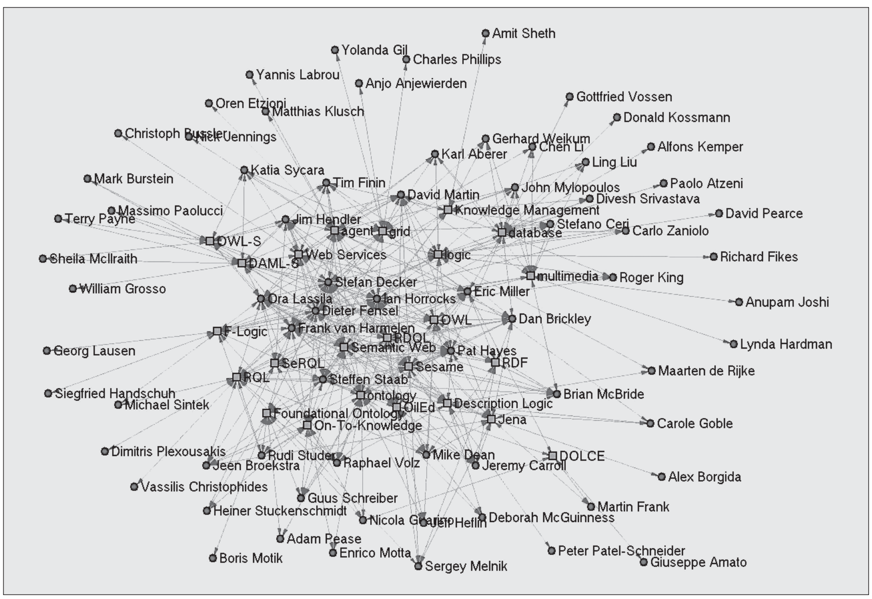


Figure 5. Members of the Semantic Web community and their association to research interests.

relationship means for an ontology's use and the metadata created with it. How could we move part of the social process surrounding ontologies within the system's boundaries so that we can give our machines a chance to get a grip on the messy, confused world of human knowledge?

Proof of concept

Many researchers have used a Web keyword index such as Google for ontology-learning purposes, most recently in Philipp Cimiano and his colleagues' excellent work.[4] In my recent effort, I've extended this technique with the EAT idea, building a conceptual model that reflects a community's rather than the whole Web's wisdom. Besides mining social networks from the Web, as in the pioneering ReferralWeb project,[5] I queried Google with community members' names (in my case, Semantic Web researchers), along with terms from their domains, such as research topics, tool names, and so on. I've measured the associations among people and concepts by the number of pages returned—that is, the number of pages where the name and the concept co-occur—after a normalization step (see also the online demonstration at http://flink.semanticweb.org).[6,7]

By taking the most highly associated individuals as representatives of a concept or of community members, I measured the associations between concepts by the number of people shared between the two communities, much like the EAT (see Figure 5). Although the method is under evaluation, the first results suggest that an ontology obtained this way represents the community better than one obtained by mining associations on the basis of all of Google's index pages, especially when it comes to terms that have both a generic and a community-specific interpretation (see Figure 6). For me, this simple experiment is the proof of concept: elusive, fuzzy communities are a real independent variable in learning an ontology. If I'd run the same experiment with the Edinburgh master's students' names, I'm certain I would have gotten a quite different picture.

As for the Semantic Web Activity's Phase 2, the message is clear. It's time to think about that final layer of the Semantic Web architecture, the social structures and processes that one day will lead us to a sustainable worldwide ecosystem of people and semantics.

Acknowledgments

The Vrije Universiteit Research School for Business Information Sciences (VUBIS) provided funding for this work.

References

1. A. Ziberna, "Network Analysis of a Word Association Thesaurus," *Proc. XXIV Int'l Sunbelt Social Network Conf.*, Univ. of Ljubljana, 2004, pp. 131–132.

2. M. Klein et al., "Ontology Versioning and Change Detection on the Web," *Proc. 13th Int'l Conf. Knowledge Eng. and Knowledge Management* (EKAW02), LNCS 2473, A. Gómez-Pérez and V.R. Benjamins, eds., Springer-Verlag, 2002, pp. 197–212.

3. L. van Elst and A. Abecker, "Ontologies for Information Management: Balancing Formality, Stability, and Sharing Scope," *Expert Systems with Applications*, vol. 23, no. 4, Nov. 2002, pp. 357–366.

4. P. Cimiano, S. Handschuh, and S. Staab, "Towards the Self-Annotating Web," *Proc. 13th Int'l Conf. World Wide Web*, ACM Press, 2004, pp. 462–471.

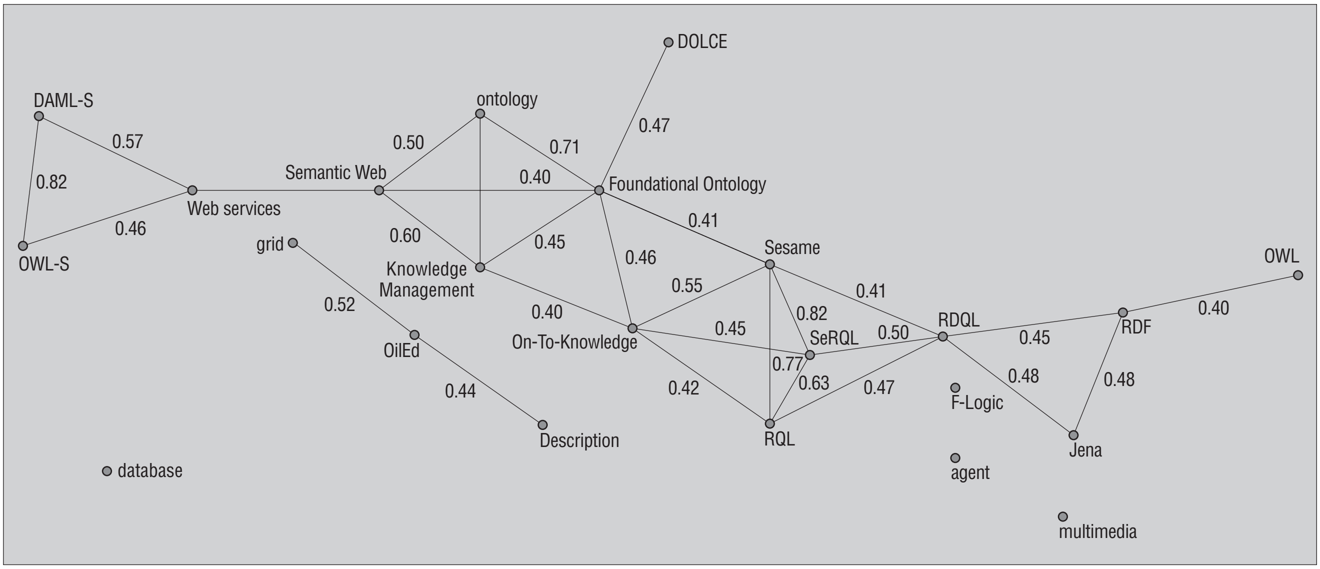


Figure 6. The resulting community ontology is the result of folding Figure 5's bipartite graph and normalizing the weights using geometric normalization.

5. H. Kautz, B. Selman, and M. Shah, "The Hidden Web," *AI Magazine*, vol. 18, no. 2, 1997, pp. 27–36.
6. P. Mika, "Bootstrapping the FOAF-Web: An Experiment in Social Network Mining,"*Proc. 1st Workshop Friend of a Friend, Social Networking and the Semantic Web*, 2004; www.w3.org/2001/sw/Europe/events/foaf-galway/papers/fp/bootstrapping_the_foaf_web.
7. P. Mika, "Social Networks and the Semantic Web: An Experiment in Online Social Network Analysis," *Proc. IEEE/WIC/ACM Int'l Conf. Web Intelligence* (WI 2004), IEEE CS Press, 2004, pp. 285–291.

Sharing and Using Links in Social Networks

Jennifer Golbeck, *University of Maryland*

Social networking has grown dramatically more popular since the release of the film *Six Degrees of Separation* and the now-infamous Kevin Bacon Game's popularization (see the Related Links sidebar). In the past few years, several Web sites have emerged to support the public interest in social networking. Some, such as LinkedIn, focus on building business relationships, while others, such as Tribe, Orkut, and Friendster, have social and entertainment motivations. Recently, a Web site even emerged to extend social networking to our dogs, though the people at Dogster prefer that you call it social petworking.

Data about millions of people and their connections is publicly available on the Web. Users spend a lot of time maintaining their data, building up networks, and browsing links. If it's possible to integrate this distributed data into centralized models, we should be able to create socially aware, intelligent applications that will let users benefit from their participation in online social networks. To achieve this goal, we must answer two questions. First, how can we merge data from separate databases into a large social network model that applications can access? Second, how can social networks benefit users?

The Friend of a Friend project

The question of making data accessible and understandable by applications has a strong foundation of work behind it. The Friend of a Friend project is one of the Semantic Web's largest and most popular. It's essentially a vocabulary for describing people and whom they know. Literally millions of FOAF files exist on the Semantic Web, some from users who've authored their own data and others from Web sites that publish data from their databases using the FOAF ontology. Because users are beginning to accept FOAF as something of a standardized ontology for representing social networks on the Semantic Web, it's a good option for Web sites that want to start sharing some of their data.

However, FOAF isn't a complete data-sharing solution. The vocabulary has a limited set of properties, and only one type of relationship exists between people: the "knows" relationship. In reality, the scale of knowing someone varies from lifelong best friend to Internet acquaintance. A specific application might want properties that aren't part of FOAF, but on the Semantic Web, that's not much of a problem. By nature, the Semantic Web lets individual projects make

Related Links

The Oracle of Bacon at Virginia: www.cs.virginia.edu/oracle
LinkedIn: www.linkedin.com
Tribe: www.tribe.net
Orkut: www.orkut.com
Friendster: www.friendster.com
Dogster: www.dogster.com
Friend of a Friend project: www.foaf-project.org
Trust and Reputation project: http://trust.mindswap.org
Friend of a Friend Relationship module: www.perceive.net/schemas/20021119/relationship

their own extensions to the vocabulary while preserving FOAF as a core. My own project, for example, studies trust in social networks. FOAF doesn't define trust, but my simple ontology extends FOAF with trust properties. Another project—the FOAF Relationship module—defines a long list of relationship types.

The benefits of using FOAF as a core are twofold. A Web site that understands FOAF can augment its own databases with information gathered from other FOAF files on the Web. Also, by publishing a set of data with a FOAF core on the Semantic Web, Web sites provide a service to users because other FOAF-based services can now use that data. (Paul Mutton describes one technique for spidering FOAF data and creating applications with it.[1])

Applications

Access to this large data model holds great promise for improving and developing intelligent systems. Based on the time they spend developing networks online, users clearly feel that there's a benefit to building social networks. The question to us as scientists is, "What can we do with the connections in these networks?" Some of the greatest potential for creating socially intelligent systems lies in being able to compose relationships within the network and integrate that information into systems.

What can we say about two people connected by intermediate friends? With the standard approach of social networking, which has a single relationship indicating some sort of connection, we have too little information to say much. By adding dimension to relationships, more possibilities open up. Once we start refining relationships, we can ask a series of questions that will affect the analysis and potential for composition. If Alice and Chris aren't directly connected in a social network, and we want to calculate a recommendation about the relationship between them, we need to know about the relationships and intermediate people connecting them.

Using trust as an example, say Alice highly trusts Bob, and Bob highly trusts Chris. Can we recommend that Alice should have some level of trust for Chris? It's debatable, but in practice, we use this sort of social logic every day. Asking for a recommendation about a mechanic or restaurant employs exactly this type of computation, taking into account the trust we have for the person we're asking and their trust of the mechanic or restaurant. Blind dates are another (albeit often unsuccessful) application of the same logic.

With algorithms that can accurately infer relationships between people—be they trust relationships or other types—using those inferences in applications is a natural next step. Continuing with the trust example, we can integrate recommendations we make about trust using a social network in many contexts: using ratings to allow access to personal information on the Web, using them as a filter for Web-based information, or, as in one of my projects, integrating ratings into an email client. In email, we could use ratings as scores for email messages, indicating how much the user should trust the sender—even if they've never met.[2] That essentially creates a social email filter, benefiting from algorithms applied to large social networks.

Certainly, we can use values in other ways and compose other types of relationships. The core point is that we must be able to bring something useful to systems from social networks. Telling users the number of relationships and number of friends-of-friends they have might be entertaining, but falls short of offering any real assistance. The Web-based infrastructure exists that will let people and Web sites share social network data in a distributed way, and that will let services and applications aggregate that large set of data into queriable models. Our work must focus on doing something with the social network, like composing relationships. In that type of analysis lies the real ability for creating intelligent, socially aware systems that integrate the social and computational.

References

1. P. Mutton, *IRC Hacks*, O'Reilly, 2004.
2. J. Golbeck and J. Hendler, "Reputation Network Analysis for Email Filtering," *Proc. 1st Conf. Email and Anti-Spam*, CEAS, 2004; www.ceas.cc/papers-2004/177.pdf.

Analyzing Social Networks on the Semantic Web

Li Ding, Tim Finin, and Anupam Joshi, *University of Maryland, Baltimore County*

The past year has seen a dramatic increase in the amount of social information published in RDF documents. Our investigations show that the Friend of a Friend ontology (http://xmlns.com/foaf/0.1) is among the most-used Semantic Web ontologies.[1,2] This is true if we measure the number of Semantic Web documents (SWDs) that use the FOAF namespace, as Table 1 shows, or the number of triples using FOAF terms. The Swoogle Ontology Dictionary (http://swoogle.umbc.edu) shows that the class foaf:Person (the qualified name, or QName, of http://xmlns.com/foaf/0.1/Person) has nearly one million instances spread over about 45,000 Web documents. The FOAF ontology isn't the only one people use to publish social information on the Web. For example, Swoogle identifies more than 360 RDF Schema or OWL classes defined with the local name "person."

The Semantic Web and social-network models support one another. On one hand, the Semantic Web enables online and explicitly represented social information; on the other hand, social networks, especially trust networks,[3] provide a new paradigm for knowledge management in which users "outsource" knowledge and beliefs via their social networks.[4] To turn these objectives into reality, we need to address many challenging issues such as

- *Knowledge representation*. Although various ontologies capture rich social concepts, we don't need hundreds of dialectic ontologies defining the same concept. How can we coalesce around a small number of common, comprehensive ontologies?
- *Knowledge management*. Compared to the entire Web, the Semantic Web is fairly well connected at the RDF-graph level but poorly connected at the RDF-document level. The Semantic Web's open, distributed nature introduces other issues. How do we provide efficient, effective mechanisms for accessing knowledge, especially social networks, on the Semantic Web?
- *Social network extraction, integration, and analysis*. Even with well-defined ontologies for social concepts, extracting social networks correctly from the noisy, incomplete knowledge on the Semantic Web is difficult. What are good heuristics for integrating and fusing social information, and what metrics are useful for the results' credibility and utility?
- *Provenance- and trust-aware distributed inference*. Provenance associates facts with social entities that are interconnected

in social networks. We can derive trust among social entities from social networks. How do we manage and reduce the complexity of distributed inferences by using provenance of knowledge in the context of a given trust model?

Datasets

To understand how social networks on the Semantic Web are being modeled, we collected two datasets: DS-SWOOGLE and DS-FOAF. (We noticed that these datasets are the largest among related works.[1,5,6]) We used Swoogle, a crawler-based indexing and retrieval system for Semantic Web documents, to collect the first dataset,[2] which provides a baseline model of the ontologies and information encoded in RDF on the Web. The dataset shows that the terms in the FOAF ontology, especially foaf:Person, are among the most used and populated. (We say that a class or property is populated when it has direct instances.) We assume that it's reasonable to use the foaf:knows property to connect people forming social networks. Therefore, we collected the second dataset for the SemDis project[1] to focus on available FOAF documents containing instances of foaf:Person. We collected both datasets from conventional Web search engines, user-supplied URLs, and our Semantic Web crawlers.

DS-SWOOGLE

At the time of this writing, DS-SWOOGLE represented more than 335,000 valid SWDs (that is, online RDF documents in formats such as RDF/XML and N3). The SWDs contain about 47,000,000 RDF triples and are hosted by about 125,000 Web sites. (Swoogle is running continuously, and its database grows as new SWDs are added to the Web.) Swoogle samples at most 10,000 documents from each Web site to avoid being overwhelmed by Web sites with millions of RDF documents. Swoogle Ontology Dictionary and Swoogle Statistics are based on this dataset.

DS-FOAF and DS-FOAF-VAR

The DS-FOAF dataset has over one million valid online FOAF document URLs from over 1,800 sites. We consider a FOAF document to be any RDF document that has at least one instance of the foaf:Person class. We count Web sites by DNS (Domain Name System) name in DS-SWOOGLE and by IP (Internet Protocol) address in DS-FOAF.

Table 1. The seven namespaces most frequently used in RDF documents (from Swoogle).

	Namespace URI	Number of documents
1	www.w3.org/1999/02/22-rdf-syntax-ns#	200,097 (96.9%)
2	http://purl.org/dc/elements/1.1/	146,923 (71.2%)
3	http://purl.org/rss/1.0/	111,595 (54.0%)
4	http://webns.net/mvcb/	68,330 (33.1%)
5	http://xmlns.com/foaf/0.1/	49,504 (24.0%)
6	www.w3.org/2000/01/rdf-schema#	44,656 (21.6%)
7	http://purl.org/rss/1.0/modules/content/	28,607 (13.9%)

Table 2. Statistics of DS-FOAF and DS-FOAF-VAR.

	DS-FOAF			DS-FOAF-VAR		
	max	avg	std	max	avg	std
Persons/doc	2,216	30.5	52.3	2,196	5.1	49.4
SeeAlso/doc	2,238	29.3	51.8	2,066	1.9	36.7
Triples/person	—	—	—	3,192	5.5	36.1

Five major blog sites, which use limited vocabularies and fixed structures in describing personal profiles, host more than 95 percent of the URLs. To reduce these sites' impact, we studied a smaller dataset, DS-FOAF-VAR, which considers Web sites that host at most 1,000 FOAF documents. This dataset has over 7,000 FOAF documents drawn from 1,065 Web sites that define nearly 37,000 instances of foaf:Person. These include 4,158 strict FOAF documents, intended to describe one person and his or her acquaintances. Table 2 shows the two datasets' detailed statistics.

Building a common social ontology

The Semantic Web provides a powerful distributed mechanism to represent and publish social network information. Although FOAF terms are widely used to encode social relations, other ontologies show up as well. We expect these to coalesce and merge as they evolve. In light of the statistical approach to finding common terms,[1,5] we studied a particular class: foaf:Person, the most frequently used class for describing personal profiles, according to our datasets. The definition of foaf:Person comes from three sources: its ontology definition, which relates it to other classes; the ontological properties that relate to it via rdfs:domain relation; and empirical properties, which correlate with it by modifying its instances (see Figure 1).

The DS-SWOOGLE dataset includes 17

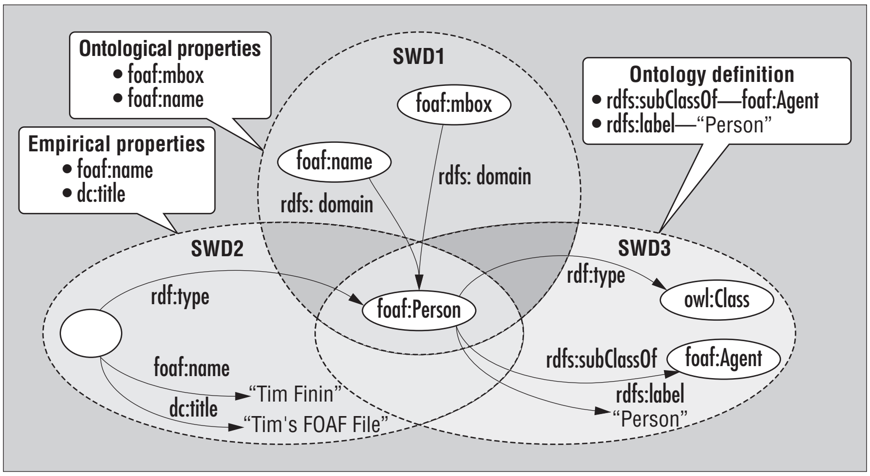


Figure 1. Three sources of term definition.

Table 3. Top 10 empirical properties of foaf:Person in DS-Swoogle.

	Property	Max. cardinality	Min. cardinality	Documents Number	Documents Percent
1	foaf:mbox_sha1sum	12	1	41,403	95%
2	foaf:nick	7	1	36,095	83%
3	foaf:weblog	5	1	35,303	81%
4	rdfs:seeAlso	329	1	27,838	64%
5	foaf:name	4	1	26,749	62%
6	foaf:knows	3,187	1	25,736	59%
7	foaf:homepage	3	1	17,616	41%
8	foaf:dateOfBirth	1	1	12,783	29%
9	foaf:page	3	1	11,255	26%
10	foaf:interest	300	1	10,314	24%

Table 4. Top 10 empirical properties of foaf:Person in DS-FOAF-VAR.

	Property usage per document		Property usage per instance	
1	foaf:name	80%	foaf:name	65%
2	foaf:mbox_sha1sum	70%	foaf:mbox_sha1sum	60%
3	foaf:nick	51%	rdfs:seeAlso	37%
4	foaf:homepage	40%	foaf:nick	24%
5	foaf:depiction	35%	foaf:homepage	16%
6	foaf:weblog	30%	foaf:mbox	14%
7	foaf:knows	28%	foaf:weblog	14%
8	foaf:surname	27%	foaf:firstName	12%
9	foaf:firstName	27%	foaf:surname	12%
10	rdfs:seeAlso	26%	foaf:depiction	9%

ontologies that add to the foaf:Person definition. For example, it is defined both as an owl:Class and as an rdfs:Class and has the named super-classes foaf:Agent, wordnet:Person, geo:SpatialThing and con:Person (the geo prefix refers to www.w3.org/2003/01/geo/wgs84_pos# and the con prefix refers to www.w3.org/2000/10/swap/pim/contact#). DS-Swoogle reveals 162 ontological properties of foaf:Person, most representing social relations. Seventy-four properties exist whose rdfs:domain and rdfs:range are both foaf:Person. DS-Swoogle also finds 558 empirical properties of foaf:Person that are populated with instance data. Tables 3 and 4 list the 10 most frequently used empirical properties, which suggest that people publishing personal information are concerned about privacy. They use the property foaf:mobx_sha1sum (which hides the true email address) much more frequently than foaf:mbox.

The empirical cardinality also shows how users organize their profiles. The high value for maximum cardinality results from an unusual usage of FOAF vocabulary to build a collection of FOAF documents. In Table 4, the properties that documents (but not instances) use frequently tend to describe the strict FOAF documents' owner.

Extracting social networks

Extracting social networks from noisy, real-world data is challenging, even if the information is already encoded in RDF using well-defined ontologies. The process has three steps: discovering instances of foaf:Person, deciding which instances denote the same person and merging their information, and linking people through various social-relation properties, such as foaf:knows. Determining whether two foaf:Person in-stances denote the same person is the critical, but difficult, step. The semantics of FOAF vocabulary suggest several heuristics to answer this question, such as

- *Named URI*. Non-anonymous individuals using the same URI denote the same person.
- *Inverse-functional properties*. Inverse-functional properties such as foaf:mbox and foaf:homepage identify unique individuals. In practice, we can use one or more properties that aren't strictly inverse functional, such as foaf:name and foaf:nick, in conjunction with other properties such as foaf:phone to identify individuals with high probability.
- *Semantic equality*. When two or more values of an inverse-functional property coexist in an individual's description, they're semantically equivalent to identifying the same individual.
- rdfs:seeAlso. This property almost always links to a strict FOAF document where the root and referrer persons are the same.

In our preliminary study of DS-FOAF-VAR, we applied the first three heuristics and only consider foaf:mbox_sha1sum and foaf:mbox as inverse-functional properties. We found 18,603 merged people, but only 10,247 have unique identifiers. Figure 2 shows that the cumulative distribution of group size follows a Zipf distribution. Here, group refers to the collection of individuals being merged as a person.

These heuristics for merging individuals can fail in two ways: inconsistency and separation. OWL gives one inconsistency criterion, where cardinality constraints limit a property's semantically distinct values. For example, when property P has an owl:cardinality of one when modifying class C, all P's values for an individual of type C should be semantically equivalent. In practice, because most people only have one name, we derive a cardinality constraint over foaf:Person. We can validate a person's semantic consistency by checking whether it has two different names. Separation occurs when a person's information remains in two disjoint groups after merging. This gives rise to a dilemma—applying more merge heuristics might reduce separation but increase inconsistency.

Social network analysis

Social network analysis (SNA) is a large, active research area, so we've limited our work to studying some of the extracted social network's basic graph features. Peter Mika shows more applications of basic SNA measures on a smaller social network ($n = 167$) extracted from FOAF and other Web sources.[7]

Degree analysis

Degree analysis is an important measure in analyzing social networks. Our analysis of 14,164 distinct "knows" relations in DS-FOAF-VAR shows that both in-degree and out-degree follow a Zipf distribution (see Figure 3). We further put person into four categories: in only (51.8 percent), out only (5.8 percent), in and out (5.4 percent), and isolated (37.1 percent), according to their in-degree and out-degree. Such a social network isn't well connected because only a few individuals (in and out) lie between others. Ninety-four percent of "in only" persons are known by only one person.

Patterns of connected components

We've discovered 834 connected components and 6,904 isolated persons. The connected components exhibit interesting graphical patterns: six singletons that link to themselves, a giant component that has 6,053 connected persons, and several stars with many out-links (the average out-degree for such nodes is 6.8). Figure 4 shows a selection of connected components. We hypothesize that the FOAF network topology evolves over time; a FOAF network starts from some disjointed star-like connected components, which link together to form trees, forests, and eventually a scale-free network.

Our research uses real-world data in an open, distributed context to provide a data-digest service for efficient data access on the Semantic Web and reasoning over the knowledge encoded in Semantic Web languages. We are also working on modeling trust across multiple social networks and building a general architecture for provenance- and trust-aware inference in open, distributed, and heterogeneous environments, such as the Web and multiagent systems.

Figure 5 illustrates our ongoing work on reasoning trust across multiple social networks and reputation systems. To improve coverage and connectivity, we integrate social networks and reputation systems by mapping people to better derive and propagate trust relations through social relations. As Figure 5 shows, the gap between the two people *P. Kolari* and *A. Sheth* is connected by mapping the person *T. Finin* between two social networks. The reputation systems might offer default trust to social entities.

Our ongoing work focuses on improving the efficiency and effectiveness of data-

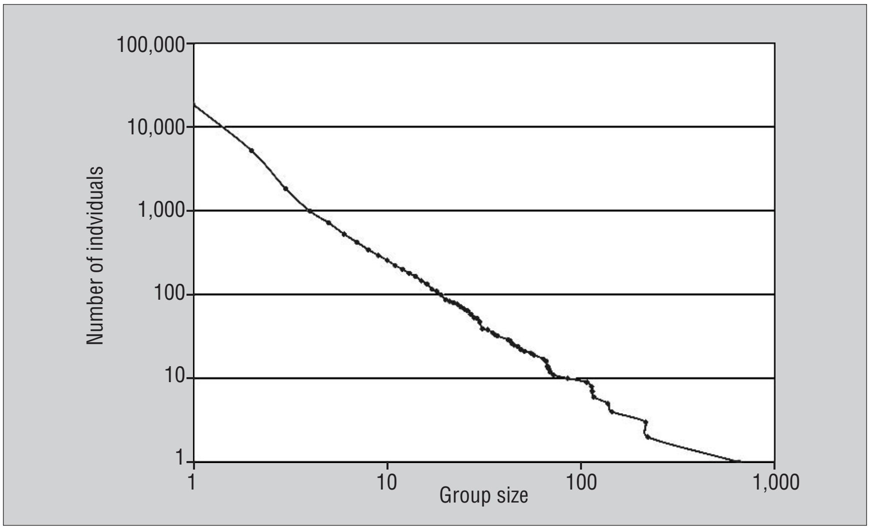


Figure 2. Cumulative distribution of group size.

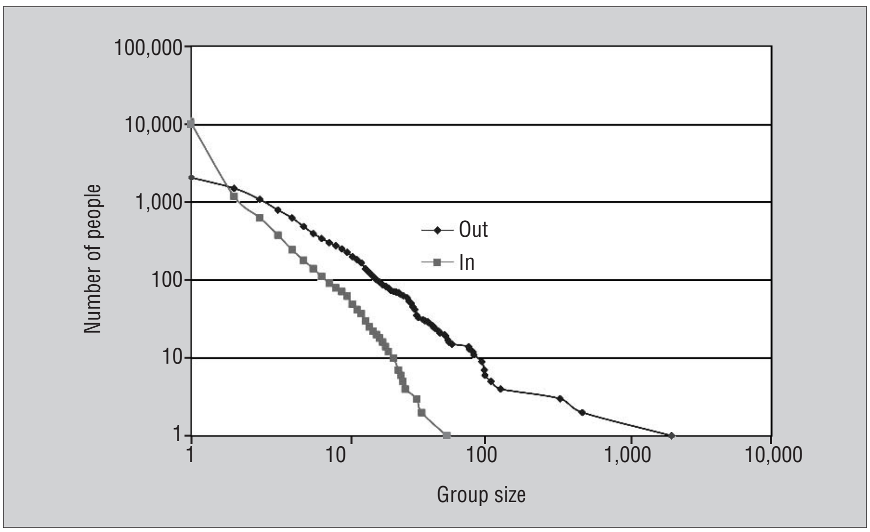


Figure 3. Cumulative distribution of in-degree and out-degree.

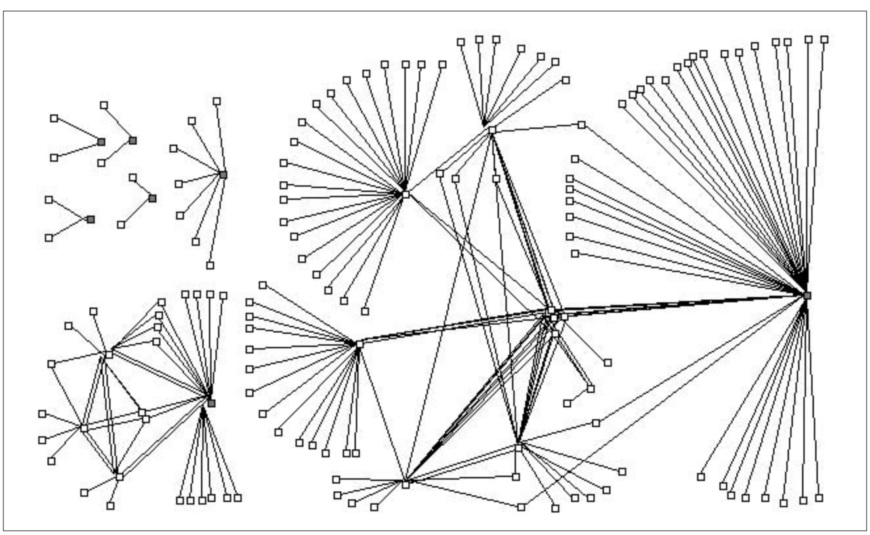


Figure 4. Some connected components in FOAF network.

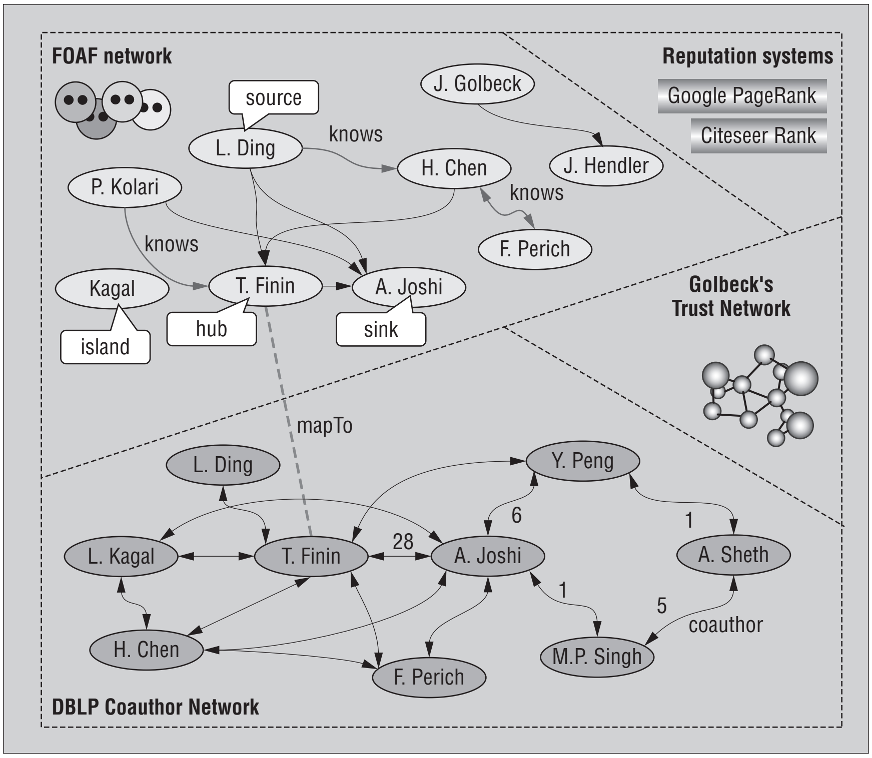


Figure 5. Deriving and propagating trust from multiple sources.

digest services, social-network extraction and integration, and modeling provenance and trust for distributed inference services.

Acknowledgments

DARPA contract F30602-00-0591(DAML) and NSF awards ITR-IIS-0326460(SPIRE) and ITR-IIS-0325464(SEMDIS) provided partial research support.

References

1. L. Ding et al., "How the Semantic Web Is Being Used: An Analysis of FOAF," *Proc. 38th Ann. Hawaii Int'l Conf. System Sciences* (HICSS 05), CD-ROM, IEEE CS Press, 2005.
2. L. Ding et al., "Swoogle: A Search and Metadata Engine for the Semantic Web," *Proc. 13th ACM Conf. Information and Knowledge Management* (CIKM 04), ACM Press, 2004, pp. 652–659.
3. J. Golbeck, B. Parsia, and J. Hendler, "Trust Networks on the Semantic Web," *Proc.7th Int'l Workshop Cooperative Intelligent Agents* (CIA 2003), LNCS 2782, Matthias Klusch et al., eds., Springer-Verlag, 2003, pp. 238–249.
4. L. Ding, L. Zhou, and T. Finin, "Trust Based Knowledge Outsourcing for Semantic Web Agents," *Proc. IEEE/WIC Int'l Conf. Web Intelligence* (WI 03), IEEE CS Press, 2003, pp. 379–387.
5. J.C. Paolillo and E. Wright, "The Challenges of FOAF Characterization," *Proc. 1st Workshop Friend of a Friend, Social Networking and the Semantic Web*, 2004, www.w3.org/2001/sw/Europe/events/foaf-galway/papers/fp/challenges_of_foaf_characterization.
6. G.A. Grimnes, P. Edwards, and A. Preece, "Learning Meta-descriptions of the FOAF Network," *Proc. 3rd Int'l Semantic Web Conf.* (ISWC 2004), LNCS 3298, S.A. McIlwraith, D. Plexousakis, and F. van Harmelen, eds., Springer-Verlag, 2004, pp. 152–165.
7. P. Mika, "Bootstrapping the FOAF-Web: An Experiment in Social Network Mining," *Proc. 1st Workshop Friend of a Friend, Social Networking and the Semantic Web*, 2004, www.w3.org/2001/sw/Europe/events/foaf-galway/papers/fp/bootstrapping_the_foaf_web.

Information and Influence in the Construction of Shared Reality

Andrzej Nowak, *Warsaw University*
Robin R. Vallacher, *Florida Atlantic University*

Information isn't hard to come by in today's world. Evaluating that information, however, is a daunting task for individuals, groups, and institutions. How do you determine which information is important? How do you resolve conflict among different pieces of information? How should you assemble information into a meaningful higher-order structure? These are difficult enough issues for the individual; their difficulty multiplies when a group must coordinate individuals' decisions and act on the basis of socially validated information.

Research in social psychology suggests that individuals interact, in large part, to construct a shared reality that consists not only of shared information, but also of agreed-upon opinions. In this process, they don't simply transmit information. More importantly, they influence one another to arrive at a common interpretation of information. Most research on social networks is concerned with information transmission.[1–3] We aim to supplement the social-network perspective by incorporating mechanisms that govern social influence.

Dynamics of social influence

Models of social influence focus on the effects that people's words, actions, or presence have on other people's thoughts, feelings, attitudes, and actions. Numerous experiments have shown that three critical factors determine social influence's impact: the number of sources exerting the influence, the sources' immediacy to the targets, and the sources' strength. We can describe these variables' joint effect in terms of a multiplicative function.[4] Whether the issue is conformity, stage fright, or interest in news events, influence grows approximately as a square root of the number of people involved, decreases with the square of the distance between the source and target, and is proportional to the sources' strength (for example, social status or credibility). In reality, of course, each individual influences and is influenced by other individuals, which provides the basis for dynamic social impact theory.[5] The formula below describes the mutual influence among individuals who differ in strength and who occupy different locations in social space, where I_i denotes total influence, s_j corresponds to each individual's strength, and d_{ij} corresponds to the distance between individuals i and j.

$$I_i = \left(\sum_{1}^{N} \left(\frac{s_j}{d_{ij}^{2}} \right)^2 \right)^{1/2}$$

The formula indicates that individuals

who are strongest and nearest to the recipient have the greatest impact. With respect to strength, research has established leaders' crucial importance for group-level phenomena. Some individuals are clearly more influential than others, either because of stable characteristics (such as credibility, expertise, or social status) or transient states (such as motivation to influence).

With respect to locality of influence, research has shown two effects. First, the probability of interaction (and hence of influence) decreases with the square of the distance between individuals. In both China and the US, for example, the probability that two individuals will discuss matters of mutual importance decreases as a square of the distance between their residences.[6] Second, information's impact decreases with the distance between source and recipient.[4] For example, information concerning a traffic accident is more likely to attract attention if it occurs in your town than if it occurs in a distant one.

Individual variations in the influence function's strength and locality are critical factors that dictate the dynamics of information in a social network.[7] A model showing how attitudes combine to form public opinion initially demonstrated these factors' importance.[5] The social group, modeled in a manner resembling cellular automata, consists of n individuals on a 2D grid. Each individual has an opinion (such as pro or con in the simplest case) and a value of persuasive strength. To simulate influence in social interactions, each individual assesses the degree of support for each position on the issue according to the formula. (In most simulations, we also assume a small noise, represented as a small random number that we add to a position's resultant influence.) If any opinion's resultant strength is greater than the individual's current opinion's strength, his or her opinion changes to match the prevailing opinion. The model performs this process for each person in the group until no further opinion changes occur.

A typical simulation starts with a clear majority and a clear minority (such as 60 and 40 percent), randomly distributed in the grid. After several rounds of discussion, the group reaches an equilibrium that typically entails a larger majority and a smaller minority (such as 90 and 10 percent). Opinions are no longer distributed randomly; rather, minority opinions survive in clusters of like-minded people, often formed around strong individuals.

Local effects

Local effects drive the model's global dynamics. In the initial random configuration, more majority than minority members surround the average group member. This results in more minority members converting to the majority opinion than vice versa. Research concerning attitude change in groups[8] and opinion swings on the eve of an election[9] has observed such polarization. The other global property—clustering—reflects the influence function's locality (that is, neighboring individuals have the strongest influence). In a clustered society, social interaction provides a biased estimate of your opinions' prevalence. Opinions in the global minority thus form a local majority. Many social phenomena—from social attitudes to farming techniques and clothing fashions—display pronounced clustering.

Computer social-change models based on the social-influence formula indicate that transitions occur as clusters of new that appear and grow in the sea of old.

Locality of the influence function

The influence function's locality also shapes social change's dynamics. Computer social-change models based on the social-influence formula indicate that transitions occur as clusters of new that appear and grow in the sea of old. From this perspective, interacting groups rather than isolated individuals are subject to change. Clustering of change is pervasive across a variety of social and economic phenomena.[10] Research has demonstrated the equivalence of simulation and empirical data with respect to the economic and political transformations in Poland following communism's collapse in the late 1980s.[11,12] In the social-interaction process, individuals define and shape their common social reality. Because this happens on a local level, it promotes different social realities that are separated in space. Social transitions occur as the new reality gains at the old reality's expense.

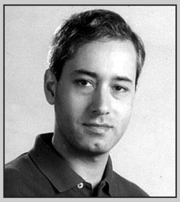
Social-space topology

The topology of the social space defining interaction patterns is critical for social-influence dynamics.[7,12] Social space's topology constrains interactions' spatial structure and thus dictates cluster formation's nature, the resultant clusters' shapes, and the probability of their survival or decay. Locality is essentially absent when the communication pattern resembles a random structure. Under these conditions, minority opinion rapidly decays because minority opinions can't form clusters.

You can interpret one-dimensional topology, in which people interact primarily with people on their left and right, as a village stretching along a road. Because of well-pronounced local interactions, this topology induces strong clustering but no polarization. In a hierarchical topology, where people are divided into groups, subgroups, and so forth, the distance between two individuals depends on the level at which they belong to a common unit. Two members of the same research unit in a university, for example, are closer than two members of the same department, who in turn are closer than two scholars from different departments. In hierarchical topology, borders of clusters will usually coincide with subgroups at some level. The opinion structure thus follows social-interaction patterns. There is pronounced polarization, the degree of which depends mainly on the groups' size on the lowest level, initial proportion of minority, and the distribution of individual differences in strength. Once formed, clusters in such a topology are quite stable. Clearly, we can envision more elaborate social-space geometries, each likely to be associated with specific dynamics.[13]

Implications for artificial intelligence

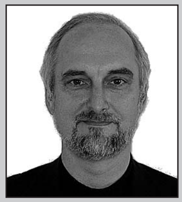
The process by which humans construct social reality might prove informative for designing rules for interaction among intelligent agents. The present model's primary implication is that information isn't merely acquired, but also evaluated and negotiated in a social context. An individual is likely to receive many pieces of mutually contradictory information relevant to almost any issue. Individuals resolve such conflicts not only by cognitive means, but also by social-influence mechanisms. Mimicking this


Pedro Domingos is an associate professor in the University of Washington's Department of Computer Science and Engineering. Contact him at pedrod@cs.washington.edu.


Peter Mika is a PhD student in the Free University Amsterdam's Department of Business Informatics and Department of Management and Organizations. Contact him at pmika@cs.vu.nl.

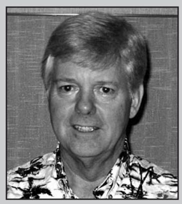

Jennifer Golbeck is a PhD candidate at the University of Maryland, College Park. Contact her at golbeck@cs.umd.edu.


Li Ding is a PhD student and a graduate research assistant at the University of Maryland, Baltimore County. Contact him at dingli1@csee.umbc.edu.


Tim Finin is a professor in the University of Maryland, Baltimore County's Department of Computer Science and Electrical Engineering. Contact him at finin@cs.umbc.edu.


Anupam Joshi is an associate professor of computer science and electrical engineering at the University of Maryland, Baltimore County. Contact him at joshi@umbc.edu.


Andrzej Nowak is a psychology professor at the University of Warsaw and a professor at the Warsaw School of Psychology and Florida Atlantic University. Contact him at nowak@fau.edu.


Robin R. Vallacher is a psychology professor at Florida Atlantic University and a research affiliate at Warsaw University's Center for Complex Systems. Contact him at vallacher@fau.edu.

process might let artificial agents deal with abundant information of varying quality and degree of conflict. This would require establishing mechanisms by which agents mutually negotiate different informational elements' importance and validity. Each agent evaluates information, and based on this evaluation, decides on its further transmission.

The magnitude of influence determines which information individuals will adopt and transmit. It's also conceivable that the sheer volume of information, weighted by its sources' strength, and the information's coherent versus contradictory nature might serve as criteria for evaluating information flow. Depending on the information's coherence, the agents might interact further to evaluate the information or reach a judgment on the basis of information that can provide a platform for action.

In this process, agents differ in their strength, which is equivalent to social status or credibility. To some degree, network designers might predefine strength. Designers specify some nodes, for example, to be more important due to their greater processing capacity (supercomputers versus desktop computers) or their superior access to information (such as more precise sensors). Strength can also change in the interaction process. Nodes broadcasting early information that turns out to be important and valid are likely to acquire higher status than nodes that are slow to broadcast or that broadcast unimportant or invalid information.

In larger networks, the abundance of information (including noise) might prove overpowering to agents. Moreover, information's importance and validity are likely to vary depending on its location in social space. Information might be important and valid in one location but irrelevant or false in another. Interactions' locality enables social means of validating information while allowing for diverse decisions in different locations. The possibility of local events in the network depends on the interaction space's topology. By choosing the appropriate topology of interaction patterns among artificial agents, you can choose the most appropriate dynamics to handle different tasks. To ensure that all agents preserve safety standards, for example, you might design a network with no topology, thereby making it impossible for a local group of agents to negotiate a lower safety standard. If an adaptation is crucial for a local changing environment, however, interaction based on proximity in that environment might be most beneficial. Such an interaction structure would not only let different subgroups of agents negotiate the most appropriate information for individual and group action, but would also let them try different adaptation strategies in different locations. The rest of the network's agents might then adopt the adaptation strategies that work out best.

Connections in social networks provide links for information transmission and social influence. For adaptations in the real world, both processes are vital. Different network features dictate the dynamics of information transmission and social influence. Social-network formalisms are flexible. In principle, we can accommodate the social-influence features we've described in social-network models,[14] thereby making them similar to attractor neural networks.[15]

Although such a description is well suited to capture social interactions' existing structure, it provides few clear cues about influence's dynamics in the network.

Networks of artificial agents, by mimicking social interaction, should be capable of dealing with real-world information. Social psychology might prove informative in designing interaction rules for such agents in the same way that cognitive psychology proved instructive for constructing individual agents.

References

1. A.L. Barabasi, *Linked: How Everything Is Connected to Everything Else and What It Means*, Blume Books, 2003.
2. S. Wasserman, K. Faust, and M.R. Granovetter, eds., *Social Network Analysis: Methods and Applications*, Cambridge Univ. Press, 1994.
3. D.J. Watts, *Six Degrees: The Science of a Connected Age*, W.W. Norton, 2003.
4. B. Latané, "The Psychology of Social Impact," *American Psychologist*, vol. 36, no. 4, 1981, pp. 343–356.
5. A. Nowak, J. Szamrej, and B. Latané, "From Private Attitude to Public Opinion: A Dynamic Theory of Social Impact," *Psychological Rev.*, vol. 97, no. 3, 1990, pp. 362–376.
6. B. Latané et al., "Distance Matters: Physical Space and Social Influence," *Personality and Social Psychology Bull.*, vol. 21, no. 8, 1995, pp. 795–805.
7. M. Lewenstein, A. Nowak, and B. Latané, "Statistical Mechanics of Social Impact," *Physical Rev. A*, vol. 45, no. 2, 1993, pp. 703–716.
8. S. Moscovici and M. Zavalloni, "The Group as a Polarizer of Attitudes," *J. Personality and Social Psychology*, vol. 12, no. 1969, pp. 124–135.
9. E. Noelle-Neumann, *The Spiral of Silence: Public Opinion—Our Social Skin*, Univ. of Chicago Press, 1984.
10. F. Perroux, "Economic Space: Theory and Application," *Quarterly J. Economics*, vol. 64, Feb. 1950, pp. 89–104.
11. A. Nowak et al., "Simulating the Coordination of Individual Economic Decisions," *Physica A*, vol. 297, 2002, pp. 613–630.
12. A. Nowak and M. Lewenstein, "Modeling Social Change with Cellular Automata," *Modeling and Simulation in the Social Sciences from the Philosophy of Science Point of View*, K.T. Hegselmann and U. Müller, eds., Kluwer Academic, 1996, pp. 249–285.
13. A. Nowak, B. Latané, and M. Lewenstein, "Social Dilemmas Exist in Space," *Social Dilemmas and Cooperation*, U. Schulz, W. Albers, and U. Müller, eds., Springer-Verlag, 1994, pp. 114–131.
14. A. Nowak, R.R. Vallacher, and E. Burnstein, "Computational Social Psychology: A Neural Network Approach to Interpersonal Dynamics," *Computer Modeling of Social Processes*, W. Liebrand, A. Nowak, and R. Hegselman, eds., Sage, 1998, pp. 97–125.
15. J.J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities," *Proc. Nat'l Academy of Sciences*, vol. 79, no. 8, 1982, pp. 2554–2558.

NEXT ISSUE

March/April: Planning with Templates

Despite the promise of intelligent systems technology, interactive intelligent systems are still hard and expensive to build, hard to understand, hard to use, and hard to evolve and modify through their use. One promising new technology area is agent-based intelligent forms as typified in DARPA's research and demonstration program in Active Templates. Such forms let people create and use applications with simple, yet smart interfaces for getting the information they need and for capturing decisions they make on the basis of this information. This special issue will present the state of the practice, descriptions of research and evolving technology, application-focused case studies, and a roadmap for required research. In addition, this issue will document the substantial progress that has taken place in the past 10 years. The February 1995 *IEEE Expert* (now *IEEE Intelligent Systems*) described the state of the practice in knowledge-based planning and scheduling with case studies of application in the military context and to user-centered software engineering. Significant progress since then has yielded powerful, yet underutilized new capabilities that can improve intelligent systems applications in many areas.