

МЕТОДЫ ПОСТРОЕНИЯ СОЦИОСЕМАНТИЧЕСКИХ СЕТЕЙ ЗНАНИЙ

Сергеев Н.Е.¹, Целых А.А.¹

¹ФГАОУ ВПО «Южный федеральный университет», Ростов-на-Дону, Россия (347928, г. Таганрог, пер. Некрасовский, 44), e-mail: nesergeev@sfedu.ru

Авторы проанализировали использование знаний в таких, на первый взгляд, различных областях, как управление знаниями промышленных корпораций, в поисковых системах и системах поддержки принятия решений. Полученные результаты натолкнули на необходимость создания единой модели представления знаний, общей для всех трех областей. Знания в современном мире социальны – и в смысле распределенности между индивидами, находящимися в социальных отношениях, и в том смысле, что решать современные проблемы в этих областях можно только в социальном взаимодействии. В основе предлагаемой модели социосемантической сети лежит графогиперграфовая парадигма с тем допущением, что могут существовать отношения не только между вершинами в виде ребер, но и отношения различного типа между группами вершин, ребрами и вершинами и между самими ребрами. Такой подход наиболее точно отражает социальность знаний. Предлагаемая модель не исключает выделения в ней процессов трансформации знаний для управления знаниями и возможностей для интеллектуального индексирования в поисковых системах, а также использования фреймов, продукций и семантических сетей для систем поддержки принятия решений.

Ключевые слова: социосемантическая сеть, сеть знаний, гиперграф.

METHODS OF BUILDING SOCIOSEMANTIC KNOWLEDGE NETWORKS

Sergeev N.E.¹, Tselykh A.A.¹

¹Southern Federal University, Rostov-on-Don, Russia (347928, Taganrog, 44, Nekrasovsky Street), e-mail: nesergeev@sfedu.ru

Authors have analyzed the usage of knowledge in such at first glance different areas as knowledge management in a corporate sector, as well as information retrieval systems and decision support systems. The obtained results revealed the necessity to create shared model of knowledge representation that will be adequate for all the three knowledge areas. In the contemporary world, knowledge is social in the sense of its distribution between human beings engaged in social relations and in the sense of understanding that modern problems in these areas could be solved only by means of social interaction. The suggested model of the socio-semantic network is based on a graph-hypergraph paradigm with an assumption that there are relations not only between nodes in the form of edges, but also between groups of nodes, edges and nodes as well as between edges themselves. This approach depicts more adequately social character of knowledge. The suggested model does not exclude accentuation of transformation processes for knowledge management and knowledge-based indexing in information retrieval systems as well as usage of frames, production and semantic networks for decision support systems.

Keywords: sociosemantic network, knowledge network, hypergraph.

Введение

Роль знаний в современном мире необычайно возросла. Представление и оперирование знаниями лежит в основе и прикладных систем, основанных на знаниях, и экономических систем в целом. Современная экономика характеризуется как экономика, основанная на знаниях. В этой связи постоянного внимания исследователей требуют задачи представления знаний от уровня прикладной информационной системы до уровня представления и управления знаниями корпорации.

Авторы проанализировали использование знаний в таких, на первый взгляд, разных областях, как управление знаниями промышленных корпораций, в поисковых системах и системах поддержки принятия решений. Полученные результаты натолкнули на

необходимость создания единой модели представления знаний, общей для всех трех областей.

Трансформации знаний в организационных системах

Приведем с небольшими изменениями представленный в [2] «градиент» возрастания общности принадлежности знаний: индивидуальные, групповые, организационные, отраслевые, общеонтологические. Изменения заключаются в том, что мы заменили «межотраслевой уровень» общеонтологическим, имея в виду, что знания в конкретной организации или предметной области часто становятся полезными не только для других организаций, но и для бытового применения, и для человечества в целом. Например, Интернет из инструмента обмена данными в военной области стал основой для многих социальных процессов.

Другой аспект создания знаний – эпистемологический. Он представлен двумя полюсами – «формализованное» и «неформальное» знание. Большинство авторов [2; 4], исследующих проблему управления знаниями, выделяют следующие процессы преобразования неформальных знаний в формальные и обратно: социализация (неформальные в неформальные – приобретение опыта в процессе общения), экстернализация (неформальные в формальные – создание (рождение) концепций, курс введения в специальность), комбинация (формальные в формальные – детальное проектирование изделия, выполнение учебного проектирования), интернализация (формальные в неформальные – обучение, в том числе на практике).

Эти трансформации описаны для уровня организации, которая выполняет свойственные ей целевые функции – фирма производит, учебное заведение обучает. Однако и знания отдельного индивида могут подвергаться этим трансформациям.

В этой связи желанным становится инструмент, в котором могут существовать и совокупность знаний конкретной предметной области, и знания отдельного специалиста, и знания обучаемого, не мешая и не искажая друг друга.

Дадим несколько замечаний по поводу четкого деления на трансформации. На самом деле, трудно отделить процессы решения одной задачи от других. И отдельный индивид, и организация решают одновременно несколько задач, и даже для отдельно взятой задачи существуют параллельные процессы, пусть с разделением времени. Таким образом, можно констатировать наличие всех четырех процессов трансформации знаний одновременно. Собственно процесс формализации знаний является процессом отчуждения знания индивида. Формализованные знания ему уже не принадлежат или принадлежат не только ему. Однако исключением может быть «формализация для себя» – записи, пометки дневники и т.д.

Простое суммирование индивидуальных знаний дает значительно меньший объем знаний не только за счёт ненулевого пересечения знаний индивидов, но и из-за того, например, что неформальные знания складываются в конкретном знании индивида, значит, он должен овладеть, признать своими обе слагаемых компоненты (свои и чужие). Но, как правило, это рождает некоторое третье знание, которое владельцы слагаемых знаний не обязательно признают своим.

Не факт, что процесс превращения неформальных знаний в формальные является процессом созидательным, иногда, чтобы из абстрактного знания сделать конкретное знание, нужно многое загрузить, от многого отказаться. Делание выводов и обобщений одним человеком производит информацию более высшего уровня, но при этом вынуждает другого человека принять эти выводы, а не воспользоваться первоисточником. Конкретные работы делают одни знания востребованными, но они же и делают ненужными невостребованные в конкретной работе знания.

Да и такое деление знаний на два плюса – формализованное и неформализованное знание – не продуктивно в том смысле, что некоторое полезное знание может находиться частично в неформализованном, а частично в формализованном виде, также частично в индивидуальном, групповом, организационном, отраслевом и общеонтологическом состоянии.

Из вышесказанного мы вынуждены признать одновременное существование различных видов знаний, находящихся в различных процессах трансформации. Кроме этого, все виды трансформаций происходят в процессе общения, «совокупное знание» организации является распределенным между знаниями индивидов – знанием совокупности индивидов (индивидуумов), находящихся в процессе социализации.

Таким образом, все процессы трансформации по сути своей являются различными составляющими процесса социализации. Эра информационных технологий характеризуется значительным влиянием на социальные процессы (речь не только о социальных сетях в Интернете). Социальная сеть в силу наличия информационных и социальных связей является хранителем, генератором и трансформатором знаний. Поэтому большую популярность приобретают социальные программные продукты [1]: Wiki-технологии, технологии дистанционного обучения (MOOC, P2PU).

Авторы в настоящей работе предлагают модель для представления знаний, в которой можно отразить и сами элементарные знания (в виде фреймов, продукционных сетей, семантических сетей), и их принадлежность индивидам, и социальную распределенность, а также актуальность и изменения во времени.

Представление знаний в поисковых системах

Разрабатываемая модель должна позволять представлять знания различной степени общности. В том числе элементы метаданных, классифицирующие и обобщающие. Структуры, используемые для этого в настоящее время, практически во всех случаях состоят из поименованных элементов и связей между ними. Это таксономии, фолксономии и онтологии.

Таксономии представляют собой древовидные структуры, которые используются для представления иерархий, схем деления и классификации. Каждый объект в таксономии представлен в одном месте и одном узле. Корневой узел (корень) представляет наибольшую степень общности. Свойства корневого узла наследуются узлами нижних уровней.

Онтологии также представляют объекты или субъекты и отношения между ними в виде узлов и связей, но об иерархии можно говорить только в контексте направленности связей. В онтологиях связи не носят общий для всей модели смысл, например «состоять из...», а могут иметь индивидуальный смысл. Каждая связь имеет свой идентификатор, который отсылает нас к некоторой индивидуальной модели отношений, которая может также представляться онтологией или быть интуитивно понятной.

Местом рождения понятия «фолксономия» принято считать социальные сети, однако корнями его можно также считать привычки читателей делать пометки в книгах. Понятие фолксономия впервые употребил Томас Вандер Вал [1]. В общем случае фолксономия оперирует такими понятиями, как тег и объект. Сам объект может находиться в других более сложных отношениях, но некий субъект, имеющий на это полномочия, может присвоить (прикрепить) объекту некоторый тег. Объекту может принадлежать несколько тегов, теги могут содержать элементы авторизации, могут выбираться из стандартного (предоставленного) набора либо создаваться самостоятельно. Наличие тегов представляет собой более гибкий подход к классификации.

Заметим, что это еще не все типы структур. Нет сомнения, что Интернет еще представит нашему вниманию новые.

Авторы поставили перед собой задачу предоставления возможности описания максимально возможного числа структур отношений между элементами и их группами в социосемантических системах.

Классификация типов отношений в социосемантической модели

Классификация типов отношений между элементами в социосемантической сети основана на модификации графогиперграфовой парадигмы. Предложенная в [3; 5; 6], она позволяет отображать разнотипные отношения объектов.

1. Связи между простыми компонентами данных (неориентированные). Две вершины связаны неориентированным ребром, все связи в модели имеют одинаковый смысл

(например, возможные пути распространения информации, «петля» – пользователь переслал информацию на свой другой адрес), связь типа «вершина-вершина», представляется на нижнем уровне при помощи матриц (таблиц) инцидентности ($G_v \times G_e$ определяет, какое ребро какие две вершины связывает, G_v – множество вершин графа, G_e – множество рёбер графа), либо смежности ($G_v \times G_v$ определяет, какая вершина с какой другой вершиной связана, в неориентированном графе матрица смежности симметрична относительно диагонали).

2. Ориентированные связи между простыми компонентами данных. Две вершины связаны ребром (связь типа «вершина → вершина»), все связи в графе имеют одинаковый смысл (например, указана последовательность распространения информации или последовательность запуска приложений одним пользователем), представляются также матрицами $G_v \times G_e$ и $G_v \times G_v$, но матрица смежности уже не симметрична.

3. Ориентированные и неориентированные связи между простыми компонентами данных. В графе присутствуют группы связей различных типов, представленных выше. Например, на одном и том же множестве вершин типом связей «вершина-вершина» обоюдные одинаковые возможности, а «вершина → вершина» представляет разные возможности.

4. Каждая ориентированная или неориентированная связь графа имеет своё смысловое значение, каждая связь должна быть идентифицирована. Например, парные отношения узлов имеют различный тип или указывают, каким образом узел перемещает информацию между приложениями. В графе присутствуют группы связей различных типов: ориентированные и неориентированные.

5. Неориентированные ребра гиперграфа представляют подмножества некоторого множества. Связь ребра и вершины типа « $h_v - h_e$ », где h – hypergraph, представляется на нижнем уровне при помощи матриц инцидентности $H_v \times H_e$ – определяет, какое ребро какие вершины связывает.

6. В ориентированных ребрах гиперграфа выделяется одна из вершин (например, $*x_g$), объединенных ребром, которая имеет особый статус. Связи внутри ребра типа «все к одному» (например, помеченная вершина указывает приоритетный элемент).

7. Сложные связи внутри рёбер гиперграфа (например, описание приоритетов некоторых элементов). «Графовые» связи внутри ребра здесь не следует рассматривать как самостоятельные, а только как описывающие некоторые свойства конкретного ребра. Такие отношения представляются матрицей инцидентности гиперграфа, содержащей ссылки на матрицы смежности, описывающие отношения вершин внутри конкретного ребра.

8. «Графовые» связи между вершинами, входящими в разные ребра гиперграфа (например, обозначение скрываемых отношений).

9. Связи между ребрами гиперграфа (например, связи приоритета одной группы элементов над другой) представляются несколькими матрицами.

10. Связь «вершина одного гиперграфа – ребро гиперграфа» (отношение одного элемента в целой группе элементов другого ребра).

Таким образом, несмотря на то, что понятие гиперграфа уже представляет собой обобщение графовой парадигмы, предложенная модель обобщает понятие гиперграфа – в пределах одного ребра гиперграфа могут существовать связи других функциональных типов, которые будем представлять в виде графов. Можно также сказать, что такая модель представляет собой иерархический гиперграф, то есть гиперграф, в котором и вершины, и ребра могут входить в различные иерархические группы для разных режимов функционирования модели. На рис. 1 представлены все типы отношений иерархического гиперграфа социосемантической сети. Теги, свойственные фолксономии, могут присваиваться любым элементам и их группам.

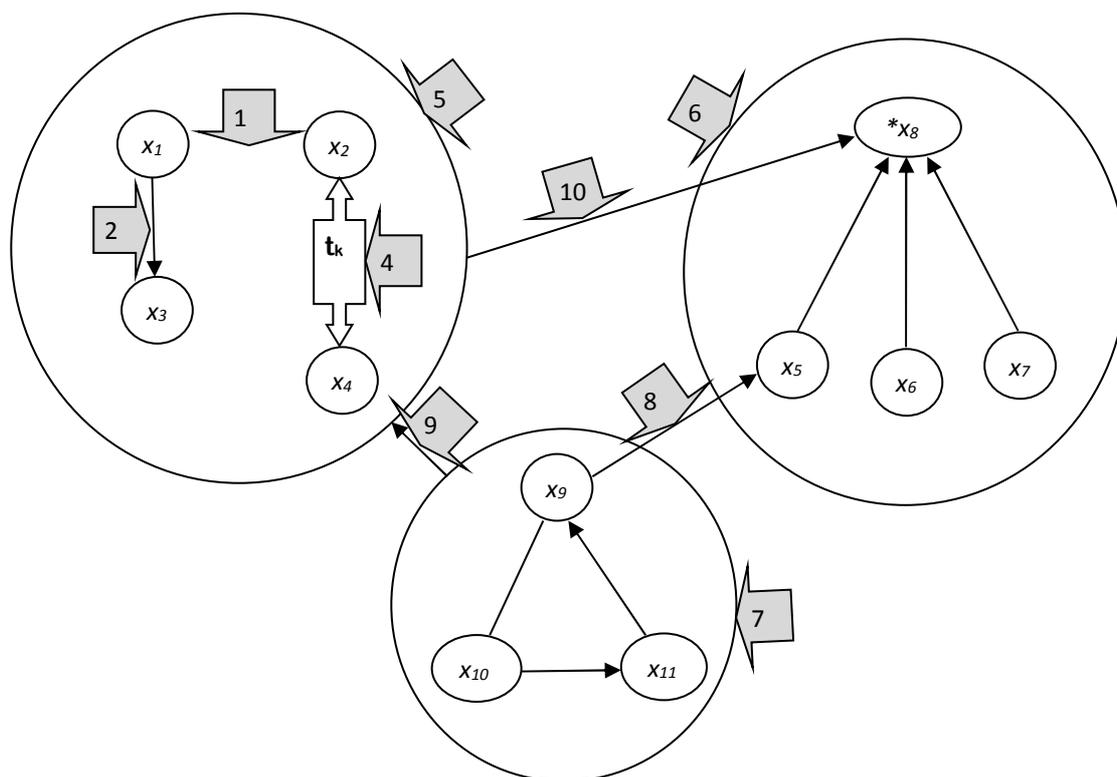


Рис. 1. Типы отношений в социосемантической сети

Заключение

Изложенный авторами подход наиболее точно отражает социальность знаний. Предлагаемая модель социосемантической сети не исключает выделения в ней в дальнейшем процессов трансформации знаний для управления знаниями и возможностей для интеллектуального индексирования в поисковых системах, а также использования фреймов, продукций и семантических сетей для систем поддержки принятия решений.

Список литературы

1. Морвиль П. Тотальная видимость. – СПб. : Символ-Плюс, 2008. – 272 с.
2. Нонака И., Такеучи Х. Компания – создатель знаний. Зарождение и развитие инноваций в японских фирмах. – М. : Олимп-Бизнес, 2003. – 384 с.
3. Сергеев Н.Е., Целых Ю.А. GN-модели социальных сетей // Известия Южного федерального университета. Технические науки. – 2009. – № 1. – С. 90-95.
4. Управление знаниями / Высшая школа менеджмента СПбГУ [под ред. Т.Е. Андреевой, Т.Ю. Гутниковой]. – СПб. : Высшая школа менеджмента, 2010. – 514 с.
5. Целых А.А. Графогиперграфовая модель семантической социальной сети // Известия Южного федерального университета. Технические науки. – 2012. – Т. 129, № 4. – С. 225-229.
6. Sergeev N.E., Tselykh A.N., and Tselykh A.A. Generalized Approach to Modeling User Activity Graphs for Network Security and Public Safety Monitoring // Proceedings of the 4th International Conference on Security of Information and Networks. SIN 2013, Aksaray, Turkey, November 26-28, 2013. – URL: <http://dx.doi.org/10.1145/2523514.2523534>.

Рецензенты:

Ромм Я.Е., д.т.н., профессор, заведующий кафедрой информатики ФГБОУ ВПО «Таганрогский государственный педагогический институт имени А.П. Чехова», г. Таганрог.

Курейчик В.В., д.т.н., профессор, заведующий кафедрой систем автоматизированного проектирования ФГАОУ ВПО «Южный федеральный университет», г. Таганрог.

Антонов А.В., д.т.н., профессор, декан факультета кибернетики, Обнинский институт атомной энергетики Национального исследовательского ядерного университета МИФИ Министерства образования и науки РФ, г. Обнинск.