

ПРЕДСТАВЛЕНИЕ СТРУКТУРЫ ПРЕСТУПЛЕНИЯ В ВИДЕ СЕМАНТИЧЕСКОЙ СЕТИ

В статье рассматривается модель преступления, построенная на основе одного из вариантов расширенных семантических сетей. Для этого предлагается новый для криминалистики подход к представлению преступления в виде совокупности взаимосвязанных событий. Приведена технология выявления событий и пример представления семантической сетью подделки документов.

Проблема построения модели широко рассматривается в науке. В криминалистике развиваются ряд подходов к описанию преступления в рамках фазового и элементного подходов, однако все они носят содержательный характер и далеки от своего формального завершения. Представляется, что нужна новая концепция описания предметной области, построенная на несколько иных основаниях. Последние достижения науки в области описания сущности предметной области достигнуты в области искусственного интеллекта. Развитие этой отрасли показало, что наиболее оптимальным является ориентация на модели знаний.

При всем многообразии работ в области искусственного интеллекта, выбор модели для представления знаний представляется небольшим: наиболее распространены продукционные модели, фреймы Минского, семантические сети. По всей видимости, все они имеют право на существование, однако в настоящей работе остановимся на модели в виде семантической сети по ряду соображений. Во-первых, семантическая сеть в большей степени соответствует описанию на естественном языке; во-вторых, легко позволяет производить детализацию модели, т. е. при необходимости более детально описывать объект; в-третьих, она соответствует современным представлениям об организации долговременной памяти человека. Более того, многие ученые приходят к выводу, что комплексное представление информации в компьютере будет развиваться именно в направлении семантических сетей¹. И, на наш взгляд, основное преимущество этой модели состоит в том, что она соответствует еще и сложившимся сейчас представлениям о механизме преступления как совокупности элементов и связей между ними. Хотя отмечается и недостаток - сложность поиска решения на семантической сети².

Семантические сети представляют богатые возможности по описанию предметной области,

поскольку используют семантические языки, максимально приближенные к естественному, имеют вполне доступный и понятный вид, оперируют со структурными объектами, приближенными к объектам предметной области.

Начало использования семантики для описания предметной области при применении вычислительной техники относится еще к 60-м годам. Сейчас подход активно эксплуатируется для создания кибернетических систем искусственного интеллекта, систем автоматизированного перевода, информационных систем, общающихся на естественном языке.

Семантическая сеть – ориентированный граф, вершины которого соответствуют понятиям, а дуги отношениям между ними. Семантическая сеть конструируется таким образом, чтобы отражать сущность моделируемой области. Однако такие сети, отражающие вершины-объекты и связывающие их дуги-отношения, обладают ограниченными возможностями для представления знаний. В частности, возникают трудности для реализации обобщенной информации, отношений между самими отношениями и т. д. Указанный недостаток вполне очевиден, так как дуга однозначно определяется лишь тремя параметрами: имя дуги, входящий узел, исходящий узел. То есть дуга не может моделировать ситуацию, зависящую более чем от трех аспектов.

Для устранения этого недостатка вводится понятие расширенных семантических сетей, в которых вводятся так называемые «развязывающие элементы» – вершины связи. Эти вершины отображают сложные отношения между понятиями, например, причинно-следственные, временные динамические отношения³.

Один из таких вариантов языка расширенных семантических сетей приведен в известной работе Э.В. Попова. В данном представлении основ-

¹ Шемакин Ю.И., Романов А.А. Компьютерная семантика. – М.: «Школа Китайгородской», 1995. – С. 116. Одной из лучших форм представления знаний называют семантические сети. Э.В. Попов (Попов Э.В. Общение с ЭВМ на естественном языке. – М.: Наука, 1982. – С. 163-164), Ю.Я. Любарский (Любарский Ю.Я. Интеллектуальные информационные системы. – М.: Наука, 1990. – С. 30).

² Информатика. Учебник // Под. ред. Н.В. Макаровой. – М., 1997.

³ Кузнецов И.П. Семантические представления / Отв. ред. Золотов Е.В. – М.: Наука, 1986. – С. 18.

ным элементом является ситуация, для отображения которой вводится специальный узел. Количество входящих и исходящих из данного узла дуг может быть произвольным и зависеть лишь от характера ситуации⁴. В криминалистике термин «ситуация» ассоциируется с следственной ситуацией, поэтому в дальнейшем «ситуацию» в интерпретации Э.В. Попова будем обозначать термином «событие».

Следует отметить, что подход по представлению преступления в виде совокупности событий достаточно нов для криминалистики. Чаще всего криминалисты рассматривали преступление как совокупность этапов (фаз), показывающих формирование и уточнение преступного умысла; создание прогностической модели преступления; создание преступной группы; реализацию преступного замысла⁵. Но предлагаемая трактовка преступления через понятие «событие», в отличие от этапов, кроме всего прочего предполагает ряд чисто формальных преимуществ. Так, событие легко привязать ко времени, месту, конкретным объектам и субъектам деятельности, то есть соединить с конкретным преступным действием. Другое преимущество состоит в возможности легко установить иерархию событий. В предлагаемом Э.В. Поповым варианте предусмотрены построения процессов, подсоединение к модели внешних знаний о предметной области в виде обычных баз данных⁶.

Хотя и относительная, но все же новизна подхода в определении преступления через последовательность взаимосвязанных событий вызывает необходимость исследований содержательной структуры преступления. Технология построения семантической модели заключается в определении множества понятий, необходимых для описания преступления, которые объединяются в близкие по семантике группы. Для многих областей структурирование не составляет больших проблем, поскольку формализуются достаточно изученные области, имеющие хорошо проработанную терминологию и иерархию понятий. Но, не смотря на то, что изучение преступления составляет одну из основных задач криминалистики, представляется, что область преступной деятельности семантически не проработана до такой степени, которая бы позволила легко структурировать все входящие в нее понятия.

Кроме этого, требуется также определиться с совокупностью отношений между понятиями и, пожалуй, самое важное, провести анализ сущности событий, их выделение из структуры преступления, да и формулирование простого перечня таких событий.

Классический подход криминалистов к решению таких проблем состоит в глубоком содержательном анализе практики расследования преступлений, выделении типичных обстоятельств и характеристик преступления, основанных на выявлении механизма этих действий. По существу, результаты таких исследований можно найти в частных методиках расследования отдельных видов преступлений. Но произвольная форма таких методик, их субъективное содержание, затрудняет подобную работу.

Для выявления семантической структуры преступления (понятий, отношений и событий) в условиях неопределенности существует практически один путь, заключающийся в анализе самой предметной области. Таковой может выступить только следственная практика, результаты ее обобщения. Документированное представление этих результатов отображено в материалах уголовных дел, отдельных процессуальных, оперативных и управленческих документах.

Учитывая приведенные соображения по структуре семантической сети, конкретная реализация была осуществлена на примере преступлений, связанных с подделкой документов. Предметом исследования послужили тексты обвинительных заключений по ряду уголовных дел этой категории. Технология исследования чисто формальная: семантические единицы определялись исходя из частоты встречаемости в исходных текстах. Для этого использовались данные, выделенные при помощи формальной процедуры, предлагаемой программой «TextAnalyst 2.0». (Была использована ознакомительная версия, взятая на сайте компании «Микросистемы» по адресу: www.textanalyst.ru).

Технология анализа требует некоторого пояснения. Программа «TextAnalyst» формирует результаты в виде уже готовой семантической сети. Однако проведенная серия исследований различных документов разного содержания показала, что полученная семантическая сеть позволяет выделить основные понятия, но сама сеть мало отражает содержательную структуру области. Причиной это-

⁴ Попов Э.В. Общение с ЭВМ на естественном языке. – М.: Наука, 1982. – С. 165-166.

⁵ Каминский М.К. Криминалистическая характеристика деятельности по выявлению, раскрытию и расследованию преступлений // Правовые и общественно-экономические науки и борьба с хищениями социалистического имущества. Труды Горьковской высшей школы МВД СССР. – Выпуск 8. - Горький, 1977. – С. 154.; Горбачев А.В. Основания, содержание и значение криминалистической характеристики преступлений // Там же. – С.148.

⁶ Попов Э.В. Общение с ЭВМ на естественном языке. – М.: Наука, 1982. – С. 174-176.

го является то, что большое количество исходного материала усредняет семантические связи, и в результате они оказываются плохо связанными содержательно.

Соответственно можно сделать вывод о возможности двух технологий формального анализа текста, описывающего событие преступления. Первая – выделение содержательной структуры. Она сформируется лишь при условии малого количества исходного материала, структура семантической сети будет соответствовать семантической структуре анализируемого текста. Вторая технология – работа с большими массивами текстов. Здесь выделяются наиболее существенные объекты, но связи между объектами не будут соответствовать семантическим структурам конкретных фрагментов текстов.

Очевидно, что наибольший интерес представляет модель, охватывающая не конкретное преступление, а описывающая их наибольшее количество. Поэтому результаты формального анализа использовались лишь для выделения структуры исследуемой области, а конкретное семантическое представление было получено содержательным путем. Проведенная работа достаточно объемна, поэтому привести всю технологию не представляется возможным. Отметим лишь, что последовательность операций носила следующий характер.

1. Вначале была получена общая структура (перечень) понятий, из которых сформирована иерархическая структура.

2. Выявлены отношения, связывающие основные понятия.

3. Из выявленных отношений сформирована система событий, характерных для рассматриваемой области.

4. Описана система событий, каждое из них связано с существенными для данного события понятиями с помощью соответствующих отношений.

5. Полученный результат записан в виде семантической сети.

В качестве иллюстрации на рис. 1 приведен фрагмент семантической сети, описывающий процесс подделки документа и разработанный по материалам одного из уголовных дел. Подделка документа описывается тремя событиями, названными ВЫПОЛНЕНИЕ. Эти события отражают реальные действия по нанесению текста, подписи и оттиска печати.

Очевидно, что каждые из присутствующих понятий и событий в компьютерном представлении подробно описываются с помощью иерархической системы с рядом атрибутов, характерных для кон-

ретного понятия и понятия верхних уровней. Поэтому схема, приведенная на рис. 1, упрощенная, поскольку на ней иерархия понятий и событий предметной области отражена не в полном объеме. Сделано это для того, чтобы не загромождать схему.

Приведенный пример, по существу, лишь демонстрирует возможности метода. Так, для практики расследования существенное значение имеют, например, способ выполнения поддельной подписи, способ изготовления печати и многое другое. Отсутствие этого в приведенном примере объясняется тем, что указанные элементы отсутствовали в анализируемых первичных документах. Требуется детальная проработка и иных проблем сетевого описания преступления.

В заключении следует отметить, что предлагаемая модель преступления может найти применение в широких областях информационного обеспечения раскрытия и расследования преступлений. В частности, подобные модели являются ядром баз знаний экспертных систем, могут использоваться при создании естественно-языкового интерфейса к ИС обычного типа. Отдельно следует отметить применение сетевых моделей для создания интеллектуальных информационных систем, содержащих наиболее полное описание уголовных дел. Представляется, что разработка формализованного описания информационной структуры преступления является одной из актуальных криминалистических проблем, решение которой позволит поднять информационное обеспечение расследования на новый современный уровень.



Рисунок 1. Представление процесса подделки документов в виде семантической сети.