

автоматизированная обработка текстов протоколов допроса

Н.В. Деева

В статье рассматриваются типы документов дела по расследованию преступления. В материалах дела выделяются типы документов, отражающие следственные действия и документы, в которых представлена информация о происшествии. Приведена классификация документов допросов. Выделены их смысловые части и построена модель решения задачи анализа текста допросов. Предложены основные методы, на которых базируется предлагаемая модель.

Введение

На современном этапе развития информационных технологий представляется возможной реализация различных информационных систем, которые позволяют оперативно получать информацию необходимую при расследовании преступлений. Так, например, уже существует множество специализированных банков данных в таких направлениях, как дактилоскопия, баллистика, досье на преступников, фото и видео данные. Кроме того, существует множество систем для быстрого проведения дактилоскопической и баллистической экспертиз и т.д. Вопросы же автоматизации анализа материалов дела и документирование этих материалов остаются пока не решенными и выполняются вручную оперативным работником (следователем).

Таким образом, актуальной представляется задача обработки и анализа документов дела, например, анализ текстов допросов, анализ данных экспертиз и т.д., что окажет неоценимую помощь специалисту в расследовании преступления.

Типы документов дела

Процесс расследования полностью документируется, материалы помещаются в специальную папку и формируют, так называемое, дело. Авторами документов могут выступать как следователи, так и судмедэксперты, прокуроры, руководители организаций, представляющие некоторую справочную информацию и т.д. Например, для задержания подозреваемого, готовится постановление о задержании за подписью прокурора, затем на основании задержания пишется протокол задержания непосредственно следователем.

Таким образом, можно выделить два класса документов: документы, несущие информацию непосредственно о происшествии (протоколы допросов, материалы экспертиз и т.д.) и документы, регламентирующие

мероприятия по расследованию происшествия (постановления, ходатайства, запросы и т.д.).

Естественно, для описания сути рассматриваемого дела важно детально проанализировать документы, которые были отнесены к первому классу (типу).

Для формирования списка документов, отнесенных к первому классу, был проанализирован сборник образцов уголовно-процессуальных документов [1], что позволило выделить следующие типы документов: протоколы, объяснения, рапорты, заявления и ответы на запросы. Каждый тип может иметь набор подтипов. Что касается протоколов, можно выделить следующие типы документов: протокол допроса, протокол осмотра, протокол очной ставки, протокол обыска и т.д.

Предметом исследования в данной статье являются протоколы допросов. Данные таких протоколов можно охарактеризовать как отражение субъективной картины происшествия, с точки зрения допрашиваемого.

Протокол допроса

Допрос – это следственное действие, предусмотренное и регламентированное УПК, состоящее в получении уполномоченным органом показаний от лиц, располагающих сведениями об обстоятельствах, имеющих значение для дела и подлежащих доказыванию [2].

Так в [2] рассмотрена классификация допросов. По процессуальному положению допрашиваемого выделяют допрос свидетеля, допрос потерпевшего, допрос подозреваемого, а также допрос обвиняемого, допрос подсудимого и допрос эксперта. По возрастным особенностям допрашиваемого выделяют допрос взрослого, допрос несовершеннолетнего и допрос малолетнего. По очередности: допрос первичный и допрос вторичный. По позиции, занимаемой допрашиваемым выделяют допрос лица, дающего правдивые показания; допрос лица, дающего заведомо ложные показания и допрос добросовестно заблуждающегося лица.

В протоколе допроса можно выделить несколько смысловых частей:

1. информация о протоколе (например, протокол допроса свидетеля);
2. дата и место, проведения допроса;
3. информация о лице, производящем допрос (например, следователь прокуратуры Иванов А.В.);
4. информация о допрашиваемом (ФИО, дата рождения, место рождения, место проживания и т.д.);
5. перечень статей о правах и обязанностях допрашиваемого;

6. информация о языке, на котором будет проводиться допрос;
7. предупреждение об уголовной ответственности за дачу заведомо ложных показаний, уклонении или отказе от дачи показаний;
8. непосредственно текст допроса, который может носить повествовательную и вопросно-ответную формы или чередование их.

Для оптимизации процесса допроса наиболее важно разработать шаблон к пунктам 1-7 приведенного перечня, в который автоматически могут вноситься данные о следователе, месте и дате проведения допроса, данных статей и т.д. Посредством такого шаблона легко можно выбрать данные о допрашиваемом и использовать их в дальнейшем, например, для автоматической вставки в другие документы.

Анализ восьмой части протокола – текста допроса, который бесспорно самый важный, сводится к задаче анализа текстов на естественном языке. Именно эта часть документа является отражением субъективной модели происшествия с точки зрения допрашиваемого. Если на основании анализа текста построить некоторую субъективную модель происшествия, то анализ набора таких моделей представляет возможным выделить:

- списки фактов, которые имели место во всех моделях и предположительно правдивые;
- списки несоответствий, позволяющее установить факт лжи либо заблуждения;
- набор вопросов недостаточно полно раскрытых, но существенных для расследования происшествия.

Значительную помощь в раскрытии преступления для следователя представит отчет об анализе протоколов допроса. Такой отчет видится в форме сводки, в которой отдельными частями будут выделены:

- списки фактов и их источников (с приведением тех мест, протокола, где дается пояснение по этому факту);
- списки несоответствий и их источников (с приведением частей текста оригинала);
- списки вопросов, о которых было упомянуто в текстах допроса.

Анализ текста допроса

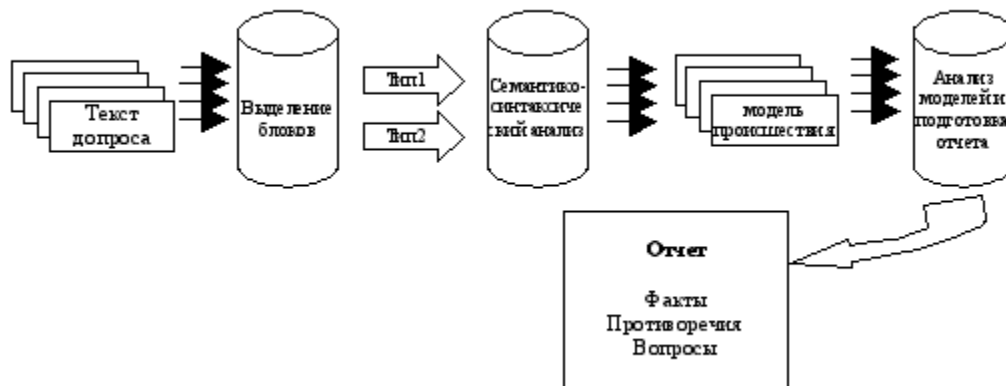
Как упоминалось выше, при анализе текста показаний выявлено, что информация может быть нескольких видов: повествование (монолог), которое, как правило, записывает следователь со слов допрашиваемого и вопросно-ответная часть (диалог), где фиксируется и вопрос следователя, и

ответ допрашиваемого. Для текстов первого вида характерен строгий стиль изложения, текст не нагружается различными конструкциями (вводными, оборотами и т.д.). Повествование и вопросно-ответные блоки могут чередоваться. Текст нагружен именами собственными, датами, названиями объектов недвижимости, фирм и т.д.

Для решения задачи обработки текста допроса выделим ряд подзадач:

1. Предварительное разбиение текста на блоки повествования и запросно-ответные, так как анализ каждого блока проводится исходя из его вида разными модулями.
2. Выделение в тексте дат, фамилий, названий и т.д.
3. Синтактико-семантический анализ входного текста.
4. Построение субъективной модели происшествия.
5. Анализ набора моделей.
6. Подготовка отчета для пользователя.

Созданную модель решения предлагаемой задачи можно представить в следующем в



иде:

Рис. 1. Модель анализа текста допросов

Для выделения блоков текста предлагается использовать обычный метод последовательного поиска слов “вопрос”, “ответ”. Для анализа повествовательного блока предлагается использовать семантико-синтаксический анализ каждого предложения, для которого строится поверхностно-семантический граф с корнем-глаголом. Затем строится субъективная модель происшествия в виде семантической сети, где узел – это граф, отражающий некоторое действие и его характеристики, а дуга – связь между действиями. Для вопросно-ответных блоков строятся поверхностно-семантические графы для вопроса и ответа, а затем строится фрагмент

семантической сети, для вопроса (в повествовательной форме), и ответа, которые будут, как сказано ранее, - узлами, а дуга отражает суть вопроса (например, почему, зачем, сколько и т.д.). Затем строится обобщенная семантическая сеть. Анализ набора таких семантических сетей происходит на базе выделенных ключевых действий, совершенных некоторыми субъектами над субъектами/объектами. В ходе анализа находят факты, отраженные во всех анализируемых семантических сетях, и противоречия, а также выделяется перечень вопросов, по которым информация предоставлена, но не всеми или не совсем полно. Понятно, что таких вопросов может быть масса, и далеко не все они будут полезны следователю. Фильтрация таких вопросов является предметом будущих исследований.

Литература

1. Сборник образцов уголовно-процессуальных документов с комментариями. Возбуждение уголовного дела и предварительное расследование: учебн.-практ.пособие / авт.-сост. Г.Н. Васильев, М.И. Емшанский, Е.И. Климова и др.; под рук. И науч. Ред. Проф. М.А. Шостака. – Минск : Амалфея, 2006. – 704с.
2. Зорин, Г.А. Многовариантные программы допросов: технологии построения и применения / Зорин Г.А. - М.: Изд-во деловой и учебной литературы, 2005. - 336 с.

Наталья Владимировна Деева, ассистент кафедры информатики и выч. техники Гродненского государственного университета им. Янки Купалы, ndeeva@grsu.by, аспирантка кафедры математического обеспечения ЭВМ факультета прикладной математики и информатики Белорусского государственного университета, научный руководитель – Михаил Константинович Буза, доктор технических наук, профессор кафедры МО ЭВМ ФПМИ БГУ.

** уголовно-процессуальный кодекс (УПК)*