УДК 343.98 DOI: 10.18698/2541-8009-2023-8-925

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ЭКСПЕРТНОГО ИССЛЕДОВАНИЯ ОФИСНЫХ ФАЙЛОВ ФОРМАТА XML

Н.И. Лядовская lyadovskayan@student.bmstu.ru

МГТУ им. Н.Э. Баумана, Москва, Российская Федерация

Аннотация

В настоящее время компьютерные технологии так или иначе внедрены во все сферы жизни общества. Однако, к сожалению, применяться они могут не только в целях упрощения некоторых аспектов жизнедеятельности, но и в противоправных целях, когда какое-либо достижение в сфере компьютерных технологий либо используется как средство совершения преступления, либо выступает его предметом. Для работы с подобными доказательствами требуется привлечение судебного компьютерно-технического эксперта, обладающего всем спектром необходимых для этого специальных знаний. В статье исследованы так называемые офисные файлы формата XML, представляющие собой популярные объекты судебной компьютерно-технической экспертизы, на предмет доступной криминалистически значимой для уголовного дела информации в целях раскрытия, расследования и предупреждения преступлений.

Ключевые слова

Теория судебной экспертизы, судебная экспертиза, судебная компьютерно-техническая экспертиза, электронные носители информации, криминалистически значимая информация, офисные файлы, формат файла, формат XML

Поступила в редакцию 21.05.2023

При рассмотрении особенностей экспертного исследования какого-либо объекта необходимо установить, во-первых, понятие объекта, а во-вторых, его место внутри существующих классификационных групп в рамках конкретной судебной экспертизы [1]. В общей теории судебной экспертизы под *объектом судебной экспертизы* подразумевается материальный носитель информации, содержащий сведения о факте или событии, необходимые для решения экспертных задач в рамках проводимого экспертного исследования. Подобные положения с поправкой на родовую и видовую специфику реплицируются и на определение объекта судебной компьютерно-технической экспертизы (СКТЭ). Под *родовым* или *видовым объектом* следует понимать определенную группу материальных объектов, обладающих определенными общими для рода в целом или для вида в частности свойствами, характеристиками и признаками, позволяющими относить их к электронным носителям информации (ЭНИ). В свою очередь, ЭНИ, обладающий рядом уникальных для него свойств, характеристик или признаков, исследуемый в рамках конкретной экспертизы, будет относиться к группе конкретных объектов.

Все объекты СКТЭ, являющиеся ЭНИ*, обладают рядом общих признаков:

– представление информации на носителе в закодированной на машинном (компьютерном) языке форме, электронной форме;

– интерпретация закодированной (цифровой) информации опосредована через материальный электронный носитель, вне которого такая информация не существует физически;

– режим доступа к информации носителя может быть многопользовательским — осуществляться несколькими субъектами;

– информация оперативно может преобразовываться в неэлектронные формы и обратно (распечатываться/сканироваться);

– копирование информации возможно на различные ЭНИ, а распространение на любые расстояния, ограниченные только радиусом действия современных средств электронной связи;

– сбор и исследование ЭНИ в целях уголовного судопроизводства осуществляются только с помощью специальных научно-технических средств собирания, обработки, хранения и исследования компьютерной информации и информационно-телекоммуникационных сетей и др. [2].

К ключевым признакам ЭНИ, характеризующим их как надежные источники доказательственной информации в рамках экспертного исследования, относятся:

– доступность при получении, в том числе с территориально удаленных облачных хранилищ, серверов;

– интерпретируемость средствами компьютерной техники, в том числе и в случаях применения правонарушителем криптографических средств;

– понятность правоприменителю, в том числе посредством содействия специалистов (экспертов), направленных на интерпретацию содержащейся информации посредством преобразования машинного языка на доступный для правоприменителя, а также разъяснения логических связей между действиями человека, работой компьютерных средств и следообразованием (в рамках экспертных выводов);

– возможность обеспечить долговременное хранение информации в аутентичной форме, то есть форме, пригодной для использования такой информации в качестве доказательств;

– обеспечение идентифицируемости цифровых документов, являющихся доказательственной информацией, в массиве им подобных (пути, способы получения информации) и др. [3].

Все объекты, которым присущи вышеуказанные свойства и характеристики, рационально дифференцировать на более мелкие группы для детальной прора-

* ГОСТ 2.051–2013. Единая система конструкторской документации. Электронные документы. Общие положения. М.: Стандартинформ, 2013. IV, 15 с.

ботки качественных и количественных показателей осуществляемых с ними экспертиз [4]. При этом следует учитывать функциональные элементы носителей, объем предмета СКТЭ, область специальных знаний, необходимых для исследования, способы оптимизации нагрузки на судебно-экспертные учреждения (СЭУ), лабораторий и отдельных экспертов, а также повышение эффективности экспертных исследований в конкретной области. Классификация в судебной экспертизе всегда являлась дискуссионным вопросом, особенно в отношении объектов, обладающих сложными смешанными свойствами. В СКТЭ это выражается в невозможности строго отнести тот или иной объект к определенной группе.

В процессе становления и развития СКТЭ подходы к классификации объектов эволюционировали следующим образом:

1) специалист в области криминалистики и судебной экспертизы, ученый-юрист Е.Р. Россинская предложила существование двух групп объектов:

– ЭНИ и их составляющие;

– программное обеспечение;

2) специалист в области методов судебно-экспертных исследований, юрист-криминалист Т.А. Аверьянова дополнила классификацию сетевыми объектами (ЭНИ, функционирующие в сети);

3) специалист в области уголовного процесса, криминалистики и судебной экспертизы А.И. Усов представил наиболее полную, по мнению автора настоящей статьи, классификацию, включающую:

– аппаратные объекты (персональные компьютеры, периферийные устройства, серверы, мобильные устройства, микроконтроллеры и т. п.), включающие наиболее значимый с позиции криминалистически значимой и доказательственной информации устройства памяти ЭНИ;

– программные объекты (системное (операционная система, среды разработки программ и т. п.) и прикладное (текстовые, графические редакторы и т. п.) программное обеспечение);

– информационные объекты (текстовые, графические документы, базы данных и т. п.);

– сетевые объекты (серверы, рабочие станции сети) [5].

Одними из информационных объектов СКТЭ из приведенной классификации являются текстовые файлы, в общем смысле определяемые как компьютерные файлы, содержащие текстовые данные, среди которых значительную часть занимают так называемые офисные файлы, или файлы офисных форматов, имеющие структуру XML. Под файлами офисных форматов подразумеваются файлы, создаваемые так называемыми офисными приложениями, которые входят в некоторый офисный пакет. К известным офисным пакетам относятся, например, Microsoft Office, OpenOffice и др. Под форматом файла принято по-

нимать определенные правила размещения в теле файла пользовательской и дополнительной служебной информации, или метаданных. Рассматриваемые объекты также обладают сложной структурой, поскольку содержат как текстовую, так и программную составляющие, что усложняет их отнесение к той или иной классификационной группе.

Документы, в основе которых лежит XML, фактически представляют собой ZIP-архивы, содержащие различные XML-файлы (части), организованные в единый пакет. XML-файл — это текстовый файл, в котором с помощью специальных маркеров создаются элементы данных, последовательность и вложенность которых определяет структуру документа и его содержание. Например, docx-файлы содержат XML-файл с описанием возможных типов содержимого и три каталога с различными XML-файлами: docProps, word, и _rels, которые включают в себя свойства документа, его содержание и отношения между файлами (рис. 1). Согласно данным официальной страницы Microsoft, такие документы сжимаются автоматически и в некоторых случаях могут быть на 75 % меньше исходных, что способствует существенной экономии места, необходимого для хранения файлов, а также уменьшению пропускной способности, необходимой для отправки файлов по сети Интернет, и др. При открытии файла он автоматически обновляется, а при его сохранении автоматически сжимается снова [6].

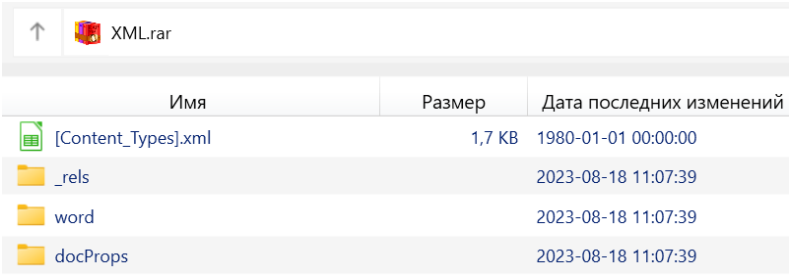

Рис. 1. Окно программы BreeZip, демонстрирующее содержимое документа XML

Исследование рассматриваемых файлов проводится на предмет наличия криминалистически значимой информации. Известно, что криминалистически значимую информацию составляют данные, имеющие отношение к уголовному делу и способствующие раскрытию и расследованию преступления [7]. Источником криминалистически значимой информации могут служить различного вида следы. В контексте настоящей статьи особый интерес представляют оставляемые цифровые следы как отображения действий пользователя с материальным носителем (ЭНИ), связанных с записью, хранением, обработкой и передачей информации. Представляться такие следы могут в виде команд или данных

в операционной системе, отдельной компьютерной программы, файлов с текстовой, графической, символьной информацией, баз данных, сообщений электронной почты (мессенджера, чата, форума), сведений о работе устройства, сайтов в сети Интернет. Из сущности XML-файлов, определяемых как структурированные иерархические текстовые файлы с различным содержанием, следует, что их создание, изменение и обработка осуществляются на электронном носителе информации и закономерно образуют цифровые следы, содержащие в себе криминалистически значимую информацию, в частности, текстового и символьного вида.

Все цифровые следы могут быть классифицированы по различным основаниям [8], например, в зависимости от инициатора каких-либо действий на ЭНИ. В этом случае исследование данных должно точно определить, какой процесс оставил тот или иной след, откуда пришла команда, либо установить, что это промежуточная точка (т. е. устройство потерпевшего использовалось как передаточное звено в цепочке событий преступного деяния), кто выполнил то или иное действие. Эти данные позволят достоверно установить причинно-следственную связь случившегося, определить механизм действий.

Особое значение такая классификация имеет для исследования XML-файлов, поскольку они описывают структуру основного файла, могут содержать информацию о преднамеренном воздействии на него, наличии скрытого содержимого, следах его возникновения. Кроме того, могут быть установлены значимые причинно-следственные связи работы с содержимым файла.

Следующим важным основанием классификации является доступность сведений для обнаружения их на ЭНИ и последующего исследования. Согласно такой классификации следует выделять следующие группы информации:

1) открытая — доступная без каких-либо сложных способов ее извлечения, находится «на виду»;

2) скрытая — в случаях, когда преступник предпринял какие-то действия, направленные на сокрытие следов своей деятельности (уничтожение, использование стеганографических методов, подделка и т. п.), либо сведения необходимо восстанавливать после удаления или применять специальные программные решения для ее нахождения;

3) зашифрованная — при применении средств криптографической защиты данных, при сложных алгоритмах которых получить семантическое содержание такой информации практически невозможно.

Применительно к файлам формата XML следует отметить возможность сокрытия в структуре файла скрытых потоков данных. Соответственно, при исследовании структуры документа эксперту нужно обращать внимание на все его элементы в отношении возможности их наличия в документе. Например, если при открытии документа рассматриваемым образом в папке Word обнаружива-

ется раздел media, содержащий конкретные изображения, а при открытии файла после его модификации в docx изображений не обнаруживается или их меньше, можно установить факт стеганографии, т. е. передачи или хранения информации (в данном примере графической) с учетом сохранения в тайне факта такого хранения или передачи. В качестве эксперимента в файл было добавлено новое изображение images.jpg, которое после изменения расширения файла на .docx не отобразилось в нем (рис. 2).

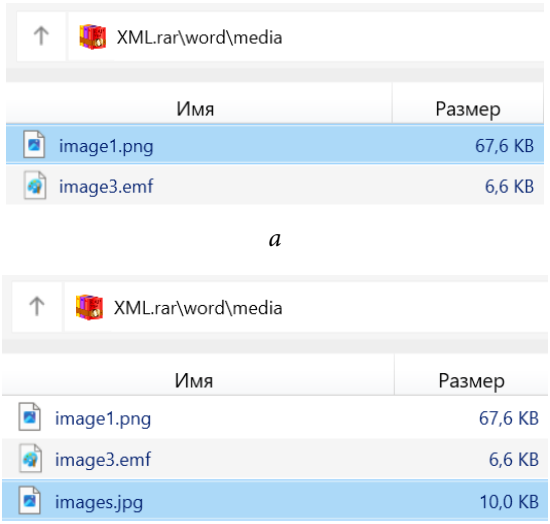

Рис. 2. Окно программы BreeZip, демонстрирующее содержимое папки media:
а — изображения image1.png и image3.emf; *б* — изображения image1.png и image3.emf, а также скрытое изображение images.jpg

Примечательным классификационным основанием также является наличие или отсутствие у исходного обнаруженного цифрового следа упомянутого семантического содержания. Так, в первом случае речь идет о первично интерпретируемой человеком информации, когда для уяснения сущности достаточно прибегнуть к непосредственному восприятию органами чувств (текстовый документ, аудиофайл и др.). Во втором случае идет речь о следах, для интерпретации которых необходима предварительная обработка обнаруженного объекта персональным компьютером или иным электронным носителем (битовый поток данных, управляющие команды операционной системы и т. п.).

В рамках данной классификации представляется верным определять место XML-файлов во второй классификационной группе, поскольку для анализа их содержимого, определяющего структуру документа, необходимо извлечь (получить) файл, например, из файла с расширением .docx.

В целях обнаружения указанной криминалистически значимой информации на электронных носителях информации целесообразно применять экспертные компьютерно-технические знания [9]. Таким образом, наряду с установлением форматов файла, содержимого носителя информации, следов намеренного изменения изображения в общем случае производства компьютерно-технического исследования, в рамках производства экспертизы файлов формата XML могут быть решены следующие задачи по получению криминалистически значимой информации:

– установление факта наличия вредоносного содержимого и его влияния на возникновение последствий, имеющие материальное отображение на носителе информации;

– установление факта изменения файла пользователем;

– выявление исходной структуры документа;

– обнаружение и фиксация привнесенных элементов и др.

Подводя итог, отметим, что цифровые следы (в частности, возникающие при создании, модификации и обработки офисных файлов формата XML) как источник криминалистически значимой информации в уголовном судопроизводстве могут иметь различные классификационные основания. Однако в процессе производства СКТЭ ключевым для эксперта является круг вопросов, которые перед ним поставил следователь (дознаватель) в постановлении о назначении судебной экспертизы. Чтобы ответить на эти вопросы, эксперт и будет выбирать то или иное направление в исследовании и поиске непосредственно следов. Для этого ему необходимо обладать специальными знаниями, включающими специфику строения как XML-файлов, так и документов, в основе которых лежит настоящий формат.

Литература

[1] Манучарян А.К. *Комплексные аспекты компьютерно-технической экспертизы*. Москва, МГТУ им. Н.Э. Баумана, 2020, 28 с.

[2] Вехов В.Б., Зуев С.В., ред. *Цифровая криминалистика*. Москва, Юрайт, 2023, 417 с.

[3] Вехов В.Б. Электронные доказательства: проблемы теории и практики. *Правопорядок: история, теория, практика*, 2016, № 4 (11), с. 46–50.

[4] Зуев С.В., Черкасов В.С. Новые правила изъятия электронных носителей и копирования информации (статья 164.1 УПК РФ): преимущества и недостатки новеллы. *Сибирское юридическое обозрение*, 2019, т. 16, № 2, с. 193–197. https://doi.org/10.19073/2658-7602-2019-16-2-193-19

[5] Усов А.И. *Концептуальные основы судебной компьютерно-технической экспертизы*. Дис. … д-ра юрид. наук. Москва, 2002, 402 с.

[6] *XML для начинающих*. URL: https://inlnk.ru/68z4pm (дата обращения 26.04.2023).

[7] Цимбал В.Н. Понятие и научное значение криминалистически значимой информации. *Общество и право* , 2014, № 4 (50), с. 239–243.
[8] Цимбал В.Н. Виды цифровой криминалистической значимой информации, получаемой в ходе расследования преступлений. *Вестник ВГУ. Серия: Право*, 2021, № 3 (46). https://doi.org/10.17308/vsu.proc.law.2021.3/3554
[9] Лядовская Н.И., Писанова В.А. Особенности специальных знаний при производстве судебной компьютерно-технической экспертизы. *Политехнический молодежный журнал*, 2022, № 10 (75). https://dx.doi.org/10.18698/2541-8009-2022-10-830

Лядовская Нелли Игоревна — студентка кафедры «Безопасность в цифровом мире», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

Научный руководитель — Манучарян Аветис Каренович, старший преподаватель кафедры «Безопасность в цифровом мире», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

Ссылку на эту статью просим оформлять следующим образом:
Лядовская Н.И. Теоретические основы экспертного исследования офисных файлов формата XML. *Политехнический молодежный журнал*, 2023, № 08 (85).
http://dx.doi.org/10.18698/2541-8009-2023-8-925

THEORETICAL FOUNDATIONS OF EXPERT STUDY OF THE XML OFFICE FILES

N.I. Lyadovskaya lyadovskayan@student.bmstu.ru

Bauman Moscow State Technical University, Moscow, Russian Federation

Abstract

As of today, computer technologies are in one way or another introduced into all spheres of the society. Unfortunately, they could be used not only to simplify certain aspects of life, but also in the illegal purposes, when any achievement in computer technology is either used as a means of committing a crime or acts as its subject. To work with such evidence, it is necessary to involve a forensic computer and technical expert, who has the entire range of special knowledge required for this purpose. The article examines the so-called XML office files, which are the popular objects of forensic computer and technical expertise to acquire information available and forensically significant in a criminal case in order to detect, investigate and prevent crimes.

Keywords

Theory of forensic expertise, forensic expertise, forensic computer and technical expertise, electronic data carriers, forensically significant information, office files, file format, XML format

Received 21.05.2023

References

[1] Manucharyan A.K. *Kompleksnye aspekty komp'yuterno-tekhnicheskoy ekspertizy* [Complex aspects of computer-technical expertise]. Moscow, BMSTU Press, 2020, 28 p. (In Russ.).

[2] *Tsifrovaya kriminalistika* [Digital forensics]. Ed. Vekhov V.B., Zuev S.V. Moscow, Yurayt Publ., 2023, 417 p. (In Russ.).

[3] Vekhov V.B. Electronic evidence: problems of theory and practice. *Pravoporyadok: istoriya, teoriya, praktika*, 2016, no. 4 (11), pp. 46–50. (In Russ.).

[4] Zuev S.V., Cherkasov V.S. New Rules on Confiscation of Electronics and Copying Information (Article 1641 of the Code of Criminal Procedure Law): The Advantages and Disadvantages of the Novel. *Siberian Law Review*, 2019, vol. 16, no. 2, pp. 193–197. (In Russ.). https://doi.org/10.19073/2658-7602-2019-16-2-193-19

[5] Usov A.I. *Kontseptual'nye osnovy sudebnoy komp'yuterno-tekhnicheskoy ekspertizy* [Conceptual Foundations of Forensic Computer Forensics]. Dr. Sci. Diss. Moscow, 2002, 402 p. (In Russ.).

[6] *XML dlya nachinayushchikh* [XML for Beginners]. URL: https://inlnk.ru/68z4pm (accessed April 26, 2023).

[7] Tsimbal V.N. Concept and scientifi c importance of forensically important information. *Obshchestvo i pravo*, 2014, no. 4 (50), pp. 239–243. (In Russ.).

[8] Tsimbal V.N. Types of digital forensic significant information obtained during the investigation of crimes. *Vestnik VGU. Seriya: Pravo*, 2021, no. 3 (46). (In Russ.). https://doi.org/10.17308/vsu.proc.law.2021.3/3554

[9] Lyadovskaya N.I., Pisanova V.A. Opportunities for minors to file a statement of claim. *Politekhnicheskiy molodezhnyy zhurnal*, 2022, no. 10 (75). https://dx.doi.org/10.18698/2541-8009-2022-10-830

Lyadovskaya N.I. — Student, Department of Security in the Digital World, Bauman Moscow State Technical University, Moscow, Russian Federation.

Scientific advisor — Manucharyan A.K., Senior Lecturer, Department of Security in the Digital World, Bauman Moscow State Technical University, Moscow, Russian Federation.

Please cite this article in English as:
Lyadovskaya N.I. Theoretical foundations of expert study of the XML office files. *Politekhnicheskiy molodezhnyy zhurnal*, 2023, no. 08 (85). (In Russ.). http://dx.doi.org/10.18698/2541-8009-2023-8-925