ма предотвращения утечки информации) поможет автоматизировать этот процесс.

DLP-система создает защищенный цифровой контур вокруг организации, анализируя всю исходящую, а в ряде случаев и входящую информацию. Контролируемым является информационный поток, состоящий из документов, которые выносятся за пределы защищаемого контура безопасности на внешних носителях, распечатываются на принтере, отправляются по сети и по почте и т. д.

Все DLP-системы по способности блокирования конфиденциальной информации можно разделить на активные и пассивные. Первые умеют блокировать передаваемую информацию, вторые, соответственно, такой способностью не обладают. Первые системы гораздо лучше борются со случайными утечками данных, но при этом способны допустить нечаянную остановку передачи важного документа в организации, вторые же безопасны, но подходят только для борьбы с систематическими утечками.

Как правило, по сетевой архитектуре DLP-системы используют совместно шлюзовые и хостовые компоненты (серверная часть и агенты, работающие на рабочих станциях сотрудников). Агенты только передают всю информацию серверной части. Поступившая от агентов информация анализируется ресурсами сервера.

Таким образом, внедрение DLP-системы в таможенные органы Республики Беларусь позволит отследить и проанализировать основные каналы передачи конфиденциальной информации и выявить факты нарушения информационной безопасности.

УДК 004.832+351.759.6

Ю.Б. Савва

МЕТОДИКА ВЫЯВЛЕНИЯ СРЕДСТВАМИ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ ПРОТИВОПРАВНЫХ ДЕЙСТВИЙ, СОВЕРШАЕМЫХ УЧАСТНИКАМИ ВИРТУАЛЬНЫХ СОЦИАЛЬНЫХ СЕТЕЙ

В виртуальных социальных сетях (ВСС), завоевавших пользователей интернета в последнее десятилетие, нашли свое отражение как положительные, так и негативные черты современного общества. К числу последних относятся: пропаганда терроризма и экстремизма, привлечение к употреблению наркотических и психотропных веществ, вовлечение в секты, понуждение к суициду и другие противоправные действия, направленные на деструктивное воздействие на участников ВСС. В связи с этим перед органами правопорядка встала задача выявления и пресечения противоправного поведения и деструктивной деятельно-

сти в ВСС, эффективно решить которую возможно только посредством использования средств современных информационных технологий по соответствующим методикам.

Для решения данной задачи нами разработана автоматизированная система мониторинга и анализа сообщений участников ВСС, структура которой приведена на рис. 1. Методика выявления противоправных действий участников ВСС средствами информационных технологий основывается на лингвистическом анализе текстов сообщений, как размещаемых ими на «стенах», так и тех сообщений, которыми они обмениваются между собой в личной переписке.

При сканировании ВСС с использованием специально разработанной программы «Краулер» модель ВСС представляется в виде графа $G = (V, E)$, где V – это множество узлов, представляющих участников ВСС, а E – множество дуг, обозначающих отношения между этими участниками.

Сканирование графа ВСС начинается с одного или нескольких узлов (система позволяет вести сканирование параллельно как в одной ВСС, так и в нескольких сетях одновременно). При посещении одного узла осуществляется сбор отправленных с него текстовых сообщений на «стену». Эти сообщения собираются в пакеты, которые подвергаются компьютерному лингвистическому анализу. При этом формируется список соседей этого узла с целью выявления его контактов, что позволяет получать тексты сообщений между участниками ВСС при их прямом общении. Эти сообщения также собираются в пакеты для последующего компьютерного лингвистического анализа.

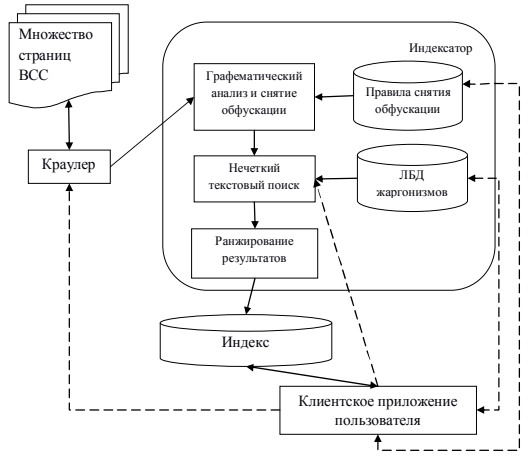

Рис. 1. Структура автоматизированной системы мониторинга и анализа сообщений участников виртуальных социальных сетей

Для сокрытия своих истинных намерений участники ВСС, ведущие противоправную деятельность в сетях, используют как специфическую терминологию – жаргонизмы, так и обфускацию текстов (рис. 2) – намеренное искажение написания слов с целью затруднения получения информации третьими лицами при проведении компьютерного лингвистического анализа. Поскольку обфусцированные тексты сообщений не поддаются простому лингвистическому анализу с помощью поиска ключевых слов, в рассматриваемой автоматизированной системе используются специально разработанные базы данных жаргонизмов и лингвистический процессор.

```
====================== RESTART: C:/Python34/deobf.py ======================
>>> D=Deobf(connection)
>>> D.deobfuscate(tree,"""schanc
так кт о ж мы, на конец? уа - chасть той силы, чтоо вehно hочет зла и вечно сов
ершает бллаг^о
Товарищ, друг, не скупись, купи немножко конопли
Ой m@ma mь|ла rra3y лала voda водафон
я-пришёл-к-тебе-с-приветом
это мыло давно и неправда н е п р а в д а
ч
у
ш
ь
весь коотрый соабка рпишёл дмой анольд коова кроова
жи3нь - б0/\ь и прочие радости страдание исчо""")
шанс так что мы наконец я часть той силы что вечно хочется и вечно совершает бл
аго товарищ друг не скупись купе немножко коноплиой м мыла разу вода вода он я
пришел тебе с приветом это мыло давно и неправда не правда чушь весь который со
бака решил домой анольдкоова крова жизнь больше и прочие радости страдание исче
>>>
```

Рис. 2. Пример вскрытого обфусцированного текста сообщения участника ВСС «ВКонтакте»

Базы данных жаргонизмов аккумулируют в себе соответствующую лексику, предоставляют возможность пополнения словарей жаргонизмами сфер незаконного оборота наркотических средств и психотропных веществ, пропаганды терроризма и экстремизма, вовлечения в секты, понуждения к суициду.

Лингвистический процессор производит:

графематический анализ текстов сообщений и снятие с них обфускации в соответствии с правилами, основанными на использовании скрытой марковской модели и методе N-грамм;

нечеткий текстовый поиск жаргонизмов в текстах сообщений и интерпретацию лингвистического анализа этих сообщений;

ранжирование результатов – распределение текстов сообщений и их авторов по тематике противоправных действий.

Также извлекаются профили авторов сообщений, отнесенных к категории противоправных действий, собирается информация об их активности: время нахождения в сети и совершаемые ими контакты вне зависимости от того, с какого устройства (персональный компьютер, планшет, мобильный телефон) они заходили в сеть (рис. 3).

Выявление устойчивых групп участников ВСС производится на основе построения графа их контактов и формировании истории активности выбранных членов этих групп: дата и время активности, статус и устройство (в том числе его тип), с которого было зафиксировано посещение персональной страницы.

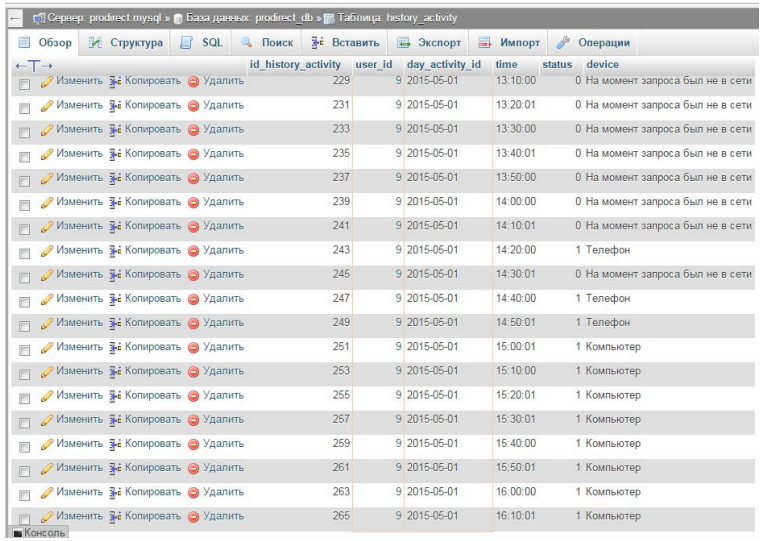

Рис. 3. Скриншот выборки из базы данных активности одного из участников ВСС

Разработанная автоматизированная система мониторинга и анализа сообщений участников ВСС позволяет решать проблему контроля за противоправной деятельностью лиц в этих сетях. Опытная эксплуатация данной системы в ряде уполномоченных органов показала ее эффективность. В настоящее время ведутся работы по следующим направлениям:

идентификация участников сетей, использующих браузер Tor, а также пиринговые сети;

чтение и определение содержания текстов, размещенных участниками сети на фотографиях (в том числе пропаганда ИГИЛ и т. п.).