Анализ данных приложения Snapchat на iOS

О. Ю. Гогина[1]
Ю. В. Литвинов[2]
Н. М. Тимофеев[3]

[1]программная инженерия, математико-механический факультет, СПБГУ
[2]научный руководитель, к. т. н., доцент кафедры системмного программирования
[3]консультант, руководитель отдела разработки ПО, ООО “Белкасофт”

15 ноября 2020 г.

Введение

- Особый интерес у криминалистов вызывает анализ данных приложений мобильных устройств компании Apple: iPhone и iPad
- Приложение Snapchat для iOS недавно начало использовать новый формат хранения данных — TSAF
- Этот формат не исследован, но местами похож на хорошо известный формат хранения настроек на Apple устройствах — Binary Plist

Цель

Цель работы — изучить проприетарный формат хранения данных TSAF и реализовать прототип программы для извлечения данных формата TSAF.

Задачи

1) сделать обзор нового формата данных TSAF
2) реализовать прототип парсера файла TSAF
3) разработать прототип для извлечения сообщений приложения Snapchat из снимка памяти iOS
4) аппробировать реализованный прототип на тестовом смартфоне iPhone 5S

Применимость на практике

- Цифровая криминалистика

Существующие решения

- Spoopy
  - Извлекает из chatConversationStore.plist текстовые сообщения
  - Представляет их в таблице с возможностью фильтрования и автоматического перевода на нужный язык
- The Cellebrite UFED Reader
  - Извлекает из chatConversationStore.plist текстовые сообщения в виде древовидной структуры

Этапы выполнения работы

1. Наладить процедуру снятия полного логического образа файловой системы iOS для получения файлов, в которых сохранены данные приложения Snapchat
2. Исследовать новый формат данных TSAF
3. Изучить алгоритм, используемый в Snapchat для сохранения данных
4. Создать прототип для извлечения текстовых сообщений
5. Реализовать возможность извлечь медиафайлы, стикеры, эмоджи и данные о геолокации из сообщения

Получение файла с данными Snapchat

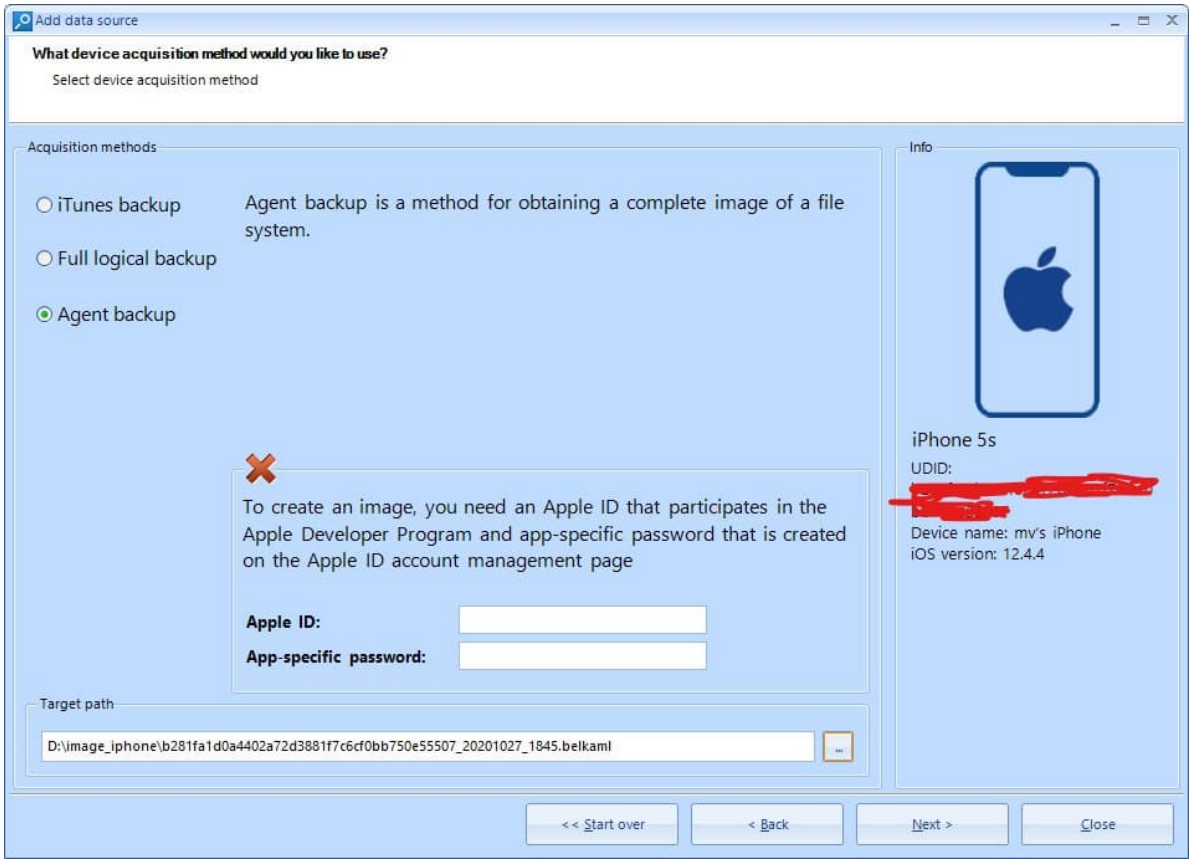


Рис.: Belkasoft Evidence Center

Формат TSAF

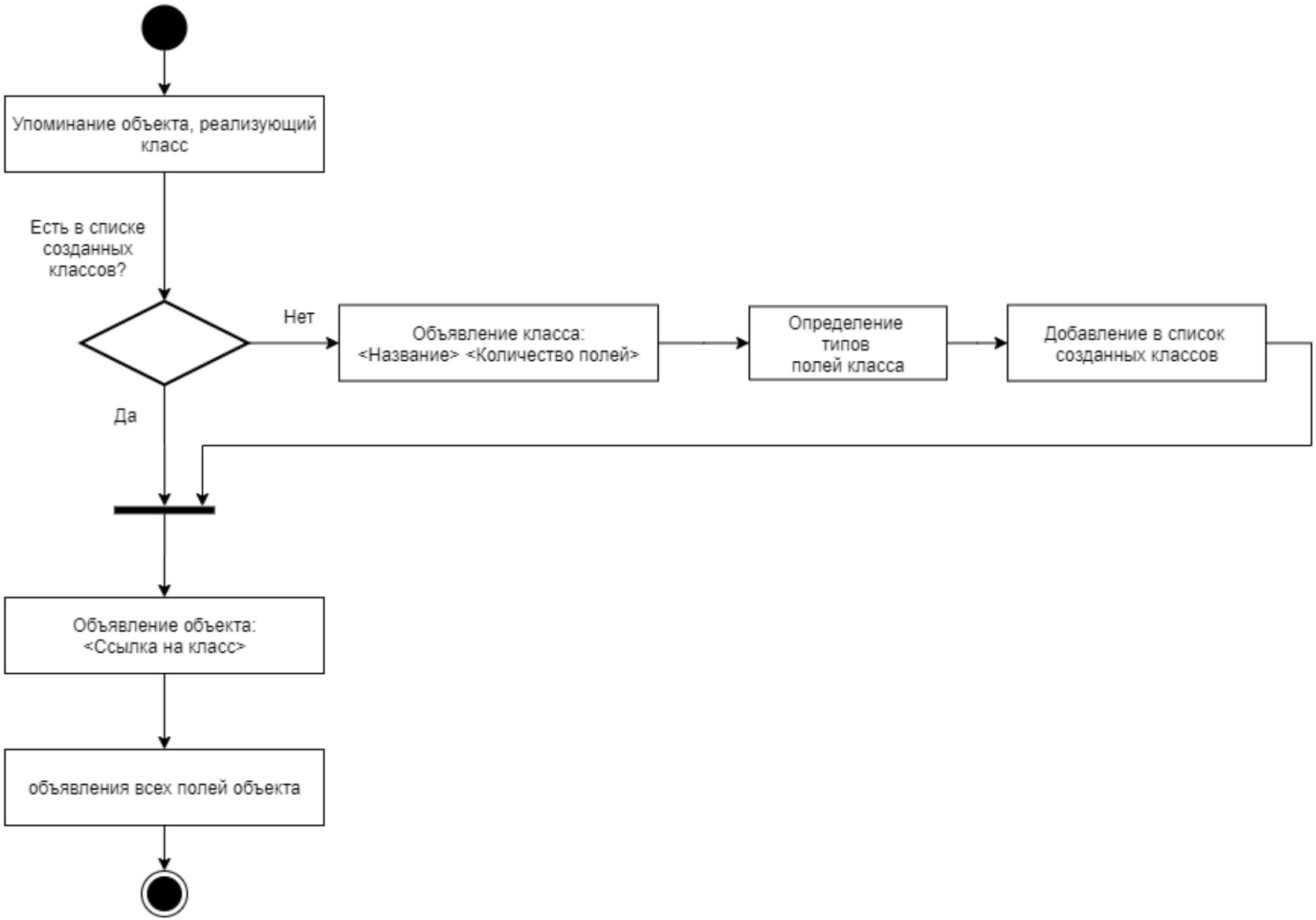
- заголовок «TSAF», за которым следуют данные приложения
- данные — снимок объекта
- объект содержит поля различных типов
- тип данных каждого поля объекта объявляется в определении класса, который реализует объект, либо непосредственно перед значением поля с помощью специальной сигнатуры

Используемые в формате TSAF типы данных

Полями объектов могут являться:

- примитивные типы данных
- указатели
- массивы
- словари
- объекты

Классы


Упоминание объекта, реализующий класс
Есть в списке созданных классов?
Нет
Объявление класса: <Название> <Количество полей>
Определение типов полей класса
Добавление в список созданных классов
Да
Объявление объекта: <Ссылка на класс>
объявления всех полей объекта

Сигнатуры данных TSAF

Как исследовался формат:

1. был реализован базовый парсер данного формата на основе статьи Ian Whiffin
2. в ходе реализации были выявлены типы, не описанные в статье
3. для классификации данных типов использовалось выравнивание полей, соответствующих неизвестному типу, в файле
4. далее точный тип таких полей устанавливался за счет определения информации, которую приложение Snapchat хранит в данных полях

	Сигнатура	Тип данных
Примитивные типы данных	0x0D (истина) 0x0E (ложь)	Логический
	0x08 (начало) 0x00 (конец)	Строка
	0x0F	Int8
	0x10	Int16
	0x04	Int32
	0x12	Int64
	0x13	Single
	0x14 0x16	Double
Указатели на строку	0x05	UInt8
	0x06	UInt16
Указатели на значения	0x02	UInt8
	0x03	UInt16
Типы данных, объявленные в определении класса	0x00	Динамический
	0x01	Логический
	0x04	Int32
	0x05	Double
	0x09	Int64
Остальные типы	0x0A	Массив
	0x09	Словарь
	0x1E	Класс
	0x1F	Объект

Структура файла для приложения Snapchat

Данные о сообщении в Snapchat:

1. информация о приложенном стикере
2. информация о приложенных медиафайлах
3. местоположение
4. имя отправителя
5. время отправки (клиентское)
6. статус сохранения сообщения у клиентов
7. время отправки (серверное)
8. текст сообщения

Использование формата TSAF приложением Snapchat

Интересующие специалистов кибербезопасности данные о сообщении:

	Тип данных	**Содержимое**
1	Объект класса «SCChatSticker»	Информация о стикере, приложенном к сообщению
2	Объект класса «SCChatMediaContent»	Информация о медиафайле, приложенном к сообщению
3	Объект класса «SCChatMessageParcel»	Геолокация, приложенная к сообщению
4	Строка	Имя отправителя
5	Число с плавающей запятой двойной точности	Время отправки, клентское
6	Словарь, сопоставляющий имена пользователей и объекты класса «SCChatMessageSavedState»	Кем было сохранено сообщение
7	Число с плавающей запятой двойной точности	Время отправки, серверное
8	Строка	Текст сообщения

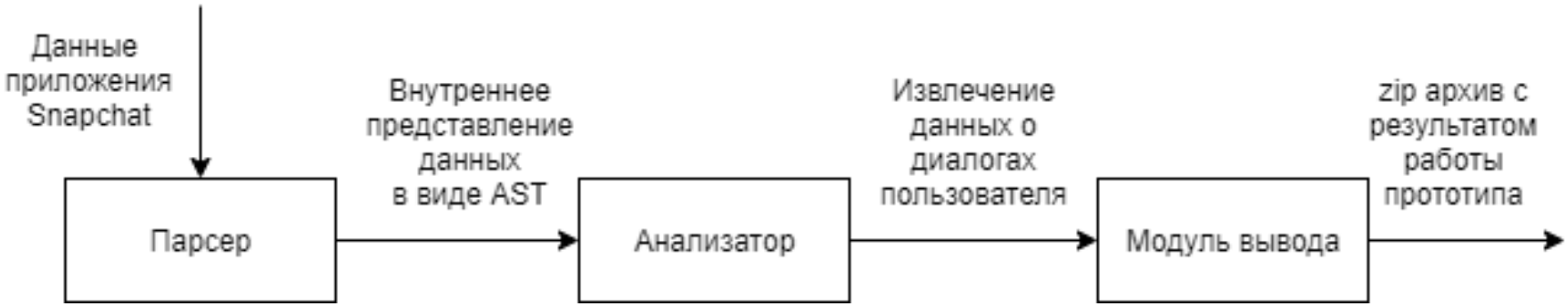
Прототип приложения для извлечения данных

Для извлечения данных приложения Snapchat из файла формата TSAF был реализован прототип на языке C#.

Данный прототип позволяет:

- разобрать файл формата TSAF
- извлечь из него информацию об имени пользователя и его диалогах

Архитектура прототипа



Апробация

Для проверки корректности работы прототипа была проведена апробация с использованием тестового смартфона iPhone 5S:

- из памяти смартфона была извлечена папка приложения Snapchat
- путь к этой папке был передан прототипу
- было подтверждено полное совпадение извлеченной информации и информации, доступной на тестовом устройстве в приложении Snapchat

Результаты

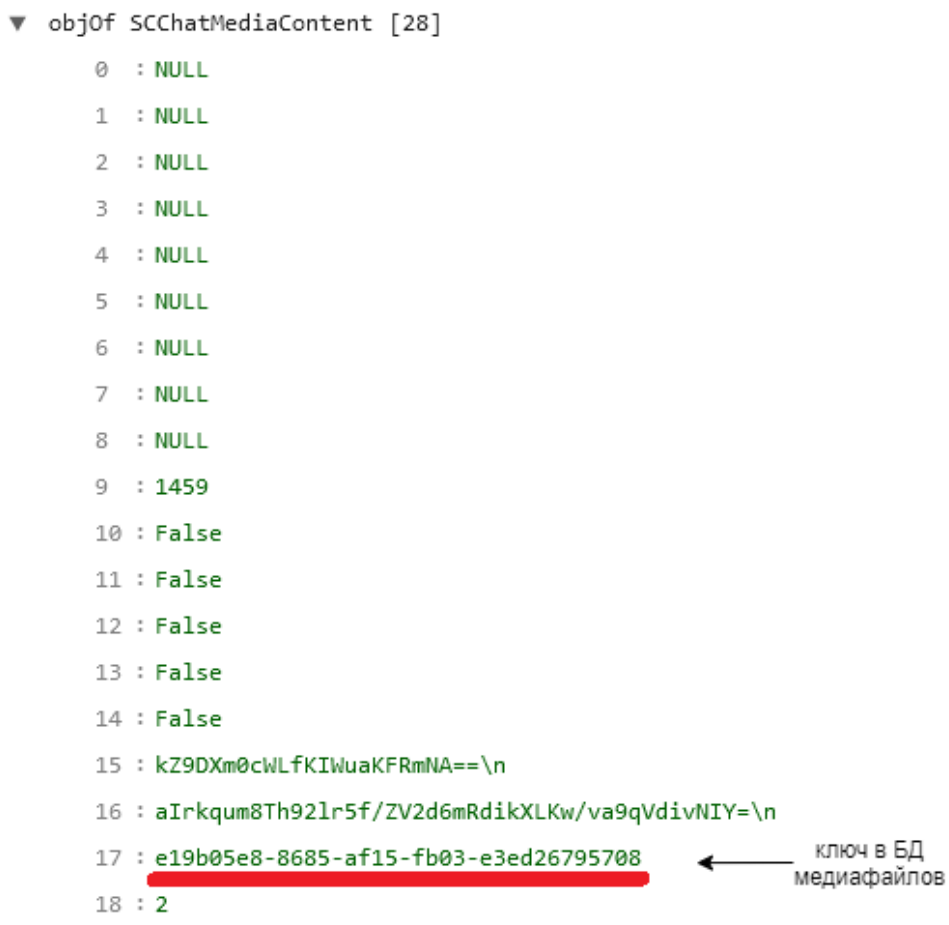
1. исследован новый формат данных TSAF и сделан его обзор
2. реализован прототип парсера файла TSAF
3. разработан прототип для извлечения сообщений приложения Snapchat из снимка памяти iOS
4. реализованный прототип апробирован на тестовом смартфоне iPhone 5S

Приложение

Извлечение медиафайлов

Информация о медиафайлах может храниться:

- в виде объекта
- в виде массива объектов, реализующих класс «SCChatMediaContent»



Извлечение стикеров

Информация о стикере, прикрепленном к сообщению, сохраняетсяв объекте, реализующем класс «SCChatSticker».

Извлечение геопозиции

Информация о геопозиции хранится в объекте, реализующем класс «SCChatMessageParcel» в поле, хранящем данные в необработанном виде.
Данные о геолокации в формате json:


```
{
  "recipient_user_id": "da31b228-7c62-4108-96bc-bc52c0e9e8c5",
  "sender_lat": 59.88172948831681,
  "sender_user_id": "cd3417a8-c57d-42f7-90ed-58053732a427",
  "sender_lng": 30.27482463981O5,
  "user_response": 0
}
```


В полях «sender_lat» и «sender_lng» хранятся координаты отправителя в момент последнего обновления диалога на устройстве, из памяти которого был извлечен исследумый файл.