Разработка эффективного метода фаззинга приложений, работающих со сложными форматами данных

Докладчик: Акользин Виталий Владимирович
Мишечкин Максим Владимирович
Курмангалеев Шамиль Фаимович

{vva1994,mish.max,kursh}@ispras.ru

11.12.2020

ИСП РАН

Критерии эффективности фаззинга

- Способность находить данные, приводящие к аварийному завершению исследуемой программы
- Формирование набора данных, приводящего к наибольшему покрытию кода

ИСП РАН

Задача

- Разработать эффективный метод фаззинга приложений, работающих со сложными форматами данных

ИСП РАН

Вызовы

- Учёт спецификации формата при фаззинге
- Баланс между корректными и некорректными данными
- Масштабирование по вычислительным мощностям
- Приоритизация полей формата для мутаций

Формат описания спецификации

- Полнота описания спецификации:
  - поля разного типа с ограничениями на них;
  - зависимости между полями: TLV, чек-сумма, шифрование.
- Описание формата с любой степенью детализации;
- Простота описания;
- Возможность переиспользовать части описания одного формата при составлении описания других форматов;
- Распространённость формата и поддержка его существующими фаззерами.

ИСП РАН

Современные аналоги ИСП Фаззера

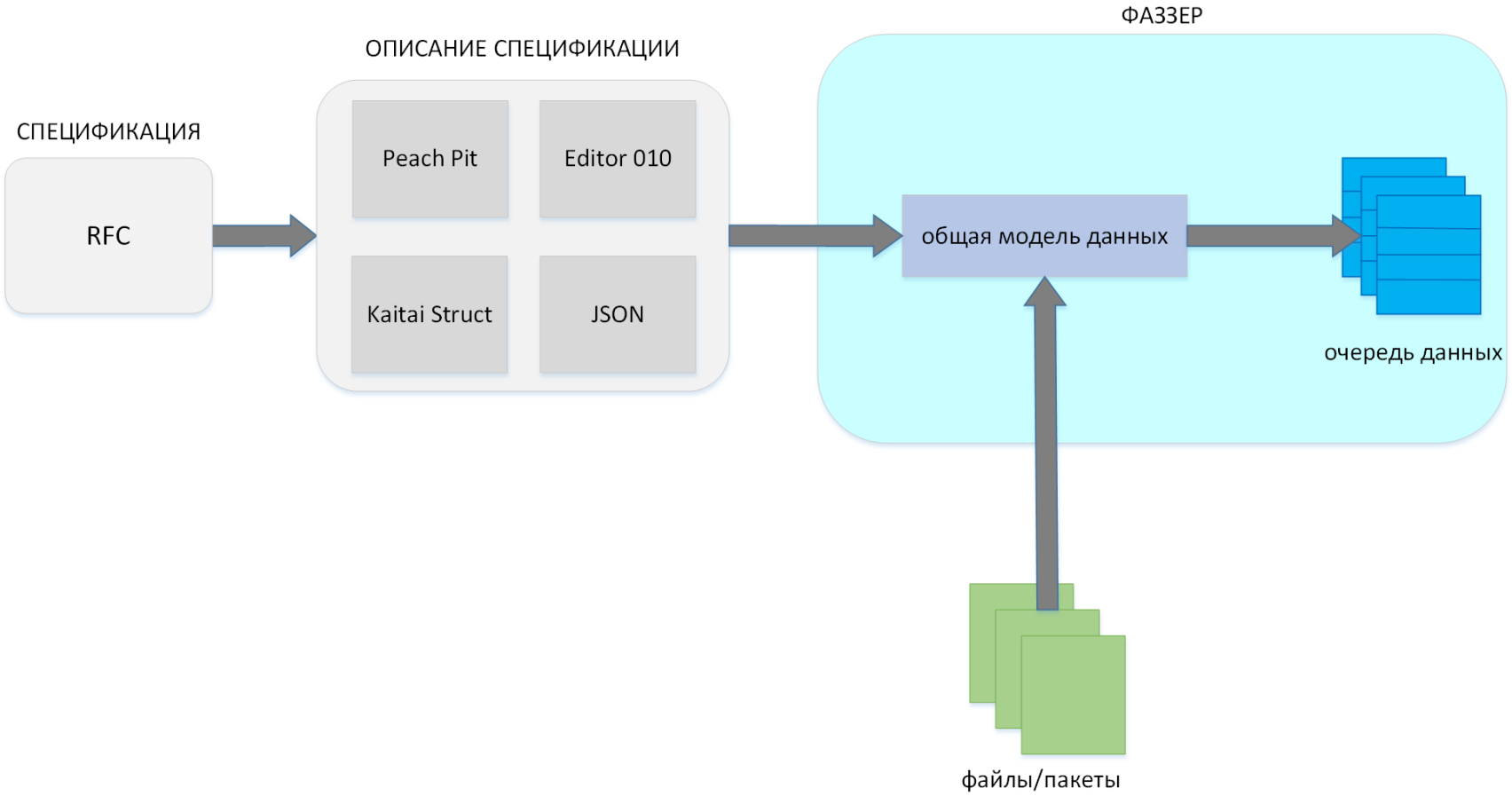
- Peach Fuzzer
- AFLplusplus
- AFLSmart

Предлагаемый метод

- **Общая модель данных** – внутреннее (в фаззере) представление структуры данных на основании спецификации формата файла/пакета.
- **Строгая модель данных** – определённый набор полей. Общая модель данных включает в себя все строгие модели данных.

Суть метода – комбинирование разных типов мутаций, сохранение баланса между корректными и некорректными данными, приоритизация мутаций полей.

Формирование начальной очереди данных



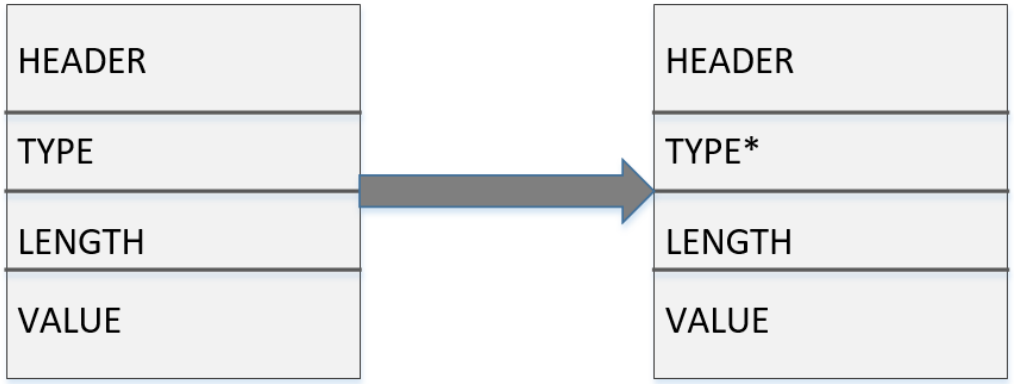
ИСП РАН

Перспективный кандидат на описание спецификации Peach Pit

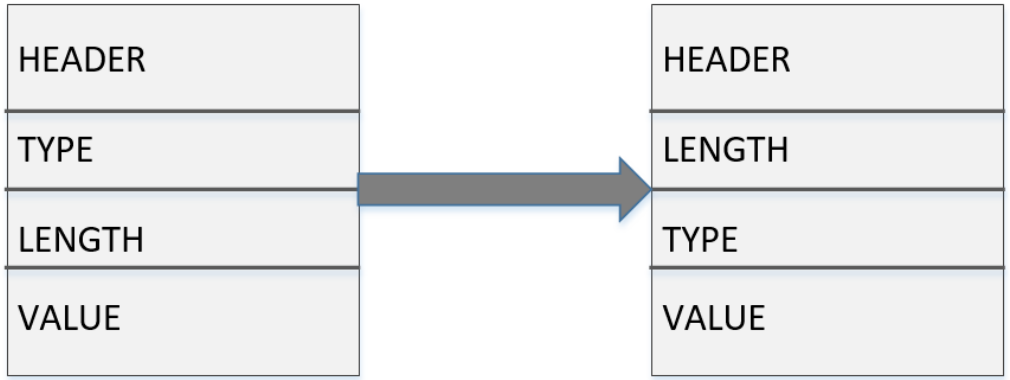
```
<Peach>

    <DataModel name="ClientHelloTemplate">
    <Block name="Headers">
        <Blob name="ContentType" valueType="hex" value="16"/>
        <Blob name="Version" valueType="hex" value="03 01" mutable="false"/>
        <Blob name="LengthHandshakeMessage" valueType="hex" value="40 00"/>
        <Blob name="LengthOfChallenge" valueType="hex" value="01 00"/>
    </Block>

    <Number name="DataLength" size="16" value="0" />
    <Blob name="Data" valueType="hex" value="00" />
    </DataModel>

    <DataModel name="FuzzDataModel" ref="ClientHelloTemplate" >
    <Number name="DataLength" size="16" signed="false" endian="big">
        <Relation type="size" of="Data"/>
    </Number>
    <Blob name="Data" valueType="hex" value="AA" />
    </DataModel>

</Peach>
```

Мутации

- Битовые



- Структурные



Планирование мутаций полей

Данные о мутациях поля:

- **Счётчик мутаций**, применённых к данному полю – mutations_count
- **Счётчик путей** - новых файлов в очереди, полученных при мутациях данного поля – paths_count
- **Приоритет поля** priority = paths_count / mutations_count

Выбор полей для мутаций:

- **Приоритетные поля**: priority >= $priority_{avg,}$ вероятность выбора – 100%
- **Немутированные поля**: mutations_count == 0, вероятность выбора P_{nomut}
- **Неприоритетные поля**: priority < $priority_{avg}$, вероятность выбора $P_{unfav} < P_{nomut}$

Масштабирование

- **Процессы 1го типа** – мутации на уровне полей и структурные мутации (без пересечений границ полей). Условия добавления файла в очередь при синхронизации:
  - прирост покрытия;
  - соответствие спецификации.

- **Процессы 2го типа** – битовые мутации всего файла как целого. Условие добавления файла в очередь при синхронизации – прирост покрытия.

Качественное сравнение фаззеров

	Peach Fuzzer	AFLplusplus	AFLSmart	ИСП Фаззер
Поддержка формата Peach Pit	+	-	+	+
Битовые мутации	+	+	+	+
Структурные мутации	+	-	+	+
Генетические алгоритмы формирования входных данных	-	+	+	+
Масштабируемость	-	+	+	+
Приоритеты полей для мутаций	-	-	-	+

ИСП РАН

Результаты

Приложение/ формат	Fuzzer	Edges	Crashes
wavpack / WAV	AFLplusplus	3425	8
	AFLSmart	3030	0
	ИСП Фаззер	3363	10
jasper / JPEG2000	AFLplusplus	8596	1
	AFLSmart	7143	0
	ИСП Фаззер	7928	0
qpdf / PDF	AFLplusplus	10914	0
	AFLSmart	10203	0
	ИСП Фаззер	10383	0
tcpdump / PCAP	AFLplusplus	16913	0
	AFLSmart	15260	0
	ИСП Фаззер	15374	0

ИСП РАН

Дальнейшая работа

- Поддержка других форматов описания спецификации (наряду с Peach Pit): Editor 010, Kaitai Struct, JSON
- Использование новых планировщиков очереди
- Исследование и разработка новых алгоритмов планирования мутаций полей
- Синхронизация с другими фаззерами

ИСП РАН

www.ispras.ru