ПОЛИТЕХ
Санкт-Петербургский политехнический университет Петра Великого

Высшая Школа Кибербезопасности и Защиты Информации

Метод формирования обучающей выборки для нейросетевой системы выявления киберугроз в промышленных сетях

Исследование выполнено в рамках стипендии Президента РФ для поддержки молодых ученых и аспирантов (СП-443.2019.5)

конференция РусКрипто

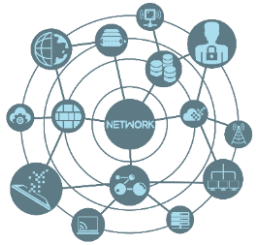
2020

Докладчик: Крундышев В.М.
Ассистент ВШ КИЗИ

Индустрия 4.0

Новые технологии

- **Интернет вещей**
- **Искусственный интеллект**
- **Облачные и сенсорные технологии**
- **Технологии виртуализации**
- **Оперативная аналитика**
- **Нейронные сети**
- **Big Data**



Сетевая инфраструктура

- Самоорганизующиеся беспроводные сети (IoT, WSN, mesh)
- Поддержка большого числа сетевых протоколов

Транспортная инфраструктура

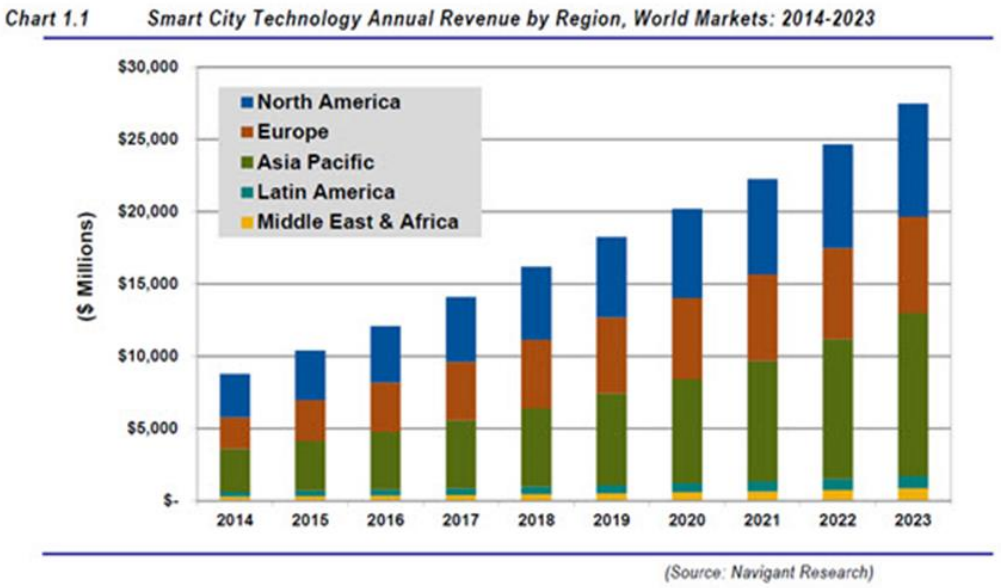
- Взаимодействие компонентов транспортной инфраструктуры (VANET, FANET, MARINET)
- Беспилотный транспорт
- Эффективная логистика (ITS)

Инфраструктура производства

- Обмен данными между компонентами производственных мощностей
- IIoT и Киберфизические системы
- Умное сельское хозяйство, smart grid

Перспективы

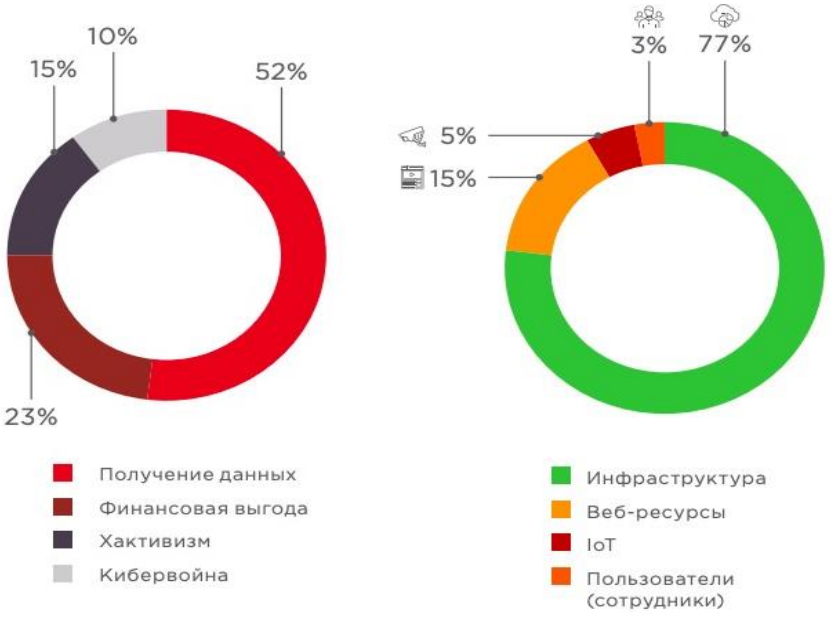
- По оценкам *Всемирного Банка* «Индустрия 4.0» может принести мировой экономике до **$30 трлн.**
- Согласно прогнозу *McKinsey* к 2025 году общий экономический эффект от промышленного интернета вещей составит до **$11 трлн. в год.**
- По оценке аналитиков *Huawei* к 2025 году будет подключено **100 млрд устройств**, используемых во всех сферах бизнеса и жизни.



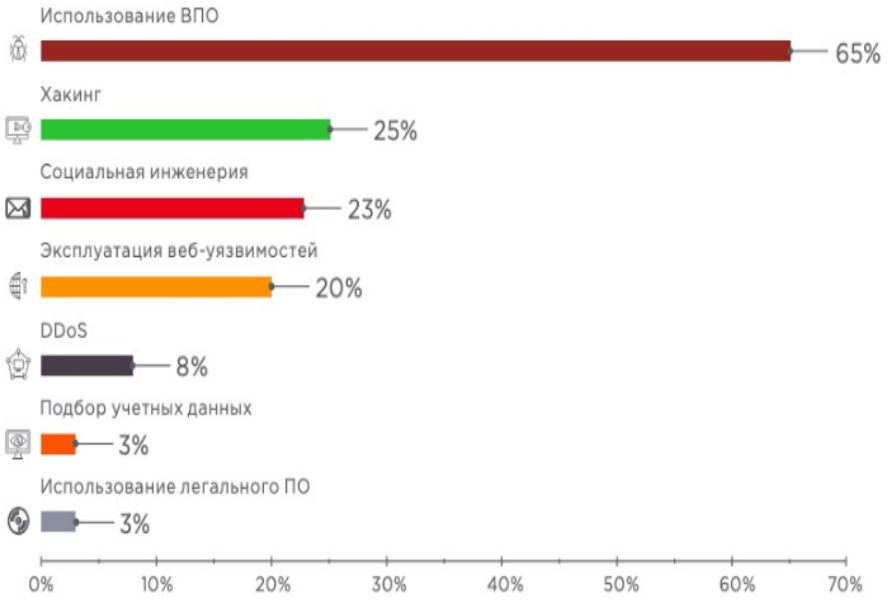
Кибербезопасность

- Согласно мнению экспертов аналитической компании *Frost&Sullivan* кибератаки только в энергетической отрасли обходятся в **$13,2 млн** ежегодно.
- Результаты исследования «*Лаборатории Касперского*» показывают, что инциденты с устройствами интернета вещей входят в тройку угроз с наибольшим финансовым ущербом для компаний, а одной из главных проблем в сфере кибербезопасности цифрового производства до сих пор остается отсутствие единых стандартов.

Мотивы злоумышленников и объекты кибератак

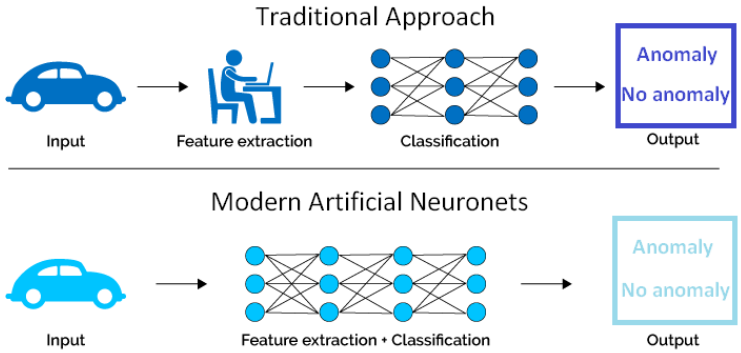


Методы кибератак



НЕЙРОСЕТЕВАЯ СИСТЕМА ОБНАРУЖЕНИЯ ВТОРЖЕНИЙ

- более высокое качество классификации
- автоматическое извлечение закономерностей (самообучение)
- более высокая скорость классификации
- готовность работать с большими данными в качестве ввода
- отсутствие необходимости формализации знаний (заменяется обучением)
- способность обучаться автоматически и в процессе работы
- вероятность обнаружения неизвестных атак
- возможность распараллеливания работы



Архитектуры нейронных сетей:

- Перцептрон
- Сверточные нейронные сети
- Рекуррентные нейронные сети
- LSTM-сети
- Генеративно-Состязательные сети
- …

ПОЛИТЕХ

Основные этапы

При работе с ИНС можно выделить следующие основные этапы:

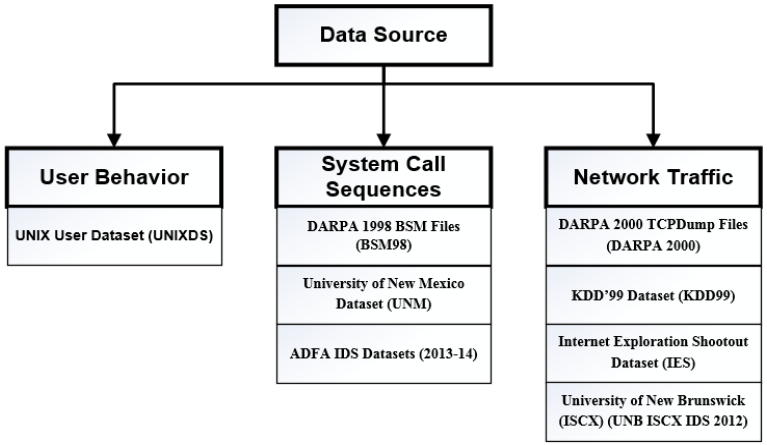
1. Постановка задачи и выбор архитектуры ИНС.
2. Формирование исходной выборки данных.
3. Настройка параметров ИНС и алгоритма ее обучения.
4. Обучение ИНС.
5. Тестирование ИНС.

Формирование исходной выборки и **подготовка** данных является одной из **сложнейших задач** при построении ИНС. Исходные данные преобразуются к виду, в котором их можно подать на входы сети, преобразованные данные называют датасетами. Правильная подготовка датасетов позволяет ИНС обучаться *быстрее*, а на выходе выдавать более *точные* результаты.

Виртуальный полигон кибератак

- *Dos/DDoS*
- *Черная дыра*
- *Серая дыра*
- *Червоточина*

Наборы данных



Дамп сетевого трафика

```
From Id To Id   Tx      Meta
9       8       8.07595  UDP 49153 > 8
9       14      8.07595  UDP 49153 > 8
29      24      8.08126  Wifi CTL_ACK RA:00:00:00:00:00:19
29      28      8.08126  Wifi CTL_ACK RA:00:00:00:00:00:19
14      9       8.08498  Wifi CTL_ACK RA:00:00:00:00:00:0a
14      13      8.08498  Wifi CTL_ACK RA:00:00:00:00:00:0a
14      19      8.08498  Wifi CTL_ACK RA:00:00:00:00:00:0a
14      9       8.08596  UDP 49153 > 8
14      13      8.08596  UDP 49153 > 8
14      19      8.08596  UDP 49153 > 8
29      24      8.092   AODV:RREP D=10.1.2.30 S=10.1.2.30 Seq=0
29      28      8.092   AODV:RREP D=10.1.2.30 S=10.1.2.30 Seq=0
19      14      8.09499  Wifi CTL_ACK RA:00:00:00:00:00:0f
19      18      8.09499  Wifi CTL_ACK RA:00:00:00:00:00:0f
19      24      8.09499  Wifi CTL_ACK RA:00:00:00:00:00:0f
19      14      8.09535  UDP 49153 > 8
19      18      8.09535  UDP 49153 > 8
```

Таблицы маршрутизации

ns-3 NETWORK SIMULATOR

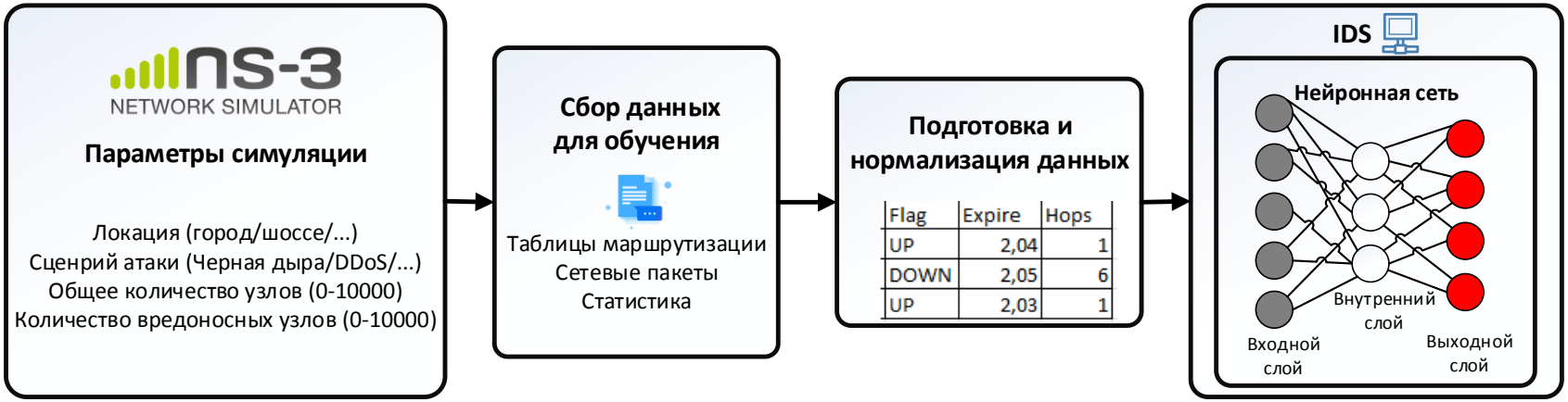
```
Node: X Time: 13s
AODV Routing table
Destination     Gateway         Interface       Flag    Expire          Hops
10.1.2.2        10.1.2.2        10.1.2.1        UP      2.05            1
10.1.2.3        10.1.2.2        10.1.2.1        DOWN    12.05           2
10.1.2.4        10.1.2.2        10.1.2.1        DOWN    12.05           3
10.1.2.6        10.1.2.6        10.1.2.1        UP      2.05            1
10.1.2.21       10.1.2.6        10.1.2.1        DOWN    12.05           4
10.1.2.30       10.1.2.6        10.1.2.1        DOWN    10.05           5
10.1.2.255      10.1.2.255      10.1.2.1        UP      9223372023.85   1
127.0.0.1       127.0.0.1       127.0.0.1       UP      9223372023.85   1

Node: Y Time: 13.00s
AODV Routing table
Destination     Gateway         Interface       Flag    Expire          Hops
10.1.2.1        10.1.2.1        10.1.2.2        UP      2.05            1
10.1.2.3        10.1.2.3        10.1.2.2        UP      2.06            1
10.1.2.4        10.1.2.3        10.1.2.2        DOWN    12.05           2
10.1.2.7        10.1.2.7        10.1.2.2        UP      2.05            1
10.1.2.21       10.1.2.1        10.1.2.2        DOWN    12.05           5
10.1.2.30       10.1.2.1        10.1.2.2        DOWN    10.05           6
10.1.2.255      10.1.2.255      10.1.2.2        UP      9223372023.85   1
127.0.0.1       127.0.0.1       127.0.0.1       UP      9223372023.85   1
```

Статистика

```
    <Flow flowId="1" timeFirstTxPacket="+6000000000.0ns" timeFirstRxPacket="+0.0ns" timeLastTxPacket="+6299519991.0ns"
timeLastRxPacket="+0.0ns" delaySum="+0.0ns" jitterSum="+0.0ns" lastDelay="+0.0ns" txBytes="40680" rxBytes="30120" txPackets="40"
rxPackets="30" lostPackets="10" timesForwarded="8">
```

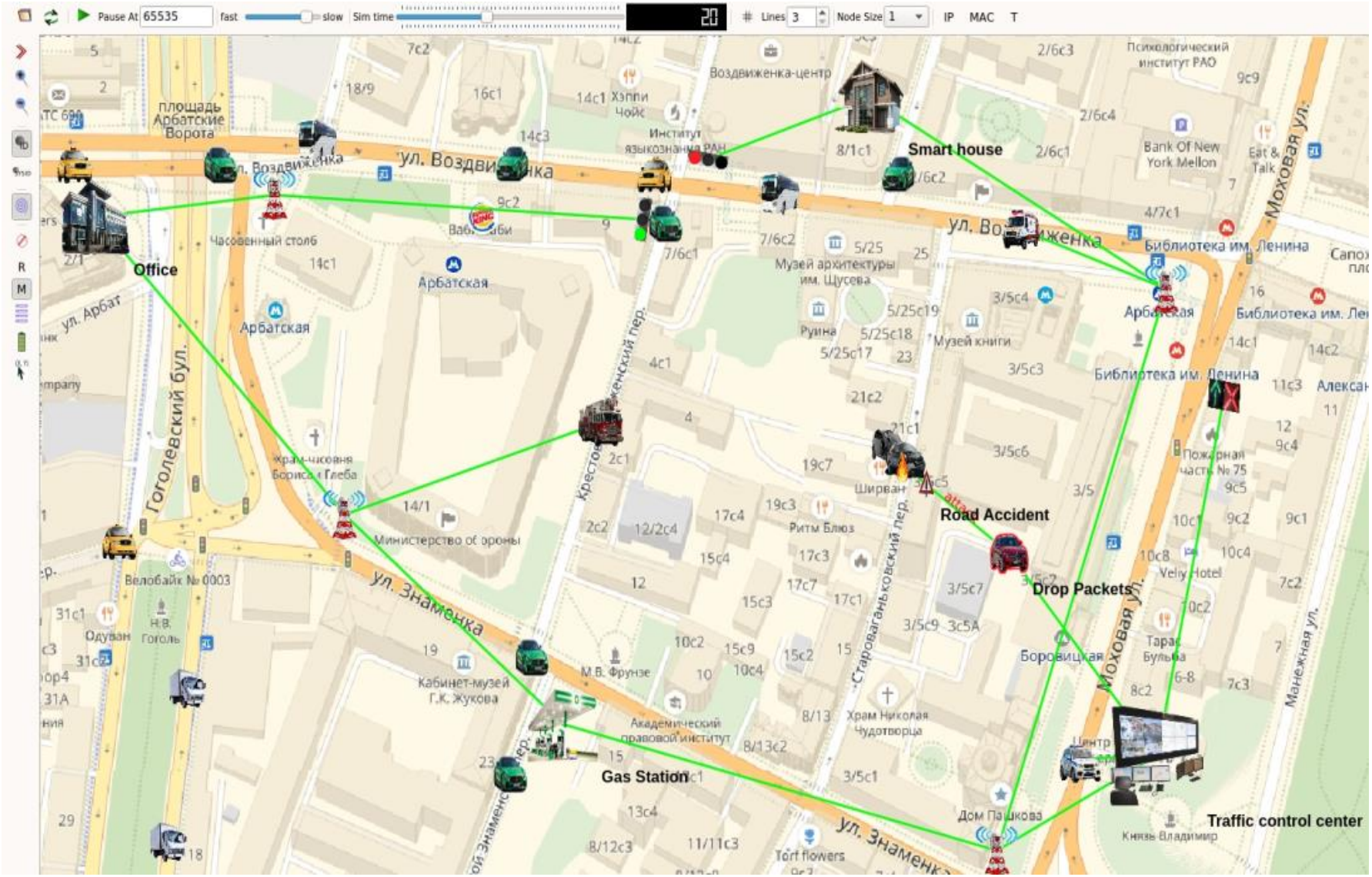
Общая схема



- СетевойТрафик:{ТаблицыМаршрутизации,Пакеты,Статистики}
- ТаблицыМаршрутизации:{*Destination,Gateway,Ineterface,Hops*}
- Пакеты{*RREP,RREQ*,…}
- Статистики{*Throughput,PacketDeliveryRatio,TimeDelay*,…}
- Атаки{*DoS,DDoS,ЧернаяДыра,СераяДыра,Червоточина*,…}
- ПризнакиАтак{*АномальныеВременныеЗадержки,СекртныеКаналыСвязи, ВысокаяДоляПотерянныхПакетов,КорреляцияRREPuRREQ*,…}

1. Проверить наличие признаков атак в сетевом трафике.
2. Если количество признаков атаки превышает пороговое значение, то обнаружена атака.
3. Определить тип атаки и атакующий узел
4. Принять необходимые меры защиты: изолировать злоумышленника и восстановить корректную маршрутизацию.

Визуализация



Заключение

- разработан метод синтетической генерации данных для обучения нейросетевой системы обнаружения вторжений
- представлены характеристики полученных наборов данных
- разработаны сценарии сетевых атак на инфраструктуру промышленных систем
- визуализация работы промышленной системы

…дальнейшие исследования:

- увеличить количество моделируемых кибератак
- дополнять признаки кибератак
- провести масштабное моделирование с использованием ресурсов суперкомпьютера

Спасибо за внимание!