Обнаружение рас..пре..де..лен..ных атак:

симбиоз экспертизы и машинного обучения

Антон Лось
аналитик ИнфоТеКС

БОЛЬШОЙ МОСКОВСКИЙ ТЕХНО infotecs ФЕСТ

Краткий план:

1. Современные вызовы
2. Задача
3. Откуда данные «растут»
4. Архитектура подхода
5. Дефрагментация распределенной атаки
6. Качество результатов
7. Повышение уверенности в результатах
8. Заключение

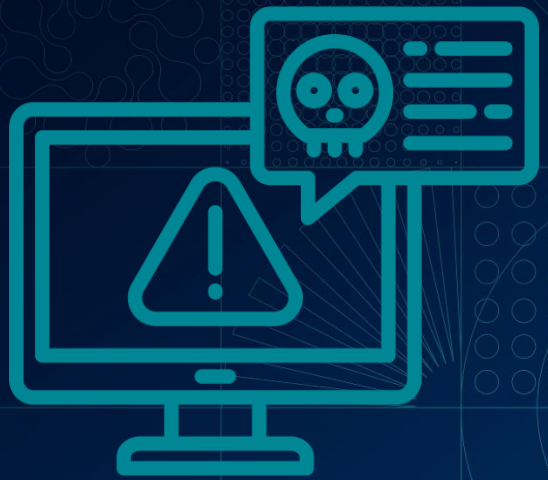
Современные вызовы

Маскировка под легитимную активность

Растяжение во времени

Распределение по узлам защищаемого контура

Задача



Обнаружение вредоносной сетевой активности

Начнем с чистого листа!

ТЕХНО infotecs ФЕСТ

А что у нас есть?

– Записи сетевого трафика активности ВПО в ходе запуска в песочнице:
  - почти 200 тысяч семплов;
  - почти все длительностью до 10 минут.

– Неограниченное количество образцов легитимного трафика.

В ML все начинается с данных

Потоки данных

Потоки данных

Цепочка событий

Последовательность сетевых событий защищаемого узла, распределенная во времени.

Описание цепочки событий

Описание события в составе цепочки:

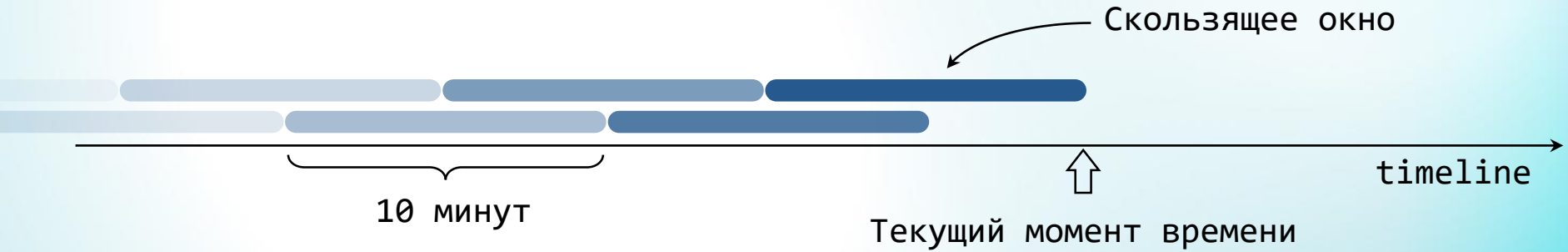
1. Идентификатор правила Snort (sid);
2. Пораженный актив (src или dst);
3. Временной шаг от предыдущего события (δt).

ML модель

Обучаем модель бинарной классификации с архитектурой bidirectional GRU.

Архитектура подхода

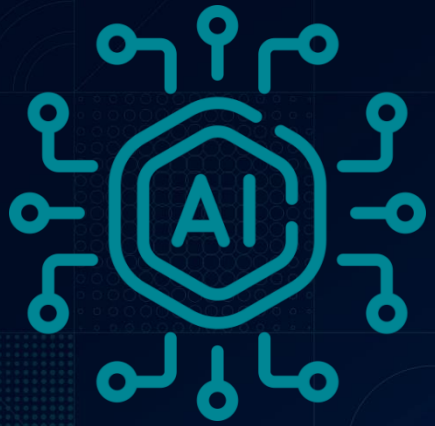
Входные данные



Формирование входных данных:

1. Распределение потока всех сетевых событий по защищаемым IP адресам;
2. Разделяем на фрагменты посредством скользящего окна длительностью 10 минут с шагом в 5 минут.

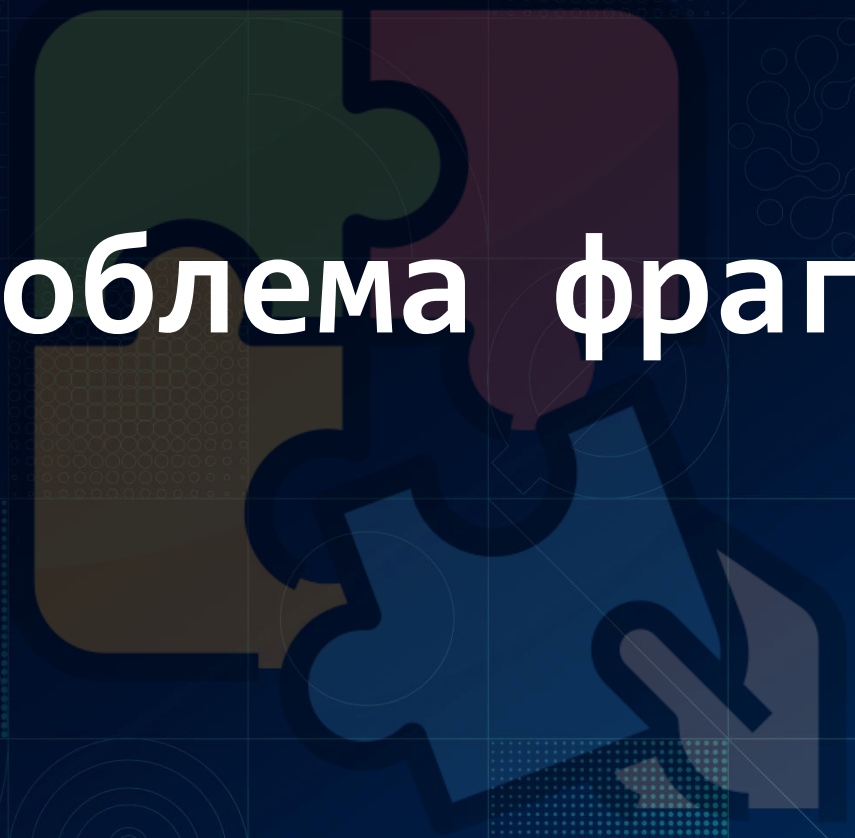
Пайплайн модуля детекции



1. Получение входной цепочки;
2. Формирование эмбеддинга;
3. Обнаружение признаков вредоносной активности;
4. Выделение событий цепочки, характеризующих вредоносную активность.

И это только начало!

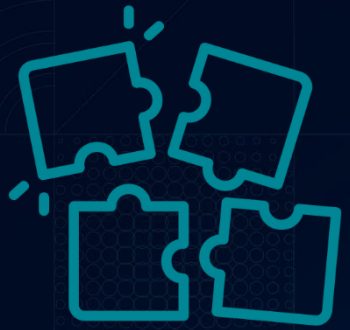
Проблема фрагментации



Фрагментация атаки

- Анализ цепочек разных адресов по отдельности;
- Малый размер скользящего окна;
- Пересечение соседних окон.

Агрегация цепочек



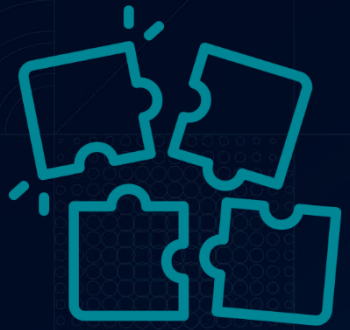
Реальность:

- Вредоносная активность на узле продолжается и по истечение 10 минут одного окна.
- Будете засыпать пользователя инцидентами?

Наш ответ:

- Агрегируем цепочки на данном узле в один инцидент, если активность продолжается в течение часа.

Агрегация цепочек



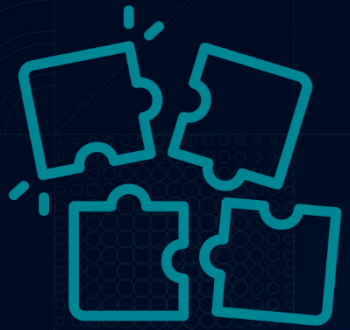
Реальность:

- Пусть ВПО распространилось по узлам защищаемого контура и проявляет активность.
- Будете расследовать одинаковые инциденты на разных узлах?

Наш ответ:

- Агрегируем цепочки, схожие по составу событий.

Агрегация цепочек



Реальность:

- ВПО на разных узлах в одно время реализует разные этапы атаки с разной активностью.
- Это одна атака с общим набором атакующих узлов.

Наш ответ:

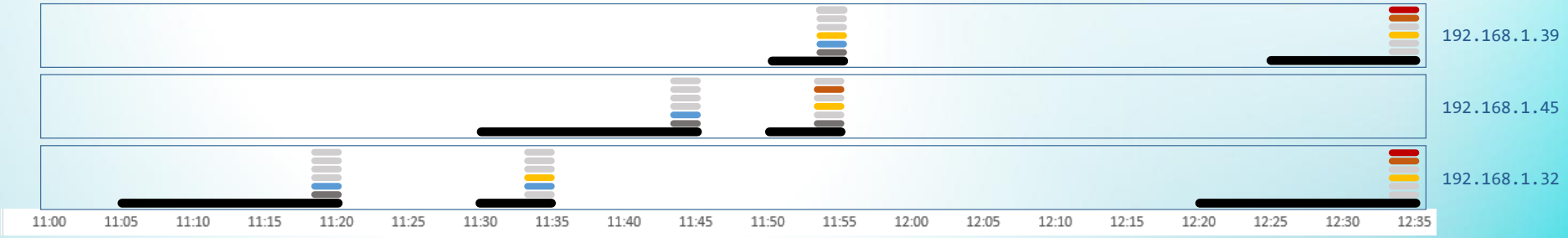
- Агрегируем цепочки, схожие по составу абонентов.

Результат агрегации

- Инциденты охватывают задействованные в атаке разные узлы защищаемого контура;
- Инциденты имеют выраженное развитие со сменой этапов атаки;
- Общее количество инцидентов снижено в 5 раз.

Графическая интерпретация

Дефрагментированный инцидент



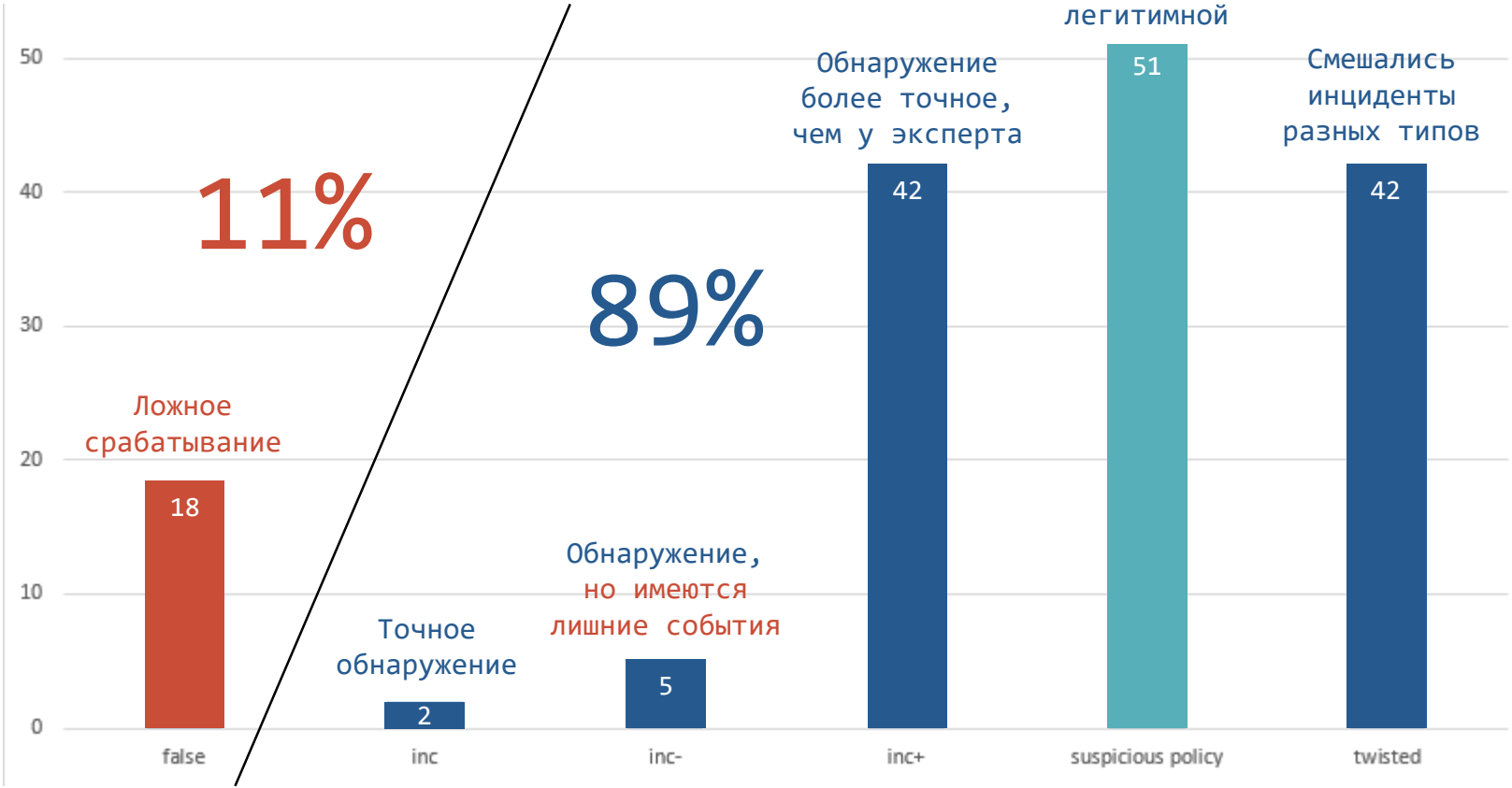
Тактики матрицы атак Mitre

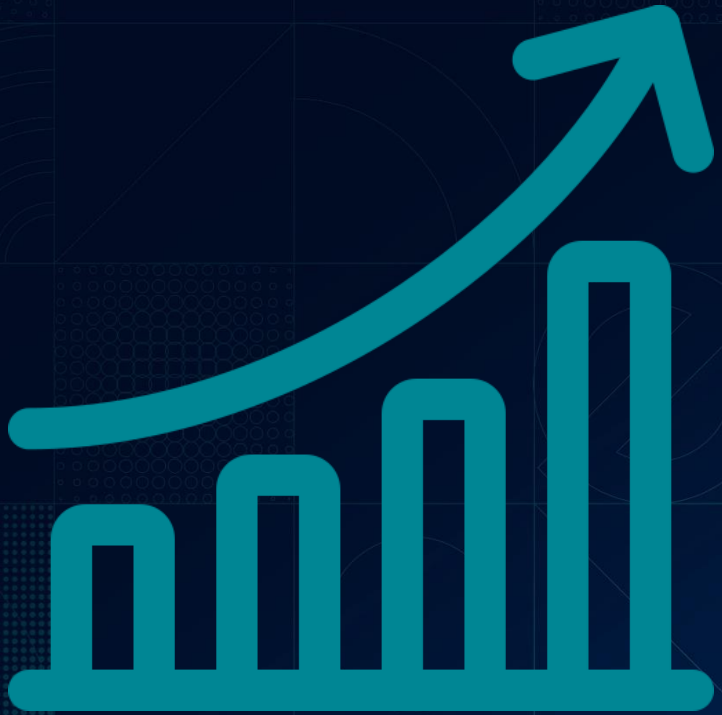
- TA0010 – Exfiltration
- TA0011 – C&C
- TA0002 – Execution
- TA0001 – Initial Access
- TA0043 – Reconnaissance

ТЕХНО infotecs ФЕСТ

Качество обнаружения

Результаты расследования



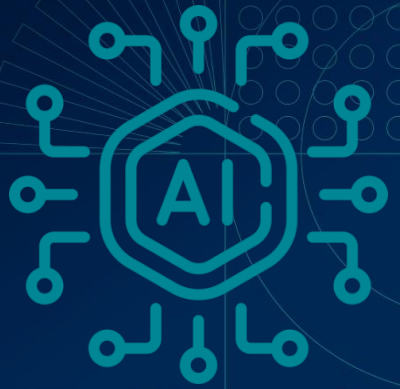


Повышение качества

А какие возможности?

Экспертиза

Машинное обучение



Привлечение экспертизы

- Обновление правил Snort;
- Сбор и запуск в песочнице новых ВПО;
- Разметка инцидентов в системе TIAS.

Применение ML

- Обнаружение аномалий сетевой активности;
- Обнаружение признаков DGA;
- Обнаружение признаков фишинга;
- Обнаружение признаков C&C;
- Обнаружение SSDEEP.

Обнаружение аномалий

- Модель – нейросеть для построения профиля хоста
- Входные данные – объём трафика, количество пакетов
- Продукт – IDS NS

Ключевая особенность модуля:

Обнаружение нетипичных объёмов трафика
для хоста с учётом его периодических характеристик.

Обнаружение DGA

- Модель – сверточная нейронная сеть (CNN)
- Входные данные – доменные имена
- Продукт – IDS NS

Ключевая особенность модуля:

Обнаружение установления связи ВПО с командным центром.

Качество обнаружения DGA

dataset	acc, thr=0.25	acc, thr=0.5	acc, thr=0.75	acc, thr=0.875	acc, thr=0.9375	count
chinad	1	1	1	1	0,996	1000
murofet	1	1	1	1	1	1000
newgoz	1	1	1	1	1	1000
sharkbot	1	1	0,993	0,986	0,972	1000
sisron	1	1	1	1	0,992	1000
qadars	1	0,995	0,995	0,995	0,995	200
verblecon	0,994	0,993	0,99	0,987	0,987	1000
shiotob	0,998	0,992	0,991	0,991	0,99	1000
bumblebee	0,99	0,99	0,99	0,99	0,99	100
monerodownloader	1	0,99	0,99	0,985	0,975	1000
reconyc	0,99	0,99	0,99	0,98	0,97	100
zloader	0,99	0,984	0,977	0,972	0,971	1000
fobber	0,985	0,9783	0,965	0,955	0,9433	600
tufik	0,984	0,976	0,972	0,962	0,96	1000
ranbyus	0,975	0,975	0,975	0,975	0,975	1000
qakbot	0,979	0,97	0,952	0,939	0,931	1000
dircrypt	0,971	0,958	0,944	0,935	0,925	1000
tinba	0,9576	0,9576	0,9476	0,9476	0,9377	401
locky	0,955	0,931	0,912	0,889	0,869	1000
ramdo	0,938	0,916	0,889	0,871	0,853	1000
necurs	0,936	0,913	0,9	0,884	0,872	1000
ramnit	0,96	0,91	0,89	0,87	0,85	100
dnschanger	0,932	0,903	0,885	0,863	0,84	1000
pykspa	0,911	0,876	0,852	0,834	0,808	1000
proslikefan	0,89	0,86	0,83	0,81	0,81	100
unnamed_javascript_dga	0,925	0,86	0,825	0,797	0,783	1000
nymaim	0,8984	0,8594	0,8047	0,7734	0,7656	128
tempedreve	0,887	0,844	0,808	0,78	0,758	1000
dmsniff	0,83	0,78	0,75	0,72	0,68	100
padcrypt	0,832	0,759	0,686	0,632	0,582	1000
orchard	0,7406	0,7406	0,7406	0,7406	0,7359	640
mydoom	0,7879	0,7071	0,6869	0,6667	0,6465	99
simda	0,744	0,653	0,59	0,548	0,522	1000
qsnatch	0,51	0,464	0,421	0,396	0,377	1000
charbot	0,444	0,389	0,362	0,342	0,329	1000
m0yv	0,5469	0,3594	0,2734	0,2188	0,1719	128
pushdo	0,319	0,253	0,214	0,187	0,171	1000
vawtrak	0,16	0,1	0,08	0,06	0,05	100
pitou	0,1107	0,0821	0,0786	0,075	0,0679	280
banjori	0,015	0,008	0,005	0,004	0,003	1000
nymaim2	0,0028	0,0014	0	0	0	704
ngioweb	0,001	0,001	0	0	0	1000
suppobox	0,001	0,001	0,001	0	0	1000
bazarbackdoor	0	0	0	0	0	1000
corebot	0	0	0	0	0	1000
fosniw	0	0	0	0	0	202
gozi	0	0	0	0	0	12
kraken	0	0	0	0	0	1000
symmi	0	0	0	0	0	64
unnamed_downloader	0	0	0	0	0	120

Из протестированных 50 семейств DGA большая часть обнаруживается с высокой вероятностью.

Некоторые семейства практически не обнаруживаются.

Плохие DGA

1. Легитимный домен второго уровня (unnamed_downloader):
   - ddknt.github.io
2. Короткий домен второго уровня (corebot):
   - c4sd1pchctqfo430k0o4mp3.ddns.net
3. Домен составленный из слов (gozi):
   - pertantumfitusu.com,
   - indulgentiarumlicet.com,
   - moriblasphemianegocii.com,
   - ptribueretnossetnonin.com

Обнаружение фишинга

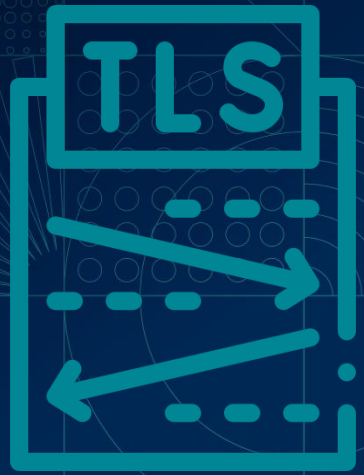
- Модель – определение степени схожести доменов
- Входные данные – доменные имена
- Продукт – IDS NS

Ключевая особенность модуля:

Уверенное обнаружение доменных имён, похожих на защищаемые или легитимные.

Обнаружение C&C

- Модель – анализ длин записей TLS Handshake
- Входные данные – HandShake TLS
- Продукт – IDS NS



Ключевая особенность модуля:

Каждое ПО имеет свои характерные особенности установления связи с сетью по протоколу TLS.

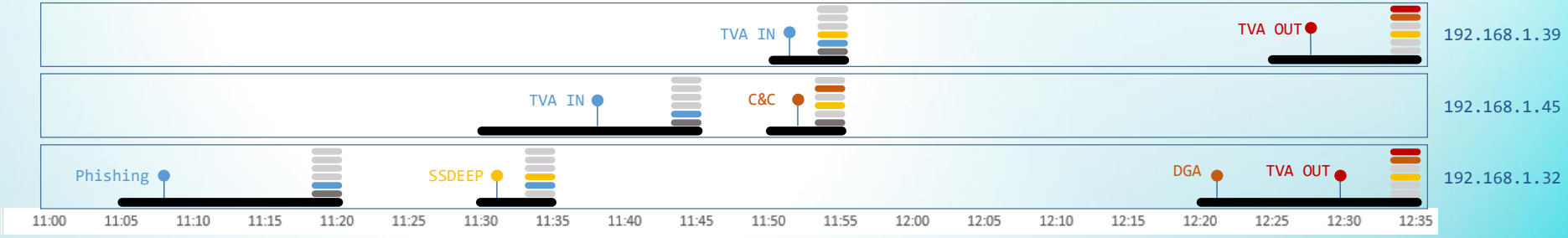
Обнаружение SSDEEP

- Модель – нечеткое хеширование алгоритмом SSDEEP
- Входные данные – сетевые потоки
- Продукт – IDS NS

Ключевая особенность модуля:

Обнаружение нового образца ВПО, полученного минимальным изменением ранее существующего.

События ML-модулей IDS NS



TVA OUT — TA0010 – Exfiltration

DGA, TLS C&C — TA0011 – C&C

SSDEEP — TA0002 – Execution

Phishing, TVA IN — TA0001 – Initial Access

Вместо заключения

ТехноИнфотекс Фест

Антон Лось, аналитик команды ML АО «ИнфоТеКС»

anton.los@infotecs.ru

Подписывайтесь
на наши соцсети,
там много интересного