Система автоматизированного массового тестирования проекта CODA

Комаров Константин, 344 группа

Научный руководитель:

ст. преп. Баклановский М.В.

СПбГУ, Математико-механический факультет, 2015

Система CODA

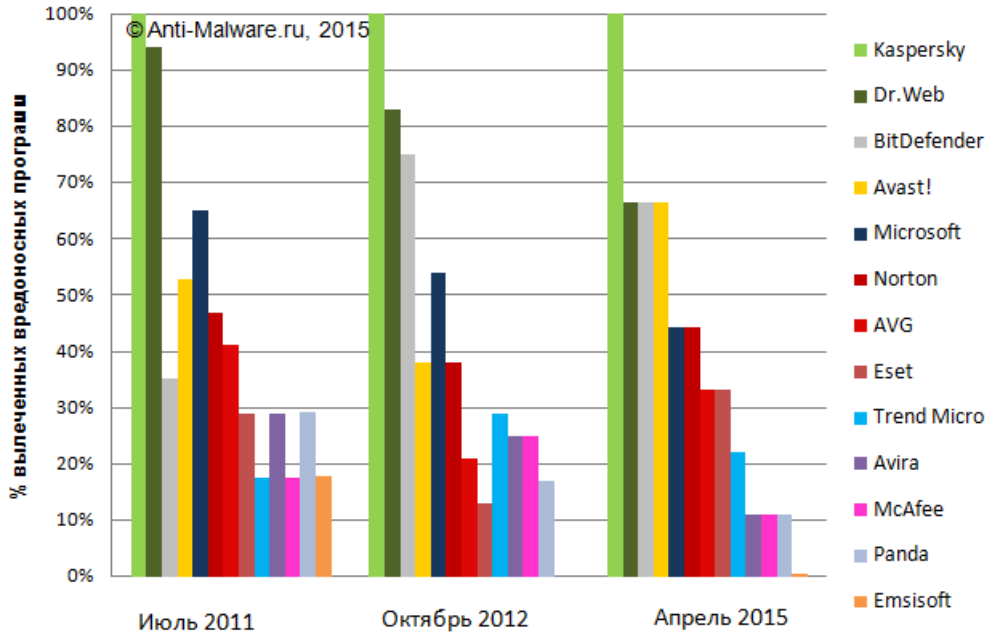
- Система противодействия вредоносным программам
- Основана на поведенческом аномальном обнаружении
- Алгоритмы нуждаются в тестировании

Задачи

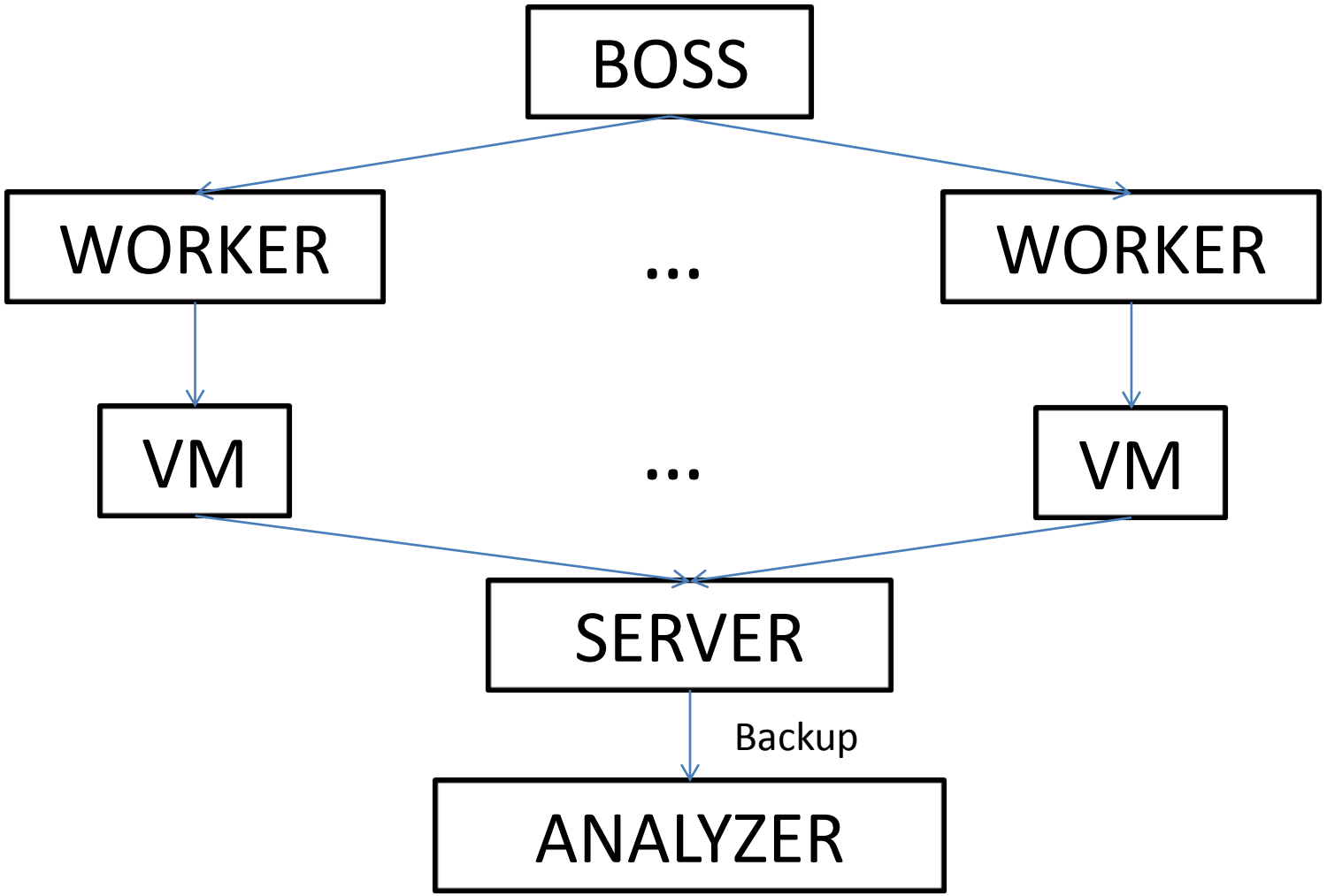
- Разработать платформу для массового запуска и тестирования вредоносных объектов
- Протестировать систему на большом количестве данных
- Оценить эффективность работы алгоритмов
- Обучение нейронной сети

Существующие аналоги

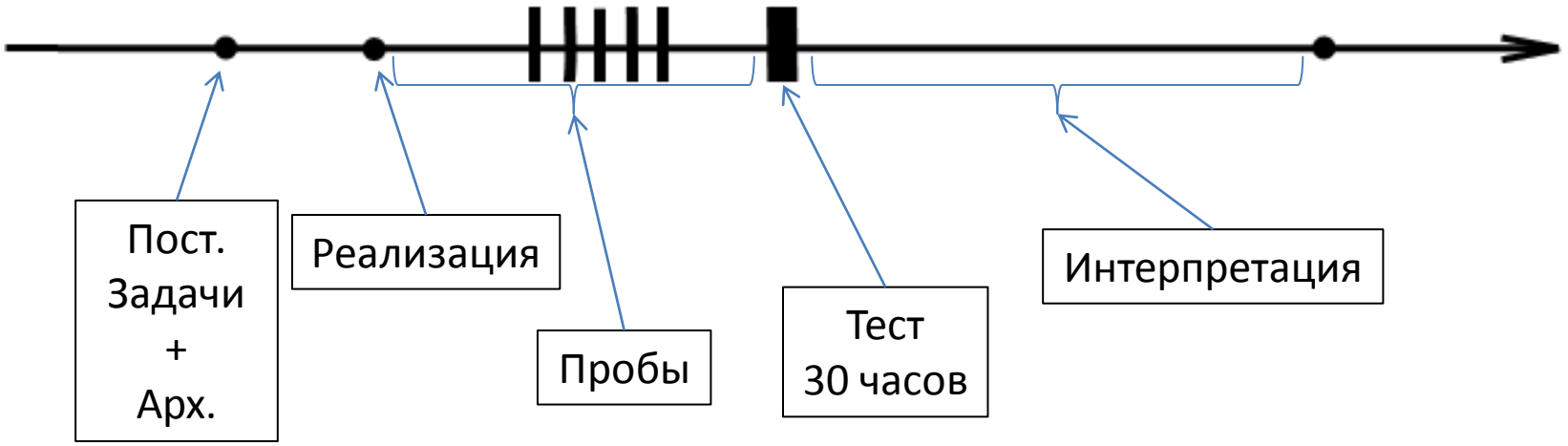
- anti-malware.ru
- av-comparatives.org



Архитектура системы



Timeline



1,8 тыс. вирусов отобрано для запуска
62% «умерли» при запуске
Запустившиеся разрушили сокетные соединения стенда, насоздавали файлов в общей папке, после этого 6% впали в спячку

Подготовка теста

- Онлайн база вредоносных программ
- Windows XP
- Переименование в «virus.exe»
- Обучающий набор из **6410785** вызовов
- Минута на тестирование
- Замерить основные метрики

Отчет

Backdoor.Win32.Hupigon.btlr
9
38 44 26 0.0000 0.0000 0 0 27 1 0 0 588
27 38 20 0.1053 0.0435 4 0 13 1 9 11 664 services.exe
84 72 56 0.0000 0.0000 0 0 11 16 29 49 676 lsass.exe
44 56 21 0.0179 0.0179 1 1 8 2 0 0 744 wuauclt.exe
76 413 274 0.0993 0.0493 41 0 76 10 0 0 832 VBoxService.exe
91 18 17 0.0000 0.0000 0 0 1 4 0 0 876 svchost.exe
58 1547 84 0.0019 0.0018 3 0 350 5 0 0 1060 svchost.exe
116 11 5 0.0000 0.0000 0 0 1 1 0 0 1684 explorer.exe
25 31 9 0.0000 0.0000 0 0 2 2 1 5 1812 VBoxTray.exe
12
69 158 30 0.0253 0.0253 4 0 12 30 0 0 552 perl.exe
38 306 124 0.0327 0.0152 10 1 217 1 316 163 588
79 37505 142 0.0075 0.0074 281 274 6237 40 0 0 612 winlogon.exe
70 354 20 0.0904 0.0443 32 7 36 15 80 20 664 services.exe
84 68 56 0.0000 0.0000 0 0 11 16 32 49 676 lsass.exe
44 59 21 0.0339 0.0339 2 2 7 2 0 0 744 wuauclt.exe
91 3240 7 0.0000 0.0000 0 0 96 18 0 0 796 AgentConsolePack.exe
91 18 17 0.0000 0.0000 0 0 1 4 0 0 876 svchost.exe
81 2639 47 0.0398 0.0304 105 8 189 10 2 5 1060 svchost.exe
81 444 39 0.0158 0.0158 7 6 121 37 0 0 1284 SVCH0ST.exe
79 2130 46 0.0333 0.0333 71 58 185 42 0 0 1328 virus.exe
104 30 14 0.0000 0.0000 0 0 5 3 0 0 1812 VBoxTray.exe

10
38 26 8 0.0000 0.0000 0 0 9 1 0 0 588
110 3867 36 0.0101 0.0098 39 37 1381 7 0 0 612 winlogon.exe
44 33 21 0.0000 0.0000 0 0 5 2 0 0 744 wuauclt.exe
68 670 187 0.0164 0.0087 11 0 275 19 1 5 1060 svchost.exe
68 398 170 0.0151 0.0103 6 4 233 19 459 170 1104
45 252670 39 0.9989 0.9989 252380 252380 96 17 0 0 1284 SVCH0ST.exe
77 798 39 0.0138 0.0138 11 11 86 19 0 0 1328 virus.exe
79 845 47 0.0036 0.0034 3 3 529 40 0 0 1760 IEXPLORE.EXE
25 19 6 0.1053 0.0645 2 2 1 2 0 0 1812 VBoxTray.exe
37 10 6 0.0000 0.0000 0 0 1 1 0 0 1832 ctfmon.exe
9
38 51 26 0.0000 0.0000 0 0 25 1 0 0 588
82 1816 376 0.0022 0.0014 4 1 966 17 1230 53 664 services.exe
84 68 56 0.0000 0.0000 0 0 11 16 29 49 676 lsass.exe
44 55 21 0.0000 0.0000 0 0 9 2 0 0 744 wuauclt.exe
91 29 17 0.0345 0.0303 1 0 1 4 0 0 876 svchost.exe
68 1340 187 0.0037 0.0034 5 0 550 19 0 0 1060 svchost.exe
68 342 170 0.0000 0.0000 0 0 230 19 917 170 1104
103 3400 359 0.0000 0.0000 0 0 691 3 4 9 1684 explorer.exe
104 35 14 0.0000 0.0000 0 0 3 3 0 0 1812 VBoxTray.exe

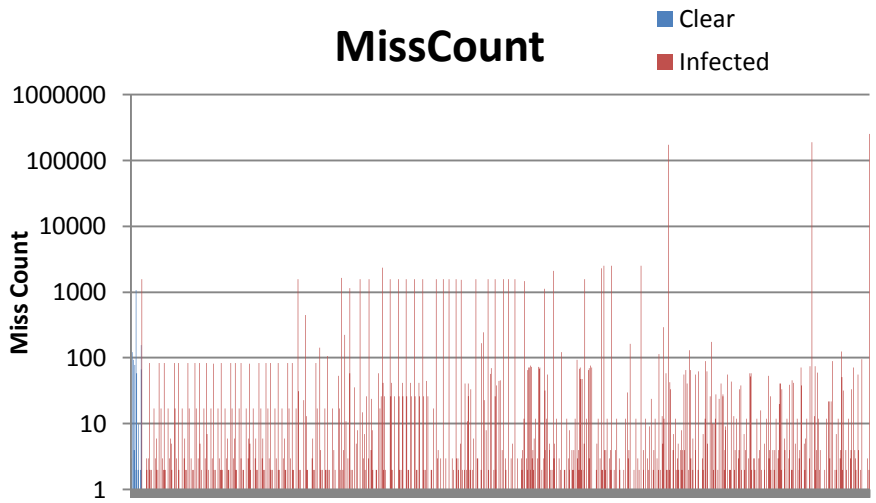
Метрики

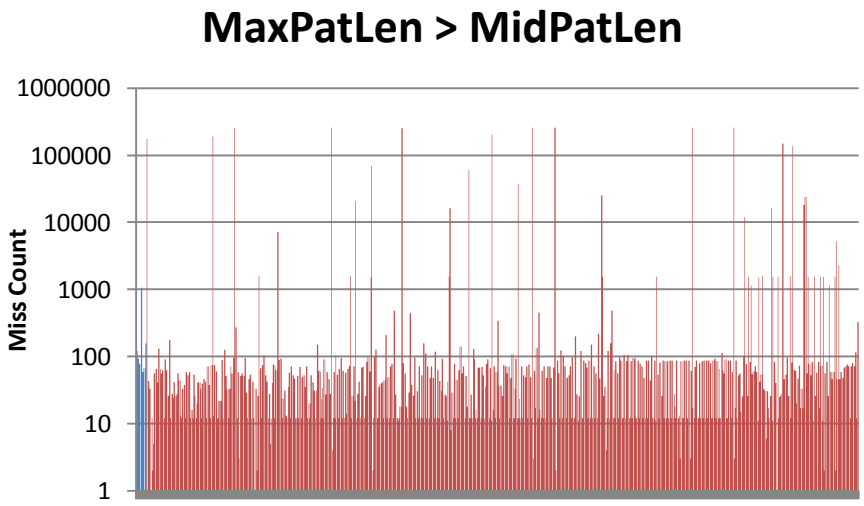
- MissCount
- PatCount
- MaxPatLen
- MidPatLen

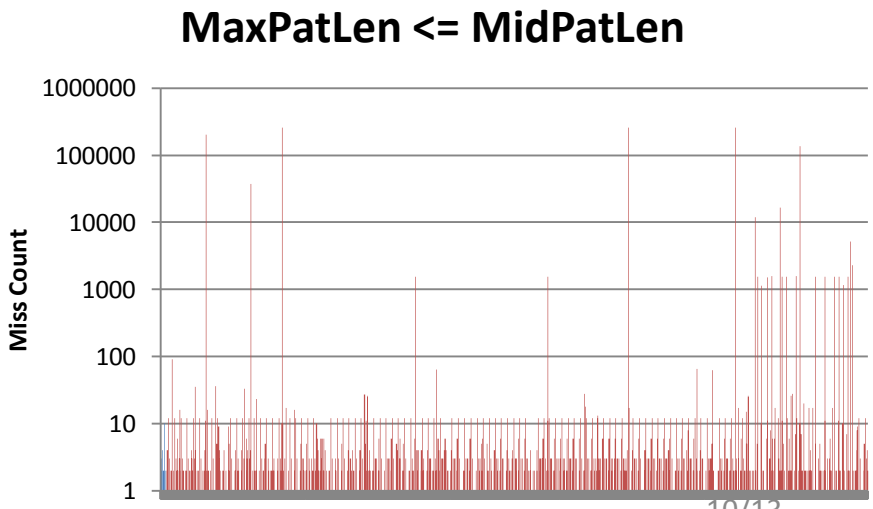
Анализ

1. MissCount > 10
   - В 685 отчетах из 690 (99,27%)
2. MissCount > 10

   && MaxPatLen <= MidPatLen
   - В 568 отчетах из 690 (82,31%)
   - В 230 отчетах выделяются процессы с неизвестными именами (33,33%)
   - Ложных срабатываний при этом нет



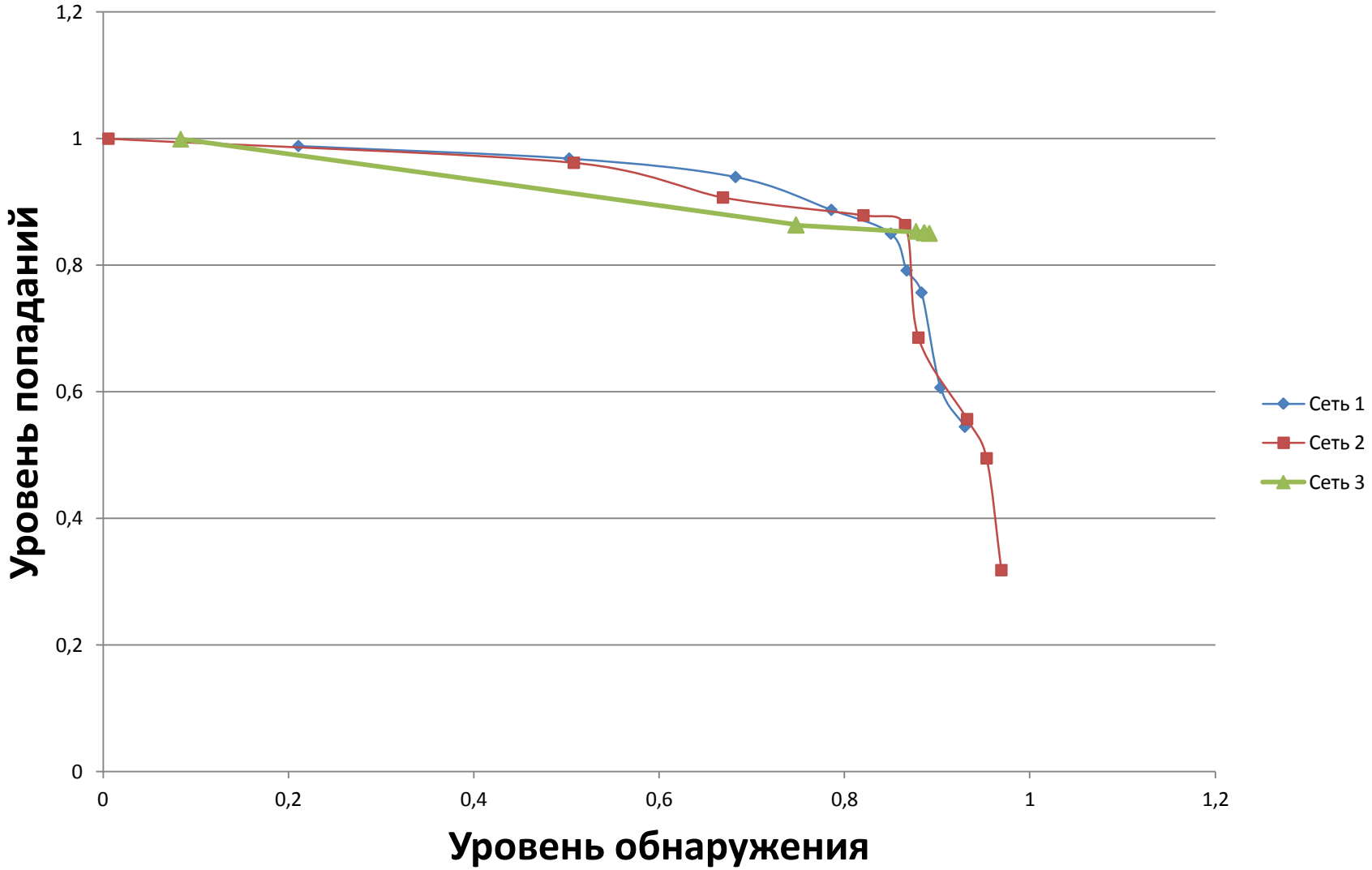




Обучение нейронной сети

- Входные вектора – отчет CODA о процессе
  - 141 легитимных образцов
  - 139 с именем virus.exe
- Тестовые данные
  - 57685 отчетов CODA о процессе
- Из них 1042 точно вирусы

Обучение нейронной сети



Результаты

- Реализована многопоточная система массового тестирования вредоносных объектов
- Проведено тестирование с использованием 1,8 тыс. вредоносных объектов (38% запустились)
- Точность алгоритма не менее 82,31% при 0 ложных срабатываний