1. 004.02
2. Построение графов-сигнатур на основе байтового кода программы для ее идентификации (Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, Санкт-Петербург)
3. К.Г. Плющ
4. И.Е. Кривцова
5. Основные части тезиса

Краткое введение:

Объект исследования – идентификация программ по байтовым сигнатурам, предмет исследования – графы-сигнатуры исполняемых elf-фаилов. Идентификация программ является одной из важных и актуальных проблем в информационной безопасности. Большинство существующих в области идентификации программ научных работ нацелено на обнаружение вредоносных программ. Чаще всего для анализа используются такие методы, как побайтовое сравнение, сравнение контрольной суммы, а также сравнение цифровой подписи. В данной работе предложен способ построения сигнатуры elf-файлов в виде графа для дальнейшей идентификации программы по ее исходному коду.

Цели работы:

1. Построить матрицу смежности специального вида для исполняемого elf-файла.
2. Представить сигнатуру исполняемого elf-файла в виде графа для дальнейшего использования в задаче идентификации программ.

Основные этапы исследования:

1) Обзор существующих методов идентификации исполняемых файлов.
2) Изучение способов построения сигнатур предыдущих исследователей и преобразование их в формат, который можно будет подать на вход алгоритма.
3) Выбор реализации алгоритма для построения графа-сигнатуры (основные параметры, системные требования, возможный формат входных данных для обучения и тестирования, формат получаемого результата).
4) Подбор параметров для дальнейшего сравнения результатов идентификации.

Промежуточные результаты:

Для каждой программы составлено ее текстовое представление в побайтовом виде с помощью hex-редактора. Такое побайтовое представление имеет 16 столбцов и *x* строк в зависимости от веса файла. Разработан алгоритм построения матрицы смежности elf-файла: полученные данные представляются в виде двумерной матрицы, затем в новую матрицу, размерностью 256 x 256, на основе исходной матрицы *X*, записывется количество связей между элементами. Таким образом, из нулевой матрицы *Y* получена матрица весов ориентированного графа, вершинами которого являются байты (от 00 до ff), признаком смежности элементов является соседство двух байтов в hex-виде программы, а весами – количество таких пар встречаемости двухбайтовых последовательностей. Полученная матрица и является сигнатурой программы.

Автор: _____/(К.Г. Плющ)

Научный руководитель: _____/(И.Е. Кривцова)

Декан: _____/(Д.А. Заколдаев)