А. О. Куценко, П. В. Бычков
(*ГГУ им. Ф. Скорины, Гомель*)

ОБНАРУЖЕНИЕ DDOS-АТАК ПРИКЛАДНОГО УРОВНЯ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ

Атака типа «отказ в обслуживании» (DDoS) – это атака на вычислительную систему, выполняемая одновременно с большого числа устройств. Целью данной атаки является доведение вычислительной системы до отказа. DDoS-атака создает условия, при которых обычные пользователи теряют возможность получения доступа к системным ресурсам (серверам).

Для того, чтобы обнаружить DDoS-атаки и иметь возможность защититься, необходимо уметь классифицировать их, так как каждая из разновидностей DDoS-атак имеет свои методы организации. DDoS-атаки прикладного уровня (L7) являются одними из самых трудно определяемых по причине того, что данные атаки не генерируют массивный трафик и работают скрытно, используя трафик, который соответствует требованиям протокола. Примером атаки прикладного уровня может служить отправка неполных http-запросов.

DDoS-атаки прикладного уровня имеют большое количество способов организации, что еще больше препятствует созданию эффективных методов защиты. Традиционные защитные решения в данном случае не будут действовать, так как трафик, создаваемый при атаке прикладного уровня, определяется как легитимный. Помочь может постоянный мониторинг систем распределения ресурсов, а также корреляционный и поведенческий анализ. Важно учитывать, что полностью автоматизировать системы защиты от атак прикладного уровня практически невозможно.

Задача предотвращения DDoS-атак прикладного уровня является достаточно неординарной. Для ее решения нужно научиться определять, какие запросы поступили от настоящих пользователей, а какие от ботов. Машинное обучение может помочь решить проблему наиболее эффективным способом.

Чтобы успешно решить поставленную задачу, алгоритм должен уметь игнорировать аномалии, встречающиеся в исходных данных, а также обучаться на наборе данных, который содержит достаточное количество примеров нормального и вредоносного трафика.

Для создания модели машинного обучения, распознающей DDoS-атаки, был использован набор данных CICIDS 2017 и алгоритмы классификации k ближайших соседей и случайный лес. Оптимизация позволила добиться 96% точности предсказаний для алгоритма k ближайших соседей и 97% точности предсказаний для алгоритма случайного леса.