

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Факультет прикладной математики и информатики
Кафедра технологий программирования

Аннотация к дипломной работе
**«Классификатор интернет-трафика на основе машинного
обучения»**

Фурс Дмитрий Сергеевич

Научный руководитель — доцент кафедры ТП,
канд. физ.-мат. наук Горячкин В. В.

Минск, 2021

РЕФЕРАТ

Дипломная работа, 53 с., 27 рис., 21 источник.

DEEP PACKET INSPECTION, МАШИННОЕ ОБУЧЕНИЕ, PYTHON, КЛАССИФИКАЦИЯ ТРАФИКА, DECISION TREE, RANDOM FOREST, NDPI, EXTREME GRADIENT BOOSTING, WIRESHARK, PYQT, TENSORFLOW.

Объект исследования — методы классификации интернет-трафика, возможность использования машинного обучения для классификации интернет-трафика в DPI системах, а также программная система для классификации интернет-трафика с использованием машинного обучения.

Цели работы — разработка программной системы для классификации интернет-трафика приложений и сервисов на основе поведенческих признаков с использованием машинного обучения.

За время работы были реализованы следующие задачи: изучена литература, посвященная DPI системам, рассмотрены примеры классических DPI сигнатур, методы классификации на основе машинного обучения, проведён анализ поведенческих особенностей интернет-трафика, на основе выбранных статистических признаков подготовлена модель машинного обучения, спроектирована и реализована программная система для классификации интернет-трафика выбранных приложений.

Работа имеет большое практическое значение для решения задач классификации интернет-трафика в современных DPI системах. Результаты дипломной работы нашли применение в проекте компании ООО «Белитсофт интернешнл».

Реализованная программная система позволяет классифицировать интернет-трафика выбранных приложений с использованием методов машинного обучения.

РЭФЕРАТ

Дыпломная работа, 53 ст., 27 мал., 21 крыніца.

DEEP PACKET INSPECTION, МАШЫННАЕ НАВУЧАННЕ, PYTHON, КЛАСІФІКАЦЫЯ ТРАФІКУ, DECISION TREE, RANDOM FOREST, NDPI, EXTREME GRADIENT BOOSTING, WIRESHARK, PYQT, TENSORFLOW.

Аб'ект даследавання — метады класіфікацыі інтэрнэт-трафіку, магчымасць выкарыстання машыннага навучання для класіфікацыі інтэрнэт-трафіку ў DPI сістэмах, а таксама праграмная сістэма для класіфікацыі інтэрнэт-трафіку з выкарыстаннем машыннага навучання.

Мэты працы — распрацоўка праграмнай сістэмы для класіфікацыі інтэрнэт-трафіку прыкладанняў і сэрвісаў на аснове паводніцкіх прыкмет з выкарыстаннем машыннага навучання.

За час працы былі рэалізаваны наступныя задачы: вывучана літаратура, прысвечаная DPI сістэмах, разгледжаны прыклады класічных DPI сігнатур, метады класіфікацыі на аснове машыннага навучання, праведзены аналіз паводніцкіх асаблівасцяў інтэрнэт-трафіку, на аснове выбраных статыстычных прыкмет падрыхтавана мадэль машыннага навучання, спраектаваная і рэалізавана праграмная сістэма для класіфікацыі інтэрнэт-трафіку выбраных прыкладанняў.

Праца мае вялікае практычнае значэнне для вырашэння задач класіфікацыі інтэрнэт-трафіку ў сучасных DPI сістэмах. Вынікі дыпломнай працы знайшлі прымяненне ў праекце кампаніі СТАА «Белітсофт інтэрнэшнл».

Рэалізаваная праграмная сістэма дазваляе класіфікаваць інтэрнэт-трафіку выбраных прыкладанняў з выкарыстаннем метадаў машыннага навучання.

ESSAY

Graduate work, 53 p., 27 illustrations, 21 sources.

DEEP PACKET INSPECTION, MACHINE LEARNING, PYTHON, TRAFFIC CLASSIFICATION, DECISION TREE, RANDOM FOREST, NDPI, EXTREME GRADIENT BOOSTING, WIRESHARK, PYQT, TENSORFLOW.

The object of research — the methods for classifying Internet traffic, the ability to use machine learning to classify Internet traffic in DPI systems, as well as a software system for Internet traffic classification using machine learning.

The Purpose of the work — development of a software system for Internet traffic classification of applications and services based on behavioral features using machine learning.

During the work, the following tasks were implemented: the literature on DPI systems was studied, examples of classical DPI signatures were considered, classification methods based on machine learning were analyzed, behavioral features of Internet traffic were analyzed, a machine learning model was prepared based on the selected statistical features, a software system for classifying Internet traffic of selected applications was designed and implemented.

The work is of great practical importance for solving the problems of Internet traffic classification in modern DPI systems. The results of the thesis were used in the project of Belitsoft International JLLC.

The implemented software system allows you to classify the Internet traffic of selected applications using machine learning methods.