УДК 519.2
DOI:10.25729/ESI.2025.40.4.004

# Оценка параметров стохастических потоков событий методами машинного обучения

**Салимзянова Дарья Дмитриевна, Лисовская Екатерина Юрьевна, Самойлов Сергей Антонович**

Томский государственный университет,
Россия, Томск, *darya2001@inbox.ru*

**Аннотация.** В данной работе рассматривается задача оценки параметров стохастических потоков событий на основе выборочных данных с применением методов машинного обучения. Потоки событий, характеризуемые случайными интервалами между моментами их наступления, широко применяются в моделировании сетевого трафика, телекоммуникаций, вычислительных систем и в теории массового обслуживания. Точная оценка параметров таких потоков имеет ключевое значение для последующего анализа, прогнозирования и управления нагрузкой в системах с неопределённой входной информацией. В качестве исходных данных для обучения моделей были использованы моменты наступления событий в двух типах потоков: пуассоновский поток (интервалы между событиями подчиняются экспоненциальному распределению), рекуррентный поток, (интервалы между событиями следуют одной из двенадцати функций распределения вероятностей: гамма, гиперэкспоненциальное, логнормальное, равномерное, обратное гамма, распределение Вейбулла, Парето, Леви, Фишера, Фреше, Ломакса и Бура XII). Выбор этих распределений обусловлен их разнообразными статистическими свойствами (наличие/отсутствие моментов, асимметрия, тяжёлые хвосты), что позволяет охватить широкий спектр применимых сценариев. Для решения задачи оценки параметров были использованы полносвязные нейронные сети и алгоритм градиентного бустинга в реализации CatBoost. Входными данными для моделей были выбраны интервалы между моментами наступления событий и числовые характеристики этих интервалов: математическое ожидание, среднеквадратичное отклонение, дисперсия, коэффициент вариации, квантили различных уровней. Для оценки качества моделей применялись классические метрики машинного обучения: $MAE$, $RMSE$, $R^2$. Также в рамках исследования была проведена оценка важности признаков, участвующих в обучении моделей. Для этого использовались встроенные механизмы интерпретации градиентного бустинга, позволяющие количественно определить вклад каждого признака в оценку параметров.

**Ключевые слова:** идентификация трафика, сетевой трафик, оценка параметров, градиентный бустинг

**Цитирование:** Салимзянова Д.Д. Оценка параметров стохастических потоков событий методами машинного обучения / Д.Д. Салимзянова, Е.Ю. Лисовская, С.А. Самойлов // Информационные и математические технологии в науке и управлении, 2025. – № 4(40). – С.38-51. – DOI:10.25729/ESI.2025.40.4.004.

**Введение.** Современные информационные и телекоммуникационные сети сталкиваются с увеличением разнообразия интернет-трафика, что обусловлено широким распространением цифровых сервисов, мультимедийного контента, а также многообразием аппаратного и программного обеспечения. Это усложняет задачу обеспечения надежности и высокой производительности передачи данных. Интернет-трафик играет ключевую роль в функционировании информационных сетей, поскольку его анализ позволяет выявлять поведенческие характеристики пользователей и оценивать параметры работы сети. Правильная оценка параметров сетевого трафика является важным инструментом для решения множества задач: моделирования нагрузки на каналы связи, оптимального распределения сетевых ресурсов, маршрутизации данных и предотвращения перегрузок. В связи с этим возрастает актуальность автоматизированных методов идентификации интернет-трафика, направленных на улучшение качества сетевых услуг.

При изучении различных статистических исследований реальных телекоммуникационных потоков было установлено, что сетевой трафик обладает следующими важными свойствами: фрактальность, самоподобие, импульсность и долговременная зависимость трафика, исходя из которых, следует, что широко применяемые

в настоящее время методы моделирования, основанные на использовании пуассоновских потоков, не дают полной и точной картины происходящего в сети. Задача исследования состоит в том, чтобы разработать программный продукт, который в режиме реального времени будет выбирать оптимальную модель системы массового обслуживания процесса на основе данных о времени поступления пакетов и оценивать параметры этой системы. Такой программный комплекс даст возможность эффективно управлять обработкой сетевого трафика, а также не упрощать математические модели при изучении новых технологий.

Свойство импульсности характеризуется наличием интервалов, в которых интенсивность трафика довольно высока, и интервалов с низкой/нулевой интенсивностью. Большинство классических работ по теории массового обслуживания и проектированию сетей связи основаны на предположении, что процесс поступления пакетов является пуассоновским [1-3]. Детальные исследования интернет-трафика показали, что процесс поступления пакетов не пуассоновский [4], а импульсный [5-7]. Длительности интервалов между двумя последовательными моментами поступления пакетов не являются независимыми и экспоненциально распределенными, а пакеты приходят пачками, поэтому интенсивность трафика колеблется во времени.

Интернет-трафик нелегко охарактеризовать из-за его неоднородной природы. Исследования различных типов трафика выявили, что сетевой трафик обладает свойствами самоподобия (фрактальности) [8], это означает, что в нем присутствуют пачки пакетов, наблюдаемые в различных временных интервалах. Также самоподобный трафик обладает свойством долговременной зависимости (зависимости между событиями через достаточно большие промежутки времени) [9]. Самоподобный трафик иногда можно описать с помощью рекуррентных потоков с использованием распределений с «тяжелыми хвостами», таких, как Парето, Вейбулла, логнормальное, гиперэкспоненциальное распределения со специальным выбором значений параметров для описания времени между поступлениями [10-12]. Использование рекуррентных потоков более точно описывает поведение сетевого трафика (локальный и глобальные сети) [9, 13], так как позволяет фиксировать зависимость между событиями на большом расстоянии и самоподобие трафика. В некоторых случаях распределения с «тяжелыми хвостами» имеют бесконечные моменты, что не позволяет использовать метод моментов для оценки параметров.

Методы идентификации сетевого трафика можно разделить на три основные категории: методы, основанные на портах; методы, использующие глубокую инспекцию пакетов; методы, основанные на машинном обучении.

Методы, основанные на портах и IP-адресах, являются самыми простыми и быстрыми, но их точность не превышает 70% [14, 15]. Современные приложения используют динамические номера портов, что снижает точность этого метода [16]. Методы, основанные на глубокой инспекции пакетов, позволяют получить точную классификацию [17-19], но требуют высокоскоростного оборудования, сложного программного обеспечения и больших вычислительных ресурсов. При шифровании приложений доступна только общая информация о пакете.

В настоящее время популярность приобретают методы машинного обучения и глубокого обучения для классификации сетевого трафика [20-26]. В большинстве работ по идентификации интернет-трафика используются наборы данных, содержащие характеристики передаваемых пакетов, и не стоит задача определения оценки интенсивности входящего потока согласно аппроксимации какой-либо из математических моделей для последующего использования.

**1. Описание исследуемых данных и используемых архитектур.** В качестве исследуемых данных были использованы выборки, содержащие моменты наступления

событий в потоке $t_i$, полученные путем моделирования [27]. В работе исследовались стационарный пуассоновский поток и рекуррентный поток с 12 функциями распределения вероятностей длин интервалов между моментами наступления событий: гамма, гиперэкспоненциальное, логнормальное, равномерное, обратное гамма, Вейбулла, Парето, Леви, Фишера, Фреше, Ломакса и Бура XII. Выборки содержат информацию о количестве моментов наступления событий, а также о параметрах, с которыми они были сгенерированы (таблицы 1-2). Количество выборок для всех потоков равно 10000, количество интервалов в выборке – 10000.

**Таблица 1.** Значения параметров пуассоновского потока

| Параметр | Значение |
|---|---|
| Значение интенсивности, $\lambda$ | $\lambda = unif(0; 20000)$ |

**Таблица 2.** Значения параметров распределений длин интервалов рекуррентного потока

| Распределение | Значение |
|---|---|
| Гамма | $\alpha = unif(0{,}5; 5)$ |
| | $\beta = unif(0{,}1; 10)$ |
| Логнормальное | $\mu = unif(-2; 2)$ |
| | $\sigma = unif(0{,}1; 1{,}5$ |
| Равномерное | $a = unif(0{,}01; 1)$ |
| | $b = unif(a; 10)$ |
| Вейбулла | $k = unif(0{,}1; 10)$ |
| | $\theta = unif(0{,}5; 3)$ |
| Леви | $\mu = unif(0{,}01; 0{,}5)$ |
| | $c = unif(0{,}1; 10)$ |
| Фишера | $d_1 = unif(0{,}01; 1)$ |
| | $d_2 = unif(1{,}5; 3)$ |
| Парето | $x_m = unif(0{,}01; 10)$ |
| | $a = unif(1{,}5; 3)$ |
| Обратное гамма | $a = unif(1{,}5; 3)$ |
| | $c = unif(0{,}01; 1)$ |
| Ломакса | $a = unif(1{,}5; 3)$ |
| | $\lambda = unif(0{,}01; 1)$ |
| Гиперэкспоненциальное | $\lambda_1 = unif(0{,}1; 10)$ |
| | $\lambda_2 = unif(0{,}01; 1)$ |
| | $p = unif(0{,}1; 0{,}9)$ |
| Бура XII | $\alpha = unif(1{,}5; 3)$ |
| | $\beta = unif(1{,}5; 3)$ |
| | $c = unif(0{,}01; 1)$ |
| Фреше | $\alpha = unif(1{,}5; 3)$ |
| | $s = unif(0{,}01; 1)$ |
| | $m = unif(0; 0{,}5)$ |

**2. Метрики качества оценки параметров.** В качестве метрик для измерения обобщающей способности моделей оценки параметров были использованы: $MAPE$, $MAE$, $RMSE$, $R^2$.

Средняя абсолютная процентная погрешность (Mean Absolute Percentage Error, $MAPE$) измеряет отклонение прогнозов от фактических значений в процентах:

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{y_i},$$

где $n$ – количество наблюдений, $y_i$ – действительное значение переменной, $\hat{y}_i$ – предсказанное значение.

Средняя абсолютная ошибка (Mean Absolute Error, $MAE$) является линейной оценкой, следовательно, все ошибки в среднем взвешены одинаково:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|.$$

Квадратный корень из среднеквадратичной ошибки (Root Mean Squared Error, $RMSE$) показывает, на сколько в среднем отклоняется прогноз от реального значения:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}.$$

Коэффициент детерминации (Coefficient of determination, $R^2$) – это наиболее интуитивная и понятная метрика, которая показывает, насколько данная модель работает лучше, чем «наивная» модель:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2},$$

где $\bar{y}$ – выборочное среднее $y_i$. Для модели с идеальной предсказывающей способностью коэффициент детерминации равен 1.

**3. Оценка параметров потоков.** Задача оценки значений параметров исследуемых потоков является задачей регрессии, так как значения параметров являются непрерывными. Под регрессией принято понимать зависимость среднего значения какой-либо величины от некоторой другой величины или от нескольких величин. Если пространство объектов обозначить как $X$, и множество возможных ответов $Y$, то существует неизвестная целевая зависимость $y^*: X \to Y$, значения которой известны только на объектах обучающей выборки $X^l = (x_i, y_i)_{i=1}^{l}$, $y_i = y^*(x_i)$. Требуется построить алгоритм, который принято называть функцией регрессии $f: X \to Y$, аппроксимирующей целевую зависимость $y^*$. В выборках для обучения содержится целевая переменная, поэтому такая задача относится к классу задач обучения с учителем.

В качестве модели для оценки параметров в пуассоновском и рекуррентном потоках был использован CatBoostRegressor, для каждого параметра была обучена отдельная модель, на трех наборах данных (НД):

1) моменты наступления событий в потоке $t_i$;
2) интервалы между моментами наступления событий в потоке $\tau_i = t_i - t_{i-1}$;
3) 10 числовых характеристик интервалов $\tau_i$ (математическое ожидание, среднеквадратичное отклонение, дисперсия, коэффициент вариации, 6 квантилей: 0,1; 0,25; 0,5; 0,75; 0,9; 0,95).

**а. Пуассоновский поток.** В работе [28] значение параметра пуассоновского потока варьировалось на интервале (0;10000), для его оценки использовались различные архитектуры нейронных сетей. Значение метрик $MAE = 226{,}271$ и $R^2 = 0{,}989$. В этой работе был увеличен интервал допустимых значений параметра (0;20000) и использован алгоритм градиентного бустинга (ГБ) CatBoostRegressor для оценки параметра простейшего потока. Также была проведена оценка параметра с помощью метода моментов (ММ), для этого потока был выбран

первый начальный момент. В таблице 3 приведены значения метрик качества для ММ и для алгоритма, основанного на градиентном бустинге.

**Таблица 3.** Значения метрик качества оценки параметров пуассоновского потока

| Модель | НД | $MAPE$ | $MAE$ | $RMSE$ | $R^2$ |
|---|---|---|---|---|---|
| ГБ | 1 | 2,574 | 82,304 | 108,852 | 0,999 |
| ГБ | 2 | 11,529 | 667,714 | 855,737 | 0,977 |
| ГБ | 3 | 1,259 | 88,525 | 124,405 | 0,999 |
| ММ | 2 | 0,808 | 78,863 | 114,461 | 0,999 |

Как видно из таблицы 3, несмотря на увеличение интервала допустимых значений параметра пуассоновского потока, по сравнению с интервалом в работе [28], метрики качества модели улучшились, а, следовательно, и предсказательная способность модели. Лучше всего параметр оценивается с помощью метода моментов, при использовании метода градиентного бустинга модель лучше обобщает данные при обучении на 3-м наборе данных, который содержит числовые характеристики интервалов $\tau_i$.

**б. Рекуррентный поток.** Для рекуррентных потоков с функциями распределения вероятностей (ФРВ) длин интервалов: гамма, равномерная, логнормальная, гиперэкспоненциальная, Вейбулла параметры были оценены с помощью метода моментов (ММ), для остальных ФРВ с помощью метода наименьших квадратов (МНК). Метод наименьших квадратов использовался для тех ФРВ, в которых некоторые моменты не существуют или равны бесконечности, поэтому использование метода моментов для таких ФРВ невозможно.

В работе [29] были рассмотрены оценки параметров для некоторых ФРВ длин интервалов рекуррентного потока, в этой работе были использованы другие интервалы допустимых значений параметров, которые наиболее приближены к реальным параметрам потоков в сетевом трафике.

В таблицах 4-15 представлены значения метрик качества оценки параметров рекуррентного потока с разными функциями распределения длин интервалов между моментами наступления событий.

**Таблица 4.** Значения метрик качества оценки параметров рекуррентного потока с гамма ФРВ

| Модель | НД | Параметр | $MAPE$ | $MAE$ | $RMSE$ | $R^2$ |
|---|---|---|---|---|---|---|
| ГБ | 1 | $\alpha$ | 26,719 | 0,542 | 0,678 | 0,725 |
| | | $\beta$ | 24,822 | 0,832 | 1,035 | 0,867 |
| ГБ | 2 | $\alpha$ | 17,047 | 0,360 | 0,445 | 0,890 |
| | | $\beta$ | 12,208 | 0,558 | 0,769 | 0,929 |
| ГБ | 3 | $\alpha$ | 1,451 | 0,037 | 0,048 | 0,999 |
| | | $\beta$ | 2,124 | 0,080 | 0,110 | 0,999 |
| ММ | 2 | $\alpha$ | **1,362** | **0,035** | **0,047** | 0,999 |
| | | $\beta$ | **1,436** | **0,073** | **0,107** | 0,999 |

**Таблица 5.** Значения метрик качества оценки параметров рекуррентного потока с логнормальной ФРВ

| Модель | НД | Параметр | $MAPE$ | $MAE$ | $RMSE$ | $R^2$ |
|---|---|---|---|---|---|---|
| ГБ | 1 | $\mu$ | 164,918 | 0,481 | 0,638 | 0,694 |
| | | $\sigma$ | 18,482 | 0,247 | 0,306 | 0,954 |
| ГБ | 2 | $\mu$ | 36,692 | 0,107 | 0,147 | 0,984 |
| | | $\sigma$ | 8,461 | 0,158 | 0,211 | 0,978 |
| ГБ | 3 | $\mu$ | **7,388** | **0,024** | **0,034** | **0,999** |
| | | $\sigma$ | **1,183** | **0,024** | **0,032** | **0,999** |

| | | | | | | |
|---|---|---|---|---|---|---|
| ММ | 2 | $\mu$ | 1052,656 | 1,921 | 2,981 | 0,003 |
| | | $\sigma$ | 17,487 | 0,652 | 0,961 | 0,543 |

**Таблица 6.** Значения метрик качества оценки параметров рекуррентного потока с равномерной ФРВ

| Модель | НД | Параметр | $MAPE$ | $MAE$ | $RMSE$ | $R^2$ |
|---|---|---|---|---|---|---|
| ГБ | 1 | $a$ | 152,170 | 0,233 | 0,273 | 0,117 |
| | | $b$ | 6,880 | 0,240 | 0,283 | 0,989 |
| ГБ | 2 | $a$ | 34,250 | 0,061 | 0,079 | 0,926 |
| | | $b$ | 3,126 | 0,142 | 0,037 | 0,995 |
| ГБ | 3 | $a$ | 7,696 | 0,017 | 0,023 | 0,994 |
| | | $b$ | 0,809 | 0,028 | 0,036 | 0,999 |
| ММ | 2 | $a$ | **6,751** | **0,014** | **0,020** | **0,995** |
| | | $b$ | **0,251** | **0,014** | **0,020** | **0,999** |

**Таблица 7.** Значения метрик качества оценки параметров рекуррентного потока с ФРВ Вейбулла

| Модель | НД | Параметр | $MAPE$ | $MAE$ | $RMSE$ | $R^2$ |
|---|---|---|---|---|---|---|
| ГБ | 1 | $k$ | 5,162 | 0,215 | 0,418 | 0,979 |
| | | $\theta$ | 26,153 | 0,386 | 0,471 | 0,570 |
| ГБ | 2 | $k$ | 5,607 | 0,199 | 0,307 | 0,989 |
| | | $\theta$ | 10,025 | 0,154 | 0,191 | 0,929 |
| ГБ | 3 | $k$ | 1,371 | **0,045** | **0,063** | **0,999** |
| | | $\theta$ | 1,220 | 0,019 | **0,025** | **0,999** |
| ММ | 2 | $k$ | **0,927** | **0,045** | 0,228 | 0,994 |
| | | $\theta$ | **1,173** | **0,015** | 0,037 | 0,997 |

**Таблица 8.** Значения метрик качества оценки параметров рекуррентного потока с ФРВ Леви

| Модель | НД | Параметр | $MAPE$ | $MAE$ | $RMSE$ | $R^2$ |
|---|---|---|---|---|---|---|
| ГБ | 1 | $\mu$ | 146,863 | 0,021 | 0,143 | 0,487 |
| | | $c$ | 91,938 | 33,248 | 5,766 | 0,042 |
| ГБ | 2 | $\mu$ | 128,981 | 0,019 | 0,138 | 0,032 |
| | | $c$ | 11,693 | 0,223 | 0,472 | 0,973 |
| ГБ | 3 | $\mu$ | **35,601** | **0,003** | **0,054** | **0,851** |
| | | $c$ | **2,461** | **0,019** | **0,138** | **0,998** |
| МНК | 2 | $\mu$ | 150,678 | 0,438 | 0,662 | 0,005 |
| | | $c$ | 4,006 | 0,217 | 0,466 | 0,973 |

**Таблица 9.** Значения метрик качества оценки параметров рекуррентного потока с ФРВ Фишера

| Модель | НД | Параметр | $MAPE$ | $MAE$ | $RMSE$ | $R^2$ |
|---|---|---|---|---|---|---|
| ГБ | 1 | $d_1$ | 84,919 | 0,268 | 0,518 | 0,092 |
| | | $d_2$ | 29,904 | 0,479 | 0,692 | 0,177 |
| ГБ | 2 | $d_1$ | 17,915 | 0,006 | 0,080 | 0,922 |
| | | $d_2$ | 12,136 | 0,093 | 0,305 | 0,507 |
| ГБ | 3 | $d_1$ | 3,038 | **0,001** | 0,013 | 0,998 |
| | | $d_2$ | **2,972** | **0,009** | **0,096** | **0,951** |
| МНК | 2 | $d_1$ | **1,148** | **0,001** | **0,009** | **0,999** |
| | | $d_2$ | 23,668 | 46,767 | 6,839 | 0,001 |

**Таблица 10.** Значения метрик качества оценки параметров рекуррентного потока с ФРВ Парето

| Модель | НД | Параметр | $MAPE$ | $MAE$ | $RMSE$ | $R^2$ |
|---|---|---|---|---|---|---|
| ГБ | 1 | $x_m$ | 23,895 | 0,472 | 0,686 | 0,172 |
| | | $a$ | 93,331 | 0,321 | 0,567 | 0,046 |
| ГБ | 2 | $x_m$ | 9,281 | 0,056 | 0,238 | 0,693 |
| | | $a$ | 5,930 | 0,001 | 0,024 | 0,993 |
| ГБ | 3 | $x_m$ | 1,457 | 0,002 | 0,045 | 0,989 |
| | | $a$ | 1,234 | 0,001 | 0,004 | 0,999 |
| МНК | 2 | $x_m$ | **0,905** | **0,001** | **0,026** | **0,996** |
| | | $a$ | **0,120** | **0,001** | **0,001** | **0,999** |

**Таблица 11.** Значения метрик качества оценки параметров рекуррентного потока с ФРВ обратное гамма

| Модель | НД | Параметр | $MAPE$ | $MAE$ | $RMSE$ | $R^2$ |
|---|---|---|---|---|---|---|
| ГБ | 1 | $a$ | 29,693 | 0,476 | 0,690 | 0,171 |
| | | $c$ | 414,240 | 0,317 | 0,563 | 0,062 |
| ГБ | 2 | $a$ | 10,907 | 0,077 | 0,278 | 0,582 |
| | | $c$ | 11,732 | 0,006 | 0,077 | 0,931 |
| ГБ | 3 | $a$ | 1,566 | 0,002 | 0,046 | 0,988 |
| | | $c$ | **2,256** | **0,001** | **0,012** | **0,998** |
| МНК | 2 | $a$ | **1,273** | **0,002** | **0,037** | **0,993** |
| | | $c$ | 115,412 | 115,239 | 10,735 | 0,001 |

**Таблица 12.** Значения метрик качества оценки параметров рекуррентного потока с ФРВ Ломакса

| Модель | НД | Параметр | $MAPE$ | $MAE$ | $RMSE$ | $R^2$ |
|---|---|---|---|---|---|---|
| ГБ | 1 | $a$ | 19,012 | 0,203 | 0,451 | 0,476 |
| | | $\lambda$ | 93,738 | 0,320 | 0,566 | 0,046 |
| ГБ | 2 | $a$ | 13,768 | 0,125 | 0,354 | 0,325 |
| | | $\lambda$ | 23,124 | 0,011 | 0,106 | 0,860 |
| ГБ | 3 | $a$ | **3,045** | **0,009** | **0,093** | **0,954** |
| | | $\lambda$ | **4,526** | **0,001** | **0,028** | **0,990** |
| МНК | 2 | $a$ | 4,729 | 0,112 | 0,334 | 0,398 |
| | | $\lambda$ | 6,233 | 0,002 | 0,044 | 0,976 |

**Таблица 13.** Значения метрик качества оценки параметров рекуррентного потока с гиперэкспоненциальной ФРВ

| Модель | НД | Параметр | $MAPE$ | $MAE$ | $RMSE$ | $R^2$ |
|---|---|---|---|---|---|---|
| ГБ | 1 | $\lambda_1$ | 71,585 | 1,389 | 1,671 | 0,661 |
| | | $\lambda_2$ | 49,960 | 0,626 | 0,807 | 0,686 |
| | | $p$ | 53,577 | 0,160 | **0,190** | 0,335 |
| ГБ | 2 | $\lambda_1$ | 87,356 | 1,349 | 1,799 | 0,609 |
| | | $\lambda_2$ | 98,425 | 0,771 | 1,002 | 0,522 |
| | | $p$ | 62,755 | 0,182 | 0,217 | 0,159 |
| ГБ | 3 | $\lambda_1$ | 50,602 | **0,782** | **1,091** | **0,856** |
| | | $\lambda_2$ | 54,087 | **0,534** | **0,802** | **0,694** |
| | | $p$ | **50,699** | **0,144** | **0,190** | **0,357** |
| ММ | 2 | $\lambda_1$ | **48,715** | 1,979 | 2,856 | 0,045 |
| | | $\lambda_2$ | **23,980** | 0,628 | 1,090 | 0,432 |
| | | $p$ | 63,248 | 0,246 | 0,297 | 0,356 |

**Таблица 14.** Значения метрик качества оценки параметров рекуррентного потока с ФРВ Бура XII

| Модель | НД | Параметр | $MAPE$ | $MAE$ | $RMSE$ | $R^2$ |
|---|---|---|---|---|---|---|
| ГБ | 1 | $\alpha$ | 23,443 | 0,466 | 0,683 | 0,177 |
| | | $\beta$ | 17,303 | 0,183 | 0,428 | 0,014 |
| | | $c$ | 89,139 | 0,304 | 0,551 | 0,062 |
| ГБ | 2 | $\alpha$ | 8,807 | 0,052 | 0,228 | 0,717 |
| | | $\beta$ | 15,845 | 0,160 | 0,401 | 0,137 |
| | | $c$ | 11,121 | 0,005 | 0,068 | 0,943 |
| ГБ | 3 | $\alpha$ | **1,864** | **0,003** | **0,057** | **0,982** |
| | | $\beta$ | **6,822** | **0,037** | **0,193** | **0,799** |
| | | $c$ | **4,734** | **0,001** | **0,029** | **0,990** |
| ММ | 2 | $\alpha$ | 153,388 | 718,761 | 26,810 | 0,004 |
| | | $\beta$ | 25,688 | 1,423 | 1,193 | 0,001 |
| | | $c$ | 106,177 | 1,831 | 1,353 | 0,001 |

**Таблица 15.** Значения метрик качества оценки параметров рекуррентного потока с ФРВ Фреше

| Модель | НД | Параметр | $MAPE$ | $MAE$ | $RMSE$ | $R^2$ |
|---|---|---|---|---|---|---|
| ГБ | 1 | $\alpha$ | 17,142 | 0,188 | 0,434 | 0,498 |
| | | $s$ | 75,077 | 0,228 | 0,477 | 0,145 |
| | | $m$ | 190,050 | 0,105 | 0,324 | 0,019 |
| ГБ | 2 | $\alpha$ | 11,737 | 0,092 | 0,304 | 0,505 |
| | | $s$ | 23,256 | 0,008 | 0,091 | 0,900 |
| | | $m$ | 235,351 | 0,007 | 0,086 | 0,646 |
| ГБ | 3 | $\alpha$ | **4,107** | **0,017** | **0,132** | **0,907** |
| | | $s$ | **6,842** | **0,001** | **0,030** | **0,989** |
| | | $m$ | **5,188** | **0,001** | **0,027** | **0,964** |
| ММ | 2 | $\alpha$ | 12,550 | 6,235 | 2,497 | 0,004 |
| | | $s$ | 15,850 | 0,003 | 0,054 | 0,963 |
| | | $m$ | 63,719 | 0,003 | 0,062 | 0,815 |

Для рекуррентного потока со следующими ФРВ: гамма, равномерная, Парето метод моментов и метод наименьших квадратов лучше оценивают параметры, исходя из метрик качества. Стоит отметить, что оценивание параметров с помощью алгоритма градиентного бустинга на 3-м наборе данных для этих ФРВ также показывает стабильные и точные результаты, с незначительным отставанием по метрикам, что подтверждает его практическую применимость. Для всех остальных ФРВ модель алгоритма градиентного бустинга, обученная на числовых характеристиках случайной величины длина интервалов, показывает лучшие метрики при сравнении с другими моделями. В отдельных случаях наблюдается незначительное преимущество методов ММ или МНК, однако оно, вероятно, не является устойчивым и может не воспроизводиться при использовании других наборов данных.

**4. Оценка важности используемых признаков.** Одним из ключевых этапов при построении моделей машинного обучения является выбор информативных признаков. В этой работе в качестве признаков для обучения регрессионных моделей рассматриваются числовые характеристики случайной величины. Сравнительный анализ важности этих признаков позволяет не только повысить точность моделей, но и глубже понять, какие числовые характеристики наиболее полно отражают свойства конкретных распределений.

На рисунке 1 представлены графики важности признаков для двух моделей, обученных на числовых характеристиках выборок из распределений Леви и Парето. Каждый столбец на графиках отражает вклад соответствующего признака в предсказание параметров.

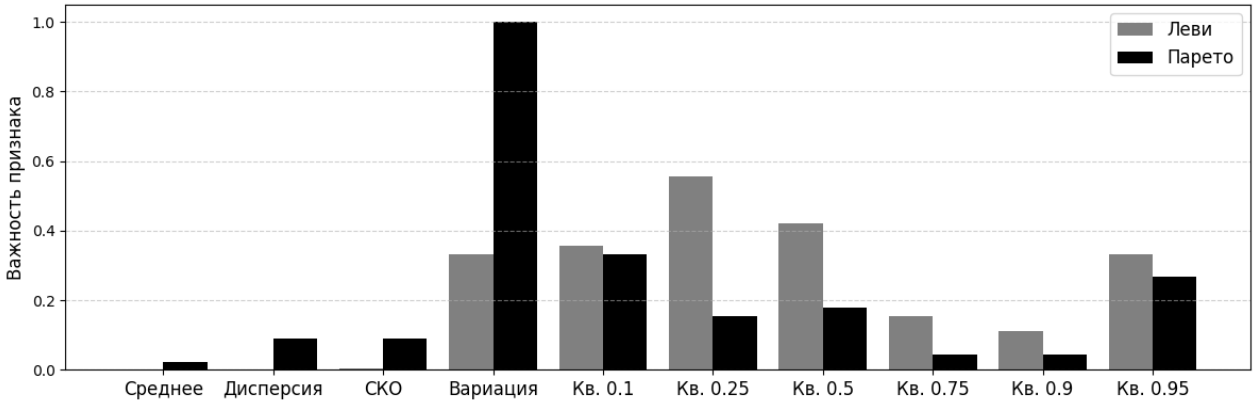


**Рис. 1.** Сравнение важности числовых признаков при оценке параметров для распределений Леви и Парето

Распределение Леви представляет собой тяжелохвостое и асимметричное распределение, для которого математическое ожидание и дисперсия часто не существуют или нестабильны. В соответствии с этим, наибольшую важность для модели представляют квантили, характеризующие поведение правого хвоста. Такие признаки, как среднее и дисперсия, оказываются слабо информативными, что соответствует теоретическим ограничениям распределения Леви.

Для распределения Парето ключевую роль играет коэффициент вариации, отражающий степень разброса значений, что согласуется с природой степенных распределений. Средними по значимости являются стандартное отклонение и квантили, особенно крайние, которые позволяют уточнять поведение хвостов при наличии конечных моментов (например, при $\alpha > 2$).

Дополнительно были проанализированы важности признаков для других распределений (рис. 2).

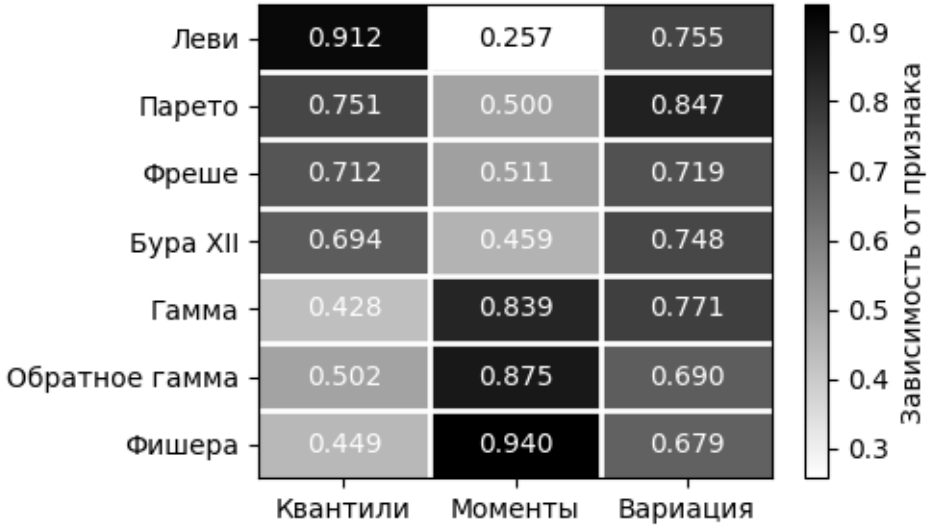


**Рис. 2.** Сравнение важности числовых признаков при оценке параметров

Полученные результаты демонстрируют, что информативность тех или иных выборочных статистик существенно зависит от природы распределения:

- для тяжелых и асимметричных распределений (Леви, Парето, Фреше, Бура XII) наибольшую значимость имеют квантили и коэффициент вариации, поскольку классические моменты либо не существуют, либо не обеспечивают стабильной оценки параметров;
- для распределений с устойчивыми моментами (гамма, обратное гамма, экспоненциальное, Фишера) наиболее полезными оказываются среднее, дисперсия и стандартное отклонение, которые напрямую связаны с параметрами этих распределений и позволяют добиться высокой точности оценки.

Таким образом, выбор признаков для оценки параметров распределений должен учитывать теоретические свойства конкретного распределения. Использование информативных характеристик позволяет существенно повысить точность моделей и интерпретируемость оценок параметров используемых моделей.

**Заключение.** В представленной работе были обучены модели машинного обучения для оценки параметров различных типов потоков: пуассоновского и рекуррентного с 12 различными распределениями интервалов между моментами наступления событий. Использован алгоритм градиентного бустинга (CatBoost), а также классические методы оценки, такие, как метод моментов (ММ) и метод наименьших квадратов (МНК).

Для оценки параметра $\lambda$ пуассоновского потока были использованы метод моментов и градиентный бустинг. Лучшие результаты показал метод моментов со значениями метрик $MAE = 78{,}863$, $RMSE = 114{,}461$ и $R^2 = 0{,}999$. Однако градиентный бустинг также продемонстрировал высокие показатели, особенно при обучении на третьем наборе данных: $MAE = 88{,}525$, $RMSE = 124{,}405$ и $R^2 = 0{,}999$.

Для рекуррентного потока лучшими оказались различные методы в зависимости от функции распределения вероятностей длин интервалов между моментами наступления событий. Для рекуррентного потока с функциями распределения вероятностей длин интервалов: гамма, равномерная, Парето метод моментов и метод наименьших квадратов лучше оценивают параметры, исходя из метрик качества, коэффициент детерминации больше $0{,}995$. Для остальных лучше всего себя показала модель градиентного бустинга, обученная на наборе данных, содержащем числовые характеристики интервалов между моментами наступления событий. В большинстве таких случаев значение $R^2$ превышало $0{,}900$, что свидетельствует о высокой точности даже для тяжелохвостых распределений.

Таким образом, использование моделей машинного обучения обеспечивает высокую точность оценки параметров потоков и может рассматриваться как надежная альтернатива или дополнение к классическим статистическим методам. Несмотря на то, что в отдельных случаях метод моментов и метод наименьших квадратов показывают незначительно лучшие результаты, модель градиентного бустинга демонстрирует универсальность, стабильность и высокое качество для всех рассмотренных распределений.

В дальнейшем планируется расширить набор исследуемых видов входящего потока, а также применить разработанные методы для анализа и прогнозирования сетевой нагрузки в реальном времени.

## Список источников

1. Ivanova D., Markova E., Moltchanov D., et al. Performance of priority-based traffic coexistence strategies in 5G mmWave industrial deployments. IEEE Access, 2022, vol. 10, pp. 9241-9256, DOI: 10.1109/ACCESS.2022.3143583.
2. Fedorova E., Lapatin I., Lizyura O., et al. Asymptotic analysis of two-phase queueing system with service rate degradation and heterogeneous customers. 5th International Conference on Problems of Cybernetics and Informatics (PCI), 2023, pp. 1-5.

3. Markova E., Moltchanov D., Pirmagomedov R., et al. Prioritized service of URLLC traffic in industrial deployments of 5G NR systems. Distributed computer and communication networks, 2020, vol 12563, pp. 497-509, DOI:10.1007/978-3-030-66471-8_38.
4. Paxson V., Floyd S. Wide area traffic: the failure of Poisson modeling. IEEE/ACM Transactions on Networking, 1995, vol. 3, no. 3, pp. 226-244, DOI: 10.1109/90.392383.
5. Johnson M., Narayana S. Descriptors of arrival-process burstiness with application to the discrete Markovian arrival process. Queueing systems, 1996, vol. 23, pp. 107-130.
6. Vazquez A., Oliveira J., Dezso Z., Goh K.-I. Modeling bursts and heavy tails in human dynamics. Physical Review, 2006, vol. 73, no. 3, DOI: 10.48550/arXiv.physics/0510117.
7. Choi J., Hiraoka T., Jo H.-H. Individual-driven versus interaction-driven burstiness in human dynamics: The case of Wikipedia edit history. American physical society, 2011, vol. 104, DOI:10.48550/arXiv.2011.01562.
8. Leland W., Taqqu M., Willinger W., et al. On the self-similar nature of Ethernet traffic. IEEE/ACM Transactions on Networking, 1994, vol. 2, no. 1, pp. 1-15, DOI: 10.1109/90.282603.
9. Cappe O., Moulines E., Pesquet J.-C., et al. Long-range dependence and heavy-tail modeling for teletraffic data, IEEE Signal processing magazine, 2002, vol. 19, no. 3, pp. 14-27, DOI: 10.1109/79.998079.
10. Arfeen M., Pawlikowski K., Willig A., et al. Fractal renewal process-based analysis of emerging network traffic in access networks. 26th International telecommunication networks and applications conference (ITNAC), 2016, pp. 265-270, DOI: 10.1109/ATNAC.2016.7878820.
11. Becchi M. From poisson processes to self-similarity: a survey of network traffic models, Washington University in St. Louis, 2008.
12. Ryu B., Lowen S. Point process approaches to the modeling and analysis of self-similar traffic .I. Model construction. Conference on computer communications, 1996, vol. 3, pp. 1468-1475, DOI: 10.1109/INFCOM.1996.493096.
13. Feldmann A., Gilbert A.C., Willinger W. Data networks as cascades: investigating the multifractal nature of Internet WAN traffic. Association for Computing Machinery, 1998, pp. 42-55, DOI:10.1145/285237.285256.
14. Moore A., Papagiannaki K. Toward the accurate identification of network applications. Passive and active network measurement: book of abstracts 6th international workshop, 2005, pp. 41-54, DOI:10.1007/978-3-540-31966-5_4.
15. Dreger H., Feldmann A., Mai M., et al. Dynamic application-layer protocol analysis for network intrusion detection. 15th USENIX Security Symposium, 2006, pp. 257-272.
16. Karagiannis T., Broido A., Faloutsos M., et al. Transport layer identification of P2P Traffic. 4th ACM SIGCOMM conference on Internet measurement, 2004, pp. 121-134, DOI:10.1145/1028788.1028804.
17. Choi T., Kim C., Yoon S., Park J. Content-aware Internet application traffic measurement and analysis. IEEE/IFIP Network operations and management symposium, 2004, vol. 1, pp. 511-524, DOI:10.1109/NOMS.2004.1317737.
18. Moore A., Zuev D. Internet traffic classification using bayesian analysis techniques. Measurement and Modeling of Computer Systems, 2005, vol. 33, no. 1, pp. 50-60, DOI:10.1145/1064212.1064220.
19. Rezaei S., Kroencke B., Liu X. Large-scale mobile app identification using deep learning. IEEE Access, 2020, vol. 8, pp. 348-362, DOI: 10.1109/ACCESS.2019.2962018.
20. Денисенко В.В. Применение искусственного интеллекта для анализа сетевого трафика / В.В. Денисенко, А.С. Ященко // Международный журнал гуманитарных и естественных наук: международный ежемесячный научный журнал, 2023. – № 1. – С. 19-22.
21. Скрыпников А.В. Использование методов машинного обучения при решении задач информационной безопасности / А.В Скрыпников, В.В. Денисенко, И.А. Саранов // Вестник Воронежского института ФСИН России, 2020. – № 4. – С. 69-73.
22. Khan M. HCRNNIDS: Hybrid convolutional recurrent neural network-based network intrusion detection system. Processes, 2021, vol. 9, no. 5, pp. 830-844, DOI:10.3390/pr9050834.
23. Zheng W., Gou C., Yan L., et al. Learning to Classify: a flow-based relation network for encrypted traffic classification. In Proc. of the Web Conference, 2020, pp. 13-22, DOI:10.1145/3366423.3380090.
24. D'Angelo G., Palmieri F. Network traffic classification using deep convolutional recurrent autoencoder neural networks for spatial–temporal features extraction. Journal of network and computer applications, 2021, vol. 173, DOI:10.1016/j.jnca.2020.102890.
25. Lotfollahi M., Shirali hossein zade R., Jafari Siavoshani M., et al. Deep packet: a novel approach for encrypted traffic classification using deep learning. Soft Computing, 2020, vol. 24, DOI:10.1007/s00500-019-04030-2.
26. Niloofar B., Weston J., Derrick L. Deep Learning for Network Traffic Classification. arXiv, 2021, DOI:10.48550/arXiv.2106.12693.
27. Марголис Н.Ю. Имитационное моделирование: учеб. пособие / Н.Ю. Марголис – Томск: Издательский Дом Томского государственного университета, 2015. – 88 с.

28. Salimzyanova D., Lisovskaya E. Identification of Network Traffic Using Neural Networks, Information Technologies and Mathematical Modelling. Queueing Theory and Applications, 2023, vol. 2163, pp. 76-90, DOI:10.1007/978-3-031-65385-8_6.
29. Салимзянова Д.Д. Классификация и оценка параметров распределений вероятностей длин интервалов рекуррентного потока / Д.Д. Салимзянова, Е.Ю. Лисовская // Информационные технологии и математическое моделирование (ИТММ-2023): Материалы XXII Международной конференции имени А.Ф. Терпугова (Томск, 4–9 декабря 2023) – Томск, 2024. – С. 64-69.

***Салимзянова Дарья Дмитриевна.*** *Томский государственный университет, ассистент, AuthorID: 1282632, SPIN: 5509-2107, ORCID: 0009-0003-8727-0918, darya2001@inbox.ru.*

***Лисовская Екатерина Юрьевна.*** *Канд. физ.-мат. наук, Томский государственный университет, доцент, AuthorID: 863821, SPIN: 9056-4621, ORCID: 0000-0001-7345-5565, ekaterina_lisovs@mail.ru.*

***Самойлов Сергей Антонович.*** *Томский государственный университет, лаборант, ORCID: 0009-0005-8318-9885, sergei.samoilov224@gmail.com.*

---

UDC 519.2
DOI:10.25729/ESI.2025.40.4.004

# Estimation of stochastic event flow parameters using machine learning methods

**Daria D. Salimzyanova, Ekaterina Yu. Lisovskaya, Sergey A. Samoilov**

National Research Tomsk State University, Russia, Tomsk, *darya2001@inbox.ru*

**Abstract.** This paper addresses the problem of estimating parameters of stochastic event flows based on sample data using machine learning methods. Event flows, characterized by random intervals between the moments of occurrence, are widely used in the modeling of network traffic, telecommunications, computing systems, and in queuing theory. Accurate estimation of such flow parameters is crucial for subsequent analysis, forecasting, and load management in systems with uncertain input information. As training data for the models, we used event arrival times from two types of streams: a Poisson flow (with inter-arrival times following the exponential distribution) and a renewal process (with inter-arrival times following one of twelve probability distributions: gamma, hyperexponential, lognormal, uniform, inverse gamma, Weibull, Pareto, Lévy, Fisher, Fréchet, Lomax, and Burr XII). These distributions were selected due to their diverse statistical properties (presence or absence of moments, asymmetry, heavy tails), which enables coverage of a broad range of applicable scenarios. To solve the parameter estimation task, we employed fully connected neural networks and the CatBoost implementation of the gradient boosting algorithm. As input features for the models, we used the inter-arrival times and their numerical characteristics: mean, standard deviation, variance, coefficient of variation, and quantiles of various levels. To evaluate the model performance, classical machine learning metrics were used: $MAE$, $RMSE$, and $R^2$. The study also included an assessment of the importance of features used in training. This was done using built-in interpretation tools of gradient boosting, which allow for a quantitative analysis of each feature's contribution to the parameter estimation.

**Keywords:** traffic identification, network traffic, parameter estimation, gradient boosting

## References

1. Ivanova D., Markova E., Moltchanov D., et al. Performance of priority-based traffic coexistence strategies in 5G mmWave industrial deployments. IEEE Access, 2022, vol. 10, pp. 9241-9256, DOI: 10.1109/ACCESS.2022.3143583.
2. Fedorova E., Lapatin I., Lizyura O., et al. Asymptotic analysis of two-phase queueing system with service rate degradation and heterogeneous customers. 5th International Conference on Problems of Cybernetics and Informatics (PCI), 2023, pp. 1-5.
3. Markova E., Moltchanov D., Pirmagomedov R., et al. Prioritized service of URLLC traffic in industrial deployments of 5G NR systems. Distributed computer and communication networks, 2020, vol 12563, pp. 497-509, DOI:10.1007/978-3-030-66471-8_38.
4. Paxson V., Floyd S. Wide area traffic: the failure of Poisson modeling. IEEE/ACM Transactions on Networking, 1995, vol. 3, no. 3, pp. 226-244, DOI: 10.1109/90.392383.

5. Johnson M., Narayana S. Descriptors of arrival-process burstiness with application to the discrete Markovian arrival process. Queueing systems, 1996, vol. 23, pp. 107-130.
6. Vazquez A., Oliveira J., Dezso Z., Goh K.-I. Modeling bursts and heavy tails in human dynamics. Physical Review, 2006, vol. 73, no. 3, DOI: 10.48550/arXiv.physics/0510117.
7. Choi J., Hiraoka T., Jo H.-H. Individual-driven versus interaction-driven burstiness in human dynamics: The case of Wikipedia edit history. American physical society, 2011, vol. 104, DOI:10.48550/arXiv.2011.01562.
8. Leland W., Taqqu M., Willinger W., et al. On the self-similar nature of Ethernet traffic. IEEE/ACM Transactions on Networking, 1994, vol. 2, no. 1, pp. 1-15, DOI: 10.1109/90.282603.
9. Cappe O., Moulines E., Pesquet J.-C., et al. Long-range dependence and heavy-tail modeling for teletraffic data, IEEE Signal processing magazine, 2002, vol. 19, no. 3, pp. 14-27, DOI: 10.1109/79.998079.
10. Arfeen M., Pawlikowski K., Willig A., et al. Fractal renewal process-based analysis of emerging network traffic in access networks. 26th International telecommunication networks and applications conference (ITNAC), 2016, pp. 265-270, DOI: 10.1109/ATNAC.2016.7878820.
11. Becchi M. From poisson processes to self-similarity: a survey of network traffic models, Washington University in St. Louis, 2008.
12. Ryu B., Lowen S. Point process approaches to the modeling and analysis of self-similar traffic .I. Model construction. Conference on computer communications, 1996, vol. 3, pp. 1468-1475, DOI: 10.1109/INFCOM.1996.493096.
13. Feldmann A., Gilbert A.C., Willinger W. Data networks as cascades: investigating the multifractal nature of Internet WAN traffic. Association for Computing Machinery, 1998, pp. 42-55, DOI:10.1145/285237.285256.
14. Moore A., Papagiannaki K. Toward the accurate identification of network applications. Passive and active network measurement: book of abstracts 6th international workshop, 2005, pp. 41-54, DOI:10.1007/978-3-540-31966-5_4.
15. Dreger H., Feldmann A., Mai M., et al. Dynamic application-layer protocol analysis for network intrusion detection. 15th USENIX Security Symposium, 2006, pp. 257-272.
16. Karagiannis T., Broido A., Faloutsos M., et al. Transport layer identification of P2P Traffic. 4th ACM SIGCOMM conference on Internet measurement, 2004, pp. 121-134, DOI:10.1145/1028788.1028804.
17. Choi T., Kim C., Yoon S., Park J. Content-aware Internet application traffic measurement and analysis. IEEE/IFIP Network operations and management symposium, 2004, vol. 1, pp. 511-524, DOI:10.1109/NOMS.2004.1317737.
18. Moore A., Zuev D. Internet traffic classification using bayesian analysis techniques. Measurement and Modeling of Computer Systems, 2005, vol. 33, no. 1, pp. 50-60, DOI:10.1145/1064212.1064220.
19. Rezaei S., Kroencke B., Liu X. Large-scale mobile app identification using deep learning. IEEE Access, 2020, vol. 8, pp. 348-362, DOI: 10.1109/ACCESS.2019.2962018.
20. Denisenko V.V., Yashchenko A.S. Primeneniye iskusstvennogo intellekta dlya analiza setevogo traffika [Application of artificial intelligence for network traffic analysis]. Mezhdunarodnyy zhurnal gumanitarnykh i estestvennykh nauk [International Journal of Humanities and Natural Sciences], 2023, no. 1, pp. 19-22..
21. Skrypnikov A.V., Denisenko V.V., Saranov I.A. Ispol'zovaniye metodov mashinnogo obucheniya pri reshenii zadach informatsionnoy bezopasnosti [Using machine learning methods for solving information security problems]. Vestnik Voronezhskogo instituta FSIN Rossii [Bulletin of the Voronezh Institute of the Federal Penitentiary Service of Russia], 2020, no. 4, pp. 69-73.
22. Khan M. HCRNNIDS: Hybrid convolutional recurrent neural network-based network intrusion detection system. Processes, 2021, vol. 9, no. 5, pp. 830-844, DOI:10.3390/pr9050834.
23. Zheng W., Gou C., Yan L., et al. Learning to Classify: a flow-based relation network for encrypted traffic classification. In Proc. of the Web Conference, 2020, pp. 13-22, DOI:10.1145/3366423.3380090.
24. D'Angelo G., Palmieri F. Network traffic classification using deep convolutional recurrent autoencoder neural networks for spatial–temporal features extraction. Journal of network and computer applications, 2021, vol. 173, DOI:10.1016/j.jnca.2020.102890.
25. Lotfollahi M., Shirali hossein zade R., Jafari Siavoshani M., et al. Deep packet: a novel approach for encrypted traffic classification using deep learning. Soft Computing, 2020, vol. 24, DOI:10.1007/s00500-019-04030-2.
26. Niloofar B., Weston J., Derrick L. Deep Learning for Network Traffic Classification. arXiv, 2021, DOI:10.48550/arXiv.2106.12693.
27. Margolis N.Yu. Imitatsionnoye modelirovaniye: ucheb. posobiye [Simulation Modeling: Textbook]. Tomsk, Izdatel'skiy Dom Tomskogo gosudarstvennogo universiteta [Publishing House of Tomsk State University] Publ., 2015, 88 p.
28. Salimzyanova D., Lisovskaya E. Identification of Network Traffic Using Neural Networks, Information Technologies and Mathematical Modelling. Queueing Theory and Applications, 2023, vol. 2163, pp. 76-90, DOI:10.1007/978-3-031-65385-8_6.

29. Salimzyanova D.D., Lisovskaya E.Yu. Klassifikatsiya i otsenka parametrov raspredeleniy veroyatnostey dlin intervalov rekurrentnogo potoka [Classification and estimation of probability distribution parameters for recurrent flow interval lengths]. Informatsionnyye tekhnologii i matematicheskoye modelirovaniye (ITMM-2023): Materialy XXII Mezhdunarodnoy konferentsii imeni A.F. Terpugova [Information Technologies and Mathematical Modeling (ITMM-2023): Proceedings of the XXII International Conference named after A.F. Terpugov]. Tomsk, 2024, pp. 64-69.

***Salimzyanova Daria Dmitrievna.*** *National Research Tomsk State University, Assistant, AuthorID: 1282632, SPIN: 5509-2107, ORCID: 0009-0003-8727-0918, darya2001@inbox.ru.*

***Lisovskaya Ekaterina Yuryevna.*** *Candidate of Science in Physics and Mathematics, National Research Tomsk State University, Associate Professor, AuthorID: 863821, SPIN: 9056-4621, ORCID: 0000-0001-7345-5565, ekaterina_lisovs@mail.ru.*

***Samoilov Sergey Antonovich.*** *National Research Tomsk State University, Laboratory Assistant, ORCID: 0009-0005-8318-9885, sergei.samoilov224@gmail.com.*

*Статья поступила в редакцию 09.08.2025; одобрена после рецензирования 29.09.2025; принята к публикации 03.10.2025.*

*The article was submitted 08/09/2025; approved after reviewing 09/29/2025; accepted for publication 10/03/2025.*