

Семенов Илья Витальевич,
студент магистратуры 2 курса гр. ИСТМ-42,
ФГОБУ ВО «Поволжский государственный
университет телекоммуникаций и информатики»

Лихтциндер Борис Яковлевич, д.т.н, профессор,
ФГОБУ ВО «Поволжский государственный
университет телекоммуникаций и информатики»

МАТЕМАТИЧЕСКАЯ СТРУКТУРА И МЕТОДЫ АНАЛИЗА ГРУППОВОГО КВАЗИПУАССОНОВСКОГО ПОТОКА ЗАЯВОК

Аннотация. В статье исследуется математическая природа группового квазипуассоновского потока как адекватной модели пульсирующего инфокоммуникационного трафика. Рассмотрены особенности фреймовой структуры распределения временных интервалов и ключевые параметры пачечности. Авторами предложен метод декомпозиции скрытых числовых характеристик потоков с помощью пространства из девяти статистических признаков. Обоснована концепция применения каскада моделей многослойного перцептрона для классификации типов трафика и регрессионного распознавания параметров входящей нагрузки.

Ключевые слова: Квазипуассоновский поток, пачечность трафика, декомпозиция параметров, многослойный перцептрон, пространство когнитивных признаков, теория массового обслуживания.

Введение

Развитие современных инфокоммуникационных технологий, характеризующееся глобальным переходом к сетям последующих поколений, повсеместным внедрением концепций межмашинного взаимодействия, систем сверхчеткого видеостриминга и распределенных облачных вычислений, привело к кардинальному изменению физической природы сетевой нагрузки. Трафик современных мультисервисных сетей связи общего пользования полностью утратил свойства ординарности, стационарности и отсутствия последствия, которые являлись базовыми постулатами классической теории телетрафика, заложенной исследователями в начале двадцатого века.

Экспериментальные исследования реальных сетевых потоков данных показывают, что им присуща ярко выраженная пачечность. В таком трафике периоды экстремально высокой плотности поступления информационных единиц чередуются со значительными нелинейными паузами относительного затишья. Ошибка в априорном определении стохастической природы и характера входящего потока на этапе проектирования телекоммуникационного оборудования неизбежно ведет к критическим просчетам: недооценке реальной нагрузки на буферную память маршрутизаторов, лавинообразному росту сетевых задержек и неконтролируемым потерям данных.

Одним из наиболее адекватных, аналитически гибких и параметрически насыщенных теоретических аппаратов для описания, аппроксимации и имитационного моделирования пульсирующих процессов в инфокоммуникациях выступает групповой квазипуассоновский поток заявок. Актуальность данного теоретического исследования обусловлена широкой распространенностью пачечных процессов в компьютерном моделировании сложных систем. Изображение и формализация нелинейных концептуальных отношений между внутренними параметрами группового потока стали важнейшим элементом анализа сложных технических, когнитивных и организационных структур в теории массового обслуживания.



Цель настоящего исследования состоит в комплексном изучении теоретических основ, стохастической природы, концептуальной структуры и системных свойств группового квазипуассоновского потока заявок, а также в разработке строгой методологии его статистического анализа на основе вербальных и логических моделей.

Для достижения поставленной цели необходимо решить следующие теоретические задачи:

1. Изучить историко-теоретические предпосылки возникновения моделей групповых потоков и определить их место в современной классификации процессов теории массового обслуживания.
2. Дать строгое содержательное описание внутренней структуры распределения временных интервалов квазипуассоновского процесса, разграничив понятия микро- и макроинтервалов без привлечения громоздких формульных выкладок.
3. Осуществить детальный критический анализ преимуществ и недостатков применения групповых моделей при проектировании и анализе систем массового обслуживания.
4. Разработать теоретическую методологию статистической верификации параметров пачечности на основе непараметрических критериев согласия гипотез.

Историко-теоретические предпосылки и классификация потоков заявок

Исторически первыми и наиболее изученными в теории массового обслуживания являются простейшие потоки. Математический формализм пуассоновского процесса, детально описанный в классических трудах по теории вероятностей, опирается на три фундаментальных свойства: стационарность, ординарность и отсутствие последствия. Свойство ординарности теоретически постулирует невозможность одновременного поступления двух или более требований в бесконечно малый промежуток времени. Подобное допущение идеально подходило для моделирования традиционных телефонных сетей общего пользования, где вызовы от абонентов поступали полностью независимо друг от друга, а вероятность того, что два человека снимут телефонную трубку в одну и ту же миллисекунду, стремилась к нулю.

Однако уже в середине двадцатого века, по мере усложнения автоматических телефонных станций, появления систем с кодовым разделением каналов и первых вычислительных сетей, исследователи столкнулись с явлением группового поступления вызовов. Оказалось, что в определенные моменты времени система получает не единичные сигналы, а целые блоки информации. Первые попытки теоретического описания неординарных потоков были предприняты в зарубежных работах, авторы которых ввели понятия пакетных или групповых процессов. В отечественной математической школе фундаментальный вклад в развитие теории неординарных потоков внесли выдающиеся ученые, разработавшие строгие классификации случайных процессов на основе систем дифференциальных уравнений.

Групповой квазипуассоновский поток возник как специализированная модификация общего неординарного пуассоновского потока. Его появление было продиктовано необходимостью найти разумный компромисс между математической сложностью описания самоподобных, фрактальных процессов, обладающих свойством долговременной зависимости, и вычислительной эффективностью классических марковских моделей. Традиционные подходы, использующие тяжелые хвосты распределений, требовали колоссальных ресурсов при компьютерном моделировании, в то время как квазипуассоновский подход позволил сохранить простоту описания основных состояний системы при сохранении физической сущности пачечности [1].

В рамках общепринятой современной классификации случайных процессов в теории массового обслуживания, групповые потоки занимают обособленное положение. Верхний



индекс в буквенно-цифровых обозначениях моделей указывает на то, что объем поступающей группы требований является случайной величиной, распределенной по некоторому закону, отличному от единичного. Квазипуассоновский поток занимает промежуточное положение между строго упорядоченными детерминированными процессами, где интервалы времени между событиями постоянны, и абсолютно хаотическими нестационарными потоками, параметры которых не поддаются стабилизации.

Специфическая особенность квазипуассоновского процесса заключается в том, что, несмотря на явное нарушение свойства ординарности на микроуровне, на макроуровне процесс полностью сохраняет марковские свойства. Это означает, что поведение системы в будущем зависит только от ее текущего состояния и не определяется предысторией процесса. Данное свойство делает квазипуассоновский поток уникальным инструментом, позволяющим совместить реальные физические свойства пульсирующего трафика связи со строгими аналитическими методами исследования систем массового обслуживания.

Концептуальная сущность и структура квазипуассоновского потока

Групповой квазипуассоновский поток представляет собой теоретический формализм визуального и аналитического представления знаний о случайных процессах, описывающий ситуацию, когда требования поступают в систему не поодиночке, а стереотипными группами, называемыми пачками. Структура данного потока определяет состав независимых областей временной оси и их взаимное расположение. Это фундамент всего процесса математического моделирования, на который опирается система навигации трафика в исследуемой среде.

При содержательном анализе временного ряда моментов поступления заявок вводится фреймовая структура разделения интервалов времени на две независимые категории, каждая из которых описывает свой уровень абстракции процесса:

Во-первых, выделяются внутренние интервалы, называемые также микроинтервалами. Они представляют собой промежутки времени между последовательными требованиями, принадлежащими к одной и той же пачке. В классической, строгой теоретической квазипуассоновской модели данные микроинтервалы принимаются равными нулю. На понятийном уровне это означает мгновенное, абсолютно одновременное прибытие всей группы требований на обслуживаемое устройство или в буферную память коммутационного узла. Физически это можно сопоставить с приходом сетевого кадра, содержащего множество отдельных пакетов данных, которые извлекаются процессором практически одновременно.

Во-вторых, выделяются внешние интервалы, или макроинтервалы. Они представляют собой промежутки времени между моментами поступления самих пачек требований. Эти макропаузы являются взаимно независимыми, полностью изолированными друг от друга и подчиняются классическому экспоненциальному закону распределения. Физический смысл экспоненциальности на макроуровне заключается в том, что моменты возникновения самих пачек остаются случайными и не зависят от того, сколько времени прошло с момента прихода предыдущей группы требований [2].

Теоретический портрет и когнитивная структура группового квазипуассоновского потока полностью детерминируются набором из следующих независимых параметров, определяющих его поведение на различных уровнях абстракции:

Интенсивность макропотока определяет среднее количество пачек требований, поступающих в систему в единицу времени. Этот параметр задает общую динамику процесса на макроуровне и характеризует внешнюю активность источников нагрузки.

Базовый размер пачки представляет собой среднее количество единичных требований, составляющих одну поступающую группу. Этот параметр описывает внутреннее содержание пачки и определяет, насколько масштабным будет одномоментный наплыв требований на обслуживаемый прибор.



Коэффициент пачечности является важнейшим безразмерным структурным параметром, который отражает долю строго нулевых интервалов в общем объеме сгенерированной выборки. Коэффициент пачечности определяет внутреннюю геометрию потока и физически строго ограничен определенным интервалом, не достигающим единицы. При приближении коэффициента к верхнему пределу поток становится экстремально пачечным, что выражается в редких, но огромных наплывах требований. Если же коэффициент пачечности принимает нулевое значение, групповой поток мгновенно трансформируется в классический ординарный простейший пуассоновский поток, где все требования приходят строго поодиночке.

Подробный системный анализ: преимущества и недостатки квазипуассоновских моделей

Применение группового квазипуассоновского потока в задачах анализа и проектирования сложных систем массового обслуживания имеет как глубокие теоретические достоинства, так и определенные ограничения. Проведем детальный системный анализ данных аспектных областей.

Первым и ключевым преимуществом является высокая адекватность отображения физической реальности процессов. В отличие от абстрактного пуассоновского потока, квазипуассоновская модель позволяет достичь высокой точности аппроксимации при описании реального нестационарного, пульсирующего трафика, характерного для видеостриминга, пиринговых сетей и телеметрии. Она учитывает групповой характер передачи данных, который является базовым для современных пакетных технологий, где информация инкапсулируется в блоки.

Вторым преимуществом выступает статическое сохранение математических свойств пауз. Важнейшим аналитическим достоинством является то, что при изменении внутренних параметров группы межпачечные макроинтервалы полностью сохраняют свое статическое состояние и экспоненциальную природу. Это дает исследователю возможность контролировать общую интенсивность системы и частично использовать хорошо проработанный марковский аппарат для макроуровня, не перестраивая всю модель целиком.

Третьим преимуществом является параметрический контроль в многосвязных структурах. В системах со сложной композицией и неоднородной структурой коэффициенты пачечности и модуляции дают разработчику уникальную возможность гибко настраивать величину полей нагрузки, осуществлять точную привязку объектов моделирования к конкретным протоколам связи без необходимости полного изменения алгоритма генерации, что существенно экономит время проектирования.

Четвертым преимуществом выступает высокая вычислительная эффективность при имитационном моделировании. Генерация квазипуассоновского процесса статистическими методами требует значительно меньше вычислительных ресурсов оперативной памяти по сравнению с моделями, основанными на фрактальном анализе или процессах авторегрессии. Это позволяет проводить высокоскоростные многокритериальные эксперименты на больших выборках данных даже на персональных компьютерах средней производительности [3].

Первым существенным недостатком является резкое ограничение переносимости аналитических решений. Использование допущения о неординарности потока полностью разрушает возможность применения простых, классических замкнутых линейных формул, полученных для базовых систем с одиночным поступлением. Переход к групповому потоку требует индивидуального аналитического вывода уравнений для каждого конкретного закона модуляции пачек, что усложняет переносимость разработанных математических моделей на смежные классы систем.



Вторым недостатком выступает громоздкость математического интерфейса идентификации параметров. На практике, при анализе сырого временного ряда интервалов, классические статистические методы часто оказываются неспособными эффективно разделить смешанные характеристики. Математический интерфейс разделения истинных нулевых микроинтервалов и близких к ним макропауз выглядит чрезвычайно громоздким, а система навигации по классическим статистическим критериям часто оказывается запутанной и приводит к ложным выводам при наличии внешних шумов.

Третьим недостатком является несовершенство мер близости при сравнительном анализе. В результате применения стандартных мер статистической близости, таких как сравнение только по первому моменту – выборочному среднему, близкими по смыслу и значению могут оказаться потоки, имеющие одинаковое суммарное число требований за большой период, но по своей внутренней пачечной структуре совершенно далекие друг от друга. Это может приводить к критическим ошибкам при расчете требуемой емкости буфера систем массового обслуживания, вызывая либо его переполнение, либо избыточный простой оборудования.

Четвертым недостатком выступает идеализация мгновенного прибытия пакетов. Допущение о том, что внутренние интервалы строго равны нулю, является теоретической абстракцией. В реальных физических каналах связи пакеты внутри пачки все же разделены минимальными временными промежутками, обусловленными тактовой частотой сетевых адаптеров и физическими ограничениями среды распространения сигнала. Это вносит определенную погрешность в модели при анализе систем, функционирующих на сверхвысоких скоростях передачи данных.

Методология статистической верификации и декомпозиции потока

Одной из ключевых теоретических проблем при работе с эмпирическими данными, полученными в ходе моделирования или мониторинга реальной системы, является верификация соответствия потока заявленному квазипуассоновскому закону. Так как на практике из-за системных задержек, флуктуаций таймеров операционной системы и аппаратного джиттера микроинтервалы внутри пачки могут быть не строго равны нулю, а принимать некоторые бесконечно малые значения, возникает задача их корректного разделения.

Для решения этой задачи в методологии исследования вводится концепция адаптивного порога отсеечения. Все интервалы времени между последовательными заявками подвергаются процедуре фреймовой фильтрации. Если текущий интервал оказывается меньше или равен установленному порогу, система классифицирует его как микроинтервал, относящийся к внутренней структуре пачки. В противном случае, если интервал строго превышает пороговое значение, он признается макроинтервалом, описывающим межпачечную паузу [4].

Выбор величины адаптивного порога базируется на детальном анализе гистограммы плотности распределения интервалов и поиске точки перегиба функциональной кривой. В этой точке происходит качественный переход от микропроцессов к макропроцессам. После удаления всех микроинтервалов формируется очищенный массив макропауз, который в дальнейшем должен быть проверен на соответствие теоретическому экспоненциальному закону.

Для подтверждения гипотезы о том, что извлеченные макроинтервалы действительно подчиняются экспоненциальному распределению с заданной интенсивностью, применяется непараметрический критерий согласия Колмогорова-Смирнова. Этот метод основан на определении максимального абсолютного расхождения между эмпирической функцией распределения, построенной по экспериментальным данным, и теоретической функцией распределения.

Эмпирическая функция распределения формируется на основе накопленной частоты появления интервалов в очищенной выборке. Теоретическая функция представляет собой



идеализированную экспоненциальную кривую, определяемую интенсивностью макропотока. Статистика критерия Колмогорова фиксирует наибольшее вертикальное расстояние между этими двумя функциями во всех точках непрерывной оси.

В соответствии со статистическими теоремами, при неограниченном росте объема выборки закон распределения величины расхождения сходится к строгому закону Колмогорова. Вычисление достигаемого уровня значимости позволяет принять или отвергнуть исходную статистическую гипотезу. Если полученный уровень значимости превышает заранее выбранный критический порог, то гипотеза о квазипуассоновской структуре потока безоговорочно принимается. В противном случае гипотеза отвергается, и делается вывод о том, что поток имеет иную, более сложную стохастическую природу.

Данный методологический подход позволяет наглядно визуализировать знания о потоке. Построение графиков сравнения эмпирической и теоретической функций, а также наложение гистограмм плотности на эталонную кривую предоставляют исследователю мощный аналитический инструмент для качественной оценки результатов имитационного моделирования.

Применение квазипуассоновских моделей для проектирования и расчета систем массового обслуживания

Игнорирование группового характера потока требований и необоснованная подмена его ординарным пуассоновским процессом при проектировании буферных накопителей телекоммуникационного оборудования приводит к катастрофическим последствиям на практике. Когда пачка пакетов мгновенно поступает в систему массового обслуживания, обслуживающее устройство способно обрабатывать требования только последовательно, одно за другим, в соответствии с установленной дисциплиной обслуживания.

Это приводит к тому, что в момент прихода пачки длина очереди скачкообразно увеличивается на величину, равную размеру этой пачки. Теоретический анализ процессов в таких системах показывает, что среднее время ожидания требований в очереди напрямую зависит не только от интенсивности поступающих пачек, но и от степени изменчивости их размеров. Чем выше коэффициенты пачечности и модуляции, тем сильнее функциональная кривая средней длины очереди сдвигается вверх при тех же значениях общего коэффициента загрузки системы. При абсолютно одинаковой средней интенсивности поступающего трафика, групповой квазипуассоновский поток генерирует среднюю длину очереди и временные задержки, в несколько раз превышающие аналогичные показатели для простейшего пуассоновского потока, что наглядно иллюстрирует опасность недооценки пачечности [5].

Использование фреймового анализа квазипуассоновских процессов позволяет инженерам и системным аналитикам эффективно решать ключевые задачи обеспечения заданного уровня качества обслуживания в высоконагруженных сетях связи.

Во-первых, появляется возможность оптимизации емкости буферной памяти. Зная теоретические параметры коэффициента пачечности и базового размера группы требований, можно рассчитать минимальный объем памяти коммутатора, необходимый для предотвращения потерь пакетов при заданном верхнем пороге вероятности переполнения системы в условиях пиковых нагрузок.

Во-вторых, на основе декомпозиции параметров потока становится возможной разработка перспективных алгоритмов динамического управления трафиком и сглаживания пульсаций. Эти алгоритмы искусственно задерживают часть пакетов из пришедшей пачки, преобразуя групповой поток в близкий к ординарному, что существенно снижает пиковую нагрузку на магистральные каналы связи и повышает общую стабильность функционирования инфокоммуникационной инфраструктуры.



ЗАКЛЮЧЕНИЕ

В результате проведенного комплексного научно-теоретического исследования были подробно изучены и формализованы математическая сущность, стохастическая структура и системные свойства группового квазипуассоновского потока заявок в рамках теории массового обслуживания.

Разделение временной оси на независимые фреймовые области – микроинтервалы и макроинтервалы – позволило сформировать стройную теоретическую модель, эффективно сочетающую марковские свойства на макроуровне с адекватным отражением реального пачечного характера инфокоммуникационной нагрузки на микроуровне.

Проведенный подробный системный анализ преимуществ и недостатков квазипуассоновских моделей показал, что их высокая адаптивность, возможности параметрического контроля и вычислительная эффективность успешно компенсируют сложность теоретического вывода и громоздкость процедур разделения интервалов. Разработанная система из девяти структурных признаков, включающая робастные медианные показатели и высшие моменты распределения, формирует полный информационный паспорт трафика, позволяя полностью упорядочить значительные объемы данных.

Теоретические положения и методология статистической верификации на основе непараметрического критерия согласия Колмогорова-Смирнова, описанные в статье, представляют собой законченный аналитический базис. Он может быть непосредственно использован при проектировании высоконагруженных инфокоммуникационных узлов, оптимизации емкости буферных накопителей маршрутизаторов и разработке перспективных алгоритмов управления качеством сетевого обслуживания.

Список литературы:

1. Иванов, П.С., Козлов, А.А. Математическое моделирование систем распределения информации с пачечным характером нагрузки [Текст]: Системы управления, связи и безопасности, 2022. – Режим доступа: intelgr.com, свободный. № 4. С. 112-124.
2. Федоров, Д.В., Морозов, Н.С. Оценка показателей качества обслуживания в сетях связи с неординарным входящим трафиком [Текст]: Телекоммуникации и информационные технологии, 2023. – Режим доступа: tit-journal.ru, свободный. № 1. С. 45-53.
3. Романов, Д.С., Петрова, Е.А. Оценка параметров пачечности телекоммуникационного трафика на основе многослойного перцептрона [Текст]: Труды учебных заведений связи, 2023. – Режим доступа: tuzs-journal.ru, свободный. № 3. С. 56-64.
4. Шелухин, О.И., Костров, Д.В. Интеллектуальные методы распознавания аномалий и декомпозиции структуры сетевого трафика [Текст]: Электросвязь, 2022. – Режим доступа: elibrary.ru, свободный. № 11. С. 72-80.
5. Самойлов, А.В. Статистическая верификация и применение критериев согласия для немарковских моделей телетрафика [Электронный ресурс]: Сетевые технологии сегодня, 2024. – Режим доступа: nettech-today.ru, свободный. – Загл. с экрана.

