

Network Working Group
Request for Comments: 2330
Category: Informational

V. Paxson
Lawrence Berkeley National Lab
G. Almes
Advanced Network & Services
J. Mahdavi
M. Mathis
Pittsburgh Supercomputer Center
May 1998

Framework for IP Performance Metrics

Структура показателей производительности IP

1. Статус документа

Этот документ содержит информацию для сообщества Internet. Документ не задаёт стандартов Internet и может распространяться без ограничений.

2. Авторские права

Copyright (C) The Internet Society (1998). All Rights Reserved.

Оглавление

1. Статус документа.....	1
2. Авторские права.....	1
3. Введение.....	1
4. Критерии для показателей производительности IP.....	2
5. Термины для путей и облаков.....	2
6. Фундаментальные концепции.....	3
6.1. Показатели.....	3
6.2. Методика измерений.....	3
6.3. Измерения, погрешности и ошибки.....	4
7. Показатели и аналитическая модель.....	4
8. Эмпирически заданные показатели.....	5
9. Две формы композиции.....	5
9.1. Пространственная композиция показателей.....	5
9.2. Временная композиция формальных моделей и эмпирических метрик.....	6
10. Вопросы, связанные со временем.....	6
10.1. Проблемы часов.....	6
10.2. Время в линии.....	7
11. Одиночные измерения, выборки и статистика.....	8
11.1. Методы выборки.....	8
11.1.1. Пуассоновские выборки.....	9
11.1.2. Геометрические выборки.....	9
11.1.3. Генерация интервалов выборки Пуассона.....	9
11.2. Самосогласованность.....	10
11.3. Определение статистического распределения.....	10
11.4. Проверка пригодности.....	11
12. Отказ от стохастических показателей.....	11
13. Тип пакетов.....	12
14. Указание путей.....	12
15. Пакеты стандартного формата.....	12
16. Благодарности.....	13
17. Вопросы безопасности.....	13
18. Приложение.....	13
19. Литература.....	16
20. Адреса авторов.....	16
21. Полное заявление авторских прав.....	16

3. Введение

Целью этого документа является определение общей схемы для отдельных показателей, разработанных в рамках IETF IP Performance Metrics, начиная от рабочей группы BMWG (Benchmarking Methodology Working Group) направления Operational Requirements и продолжая группой IPPM (IP Performance Metrics Working Group) направления Transport.

Сначала определяется несколько критериев для используемых показателей. Эти критерии были разработаны для содействия усилиям IPPM по созданию максимально точного и общего понимания пользователями и провайдерами Internet аспектов производительности и надёжности сквозных путей через Internet и конкретных облаков IP, составляющих эти пути. Затем задаётся терминология, позволяющая чётко указывать компоненты Internet, такие как маршрутизаторы, пути и облака.

Затем определяются фундаментальные концепции метрики (показателя) и технологии измерения, позволяющие чётко обсуждать вопросы измерений. На основе этих концепций продолжается обсуждение важного вопроса неопределённости (погрешностей) и ошибок, а также предлагается ключевое и достаточно тонкое представление об их связи с аналитической моделью, используемой для многих аспектов инженерных дисциплин Internet. Затем вводится понятие определённых эмпирически показателей и эта часть документа завершается общим обсуждением вопроса «композиции» показателей.

В оставшейся части документа рассматриваются вопросы определения надёжных показателей и методик, учёта неточности часов, понятие «времени в линии», объединения набора одноэлементных показателей в выборки и получение из них надёжной статистики. Рассматривается, почему не следует представлять свойства Internet вероятностными параметрами (такими как вероятность отбрасывания пакета), поскольку такие термины часто неявно содержат предположения о поведении сети. Обсуждается полезность определения показателей применительно к пакетам определённого типа, преимущества использования адресов IP вместо имён DNS, а также понятие «стандартных» пакетов. В приложении обсуждается тест Anderson-Darling для оценки соответствия набора значений данному статистическому распределению и приведён код на языке C для реализации этого теста.

В некоторых разделах документа содержатся комментарии в фигурных скобках {Комментарий. ...}. Отметим, что это лишь комментарии и они не являются частью документа или предложением конкретной метрики. В некоторых случаях такие комментарии посвящены отдельным свойствам показателей, которые можно представить, но читателю следует понимать, что это обсуждение предназначено лишь для пояснения и не является предложением конкретной метрики.

4. Критерии для показателей производительности IP

Общей целью работы по показателям производительности IP является достижение состояния, когда пользователи и поставщики транспортных услуг Internet имеют единое точное представление о производительности и надёжности «облаков» Internet, которые они предоставляют или используют.

Для этого нужно разработать показатели (метрику) производительности и надёжности путей через Internet. На нескольких конференциях IETF были определены критерии для этих показателей, перечисленные ниже.

- Показатели должны быть конкретными и чётко определёнными.
- Методике для показателя следует обеспечивать повторяемость, когда при неоднократном измерении с идентичными условиями результаты также будут идентичны.
- Показатели должны обеспечивать отсутствие смещения для облаков IP с одной технологией.
- Показатели должны обеспечивать понятное и беспристрастное смещение для облаков IP с разной технологией.
- Показатели должны быть полезными для пользователей и провайдеров в понимании производительности, которая наблюдается или предоставляется.
- Показатели должны избегать искусственного влияния на производительность.

5. Термины для путей и облаков

Ниже приведены определения терминов, которые должны быть точными при разработке метрик путей. Сначала определяются термины нижнего уровня, такие как host, router, link, затем - термины path, IP cloud, exchange, которые позволяют сегментировать путь на отдельные части.

host - хост

Компьютер, способный обмениваться данными по протоколам Internet. Это понятие включает и маршрутизаторы.

link - канал, соединение

Одно соединение канального уровня между двумя (или более) хостами. Это может быть арендованная линия, Ethernet, облако Frame Relay и т. п.

router - маршрутизатор

Хост, обеспечивающий коммуникации сетевого уровня между другими хостами путём пересылки пакетов IP.

path - путь

Последовательность вида $h_0, l_1, h_1, \dots, l_n, h_n$, где $n \geq 0$, каждый из элементов h_i является хостом, каждый элемент l_i - каналом между хостами h_{i-1} и h_i , каждый хост из $h_1 \dots h_{n-1}$ является маршрутизатором. Пару $\langle l_i, h_i \rangle$ будем называть интервалом пути (пересылки) или «хопом» (hop). При подходящей настройке рабочих параметров маршрутизаторы и каналы пути обеспечивают обмен пакетами на сетевом уровне между h_0 и h_n . Отметим, что концепция задаёт односторонний путь (в одном направлении).

Subpath - субпуть

Для данного пути субпутём является любая часть данного пути, которая сама является путём (первый и последний элемент являются хостами)

cloud - облако

Неориентированный (возможно циклический) граф, вершинами которого являются маршрутизаторы, а рёбрами - каналы между парой маршрутизаторов. Формально сети Ethernet, облака Frame Relay и другие каналы, соединяющие более двух маршрутизаторов, представляются как полносвязные mesh-сети рёбер графа. Отметим, что подключение к облаку означает подключение к маршрутизатору облака через канал, который сам по себе не является частью облака.

exchange - канал обмена

Особый случай канала, служащего для прямого подключения хоста к облаку и/или одного облака к другому.

cloud subpath - облачный субпуть

Субпуть данного пути, все хосты которого являются маршрутизаторами данного облака.

path digest - дайджест пути

Последовательность вида $h_0, e_1, C_1, \dots, e_n, h_n$, где $n \geq 0$, h_0 и h_n - хосты, каждый из $e_1 \dots e_n$ является каналом обмена, а каждый $C_1 \dots C_{n-1}$ - облачным субпутём.

6. Фундаментальные концепции

6.1. Показатели

В рабочей среде Internet имеется несколько количественных показателей, связанных с производительностью и надёжностью Internet, значения которых хотелось бы знать. Когда такие величины чётко заданы их называют показателями (метрикой). Предполагается, что для каждой метрики (или группы тесно связанных показателей) будет подготовлен специальный документ RFC.

В некоторых случаях может не быть ясных способов для эффективного измерения показателей, это допустимо, а иногда может быть очень полезным. Однако нужно, чтобы спецификация метрики была как можно более чёткой в части определяемой величины. Таким образом, иногда допускаются сложности при практических измерениях, но неприемлема неоднозначность смысла.

Каждая метрика определяется с использованием стандартных единиц измерения. Применяются международные метрические системы с учётом отмеченных ниже моментов.

- Когда величины задаются расстоянием (дистанция, размер) или временем в секундах (продолжительность), приемлемы соответствующие тысячам или тысячным долям единицы. Т. е. расстояния можно указывать в километрах (км), время - в миллисекундах (мсек) или микросекундах (мксек), но нельзя указывать расстояние в сантиметрах, поскольку это не является тысячной долей метра.
- Когда величины задаются комбинацией единиц, можно применять производные единицы в тысячах или тысячных долях, но все тысячи или тысячные доли должны быть в начале. Таким образом, можно применять километры в секунду (км/с), но недопустимы метры в миллисекунду.
- Единицей информации является бит.
- При использовании метрических префиксов с битами или производными от них единицами эти префиксы имеют обычный метрический смысл (1000), а не принятый для компьютерных хранилищ коэффициент 1024. Во всех RFC, определяющих показатели с такими единицами, это соглашение будет соблюдаться и повторяться для ясности.
- При указании времени используется часовой пояс UTC (0).

Отметим, что сказанное относится к спецификациям метрики, а не к форматам пакетов, где использование октетов явно предпочтительней использования битов.

Некоторые показатели могут определяться исключительно в терминах других показателей, такую метрику называют производной (derived metric).

6.2. Методика измерений

Для данного набора чётко определённых показателей может существовать множество методик измерения. Некоторые примеры представлены ниже.

- Прямое измерение показателя производительности с использованием внедряемого трафика. Например, измерение круговой задержки пакета IP заданного размера на заданном пути в данный момент.
- Проекция показателя из измерений нижележащего уровня. Например, можно определить полную задержку в пути для пакета IP заданного размера на основе точного измерения задержки распространения и пропускной способности на всех этапах пути.
- Оценка отдельного показателя на основе набора более агрегированных измерений. Например, можно определить задержку канала на одном этапе пути (one-hop) на основе точного измерения задержки на этом этапе для пакетов IP разного размера.
- Оценка данного показателя в определённый момент на основе набора измерений связанных показателей в другое время. Например, можно оценить пропускную способность в данный момент на основе точного измерения задержки в прошлом и сейчас, а также динамической модели потока данных.

Этот список не является исчерпывающим и приведён лишь для иллюстрации разнообразия методов измерения.

Для конкретного показателя можно отмечать и обсуждать подход к измерению, однако обычно подход не является частью спецификации.

Методике для показателя следует обладать свойством повторяемости и при нескольких измерениях в идентичных условиях результаты должны быть согласованными.

Несколько отступая от слова «идентичных» в предыдущем абзаце, можно более точно использовать слово «непрерывность» для описания свойства методики - методика для данного показателя демонстрирует непрерывность, если при незначительных изменениях условий результаты измерения также не будут существенно изменяться. Выражаясь чуть точнее, следует сказать, что для каждого положительного значения epsilon имеется положительное значение delta, такое, что при двух наборах условий в диапазоне delta, результаты обоих измерений будут попадать в диапазон epsilon. На данный момент следует считать это эвристикой, управляющей интуитивным представлением о «здоровом смысле» (robustness property), а не точным понятием.

Метрику, имеющую хотя бы одну методику со свойством непрерывности, считают непрерывной.

Некоторые параметры, такие как число интервалов пересылки (hop-count) на пути, имеют целочисленные значения и не могут быть непрерывными в указанном выше смысле.

Отметим, что может оказаться непрактичным выяснение (или возможность оценки) условий, относящихся к измерению в данный момент. Например, по причине быстрых значительных вариаций нагрузки (числа обрабатываемых пакетов) в данном маршрутизаторе сильно нагруженной высокоскоростной WAN-сети внешнему наблюдателю может быть очень

сложно выполнить количественные оценки данных, а статистика может оказаться более повторяемой или непрерывной. В таких случаях для показателя следует задавать подходящую статистику.

Некоторые методики измерений могут быть «консервативными» в том смысле, что акт измерения не меняет или меняет слабо значение показателя производительности, которое методика пытается измерить.

{Комментарий. Например в высокоскоростной WAN-сети при умеренной загрузке тест с использованием нескольких мелких пакетов ping для измерения задержки не будет сильно влиять на величину задержки для остального трафика. Для тестов с большим потоком измерительных данных это явно не будет работать.}

6.3. Измерения, погрешности и ошибки

Даже самая лучшая методика измерения для самого лучшего определения поведения метрики будет давать ошибки. Однако разработчикам методик измерения следует стремиться:

- минимизировать погрешности и ошибки;
- находить и документировать источники погрешностей и ошибок;
- количественно оценивать погрешности и ошибки.

Например, при разработке метода определения задержки нужно разобраться, какие ошибки будут вносить погрешности часов и как можно точнее оценить их влияние. В некоторых случаях это приведёт к заданию требований к часам, применяемым для измерений.

В качестве другого примера рассмотрим ошибку синхронизации, связанную с издержками измерительного компьютера, а не участвующими в измерении компонентами Internet. Первое в данном случае является ошибкой измерения, а второе отражает измеряемый показатель. Отметим, что одним из способов предотвращения таких ошибок является использование фильтра и анализатора пакетов (filter/sniffer), работающего на отдельном компьютере, который записывает пакеты и точно указывает временные метки (см. обсуждение wire time ниже). Затем полученную трассировку можно проанализировать для оценки тестового трафика с минимизацией влияния задержек на измерительном хосте или хотя бы их учётом. Этот метод может оказаться полезным даже при работе фильтра и анализатора на той же машине, поскольку такие измерения обычно предоставляют временные метки на уровне ядра, а не менее точные метки на уровне приложения.

Отметим также, что производные показатели, определённые выше, или показатели с пространственной или временной композицией, определённые ниже, предоставляют особый интерес для анализа погрешностей измерения, а именно, распространение погрешностей при выводе или составлении показателя.

7. Показатели и аналитическая модель

По мере развития Internet от первых исследований коммутации пакетов в 1960-х годах инженерное сообщество Internet разработало общую аналитическую основу для концепций. Эту аналитическую модель или A-frame (analytical framework) применяют проектировщики и разработчики протоколов, участники измерений, а также исследователи производительности компьютерных сетей, используя инструменты моделирования и анализа обеспечивающие преимущества при работе. Основной целью этого документа является определение характеристик сети, согласованных с аналитической и практической точки зрения, поскольку это повысит шансы более эффективного соответствия неэмпирических исследований реальному поведению сети и их использования для лучшего понимания сетей.

Поэтому здесь предпринимаются усилия по использованию и развитию модели A-frame. Всякий раз, когда предполагается, что определяемая метрика тесно связана с концепциями A-frame, предпринимаются попытки задания метрики в терминах A-frame. Такая спецификация будет развивать A-frame за счёт точного определения концепций, требуемых для метрики, а затем A-frame будет применяться для задания метрики в терминах этих концепций.

Такие показатели будем называть заданными аналитически (analytically specified metric) или просто аналитическими.

{Комментарий. Примеры таких аналитических показателей приведены ниже.

propagation time of a link - время распространения по каналу

Время (в секундах) требуемое для прохождения 1 бита из выходного порта на одном хосте Internet по одному каналу на другой хост Internet.

bandwidth of a link for packets of size k - пропускная способность канала для пакетов размером k

Пропускная способность (бит/с) с учётом лишь битов пакета IP для пакетов размером k байтов.

Route - маршрут

Путь из точки A в точку B в данный момент, как определено в разделе 5.

hop count of a route - число интервалов пересылки на маршруте

Значение n для маршрута (пути).

}

Отметим, что здесь не составляется список концепций A-frame для включения в спецификации, но поощряется их использование и рекомендуется тщательно определять концепции, чтобы по мере разработки набора показателей развивался технически согласованный с ним набор концепций A-frame и оба были согласованы с общим пониманием концепций в сообществе Internet.

Эти концепции A-frame предназначены для абстрагирования от реальных компонентов Internet при условии, что:

- основные свойства компонента сохраняются;
- сохраняются создаваемые свойства компонента, относящиеся к показателям;
- подмножество этих компонентов потенциально определяется как аналитические показатели;
- эти свойства компонентов Internet не связаны с определениями создаваемой или отбрасываемой метрики.

Например, при рассмотрении маршрутизатора в контексте пересылки пакетов, можно моделировать маршрутизатор как элемент, принимающий пакеты на входном канале, помещающий их в очередь FIFO конечного размера с

отбрасыванием пакетов в конце очереди при её переполнении и пересылающий пакеты в выходной канал. Скорость передачи (в бит/с) и максимальный размер очереди (в битах) являются аналитическими показателями.

В некоторых случаях такие аналитические методы, применяемые для маршрутизаторов, будут тесно связаны с конкретными показателями производительности путей Internet. Например, очевидная формула $(L + P/B)$, включающая задержку в маршрутизаторе (L), размер пакета в битах (P) и скорость передачи выходного канала (B) может давать точную оценку роста задержки пакетов в результате размещения данного маршрутизатора на пути.

Подчёркивается, что хорошо подобранные и определённые концепции A-frame и их аналитические параметры будут поддерживать более общие усилия по созданию показателей менее очевидными способами.

{Комментарий. Например, при рассмотрении пропускной способности пути может иметь реальное значение возможность моделирования каждого маршрутизатора на пути пересылки пакетов, как отмечено выше. Методы оценки пропускной способности пути могут использовать максимальный размер очереди пакетов в качестве параметра неочевидными способами. Например, при росте максимального размера очереди растёт и способность маршрутизатора передавать трафик в выходной канал, несмотря на флуктуации трафика во входном канале. Оценка такого роста требует исследования.}

Отметим, что при задании концепций A-frame и аналитических показателей неизбежны упрощающие предположения. Ключевая роль этих концепций состоит в абстрагировании от свойств компонентов Internet, связанных с данной метрикой. Следует избегать допущений, которые задают «крен» работ по моделированию и показателям в сторону одного типа.

{Комментарий. Например, маршрутизаторы могут не использовать tail-drop, хотя это отбрасывание упрощает аналитическое моделирование.}

Отметим, что разные элементы в A-frame могут принимать для простоты различные допущения. Например, абстракция маршрутизатора, использованная для определения задержки в пути, может считать пакетные очереди маршрутизатора одной очередью FIFO, а абстракция маршрутизатора для определения обработки пакетов с поддержкой RSVP, может считать очередь маршрутизатора буфером с ограниченной задержкой. Эти допущения противоречат одно другому, но это не говорит о том, что принимаются противоречивые допущения в одно и то же время. Здесь речь может идти о разных частях работы, где эти допущения могут упростить базовую концепцию двумя разными способами для двух различающихся целей.

{Комментарий. В более строгих терминах можно сказать, что A-frame в целом не обязательно является самосогласованной моделью, но набор конкретных элементов A-frame, применяемых для определения конкретной метрики, должен быть согласованным.}

8. Эмпирически заданные показатели

Имеются полезные показатели производительности и надёжности, которые не вписываются точно в A-frame обычно по причине недостатка в A-frame деталей или возможностей для работы с ними. Например, для «наилучшей пропускной способности на пути с использованием протокола TCP, соответствующего RFC 2001» было бы хорошо иметь возможность измерения, но нет подходящей аналитической модели для представления пропускной способности аналитической метрикой.

Эти понятия можно определить путём чёткого описания методики их измерения.

Такие показатели будем называть эмпирически заданными (empirically specified) или просто эмпирическими.

Эмпирическим параметрам следует иметь три свойства:

- достаточно чёткое определение с точки зрения компонентов Internet;
- хотя бы один эффективный способ измерения;
- по возможности обеспечивать (обязательно неполное) понимание метрики в терминах A-frame, чтобы можно было использовать измерения для обсуждения производительности и надёжности компонентов A-frame и их агрегатов.

9. Две формы композиции

9.1. Пространственная композиция показателей

В некоторых случаях может быть реалистично и полезно определять метрику так, чтобы она отражала пространственную композицию. Пространственной композицией считается характеристика некоторых показателей пути, где показатель для (полного) пути может быть определён также для разных субпутей и соответствующие концепции A-frame для метрики предлагают полезные взаимосвязи с применимыми к этим субпутям метриками (включая полный путь, различные субпути в облаке для данного дайджеста пути и даже отдельные маршрутизаторы на пути). Эффективность пространственной композиции зависит от:

- полезности анализа этих взаимосвязей применительно к соответствующим компонентам A-frame;
- практического применения соответствующих соотношений для показателей и методики измерений.

{Комментарий. Рассмотрим в качестве примера метрику для задержки 100-байтового пакета на пути P , а затем рассмотрим дайджест пути $\langle h_0, e_1, C_1, \dots, e_n, h_n \rangle$ для P . Определение метрики может включать допущение о том, что задержка на пути P очень близка к сумме соответствующих показателей для каналов обмена (e_i) и облаков (C_i) данного дайджеста пути. Определение метрики дополнительно включает примечание о применении взаимоотношения к соответствующим компонентам A-frame на пути P , а также каналах обмена и облаках из дайджеста пути.}

Когда определение метрики включает допущение о связи метрики всего пути с метрикой субпутей, это говорит о том, что метрика поддерживает пространственную композицию. В определение следует включать:

- конкретное допущение, применяемое к метрике;
- обоснование полезности композиции на практике в части повышения точности измерений на пути;
- обоснование полезности композиции на практике в части повышения эффективности анализа пути с использованием концепций A-frame;
- анализ возможных некорректностей композиции.

9.2. Временная композиция формальных моделей и эмпирических метрик

В некоторых случаях реалистичным и полезным может быть определение показателей так, чтобы они отражали временную композицию. Временной композицией считается характеристика той или иной метрики пути, где метрика применительно к пути в текущий момент T также определена для различных моментов $t_0 < t_1 < \dots < t_n < T$, и в которой подходящие концепции A-frame для метрики предполагают полезные соотношения между показателями, определёнными в моменты t_0, \dots, t_n и показателем, применимым в момент T . Эффективность временной композиции зависит от:

- полезности анализа этих взаимосвязей применительно к соответствующим компонентам A-frame;
- практического применения соответствующих соотношений для показателей и методики измерений.

{Комментарий. В качестве примера рассмотрим метрику для ожидаемой пропускной способности на пути P в 5-минутном интервале вокруг момента T и предположим, что имеются соответствующие значения для каждого из 4 предшествующих 5-минутных интервалов t_0, t_1, t_2, t_3 . Определение такой метрики может включать допущение о том, что пропускную способность в момент T можно оценить некой экстраполяцией значений для t_0, \dots, t_3 . Определение будет также включать замечание о применении соответствующего соотношения к компонентам A-frame.

Примечание. Любая (пространственная или временная) композиция, включающая пропускную способность, вероятно будет «коварной», а временная - «коварной» пространственной, поэтому читателю следует понимать, что приведённый пример намеренно выбран примитивным.}

Когда определение метрики включает допущение о связи метрики пути в данный момент T с метрикой этого пути в другие интервалы времени, это говорит о том, что метрика поддерживает временную композицию. В определение следует включать:

- конкретное допущение, применяемое к метрике;
- обоснование полезности композиции на практике в части повышения точности измерений на пути;
- обоснование полезности композиции на практике в части повышения эффективности анализа пути с использованием концепций A-frame;

10. Вопросы, связанные со временем

10.1. Проблемы часов

Измерение времени лежит в основе многих показателей Internet. Поэтому при разработке методики измерения показателей часто решающее значение будут иметь ошибки и погрешности, связанные с несовершенными часами. В этом параграфе задаётся терминология для обсуждения характеристик часов и рассматриваются связанные с часами вопросы измерения, которые нужно решать с помощью надёжной методики.

Протокол сетевого времени (Network Time Protocol или NTP, RFC 1305) задаёт номенклатуру для обсуждения характеристик часов, которая будет применяться при необходимости [Mi92]. Основная задача NTP состоит в обеспечении точности часов в достаточно больших интервалах времени (от минут до дней), тогда как для измерений часто важнее обеспечение краткосрочной точности между началом и завершением измерений или в процессе измерения (выборки). Различие в целях иногда ведёт к различным определениям терминов, как отмечено ниже.

Сначала определим смещение (offset) часов в конкретный момент как разность между временем по часам и «истинным» временем UTC. Если часы показывают время T_c при истинном времени T_t , смещение составляет $T_c - T_t$.

Будем называть часы точными в определённый момент, если смещение равно 0, а в более общем смысле будем характеризовать точность часов величиной смещения. В NTP термин «точность» учитывает также «частоту» часов, а для измерительных целей вместо него будем применять термин «перекос» (skew), поскольку точность определяется для конкретного момента, а не для интервала. Перекос часов в определённый момент указывает разницу «частот» (первая производная смещения относительно точного времени) между часами и истинным временем.

Как отмечено в RFC 1305, в реальных часах наблюдаются вариации перекоса. Т. е. вторая производная смещения относительно истинного времени в общем случае отличается от 0. В соответствии с RFC 1305 эту величину будем называть дрейфом часов (drift).

Шаг (дискретность) часов (resolution) - это наименьший интервал, на который обновляются показания часов. Он задаёт нижнюю границу погрешности часов (отметим, что часы с мелким шагом не обязательно являются точными). Шаг часов указывается в секундах. Однако дискретность связана с сообщаемым часами, а не истинным временем, поэтому разрешение в 10 мсек означает лишь, что показания часов растут с шагом в 0,01 секунды, но это не связано с интервалом обновления показаний часов.

{Комментарий. Системы различаются способом передачи прикладным интерфейсом информации в случаях, когда показания часов не изменились с момента предыдущего вызова. Некоторые просто возвращают прежнее время, а другие могут инкрементировать возвращаемое время для поддержки монотонно возрастающих меток времени. Для систем второго типа при рассмотрении дискретности часов эти мелкие приращения не учитываются здесь. Фактически они препятствуют оценке реальной дискретности часов, поскольку единственным способом определения шага является повторяющиеся запросы для определения минимального приращения.}

Предполагается, что дискретность часов меняется очень редко (например, при обновлении оборудования).

Есть ряд интересных показателей, для которых естественные методики измерения включают сравнение времени от двух разных часов. Примером является задержка пакетов в одном направлении [AK97]. Здесь время прохождения пакета через сеть измеряется сравнением часов на одной стороне пути пакета при входе пакета в сеть с часами на другой стороне пути при выходе пакета из сети.

Поэтому представляют интерес также термины для сравнения двух часов $C1$ и $C2$. Для этого вводятся термины, связанные с указанными выше, в которых понятие истинного времени заменено показаниями часов $C1$. Например, смещение $C2$ относительно $C1$ в определённый момент составляет $Tc2 - Tc1$ (разность в показаниях часов $C2$ и $C1$ в один момент). Чтобы не возникало путаницы со смещением часов относительно истинного времени, в этом случае будем говорить об относительном смещении (расхождении) часов. Таким образом, указанное выше смещение является «относительным смещением» между часами $C2$ и $C1$.

При сравнении часов аналогом шага является не «относительный шаг», а общая дискретность (joint resolution), определяемая суммой шага часов $C1$ и $C2$. Общая дискретность (шаг) показывает консервативную оценку нижней границы точности интервалов, рассчитываемых вычитанием показаний одних часов из показаний других.

Если показания двух часов совпадают, их называют синхронизированными. Отметим, что синхронизация пары часов никак не связана с их смещением относительно истинного времени. Важно понимать, что для многих измерений в Internet синхронизация участвующих в измерении часов важнее их абсолютной точности. В определённой степени это относится и к перекосам - пока абсолютный перекося не слишком велик, важнее минимальный относительный перекося, поскольку он может вносить систематические ошибки при измерении времени прохождения пакетов путём сравнения временных меток от пары часов.

Отмеченные особенности связаны с тем, что для измерений в Internet зачастую важна разница во времени двух часов. При расчёте разницы между показаниями часов отклонения их от истинного времени можно учесть (скомпенсировать), но очень важно, чтобы ход обоих часов был как можно более близким.

Методики измерения часто начинаются с синхронизации пары часов и минимизации перекося и дрейфа.

{Комментарий. Эффективным способом выполнения этих условий (и точности часов) является использование часов, получающих значение времени от внешнего источника, а не только от часов хост-компьютера (они часто подвержены значительным ошибкам). Кроме того, предпочтительно, чтобы часы получали значение времени напрямую, например от модуля глобальной системы позиционирования GPS.}

При синхронизации часов от сети по таким протоколам как NTP возникают две проблемы.

1. Точность NTP частично зависит от свойств (в частности, задержки) путей Internet, используемых партнёрами NTP, и это могут быть именно те свойства, которые нужно измерить, поэтому будет неразумно использовать NTP для калибровки измерений.
2. Протокол NTP сосредоточен на точности часов, которая может обеспечиваться за счёт кратковременного перекося и дрейфа часов. Например, при опосредованной синхронизации часов с источником точного времени со значительным интервалом синхронизации, хост иногда будет сталкиваться с задачей корректировки своего текущего времени T_i по существенно отличающемуся более точному времени, которое он только что получил (T_a). Это можно сделать двумя базовыми способами - сразу установить для текущего времени значение T_a или корректировать частоту обновления локальных часов (и перекося), чтобы в какой-то момент локальное время T_i' соответствовало более точному времени T_a' . Первый механизм ведёт к скачкам времени и может нарушать допущение о монотонности роста временных меток. Если часы переводятся назад, это иногда достаточно просто обнаруживается, но перевод часов вперёд заметить сложнее. Перекося, вносимый вторым механизмом, может приводить к значительным ошибкам при расчёте разницы во времени, как отмечено выше.

Для иллюстрации важности проблемы перекося рассмотрим выборку задержки в одном направлении между двумя хостами Internet, выполняемую с интервалом в 1 минуту. Истинная задержка передачи для трансконтинентального пути может составить около 50 мсек. При перекося между парой часов в 0,01% (т. е. 1/10000) после 10 минут измерений ошибка достигнет 60 мсек. Если не исправлять ошибку, она полностью исказит результаты измерений задержки. Отметим, что оценка перекося между двумя несинхронизированными часами в сети требует исследования (см. [Pa97], где рассматривается обнаружение и компенсация такого типа ошибок). Подобные ошибки делают очень важной синхронизацию часов от независимого источника, такого как GPS.

10.2. Время в линии

Измерения в Internet часто осложняются использованием для измерений самих хостов Internet. Эти хосты могут вносить задержки, создавать «пробки» (bottleneck) и т. п. из-за особенностей оборудования или операционных систем, не связанные с измеряемыми показателями сети. Особенно обостряется эта проблема при задании временных меток событий в сети на уровне приложений.

Для обсуждения этой проблемы в целом вводится два толкования времени в линии (wire time), определяемые для хоста Internet H и канала Internet L в конкретном месте.

- Для данного пакета P время прибытия по линии (wire arrival time) на хост H из канала L определяется моментом T , когда какой-либо бит P появляется в точке наблюдения H на канале L .
- Для данного пакета P время выхода из линии (wire exit time) на хосте H канала L определяется моментом T , когда все биты P появились в точке наблюдения H на канале L .

Отметим, что эти определения тесно связаны с точкой наблюдения на канале. Это важно, поскольку для каналов с большой задержкой можно получить разное время в зависимости от размещения точки наблюдения. Можно разрешить установку точки наблюдения в произвольном месте канала, тем не менее, определение привязано к хосту Internet, поскольку предполагается, что на практике для показателей IPPM все временные характеристики будут привязаны к хостам Internet, а не к специальным устройствам, которые способны следить за каналом в точках, где нет хостов. Это также связано с проблемой соединений, включающих несколько физических каналов. Поскольку эти каналы не видны на уровне IP, их наблюдение в соответствии с приведёнными определениями невозможно.

Возможно, хотя это и происходит нечасто, многократное прохождение пакета Р по определённому каналу L в результате петли в пересылке. Такие прохождения могут даже перекрываться в зависимости от канальной технологии. Когда это происходит, время в линии определяется отдельно для каждого экземпляра Р, видимого на канале в позиции Н. Это определение следует выполнять, поскольку оно служит напоминанием о ненадёжности таких понятий как время прохождения пакета через определённую точку в Internet.

Термин *wire time* исторически применялся для обозначения времени, в которое пакет появляется на канале без точного указания первого бита, последнего бита или иных соображений. Это неформальное определение стало уже почти бесполезным, поскольку время в линии обычно применяется, чтобы различать, когда задержки распространения пакетов начинаются и заканчиваются из-за сети, а не конечных хостов.

Когда это уместно, показатели следует задавать в терминах времени в линии, а не времени конечных хостов, чтобы определение выводило на первый план вопрос разделения задержек в сети и на хостах.

Отметим наличие потенциальной сложности при использовании времени в линии, связанной с фрагментами IP. Может случиться так, что в результате фрагментации через точку Н пройдёт лишь часть определённого пакета. Такие фрагменты допустимы сами по себе и с ними связано чётко определённое время в линии, но для пакета, из которого созданы фрагменты, смысл времени в линии утрачивается.

Отметим также, что авторам неизвестно о введении этих терминов для трафика Internet, поэтому в будущем может потребоваться корректировка определений.

{Комментарий. Время в линии иногда трудно измерить. Одним из методов является применение фильтров для отслеживания трафика на канале. Архитектура таких фильтров часто пытается связать с каждым пакетом временную метку как можно близкую к времени в линии. Однако следует отметить, что одним из основных источников ошибок является запуск пакетного фильтра на одном из конечных хостов. В этом случае некоторые фильтры получают для отдельных пакетов временные метки, соответствующие моменту «запланированной» передачи пакета в сеть, а не времени фактической передачи (время в линии). Эти значения могут существенно различаться. Для решения этой проблемы можно запустить фильтр на отдельном хосте для пассивного мониторинга данного канала, однако для некоторых технологий связи это может быть проблематично. В работе [Pa97] рассмотрены ошибки, которые могут возникать в фильтрах пакетов. Отметим также, что фильтры часто захватывают лишь первый фрагмент пакета IP, поскольку в фильтрах используются поля транспортных заголовков, отсутствующие во всех фрагментах, кроме первого. Поскольку обычно стремятся избежать в методологии измерений сложностей, связанных с фрагментированием, одной из стратегий работы при наличии фрагментов, обнаруживаемых фильтром пакетов, является установка флага указывающего необычный трафик, с целью его исключения из анализа времени для пакетов.}

11. Одиночные измерения, выборки и статистика

Опыт показывает, что полезно разделить три связанных, но разных понятия в части метрики.

- Одноэлементной (*singleton*) метрикой называются показатели, которые по сути неделимы (атомарны). Например, единичный экземпляр «общей пропускной способности» от одного хоста к другому можно определить как *singleton*-метрику, даже несмотря на то, что измерение включает множество пакетов Internet.
- Выборкой (*sample*) называют показатель, полученный из данной одноэлементной метрики путём объединения нескольких экземпляров измерений. Например, можно считать выборкой одностороннюю задержку от одного хоста к другому, полученную при часовых измерениях в интервалах Пуассона 1 раз в секунду.
- Статистическим называют показатель, выведенный из данной выборки путём расчёта той или иной статистики для значений, определённых одиночными (*singleton*) измерениями в выборке. Например, это может быть среднее значение односторонней задержки для упомянутой выше выборки.

Согласованное применение одиночных измерений, выборки и статистики позволит использовать извлечённые уроки определения выборки и статистики для различных показателей. Ортогональность этих метрических понятий делает работу более эффективной и понятной для сообщества.

Далее в этом разделе рассматриваются некоторые вопросы выборки и статистики, сочтённые важными для определения показателей и выполнения измерений.

11.1. Методы выборки

Основной причиной применения выборок является потребность в просмотре вариаций и взаимозависимостей значений измеряемого показателя. Вариации могут быть связаны с разными точками в Internet или разными моментами измерения. При оценке вариаций в одной выборке обычно принимается допущение о несмещённой (*unbiased*) выборке, когда процесс измерения не искажает выборку так, что последующие измерения перестают точно отражать вариации и согласованность показателя.

Одним из распространённых способов выборки является измерение через равные интервалы времени - периодическая выборка. Этот метод привлекает своей простотой, но имеет две возможные проблемы, описанные ниже.

- Если поведение измеряемого показателя является периодическим, выборка будет показывать лишь часть поведения, когда периоды изменения и выборки совпадают или кратны. С этой проблемой связано представление о том, что периодическую выборку легко предсказать, а предсказуемые выборки подвержены манипуляциям, с помощью которых поведение компонентов сети может быть изменено так, что выборка покажет только изменённое поведение.
- Процесс измерения может влиять на измеряемый показатель (например, внедрение измерительного трафика в сеть меняет уровень загрузки сети), а периодические возмущения могут приводить к синхронизации (см. [FJ94]), значительно усиливая эффекты, которые сами по себе второстепенны.

Более надёжен подход на основе случайной аддитивной выборки, когда интервалы между измерениями задаются случайно с общим статистическим распределением $G(t)$ [BM92]. Качество такой выборки зависит от $G(t)$. Например, если $G(t)$ генерирует постоянное значение g с вероятностью 1, выборка сводится к измерению с интервалом g .

Случайная аддитивная выборка даёт значительные преимущества. В общем случае она позволяет предотвратить ненужную синхронизацию и даёт объективную оценку искомого свойства. Недостатки метода указаны ниже.

- Усложняется частотный анализ, поскольку выборки происходят с переменными интервалами, а методы преобразований Фурье предполагают фиксированный интервал.
- Если $G(t)$ не является экспоненциальным распределением (см. ниже), выборка остаётся в определённой степени предсказуемой, как было отмечено выше для периодической выборки.

11.1.1. Пуассоновские выборки

Можно доказать, что для экспоненциального распределения $G(t)$

$$G(t) = 1 - \exp(-\lambda * t)$$

поступление новых выборок невозможно предсказать (и выборка становится несмещенной). Кроме того, выборка асимптотически приближается к несмещенной, даже при её влиянии на состояние сети. Такие выборки называют пуассоновскими. Они не склонны к синхронизации и могут служить для точного сбора измерений периодического поведения, а также не подвержены манипуляциям за счёт предсказания времени новой выборки.

Благодаря этим ценным свойствам выборки в Internet обычно предпочитают выполнять с использованием распределения Пуассона.

{Комментарий. Отметим, что во многих случаях предпочтительно использовать другие $G(t)$. Например, экспоненциальное распределение не ограничено, поэтому могут возникать длительные интервалы между выборками. Можно ограничить максимальный интервал значением dT для более быстрого схождения оценки на основе выборки. Это можно сделать с помощью однородного распределения из интервала от 0 до dT

$$G(t) = \text{Unif}(0, dT)$$

Такая выборка становится предсказуемой, если интервал длительностью почти dT уже прошёл, а выборки ещё нет.}

В чистом виде выборка Пуассона выполняется путём генерации независимых интервалов с экспоненциальным распределением и одного измерения по истечении каждого интервала. Можно показать, что при выборке Пуассона с момента T с интервалом dT , в течение которого происходит N измерений, эти измерения будут равномерно распределены в интервале $[T, T+dT]$. Другой вариант выполнения выборки Пуассона состоит в выборе значения dT и N и проведения N выборок с равномерным распределением в интервале $[T, T+dT]$. Эти варианты эквивалентны за исключением случаев, когда значения N и dT доступны извне. В этом случае невозможность предсказания времени выборки ослабляется, но остальные свойства сохраняются. Подход N/dT имеет преимущество, поскольку применение фиксированных значений N и dT может быть проще работы с фиксированным значением λ и переменным числом измерений в интервалах переменной длительности.

11.1.2. Геометрические выборки

С выборкой Пуассона тесно связана «геометрическая выборка», где внешние события измеряются с фиксированной вероятностью p . Например, можно собирать в канале все пакеты, но записывать пакет в файл трассировки лишь при условии, что сгенерированное случайное значение из интервала 0 - 1 меньше заданного p . Геометрическая выборка имеет такие же свойства несмещённости и непредсказуемости как выборка Пуассона, поэтому надёжна для измерений в Internet, если подходит для определённой задачи. Дополнительные сведения приведены в [CPB93].

11.1.3. Генерация интервалов выборки Пуассона

Для генерации интервалов выборки Пуассона сначала определяется значение λ , задающее среднюю частоту выполнения отдельных (singleton) измерений (например, для среднего интервала выборки в 30 секунд устанавливается значение $\lambda = 1/30$, если время измеряется в секундах). Затем генерируется серия экспоненциально распределённых (псевдо)случайных значений E_1, E_2, \dots, E_n . Первое измерение выполняется в момент E_1 , второе - в момент E_1+E_2 и т. д.

Один из методов генерации экспоненциально распределённых (псевдо)случайных чисел основан на способности генерировать (псевдо)случайные числа U_1, U_2, \dots, U_n с однородным распределением в интервале от 0 до 1. Библиотеки для этого имеются на многих компьютерах. На основе таких U_i для создания E_i используется формула

$$E_i = -\log(U_i) / \lambda$$

где $\log(U_i)$ - натуральный логарифм U_i .

{Комментарий. Этот метод является представителем более широкого класса методов «инверсного преобразования» для генерации случайных чисел с заданным распределением.}

Детали реализации описаны ниже.

Имеется по меньшей мере 3 разных метода аппроксимации выборки Пуассона. Метод 1 проще всего в реализации, но имеет самые большие ошибки, а метод 3 сложнее реализовать, но ошибки в нем значительно меньше или отсутствуют.

Метод 1.

1. Генерируется значение E_1 и выполняется ожидание.
2. Происходит измерение.
3. Генерируется значение E_2 и выполняется ожидание.
4. Происходит измерение.
5. Генерируется значение E_3 и выполняется ожидание.

6. Происходит измерение ...

Проблема этого подхода состоит в том, что этап измерения тоже занимает время, поэтому выборки происходят не в моменты $E1$, $E1+E2$ и т. д., а в моменты $E1$, $E1+M1+E2$ и т. д, где M_i указывает время i -го измерения. Если значение M_i очень мало по сравнению с $1/\lambda$, вносимая этим методом ошибка также будет мала. Однако если M_i составляет заметную часть $1/\lambda$, возможная ошибка возрастает.

Метод 2 пытается исправить отмеченную ошибку, учитывая время, затрачиваемое на измерение (M_i) и корректируя интервалы ожидания:

1. Генерируется значение $E1$ и выполняется ожидание.
2. Выполняется измерение и учитывается затраченное на него время $M1$.
3. Генерируется значение $E2$ и выполняется ожидание в течение $E2-M1$.
4. Выполняется измерение и учитывается затраченное на него время $M2$...

Этот метод работает достаточно хорошо, пока $E_{i+1} \geq M_i$, но при $E_{i+1} < M_i$ невозможно ждать требуемое время (в этом случае требуется выполнить одновременно два измерения).

Метод 3 планирует интервалы измерения и придерживается этого правила.

1. Генерируются интервалы $E1, E2, \dots, E_n$.
2. Рассчитывается время измерения $T1, T2, \dots, T_n$ как $T_i = E1 + \dots + E_i$.
3. Выполняются измерения в моменты $T1, T2, \dots, T_n$.

Разрешая одновременные измерения, метод 3 избавляется от недостатков методов 1 и 2. Однако если одновременные измерения мешают одно другому, метод 3 не даёт преимуществ и может оказаться хуже методов 1 и 2.

Для Internet неизвестно заранее, в какой степени будут влиять погрешности этих методов. Если M_i много меньше $1/\lambda$, подойдёт любой из методов. Если же M_i меньше $1/\lambda$, но возможно несущественно, метод 2 предпочтительней метода 1. Если одновременные измерения не мешают друг другу, предпочтителен метод 3, но его реализация может быть значительно сложнее.

11.2. Самосогласованность

Фундаментальное требование к методике надёжных измерений состоит в том, что измерения выполняются с применением как можно меньшего числа неподтвержденных допущений. Горький опыт показал, насколько легко принять (зачастую неявно) допущение, которое оказывается неверным. Примером является включение в измерения показаний часов, синхронизированных с точным источником. Легко предположить, что эти часы обеспечивают точность, но в результате программных ошибок, потери питания эталонным источником или связи с ним может возникнуть фактическая неточность часов.

Это не означает отказа от любых допущений при измерениях, а говорит о необходимости проверки предположений, насколько это возможно на практике. Такие проверки применимы как к наблюдаемым при измерениях величинам, так и к значениям, которые использует сам измерительный процесс. Простым примером для первого является проверка среднего значения времени кругового обхода, которое должно быть положительным и отрицательное значение говорит об ошибке. Поскольку отрицательные интервалы времени не имеют физического смысла, получение такого значения должно сразу же считаться ошибкой. Такие ошибки следует исследовать. Очень важно определить в чем состоит ошибка, поскольку без этого будет сохраняться вера в непогрешимость методики. Например, время кругового обхода может стать отрицательным в результате корректировки часов от внешнего источника в процессе измерения. Другой причиной ошибки может быть считывание программой измерения неинициализированной памяти в процессе расчёта. Вторая ошибка более серьёзна, если та же программа применяется другими для их измерений, поскольку они тоже получают некорректные результаты. Обнаружив источник ошибок, можно внести соответствующие исправления.

Более тонким примером проверки самосогласованности является выборка значений односторонней задержки в Internet. При наличии большого числа измерений такой задержки может быть очень полезно, например, рассмотреть пары (задержка, время измерения), чтобы увидеть картину изменения задержки с течением времени. Однако такое изменение может быть связано с перекосом одних из использованных при измерении часов по отношению к другим. Другая причина может быть действительно связана с сетью. Обнаружение такой тенденции в сети является очень полезным. В работе [Ра97] рассматривается, как отличить влияние перекоса часов от фактических изменений в сети. Кроме того, при включении такой проверки в методику измерения обнаружение близкого к нулю наклона кривой зависимости задержки от времени является свидетельством того, что изменения скорей всего не связаны с перекосом часов.

Последний пример иллюстрирует проверку непротиворечивости самого процесса измерения. Выше были описаны методы выборки Пуассона на основе генерации интервалов с экспоненциальным распределением. Надёжная методика измерения будет включать проверку сгенерированных интервалов на предмет их экспоненциального распределения (а также возможной корреляции). В приложении к этому документу приведён код на языке C для одного из таких методов - хорошо зарекомендовавшей себя проверки пригодности, называемой тестом Андерсона-Дарлингса (Anderson-Darling).

В заключение отметим, что для пуассоновской выборки показателей Internet часто важно не время начала измерения, а время передачи данных измерительного процесса. Эти значения могут различаться из-за осложнений на хостах, участвующих в измерении. Таким образом даже при полной уверенности в используемом генераторе псевдослучайных чисел и последующих алгоритмах следует подумать о проверке допущений, которые могли быть приняты в каждой процедуре измерения.

11.3. Определение статистического распределения

Одним из способов описания набора измерений (выборки) является статистическое распределение - неформально как проценты. Имеется несколько слегка отличающихся способов сделать это. В этом параграфе даётся стандартное определение для единообразия таких описаний.

Эмпирическая функция распределения (empirical distribution function или EDF) набора скалярных измерений - это функция $F(x)$, которая для данного x возвращает долю общего числа измерений с результатом $\leq x$. Если x меньше минимального наблюдаемого значения, $F(x) = 0$, если же x не меньше максимального наблюдаемого значения, $F(x) = 1$.

Например, для 6 измерений

-2, 7, 7, 4, 18, -5

$F(-8) = 0$, $F(-5) = 1/6$, $F(-5,0001) = 0$, $F(-4,999) = 1/6$, $F(7) = 5/6$, $F(18) = 1$, $F(239) = 1$.

Отметим, что можно восстановить разные измеренные значения и количество наблюдений каждого из $F(x)$, информация о диапазонах значений не теряется. С другой стороны, представление результатов гистограммами обычно ведёт к потере информации о различных наблюдаемых значениях, поэтому применять EDF предпочтительно.

Однако при использовании EDF или гистограммы мы теряем сведения о порядке, в котором значения наблюдались. Важность таких потерь зависит от измеряемых показателей.

Термин процентиль (percentile) служит для обозначения наименьшего значения x , для которого $F(x)$ не меньше данного процента. Так 50-й процентиль в приведённом примере будет равен 4, поскольку $F(4) = 3/6 = 50\%$, 25-й будет равен -2, поскольку $F(-5) = 1/6 < 25\%$ и $F(-2) = 2/6 \geq 25\%$, 100-й - 18, а 0-й будет бесконечным, как и 15-й.

Следует соблюдать осторожность при использовании процентилей для представления выборки, поскольку они могут придать представленному результату точность выше реально доступной. В такие сводки должен включаться размер выборки N , поскольку любая разница в процентилях меньше $1/N$ выходит за пределы разрешения выборки.

Дополнительные сведения о EDF приведены в работе [DS86].

В заключение сделаем общее (и очень важное) замечание по понятию медианы. В статистике медианой распределения называют точку X , для которой вероятность наблюдения значений $\leq X$ равна вероятности наблюдения значений $> X$. При определении медианы для набора наблюдений результат зависит от числа наблюдений N :

- при нечётном N в качестве медианы используется 50-й процентиль (см. выше);
- при чётном N медианой будет среднее значение для двух центральных наблюдений, т. е. при отсортированных наблюдениях с номерами 1 - N , где $N = 2 \cdot K$, медианой будет среднее значение для наблюдений (K) и $(K+1)$.

Обычно слово «оценочная» исключается из фразы «оценочная медиана» и говорится просто о медиане.

11.4. Проверка пригодности

Для некоторых форм калибровки измерений требуется проверять согласование набора чисел с числами в конкретном распределении. Примером может служить проверка согласованности измерений, использующих процесс Пуассона, где нужно убедиться, что интервал между выборками действительно отражает экспоненциальное или равномерное распределение в интервале $[T, dT]$, если применяется подход dT/N .

{Комментарий. Можно протестировать по меньшей мер 3 возможных набора значений - запланированное время передачи пакетов, определяемое генератором псевдослучайных чисел, пользовательские временные метки, установленные непосредственно перед или после системного вызова, время в линии для пакетов, записанное с использованием фильтра пакетов. Все три набора потенциально информативны - несоответствие запланированного времени экспоненциальному распределению говорит о неаккуратности генерации случайных чисел, несоответствие пользовательских меток говорит о неточности таймеров при планировании передачи, несоответствия при использовании времени в линии указывают неточность фактической передачи пакетов, возможно из-за конкуренции за общий ресурс.}

Для выполнения таких тестов имеется множество статистических методов проверки пригодности (goodness-of-fit), описанных в [DS86]. Упомянутый документ рекомендует тест Anderson-Darling EDF в качестве универсального, который особенно хорош для обнаружения отклонений от данного распределения по краям функции EDF.

Важно понимать, что природа тестов на пригодность состоит в том, что сначала выбирается уровень значимости, представляющий вероятность того, что сначала будет ошибочно заявлено, что EDF данного набора измерений не подходит для конкретного распределения, хотя фактически измерение отражает данное распределение.

Если не указано иное, тесты на пригодность IPPM выполняются для уровня значимости 5%. Это означает, что если тест применяется для 100 выборок и 5 из них не проходят проверку, выборки считаются соответствующими тестируемому распределению. Если существенно большее число выборок не проходит проверку, допущение о соответствии выборки распределению следует отвергнуть. Если отказов существенно меньше, это может говорить о слишком тщательной обработке (подделке) выборок для такого распределения. Точно так же некоторые проверки пригодности (включая Anderson-Darling) могут определить вероятность подделки (дополнительной обработки) данной выборки. Для этого также используется уровень значимости 5%, т. е. тест сообщит, что данная выборка «слишком хороша, чтобы быть верной», поэтому появление такого сообщения чаще, чем для каждой 20-й выборки указывает на необычное поведение (подделку).

В Приложении представлен пример кода на языке C для реализации теста Anderson-Darling с обсуждением его использования.

Обсуждение тестов пригодности в контексте сетевых измерений приведено в работе [Pa94].

12. Отказ от стохастических показателей

При определении метрики для путей, субпутей, облаков и других элементов сети обычно не применяются стохастические показатели (вероятность), а используются детерминированные определения. Например, вместо определения показателя «вероятность потери пакета между A и B» определяется показатель «доля теряемых пакетов между A и B». При первом определении можно получить значение вида 0,73, а второе даст 73 пакета из 100. Подчеркнём, что отмеченное различие относится к определению показателей, но не к типам методов, используемым для анализа результатов измерений.

Причина этого различия достаточно проста. При выполнении измерений в терминах вероятности часто возникают скрытые допущения о стохастической модели исследуемого поведения. Основная цель исключения вероятности в определениях показателей заключается в предотвращении искажений, связанных с неявными допущениями.

Например, легко скрыть неявное допущение, что потерю пакетов в элементе сети в результате переполнения очередей можно описать как что-то, происходящее с любым пакетом с определённой вероятностью. Однако сегодня в Internet отбрасывание в очередях обычно является детерминированным и допущение об его вероятностном описании может скрыть важные корреляции между отбрасыванием в очереди для набора пакетов. Поэтому лучше явно отметить стохастические предположения, чем позволить их неявное включение в определения показателей.

Это не означает отказа от стохастических моделей для понимания производительности сети. Это значит лишь стремление исключить вероятностные параметры при определении показателей IP и избежать таких понятий как вероятность для отношений (proportion) и долей (rate). По-прежнему будут применяться, например, случайные выборки для оценки вероятности, используемые стохастическими моделями, связанными с показателями IP. Не исключается и возможность применения стохастических показателей, когда они действительно подходят (например, для моделирования ошибок передачи, вызываемых некоторыми типами шумов).

13. Тип пакетов

Фундаментальным свойством многих показателей Internet является зависимость значения показателя от типа пакетов IP, применяемых для измерения. Рассмотрим, например, показатель связности IP. Он будет разным для пакетов, направленных в общеизвестные порты TCP или незарезервированные порты UDP, пакетов с ошибками в контрольной сумме IP, пакетов с TTL = 16 и т. п. В некоторых случаях различия могут быть очень интересными (например, при наличии межсетевых экранов или резервирования RSVP).

Из-за отмеченных различий вводится базовое понятие пакета типа P, где P может задаваться явно (т. е. указывается тип пакета), частично (например, пакет с B октетами данных) или не указываться совсем. Таким образом, можно говорить о базовой связности IP для типа P или более конкретно о связности с портом HTTP. Некоторые показатели и методики могут быть определены для базового типа P, который затем конкретизируется при фактических измерениях.

Всякий раз, когда значение показателя зависит от типа участвующих в измерении пакетов, имя показателя будет включать конкретный тип или фразу, подобную type-P. Таким образом, определяется не показатель связности IP, а метрика связности IP для type-P и/или связности IP через порт HTTP. Такое именование служит важным напоминанием о необходимости учитывать точный тип трафика.

Отметим также, что очень полезно знать, одинаково ли данный элемент сети Internet обрабатывает класс C для разных типов пакетов. Если это так, любой из этих типов можно использовать в последующих измерениях для элемента. Это предполагает, что разработана метрика или набор показателей для определения C.

14. Указание путей

При рассмотрении метрики пути в Internet зачастую удобно представлять это как путь от хоста H1 к хосту H2. Однако это определение может быть неоднозначным, поскольку хосты Internet могут быть подключены к нескольким сетям. В этом случае результат оценки показателя будет зависеть от фактически использованных сетей.

Из-за этой неоднозначности в такие определения обычно следует включать IP-адреса. Для обычного случая одностороннего пути через Internet будем использовать Src для обозначения IP-адреса начальной точки пути и Dst - для IP-адреса конечной точки.

15. Пакеты стандартного формата

Если явно не указано иное, все определения показателей, относящиеся к пакетам IP, включают неявное допущение использования пакетом стандартного формата, в соответствии с приведёнными ниже критериями.

- Размер в заголовке IP соответствует сумме размеров заголовка IP и данных (payload).
- Пакет включает действительный заголовок IP с полем версии 4 (позднее будет описан случай для версии 6), имеет заголовок размером не менее 5 октетов и корректную контрольную сумму.
- Пакет не является фрагментом IP.
- Адреса отправителя и получателя соответствуют хостам.
- Значения поля TTL достаточно для прохождения от источника к получателю при условии декрементирования каждым узлом пересылки (hop) или TTL = 255.
- Пакет не включает явно не запрошенных опций IP.
- При наличии транспортного заголовка он содержит корректную контрольную сумму и иные действительные поля.

Если пакет описан как имеющий размер B октетов, должно выполняться условие $0 \leq B \leq 65535$, а если B задаёт размер данных в октетах, $B \leq (65535 - \text{размер заголовка IP в октетах})$.

Например, можно представить определение метрики связности IP «связность IP type-P для стандартных пакетов с полем IP TOS = 0» или более кратко «связность IP type-P с IP TOS = 0», поскольку стандартный формат неявно предполагается соглашением.

Конкретный тип стандартного пакета, который зачастую полезно учитывать, - это «минимальный пакет IP от A к B», удовлетворяющий приведённым ниже критериям.

- Пакет имеет стандартный формат.
- Пакет не содержит данных (0 октетов в payload).
- Пакет не имеет опций.

Отметим, что поле протокола не указано, поскольку разные значения поля могут менять обработку в сети.

При определении показателей IP учитывается, что пакеты меньше или проще указанного выше не будут передаваться корректно работающей сетью IP.

16. Благодарности

Спасибо Brian Carpenter, Bill Cerveny, Padma Krishnaswamy, Jeff Sedayao и Howard Stanislevic за их комментарии.

17. Вопросы безопасности

Этот документ связан с определениями и концепциями, относящимися к измерениям в Internet. Процедуры измерения рассматриваются лишь на верхнем уровне в отношении источников, которые могут быть надёжно измерены. Поэтому рассматриваемые вопросы не влияют на безопасность Internet и работающих в сети приложений.

Тем не менее, следует признать, что выполнение измерений в Internet может создавать проблемы безопасности и приватности. Активные методы с внедрением трафика в сеть можно использовать для атак на службы (denial-of-service), маскируемых под измерения. Пассивные методы, записывающие и анализирующие имеющийся в сети трафик, могут раскрывать содержимое трафика Internet посторонним лицам. Поэтому при определении каждой метрики и методики должно включаться соответствующее обсуждение вопросов безопасности.

18. Приложение

Ниже представлена программа на языке C для расчёта статистики теста Anderson-Darling (A2) с целью проверки согласованности набора значений с данным статистическим распределением. Интерес представляют две функции:

```
double exp_A2_known_mean(double x[], int n, double mean)
double unif_A2_known_range(double x[], int n,
                           double min_val, double max_val)
```

Обе процедуры в качестве первого аргумента x принимают массив и n тестируемых значений (при возврате элементы x сортируются). Остальные параметры характеризуют применяемое распределение, указывая среднее значение $mean$ ($1/\lambda$) для экспоненциального распределения или верхнюю и нижнюю границу для однородного распределения. Имена функций подчёркивают, что эти значения должны быть известны заранее, а не оцениваются из данных (например, путём расчёта среднего значения по выборке). Оценка параметров по данным меняет уровень значимости тестовой статистики. Хотя в [DS86] даны дополнительные таблицы значимости для некоторых экземпляров, где оценки параметров выводятся из данных, для наших целей предполагается, что эти оценки следует получить заранее, поскольку обычно тестируются такие значения, как время отправки пакетов, для которых нужно проверить соответствие известному распределению.

Обе функции возвращают уровень значимости, описанный выше, со значением от 0 до 1. Для корректного использования функций нужно заранее выбрать порог уровня значимости для теста, обычно этот 0,05, что соответствует 5%, как описано выше. Если программа возвращает значение строго ниже этого порога, данные считаются несоответствующими предполагаемому распределению для заданного уровня значимости. Т. е. при уровне значимости 5%, двадцатая часть данных, действительно извлекаемых из предполагаемого распределения будет ошибочно считаться несогласованной.

Важно учитывать, что при частом использовании этих функций будут возникать случайные сбои даже для безупречных данных.

Другой важный аспект, связанных с уровнями значимости, заключается в том, что их нецелесообразно сравнивать для определения «лучшего» из двух наборов значений с предполагаемым распределением. Для такой проверки лучше выполнять тесты «близости соответствия» (closeness-of-fit), такие как λ^2 из [Pa94].

Хотя эти функции предназначены для экспоненциального и однородного распределения с известными параметрами, несложно создать сравнимые функции для любого распределения с известными параметрами. Основой тестов A2 является статистика, рассчитанная для проверки соответствия набора значений однородному распределению в интервале от 0 до 1, которые обозначено $Unif(0, 1)$. Если нужно проверить соответствие набора значений X данному распределению $G(x)$, сначала рассчитывается $Y = G_{inverse}(X)$. Если значения X действительно распределены в соответствии с $G(x)$, распределение Y будет соответствовать $Unif(0, 1)$, поэтому проверяя Y на соответствие $Unif(0, 1)$, мы также проверяем X на соответствие $G(x)$.

Следует отметить, что указанный выше расчёт Y может давать значения Y за пределами интервала (0..1). Такие значения не должны возникать, если распределение X действительно соответствует $G(x)$, но это может быть не так. В последнем случае нужно избегать расчёта центральной статистики A2, поскольку могут возникать исключительные случаи для чисел с плавающей запятой, если любое из значений выпадает из интервала (0..1). Поэтому функции проверяют это и при возникновении таких ситуаций возвращают необработанную (raw) статистику A2 с уровнем значимости -1. Затем функция преобразует необработанную статистику A2 в уровень значимости и распространяет это значение, возвращая уровень -1. Поэтому при использовании этих функций должно быть готовым получить отрицательный уровень значимости.

Ещё один момент, связанный с применением статистики A2, относится к числу тестируемых значений n . При $n < 5$ тест не имеет смысла и в таких случаях возвращается уровень значимости -1. Для «реальных» данных тест обретает силу при росте n . Среди специалистов по статистике хорошо известно, что реальные данные почти никогда точно не соответствуют теоретическому распределению даже в таких случаях, как бросание игральные костей (в работе [Pa94] приведено краткое рассмотрение и ссылки). Тест A2 достаточно чувствителен, поэтому для крупных наборов данных он почти всегда будет давать отказ, поскольку способен обнаруживать незначительные отклонения от соответствия данных распределению. Например, отмечено, что при тестировании 8192 измеренных значений времени в линии для пакетов, переданных с интервалами Пуассона, тест A2 почти всегда даёт отрицательный результат. При этом тестирование для 128 с уровнем значимости 5% завершалось отказом лишь в 5% случаев. Таким образом, при отрицательном результате теста следует разбираться с причинами отказа.

Далее приведён исходный код программы на языке C, реализующей рассмотренные выше функции.

```

/* Программа для расчёта статистики Anderson-Darling A2.
 *
 * Реализована на основании "Goodness-of-Fit Techniques,"
 * R. D'Agostino and M. Stephens, editors, Marcel Dekker,
 * Inc., 1986.
 */

#include <stdio.h>
#include <stdlib.h>
#include <math.h>

/* Возвращает необработанную статистику теста A^2 для n сортированных
 * выборок z[0] .. z[n-1], для z ~ Unif(0,1).
 */
extern double compute_A2(double z[], int n);

/* Возвращает уровень значимости данных, связанных со статистикой
 * теста A^2 для значения A2, в предположении, что параметры
 * тестируемого распределения не оцениваются по данным.
 */
extern double A2_significance(double A2);

/* Возвращает уровень значимости A^2 для тестирования n наблюдений
 * x[0] .. x[n-1] при экспоненциальном распределении с данным
 * средним значением.
 *
 * Побочный эффект заключается в сортировке x[0..n-1] при возврате.
 */
extern double exp_A2_known_mean(double x[], int n, double mean);

/* Возвращает уровень значимости A^2 для тестирования n наблюдений
 * x[0] .. x[n-1] при однородном распределении [min_val, max_val].
 *
 * Побочным эффектом является сортировка x[0..n-1] при возврате
 */
extern double unif_A2_known_range(double x[], int n,
                                   double min_val, double max_val);

/* Возвращает псевдослучайное число в соответствии с экспоненциальным
 * распределением с данным средним значением.
 */
extern double random_exponential(double mean);

/* Вспомогательная функция, используемая qsort() для сортировки
 * действительных чисел с плавающей запятой.
 */
static int
compare_double(const void *v1, const void *v2)
{
    double d1 = *(double *) v1;
    double d2 = *(double *) v2;

    if (d1 < d2)
        return -1;
    else if (d1 > d2)
        return 1;
    else
        return 0;
}

double
compute_A2(double z[], int n)
{
    int i;
    double sum = 0.0;

    if ( n < 5 )
        /* Слишком мало значений. */
        return -1.0;

    /* Если любое значение выходит за рамки диапазона (0, 1)
     * возвращается отказ (и предотвращается исключительная
     * ситуация для действительных чисел в приведённом ниже коде).
     */
    for (i = 0; i < n; ++i)
        if ( z[i] <= 0.0 || z[i] >= 1.0 )
            return -1.0;

    /* Страница 101 в работе D'Agostino и Stephens. */
    for (i = 1; i <= n; ++i) {
        sum += (2 * i - 1) * log(z[i-1]);
        sum += (2 * n + 1 - 2 * i) * log(1.0 - z[i-1]);
    }
    return -n - (1.0 / n) * sum;
}

```

```

double
A2_significance(double A2)
{
    /* Страница 105 в работе D'Agostino и Stephens. */
    if (A2 < 0.0)
        return A2; /* Распространяется фиктивное значение A2. */

    /* Проверка наличия искажённых значений. */
    if (A2 <= 0.201)
        return 0.99;
    else if (A2 <= 0.240)
        return 0.975;
    else if (A2 <= 0.283)
        return 0.95;
    else if (A2 <= 0.346)
        return 0.90;
    else if (A2 <= 0.399)
        return 0.85;

    /* Проверка возможной несогласованности. */
    if (A2 <= 1.248)
        return 0.25;
    else if (A2 <= 1.610)
        return 0.15;
    else if (A2 <= 1.933)
        return 0.10;
    else if (A2 <= 2.492)
        return 0.05;
    else if (A2 <= 3.070)
        return 0.025;
    else if (A2 <= 3.880)
        return 0.01;
    else if (A2 <= 4.500)
        return 0.005;
    else if (A2 <= 6.000)
        return 0.001;
    else
        return 0.0;
}

double
exp_A2_known_mean(double x[], int n, double mean)
{
    int i;
    double A2;

    /* Сортировка первых n значений. */
    qsort(x, n, sizeof(x[0]), compare_double);

    /* В предположении экспоненциального распределения
     * преобразование в Unif(0,1).
     */
    for (i = 0; i < n; ++i) {
        x[i] = 1.0 - exp(-x[i] / mean);
    }

    /* Тест A^2 для проверки однородности. */
    A2 = compute_A2(x, n);
    return A2_significance(A2);
}

double
unif_A2_known_range(double x[], int n, double min_val, double max_val)
{
    int i;
    double A2;
    double range = max_val - min_val;

    /* Сортировка первых n значений. */
    qsort(x, n, sizeof(x[0]), compare_double);

    /* Преобразование Unif(min_val, max_val) в Unif(0,1). */
    for (i = 0; i < n; ++i)
        x[i] = (x[i] - min_val) / range;

    /* Тест A^2 для проверки однородности. */
    A2 = compute_A2(x, n);
    return A2_significance(A2);
}

double
random_exponential(double mean)
{
    return -mean * log1p(-drand48());
}

```

19. Литература

- [AK97] G. Almes and S. Kalidindi, "A One-way Delay Metric for IPPM", Work in Progress¹, November 1997.
- [BM92] I. Bilinskis and A. Mikelsons, Randomized Signal Processing, Prentice Hall International, 1992.
- [DS86] R. D'Agostino and M. Stephens, editors, Goodness-of-Fit Techniques, Marcel Dekker, Inc., 1986.
- [CPB93] K. Claffy, G. Polyzos, and H-W. Braun, "Application of Sampling Methodologies to Network Traffic Characterization," Proc. SIGCOMM '93, pp. 194-203, San Francisco, September 1993.
- [FJ94] S. Floyd and V. Jacobson, "The Synchronization of Periodic Routing Messages," IEEE/ACM Transactions on Networking, 2(2), pp. 122-136, April 1994.
- [Mi92] Mills, D., "Network Time Protocol (Version 3) Specification, Implementation and Analysis", [RFC 1305](#), March 1992.
- [Pa94] V. Paxson, "Empirically-Derived Analytic Models of Wide-Area TCP Connections," IEEE/ACM Transactions on Networking, 2(4), pp. 316-336, August 1994.
- [Pa96] V. Paxson, "Towards a Framework for Defining Internet Performance Metrics," Proceedings of INET '96, <ftp://ftp.ee.lbl.gov/papers/metrics-framework-INET96.ps.Z>
- [Pa97] V. Paxson, "Measurements and Analysis of End-to-End Internet Dynamics," Ph.D. dissertation, U.C. Berkeley, 1997, <ftp://ftp.ee.lbl.gov/papers/vp-thesis/dis.ps.gz>.

20. Адреса авторов

Vern Paxson

MS 50B/2239
Lawrence Berkeley National Laboratory
University of California
Berkeley, CA 94720
USA
Phone: +1 510/486-7504
E-Mail: vern@ee.lbl.gov

Jamshid Mahdavi

Pittsburgh Supercomputing Center
4400 5th Avenue
Pittsburgh, PA 15213
USA
Phone: +1 412/268-6282
E-Mail: mahdavi@psc.edu

Guy Almes

Advanced Network & Services, Inc.
200 Business Park Drive
Armonk, NY 10504
USA
Phone: +1 914/765-1120
E-Mail: almes@advanced.org

Matt Mathis

Pittsburgh Supercomputing Center
4400 5th Avenue
Pittsburgh, PA 15213
USA
Phone: +1 412/268-3319
E-Mail: mathis@psc.edu

21. Полное заявление авторских прав

Copyright (C) The Internet Society (1998). Все права защищены.

Этот документ и его переводы могут копироваться и предоставляться другим лицам, а производные работы, комментирующие или иначе разъясняющие документ или помогающие в его реализации, могут подготавливаться, копироваться, публиковаться и распространяться целиком или частично без каких-либо ограничений при условии сохранения указанного выше уведомления об авторских правах и этого параграфа в копии или производной работе. Однако сам документ не может быть изменён каким-либо способом, таким как удаление уведомления об авторских правах или ссылок на Internet Society или иные организации Internet, за исключением случаев, когда это необходимо для разработки стандартов Internet (в этом случае нужно следовать процедурам для авторских прав, заданных процессом Internet Standards), а также при переводе документа на другие языки.

Предоставленные выше ограниченные права являются бессрочными и не могут быть отозваны Internet Society или правопреемниками.

Этот документ и содержащаяся в нем информация представлены "как есть" и автор, организация, которую он/она представляет или которая выступает спонсором (если таковой имеется), Internet Society и IETF отказываются от каких-либо гарантий (явных или подразумеваемых), включая (но не ограничиваясь) любые гарантии того, что использование представленной здесь информации не будет нарушать чьих-либо прав, и любые предполагаемые гарантии коммерческого использования или применимости для тех или иных задач.

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru

¹Опубликовано в [RFC 2680](#). Прим. перев.