

Алгоритм оптимальной маршрутизации для системы управления трафиком

Г.А.Андрианов

Институт проблем передачи информации, Российская академия наук, Москва, Россия

Поступила в редколлегию 19.03.2010

Аннотация—На основании измерений для различных направлений трафика сети оператора связи предложен алгоритм расстановки приоритетов выбора терминирующих операторов с учетом показателей качества обслуживания при оптимизации маршрутизации трафика. Показано соответствие теоретической и экспериментальной зависимостей отношения числа повторных и первичных вызовов от потерь. Определены значения настойчивости абонента и потерь, не приводящих к повторным вызовам.

1. ВВЕДЕНИЕ

Рассмотрим сеть транзитного оператора, используемую для пропуска трафика. Оператор получает входящий трафик от оператора-клиента (или иницилирующего оператора) и отправляет исходящий трафик на оператора-провайдера (или терминирующего оператора). Трафик от нескольких клиентов может быть отправлен на одного провайдера и наоборот, трафик от одного клиента может направляться на различных провайдеров услуги дальней связи (для деталей см. [1]. Объем трафика, с которым, как правило, работают специалисты в области маршрутизации трафика, это так называемое направление, представляющее собой совокупность кода страны, CC — Country Code (E.164) и кода региона/оператора в стране, ABC/DEF кода. Если предполагается отдельная маршрутизация для роумингового диапазона, MSRN (Mobile Subscriber Roaming Numbering), данного оператора в стране, то расшифровка направления может быть еще глубже, до abx_1 : $CC - (ABC/DEF) - abx_1$. Сколько цифр В-номера будет участвовать в маршрутизации, как будут выставляться счета за вызовы по данному В-номеру, как будут насчитываться показатели качества обслуживания, KPI (Key Performance Indicators), все это определяется разбиением трафика по кодам в мастер-справочнике направлений транзитного оператора. При составлении такого мастер-справочника учитываются интересы контроля качества и взаиморасчетов всех трех сторон: иницилирующего, транзитного и терминирующего операторов.

При рассмотрении потерь трафика как до входа в сеть транзитного оператора, так и при выходе из этой сети, а также при окончательной терминации в стране и регионе/операторе, в силу изложенного выше, мы будем рассматривать направления из мастер-справочника, как объекты идентичности, то есть рассматривать потери по данному направлению на данном операторе (транковой группе между коммутаторами или паре кодов пунктов ОКС 7, которые удобно использовать для идентификации коммутаторов). Таким образом, одним из критериев, по которому производится фильтрация данных для настоящего исследования, является шаблон, накладываемый на начало В-номера — направление трафика.

Другим очевидным критерием фильтрации трафика является определенная транковая группа или пара коммутаторов (кодов пунктов этих коммутаторов), где производится измерение. Необходимо заметить, что все данные, использованные в настоящей работе, получены при помощи системы распределенного мониторинга и анализа качества обслуживания — LMS (Link

Monitoring System). Поэтому транковой группе ставится в соответствие пара кодов пунктов ОКС 7: OPC (Originating Point Code) — DPC (Destination Point Code).

Необходимо сказать несколько слов о самой системе управления трафиком – TMS (Traffic Management System). Основные данные о трафике, такие как CDR (Call Detail Records) по отвеченным вызовам, обслуженным всеми транковыми группами на всех коммутаторах сети, TMS получает напрямую от предбиллинга. На сети обеспечивается формирование CDR по неотвеченным вызовам не на всех коммутаторах (это обусловлено экономией производительности центральных процессоров коммутаторов). Поэтому полные данные CDR (как по отвеченным, так и по неотвеченным вызовам) TMS получает от LMS. Все данные поступают в TMS во времени, приближенном к реальному.

В системе управления трафиком происходят следующие процессы обработки данных. Во-первых, в TMS строятся агрегаты данных CDR из предбиллинга и из LMS, необходимые для последующего представления в виде многомерных “кубов” баз данных, содержащих объемы трафика и KPI. Во-вторых, данные агрегируются для оценки пропускной способности ближайших транковых групп (для ответа на вопрос “Можно ли направлять на эти транкрупшы трафик, хватит ли каналов, чтобы трафик был обслужен именно данным терминирующим оператором, а не поступил на путь следующего, возможно, не такого выгодного выбора?”). В-третьих, в TMS поступает коммерческая информация: тарифные планы инициаторов и терминаторов трафика; в перспективе – стоимость аренды каналов первичной сети, на которые опираются транковые группы сети вторичной; обязательства по пропуску трафика; ограничения по тому, на какого оператора можно направить данный код $CC = (ABC/DEF)$.

Одной из самых важных задач, решаемых в TMS, является расстановка операторов по ранжиру для определения первого, второго и последующих путей выбора для маршрутизации данного направления. При расстановке путей выбора для маршрутизации трафика на операторов мы должны принимать во внимание следующие основные факторы:

- Разность тарифов продажи и покупки услуги (маржа) пропуска трафика для каждого из операторов.
- KPI — значения ASR (Answer Seizure Ratio) и NER (Network Effectiveness Ratio) для каждого направления и каждого из операторов.
- Возможности операторов по ресурсам (каналам вторичной сети) для пропуска направления данного объема.
- Обязательства и ограничения по данному направлению через данного оператора.

Такой фактор как число свободных каналов у оператора действует в качестве ограничения. Кроме официальных ограничений есть еще и полученная в процессе организации маршрутизации информация, говорящая нам о том, каких операторов вообще не стоит рассматривать как варианты выбора для данного направления. В TMS есть возможность все эти ограничения и предписания учесть, определив множество рассматриваемых операторов. Но есть два критерия — тариф и качество, которые для существующего на сети огромного множества операторов и направлений могут быть оценены системой лишь в автоматическом режиме (специалист просто не справится с обработкой такого большого объема информации в требуемые сроки вручную). Оценка в автоматическом режиме должна производиться на основании некоторого алгоритма. Разработка такого алгоритма является основной целью данной работы.

2. КАЧЕСТВО ОБСЛУЖИВАНИЯ С УЧЕТОМ ПОВЕДЕНИЯ АБОНЕНТА

Новизну в предлагаемый нами алгоритм маршрутизации трафика привносит учет поведения абонента при рассмотрении таких показателей качества обслуживания, QoS (Quality of Service), как ASR и NER. ASR является важнейшим показателем качества обслуживания,

принятым во всем мире [2]. Определяется этот показатель как процентное отношение числа отвеченных вызовов к общему числу вызовов. Но при анализе трафика по информации, заложенной в таком показателе, нельзя забывать, что, если QoS на направлении не идеально, то среди общего числа вызовов, поступивших за интервал наблюдения, будет некоторое число повторных вызовов.

Повторным вызовом при дальнейшем рассмотрении мы будем называть такой вызов, который поступил с того же А-номера на тот же В-номер в течение небольшого интервала времени (10 минут) относительно основного вызова, причем причина, по которой был произведен этот повторный вызов, состоит в том, что неуспешной оказалась основная попытка. Не “неотвеченной”, а именно “неуспешной”, то есть, основная попытка завершилась результатом, не включенным в NER по E.425 [3]. Иными словами, если бы QoS на данном направлении вдруг сделалось идеальным, то выборка не содержала бы повторных вызовов.

Понятно, что поскольку многие показатели QoS представляют собой процентное отношение некоторого числа вызовов с каким-либо свойством к общему числу вызовов, все эти показатели оказываются искажены фактом наличия некоторого числа повторных вызовов. В этом случае полезно рассмотреть “чистые” показатели QoS. Например, “чистый” ASR. Назовём его CASR (Clean Answer Seizure Ratio).

Пусть $k(ASR)$ - среднее число повторных вызовов на один основной для данного направления, это число, вообще говоря, зависит от самого ASR. Тогда, если общее число поступивших вызовов — n , число основных вызовов:

$$n_{pr} = \frac{n}{k+1}.$$

В данном случае

$$ASR = \frac{n_{ans}}{n} \cdot 100\%,$$

где n_{ans} — число отвеченных вызовов, а определённый нами “чистый” ASR,

$$CASR = \frac{n_{ans}}{n_{pr}} \cdot 100\%;$$

$$CASR = ASR \cdot (k(ASR) + 1). \quad (1)$$

Показатель CASR позволяет напрямую рассчитывать недополученный по причине недостаточного QoS доход. В рекомендации E.420 [4] был введен так называемый целевой ASR — TASR (Target Answer Seizure Ratio). Целевой ASR — это идеальный для данного направления ASR, то, к чему в принципе надо стремиться и что может быть достигнуто. Не всегда этот показатель равен идеальным 60%, определенным в E.426 [5]. Для некоторых направлений такой высокий показатель никогда не достигался и не может быть достигнут. Там целевой ASR определяется по большому периоду наблюдения для надежности доступа к услуге при прохождении вызовов в данный регион через различных операторов связи. Лучшие для данного региона (для рассматриваемых первых цифр В-номера) показатели и определяют TASR. По TASR для данного направления можно определить “чистый” целевой ASR — CTASR.

При определении CTASR необходимо учитывать, что для тех направлений, где значения TASR не идеальны (например, составляют 30%), трафик тоже “загрязнен” повторными вызовами. Причем среднее число повторных вызовов на один основной зависит как от направления (кода региона), так и от самого QoS, в данном случае, от ASR. Вид функции может быть определен для каждого из направлений.

После определения по формуле (1) для данного направления значений CASR и CTASR, можно легко найти недополученный доход в виде

$$Q = \frac{CTASR - CASR}{CTASR} \cdot 100\%.$$

Такой процент своего дохода от трафика на направлении компания недополучает напрямую в случае, если на данном направлении не всё в порядке с QoS. Кроме того, есть масса косвенных неприятностей, которые влечет за собой недостаточное качество обслуживания. Здесь в первую очередь можно назвать недостаточную степень удовлетворенности клиентов, которая влечет отказ клиентов от пользования услугами провайдера и меньшее приобретение провайдером новых потенциальных клиентов, которые могли бы начать пользоваться услугами в случае более высокой степени общей удовлетворенности качеством.

3. НАСТОЙЧИВОСТЬ АБОНЕНТА И ЧИСЛО ПОВТОРНЫХ ВЫЗОВОВ НА ОДИН ПЕРВИЧНЫЙ ВЫЗОВ

Теоретическая зависимость наблюдаемого отношения среднего числа повторных вызовов на один первичный вызов имеет вид [6]:

$$k(p; H) = \frac{Hp}{1 - Hp}. \quad (2)$$

В работе [1] в рассмотрение была введена величина потерь, не влияющих на повторение вызовов абонентом $A - p_c$, и получена зависимость “наблюдаемых” (то есть, таких, которые могут быть получены прямым измерением) потерь q от p_c и p :

$$q = 1 - (1 - p_c)(1 - p). \quad (3)$$

Важнейшее для нашего исследования физическое отличие q и p от p_c в уравнении (3) состоит в том, что наблюдаемое число повторных вызовов на один первичный, k , зависит напрямую (физически) по формуле (2) только от p и никак не зависит от p_c . Постоянная составляющая наблюдаемых потерь p_c может быть определена наряду с настойчивостью H при интерполяции экспериментальных точек теоретической кривой. Найдём вид этой теоретической зависимости между двумя наблюдаемыми величинами: числом повторных вызовов на один первичный и потерями. Из (3) следует, что

$$p = \frac{q - p_c}{1 - p_c}. \quad (4)$$

Подставив (4) в (2), получим:

$$k(q) = \frac{\frac{q-p_c}{1-p_c} H}{1 - \frac{q-p_c}{1-p_c} H}. \quad (5)$$

В формуле (5) величины p_c и H постоянны для данного, рассматриваемого направления (измерений трафика с определенным началом В-номера, прошедшего через определенного оператора).

4. АНАЛИЗ РЕЗУЛЬТАТОВ ИЗМЕРЕНИЙ ПОКАЗАТЕЛЕЙ КАЧЕСТВА ОБСЛУЖИВАНИЯ

Нами разработано программное обеспечение, позволяющее производить анализ выборки CDR и получать на выходе требуемые расчетные показатели, связанные как с качеством обслуживания, так и с числом повторных вызовов на один первичный. Исходным файлом для расчетов является текстовый файл с некоторыми, наиболее важными полями CDR. Данный

файл формируется при выполнении запроса к базе данных MySQL сервера LMS. Пример исходного файла изображен на рис. 1.

iam	acm	anm	rel	opc	dpc	callingNormalized	calledNormalized	rc
2008-05-31 00:00:19.392			2008-05-31 00:00:20.230	1285	9473		74112563757	3
2008-05-31 00:00:21.391			2008-05-31 00:00:26.248	1285	9473		74112563757	3
2008-05-31 00:10:55.472			2008-05-31 00:10:56.280	1285	9473		74112563757	3
2008-05-31 00:17:49.428			2008-05-31 00:17:50.281	1285	9473		74112563757	3
2008-05-31 00:18:19.939			2008-05-31 00:18:20.502	1285	9473	74114326398	7411212563	3
2008-05-31 00:18:32.312			2008-05-31 00:18:33.201	1285	9473	74114326398	7411212563	3
2008-05-31 00:18:35.032			2008-05-31 00:18:35.921	1285	9473	74114326398	7411212563	3
2008-05-31 00:18:37.950			2008-05-31 00:18:38.511	1285	9473	74114326398	7411212563	3
2008-05-31 00:18:52.514			2008-05-31 00:18:53.093	1285	9473	79142202708	74112150808	3
2008-05-31 00:19:21.393			2008-05-31 00:19:22.190	1285	9473	74114326398	7411212563	3
2008-05-31 00:19:33.352			2008-05-31 00:19:34.171	1285	9473	74114326398	7411212563	3
2008-05-31 00:19:36.215			2008-05-31 00:19:37.071	1285	9473	74114326398	7411212563	3
2008-05-31 00:19:39.055			2008-05-31 00:19:39.912	1285	9473	74114326398	7411212563	3
2008-05-31 00:19:41.695			2008-05-31 00:19:42.542	1285	9473	74114326398	7411212563	3
2008-05-31 00:19:39.912	2008-05-31 00:19:42.142	2008-05-31 00:19:51.842	2008-05-31 00:20:06.259	1285	9473	79878484126	74113736273	16
2008-05-31 00:20:07.195			2008-05-31 00:20:08.062	1285	9473	74114326398	7411212563	3
2008-05-31 00:20:10.978			2008-05-31 00:20:10.882	1285	9473	74114326398	7411212563	3
2008-05-31 00:20:12.890			2008-05-31 00:20:13.741	1285	9473	74114326398	7411212563	3
2008-05-31 00:20:15.694			2008-05-31 00:20:16.502	1285	9473	74114326398	7411212563	3
2008-05-31 00:20:51.374			2008-05-31 00:20:52.222	1285	9473		74112563757	3
2008-05-31 00:22:23.456			2008-05-31 00:22:24.312	1285	9473		74112563757	3
2008-05-31 00:23:53.412			2008-05-31 00:23:54.261	1285	9473		74112563757	3
2008-05-31 00:25:12.576			2008-05-31 00:25:12.618	1285	9473		74112563757	3
2008-05-31 00:26:55.352			2008-05-31 00:26:56.190	1285	9473		74112563757	3
2008-05-31 00:24:14.575	2008-05-31 00:24:16.922	2008-05-31 00:24:19.811	2008-05-31 00:28:24.206	1285	9473	79878484126	74113736273	16
2008-05-31 00:28:27.416			2008-05-31 00:28:28.272	1285	9473		74112563757	3
2008-05-31 00:29:57.417			2008-05-31 00:29:58.232	1285	9473		74112563757	3
2008-05-31 00:31:29.455			2008-05-31 00:31:30.261	1285	9473		74112563757	3
2008-05-31 00:36:03.372			2008-05-31 00:36:04.171	1285	9473		74112563757	3
2008-05-31 00:37:33.453			2008-05-31 00:37:34.250	1285	9473		74112563757	3
2008-05-31 00:39:05.391			2008-05-31 00:39:06.271	1285	9473		74112563757	3
2008-05-31 00:41:21.657	2008-05-31 00:41:32.101		2008-05-31 00:42:15.381	1285	9473	79142731010	74113525019	16
2008-05-31 00:42:28.975	2008-05-31 00:42:39.471	2008-05-31 00:43:05.493	2008-05-31 00:44:46.804	1285	9473	79142731010	74113525019	16
2008-05-31 00:43:09.394			2008-05-31 00:43:10.243	1285	9473		74112563757	3
2008-05-31 00:46:39.475			2008-05-31 00:46:40.323	1285	9473		74112563757	3
2008-05-31 00:48:11.334			2008-05-31 00:48:12.142	1285	9473		74112563757	3
2008-05-31 00:49:41.416			2008-05-31 00:49:42.261	1285	9473		74112563757	3
2008-05-31 00:51:13.378			2008-05-31 00:51:14.142	1285	9473		74112563757	3
2008-05-31 00:52:43.357			2008-05-31 00:52:44.213	1285	9473		74112563757	3
2008-05-31 00:53:23.934			2008-05-31 00:53:24.742	1285	9473	79142882849	741120000026	3
2008-05-31 00:53:59.376			2008-05-31 00:54:00.233	1285	9473	79142882849	741120000026	3
2008-05-31 00:54:33.896			2008-05-31 00:54:34.742	1285	9473	79142882849	741120000026	3
2008-05-31 00:57:17.379			2008-05-31 00:57:18.183	1285	9473		74112563757	3
2008-05-31 00:58:47.317			2008-05-31 00:58:48.133	1285	9473		74112563757	3
2008-05-31 00:58:54.814			2008-05-31 00:58:55.663	1285	9473	79142882849	741120000026	3
2008-05-31 01:00:19.394			2008-05-31 01:00:20.253	1285	9473		74112563757	3
2008-05-31 01:01:55.853			2008-05-31 01:01:56.764	1285	9473	79241651182	7411213436	3
2008-05-31 01:03:21.339			2008-05-31 01:03:22.182	1285	9473		74112563757	3
2008-05-31 01:05:51.320			2008-05-31 01:05:52.734	1285	9473	79142351410	7411213133	3
2008-05-31 01:06:35.400			2008-05-31 01:06:36.233	1285	9473	79142351410	7411213133	3
2008-05-31 01:07:49.359			2008-05-31 01:07:50.172	1285	9473		74112563757	3
2008-05-31 01:08:00.900	2008-05-31 01:06:53.154	2008-05-31 01:07:05.643	2008-05-31 01:08:01.493	1285	9473	79151317360	74113631738	16
2008-05-31 01:08:41.730	2008-05-31 01:08:44.913	2008-05-31 01:08:55.184	2008-05-31 01:09:11.352	1285	9473	79151317360	74113631738	16
2008-05-31 01:09:19.315			2008-05-31 01:09:20.122	1285	9473		74112563757	3
2008-05-31 01:10:47.400			2008-05-31 01:10:48.244	1285	9473		74112563757	3
2008-05-31 01:11:45.380			2008-05-31 01:11:46.233	1285	9473		74112563757	3
2008-05-31 01:11:52.879			2008-05-31 01:11:53.723	1285	9473	79142903393	7411213322	3
2008-05-31 01:14:43.319			2008-05-31 01:14:44.182	1285	9473		74112563757	3

Рис. 1. Исходный файл для программы расчета показателей и доли повторных вызовов

В исходный файл помещены следующие поля CDR:

- iam — время прохождения исходного адресного сообщения, IAM (Initial Address Message);
- acm — время прохождения сообщения о полноте адреса, ACM (Address Complete Message);
- anm — время прохождения сообщения об ответе, ANM (Answer Message);
- rel — время прохождения сообщения об освобождении, REL (Release);
- opc — код исходящего пункта, OPC;
- dpc — код пункта назначения, DPC;
- callingNormalized — нормализованный (приведенный к международному формату) номер абонента А;
- calledNormalized — нормализованный (приведенный к международному формату) номер абонента В;
- rc — Release Cause — значение причины разъединения.

Максимальный размер файла, который удавалось обработать, составлял 50 Мбайт (около 300 000 строк). Этого, как правило, было достаточно, чтобы обработать трафик по направлению за несколько суток. Алгоритм обработки CDR следующий.

1. Рассматриваются только CDR, в которых явно присутствует как А, так и В номер.
2. Агрегация производится по гибко определяемому временному интервалу — интервалу, в который попали сообщения REL от 1000 последовательно завершившихся первичных вызовов (исходный файл отсортирован по времени завершения вызова).
3. За “хорошие” (с точки зрения работы сети) причины разъединения принимаются причины, не уменьшающие NER (по E.425) [1].

4. Чтобы считать вызов с теми же А и В номерами, как и у предыдущего, повторным, должны быть выполнены два условия: 1) предыдущий вызов должен завершиться с “плохой” причиной разъединения, 2) период времени между предыдущим и следующим за ним вызовами должен составлять не более 10 минут.
5. По каждому из гибких временных интервалов рассчитывается число первичных вызовов, повторных вызовов, отвеченных вызовов и успешных с точки зрения сети вызовов.

Время обработки файла — несколько минут. После обработки программа выдает информацию следующего вида:

NN	Prim	Rep	Answ	Good
20	1000	37	525	914
21	1000	42	546	932
22	1000	52	550	909
23	1000	33	518	911
24	1000	36	526	911
25	1000	59	497	859
45	1000	396	287	478
46	1000	443	293	502
47	1000	461	265	465
48	1000	517	274	482
49	1000	448	274	479
50	1000	310	299	526

В таблице специально приведены интервалы времени из разных областей профиля качества. Поле NN — номер по порядку. Поле Prim — 1000 первичных, основных вызовов. Поле Answ — число отвеченных вызовов. Легко убедиться, что когда доля успешных вызовов (поле Good) составляет 91-93% ($NER = \frac{Good}{Prim+Rep} 100\%$), число повторных вызовов (поле Rep) мало и составляет всего 0.03-0.04 на один первичный ($\frac{Rep}{Prim}$). В часы же наибольшей нагрузки, когда доступность направления ниже, и NER составляет всего 33-34%, число повторных вызовов на один первичный начинает достигать значений 0.4 - 0.5.

При дальнейшей обработке данных выяснилось, что поведение абонентов, производящих наборы на код +99450* (Азербайджан моб.; в примере файла результата представлены данные именно для этого направления), не одинаково в течение суток. В часы, когда загруженность сети низкая ($p_c = 0$), настойчивость абонентов тоже невысока ($H = 0.3$). Этой фазе работы сети соответствует кривая T1 на рис. 2. В период наибольшей нагрузки, когда каналов у оператора, терминирующего трафик в Азербайджане, начинает не хватать, среди причин разъединения начинает преобладать CV#34 — No circuit/channel available (нет доступного канала), а также другие причины из класса причин разъединения по недоступности ресурса сети. Причины из данного класса причин корректно интерпретируются окончательным коммутатором исходящего абонента, и он получает акустический сигнал так называемого fast busy — “частого” (или “быстрого”) занято. Абонент, получающий такой сигнал, ведет себя настойчивей,

чем при другой акустической сигнализации (наихудшим вариантом здесь является “тишина”, когда какая-либо сигнализация, оповещающая, что происходит с вызовом, вовсе отсутствует).

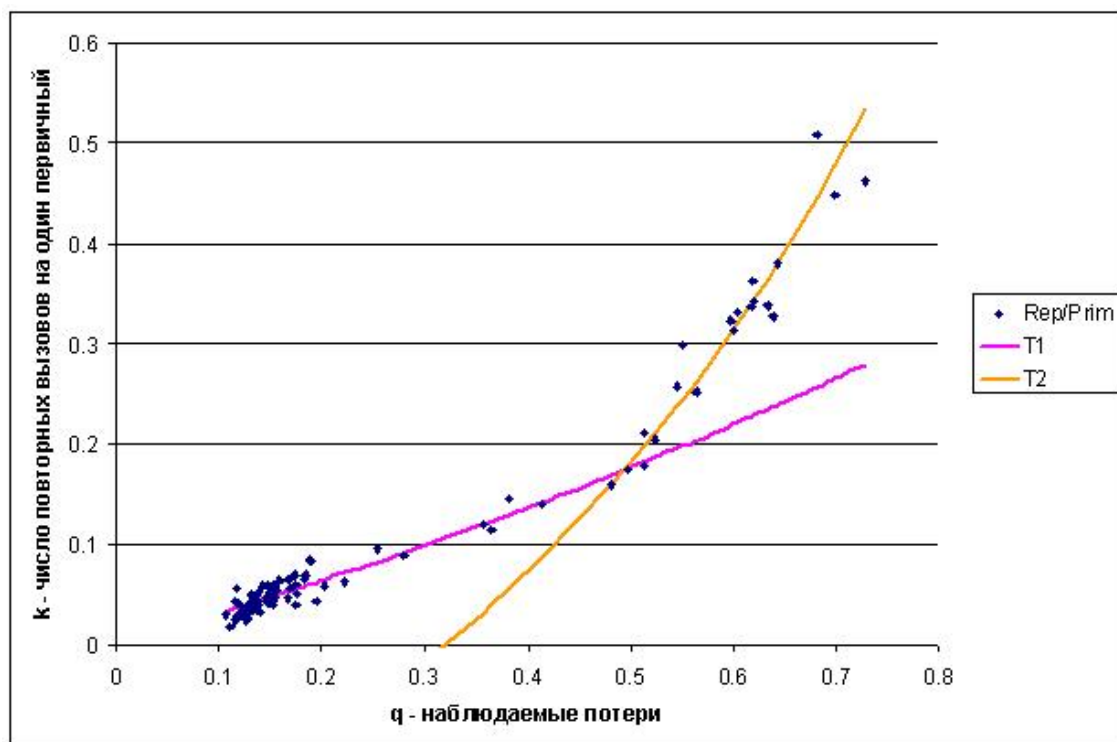


Рис. 2. Иллюстрация поведения абонентов при вызовах в Азербайджан по коду +99450*

В результате такого “фазового перехода”, происходящего на сети, изменяются усредненные характеристики настойчивости. Кривая T2 на рис. 2. соответствует уже настойчивости $H = 0.58$ при постоянной составляющей потерь у оператора $p_c = 0.32$.

Не всегда настойчивость абонента так низка в случае небольших потерь, как в примере с Азербайджаном. Если рассмотреть, например, трафик в Якутию по коду +7411* (см. рис.

3), то можно убедиться, что даже при малых потерях абонент проявляет довольно большую настойчивость. Здесь $H = 0.64$ при $p_c = 0.10$.

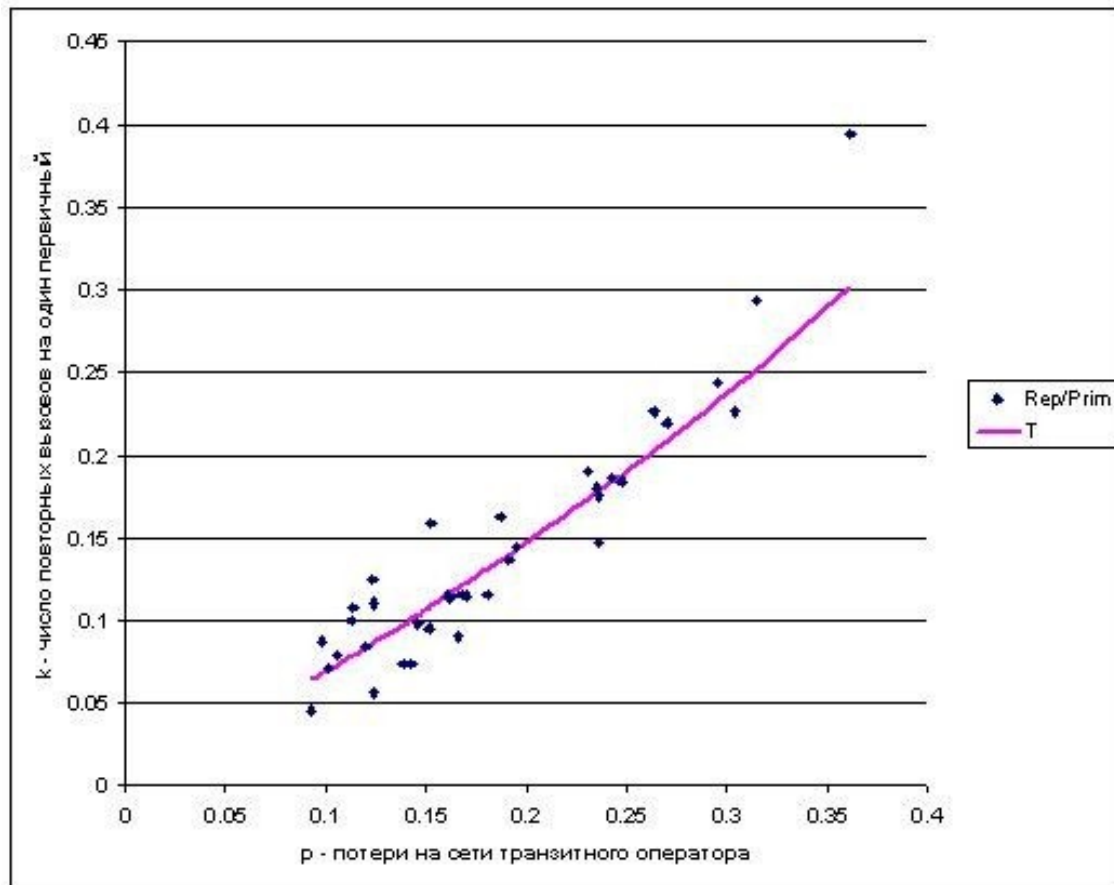


Рис. 3. Иллюстрация поведения абонентов при вызовах в Якутию по коду +7411*. Усреднение по 1000 вызовов

Чтобы не возникало никаких сомнений по поводу правомерности второй значащей цифры в приводимых значениях настойчивости и потерь, представим результаты тех же измерений, что на рис. 3, только с усреднением по пяти ближайшим точкам (см. рис. 4).

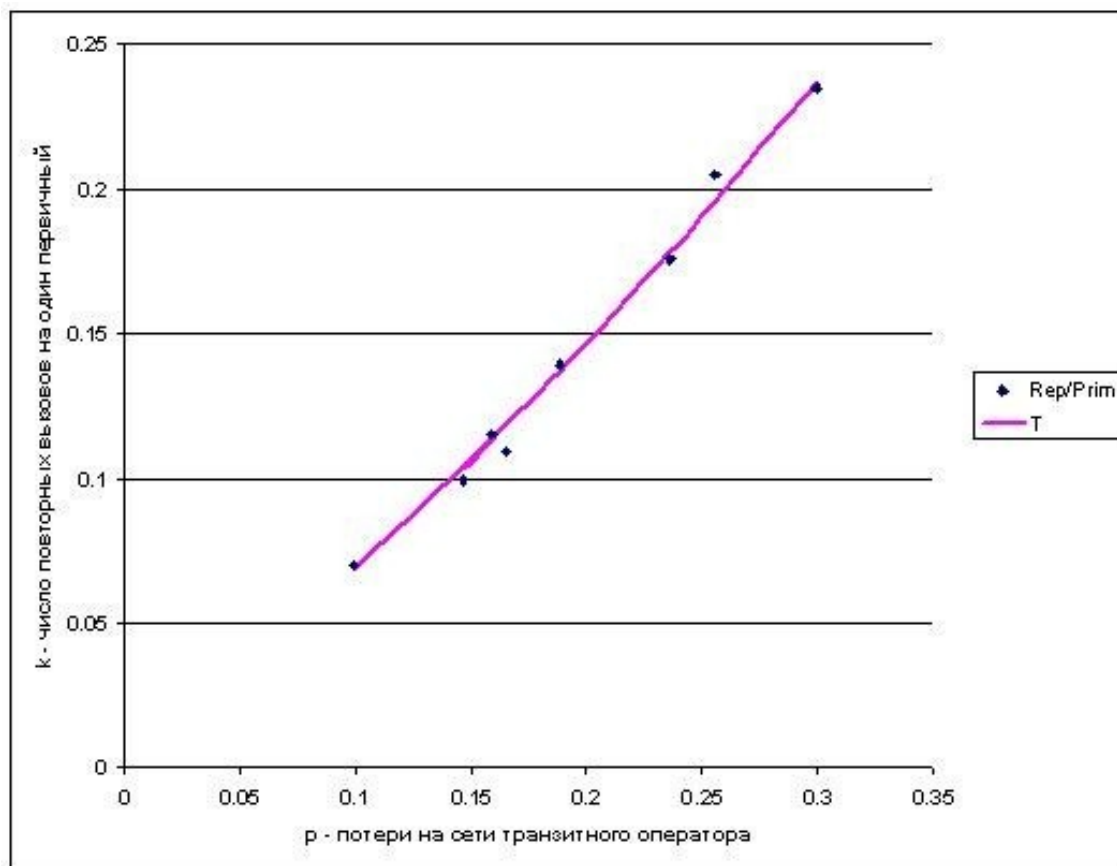


Рис. 4. Иллюстрация поведения абонентов при вызовах в Якутию по коду +7411*. Усреднение по 5000 вызовов

Как видно из графика, во всей области значений обнаруживается хорошее совпадение теоретической зависимости (5) с экспериментальными результатами.

В рассмотренных выше примерах трафик обслуживался с достаточно хорошим качеством. При этом обнаружилась некоторая тенденция к увеличению настойчивости с возрастанием потерь. Что же будет с настойчивостью абонентов, если им придётся иметь дело с направлениями, по которым доступность близка к нулю? Какова предельная настойчивость абонентов?

Для того чтобы ответить на этот вопрос, был проанализирован трафик от международных операторов на направлении Узбекистан моб. (+99893*).

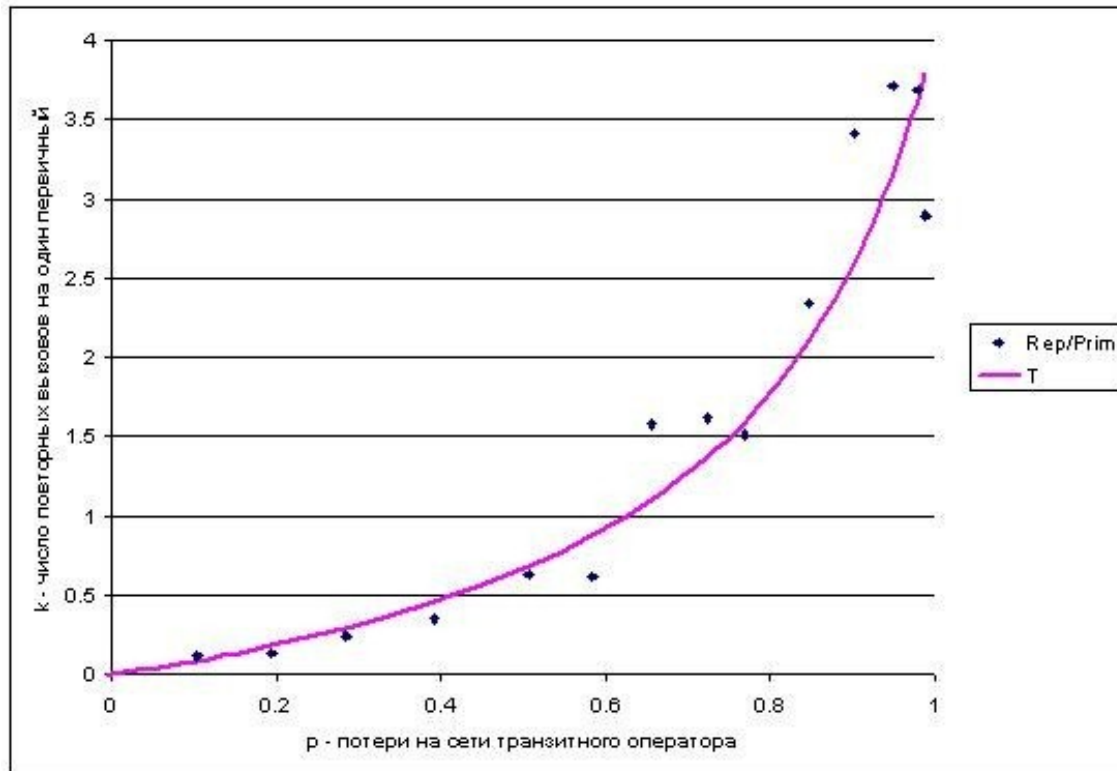


Рис. 5. Иллюстрация поведения абонентов при вызовах в Узбекистан по коду +99893*

На направлении, поведение абонентов для которого изображено на рис. 5, наблюдается почти полное отсутствие доступности в определенные часы. Это направление выбрано не случайно. Чем хуже в целом качество обслуживания на направлении, тем более высокую настойчивость абонентов можно наблюдать при вызовах по данному направлению. Для данного направления настойчивость $H = 0.8$, потери на терминирующем операторе (постоянная составляющая потерь) $p_c = 0.5$. Настойчивость, несмотря на свою тенденцию к увеличению с ухудшением доступности в среднем, в предельном случае, когда $p = 1$, а ASR и NER равны нулю, не превышает значения $H = 0.82$. Иными словами, если связи нет совсем, абонент в среднем делает четыре повторные попытки вне зависимости от направления. Этот факт представляет для автоматизации первостепенное значение, так как различия в настойчивости оказывают сильное влияние именно в области высоких значений потерь, где такому влиянию сильно подвержена целевая функция выбора оператора.

Рекомендовать выбор постоянной, не влияющей на поведение абонента, составляющей потерь $p_c = 0.2$ можно по следующим причинам. Вообще говоря, пороговые значения для показателя NER нигде не регламентированы. Но из сложившейся во всем мире практике (Белгаком, Теленор, Телия и др.) стандартным пороговым значением для NER де-факто стало значение $NER = 80\%$. Пока на сети транзитного оператора $NER > 80\%$, проблем с доступностью нет.

Если показатели меньше, то можно проявлять интерес к качеству на данном направлении. $NER = 80\%$ как раз и соответствует $p_c = 0.2$.

5. РЕАЛИЗАЦИЯ АЛГОРИТМА ПРИНЯТИЯ РЕШЕНИЯ

Итак, в формуле (5) для сети в целом мы рекомендуем использовать $H = 0.8$, $p_c = 0.2$. При этих условиях формула принимает вид:

$$k(q) = \frac{q - 0.2}{1.2 - q}. \quad (6)$$

Наблюдаемые потери q связаны с наблюдаемым ASR соотношением:

$$q = 1 - 2ASR. \quad (7)$$

Подставив (7) в (6), имеем:

$$k(ASR) = \frac{0.8 - 2ASR}{0.2 + 2ASR}. \quad (8)$$

Подставив (8) в (1), имеем для “чистого” ASR:

$$CASR = \frac{ASR}{2ASR + 0.2}. \quad (9)$$

Зависимость (9) выписана для безразмерного (в пределах от 0 до 0.4) показателя ASR. Если нам необходимо “очистить” от повторных вызовов ASR, выраженный в процентах, то можно воспользоваться зависимостью:

$$CASR = \frac{ASR}{2ASR + 20\%}. \quad (10)$$

В силу того, что мы приняли допущение (7) о том, что ASR равен половине NER, а также в силу того, что $p_c = 0.2$ постулировано на основе порогового $NER = 80\%$, в формуле (10) появляется критическое значение $ASR = 40\%$, при котором связь нормализуется и повторные вызовы исчезают. Использовать формулу (10) при $ASR > 40\%$ бессмысленно.

Окончательный алгоритм нахождения CASR по ASR следующий:

$$CASR = \begin{cases} \frac{ASR}{2ASR + 20\%}, & \text{если } ASR < 40\%, \\ 40\%, & \text{если } ASR \geq 40\%. \end{cases}$$

В системе управления трафиком для расчетов предлагается использовать не ASR, а CASR. Пусть a — чувствительность системы, причем, когда $a = 0$, нас заботит в основном качество (чтобы оно было достаточно высоким), а когда $a \rightarrow 1$ — нас беспокоит в основном цена (чтобы она была достаточно низкой). Рекомендуемое значение параметра a для каждого направления определяется маржей на этом направлении — $m, \%$ по формуле:

$$a = 1 - \frac{m}{100\%}. \quad (11)$$

Так, при средней марже по компании $m = 25\%$, рекомендуемое значение по умолчанию для параметра a составляет 0.75.

Маршрутизация трафика на оператора происходит не по одному коду, а сразу по направлению (например, Узбекистан моб. Коском), которое включает целую группу кодов, условимся

называть их “родственными” кодами. Рассмотрим цену и качество пропускания j родственных кодов регионов через каждого из i операторов-провайдеров. Маршрутизация всех j родственных кодов возможна только в целом, на одного провайдера. Найдем минимальную цену по операторам для каждого кода:

$$P_{\min j} = \min_i (P_{ij}). \quad (12)$$

Пороговое значение цены:

$$T_j = P_{\min j} a_j. \quad (13)$$

Это пороговое значение — тот единый для всех операторов порог, от которого надо отсчитывать значение цены покупки возможности пропускания трафика. Заметим, что при $a \rightarrow 1$ мы имеем частный случай — систему LCR (Least Cost Routing) без учета качества для пути первого выбора. Работа с $a = 0$ соответствует ситуации с маржей $m = 100\%$. Рассматривать более “заботящиеся” о качестве варианты не имеет смысла, так как при достаточно большой марже начинают вступать в силу различные показатели SLA (Service Level Agreement), которые накладывают свои “вето” на маршрутизацию.

“Переплата” за качество для каждого оператора будет определяться расстоянием от “порога”: $P_{ij} - T_j$, а набор коэффициентов для принятия решения определится по формуле:

$$D_i = \sum_j \frac{CASR_{ij} W_j}{P_{ij} - T_j}. \quad (14)$$

Здесь W_j — вес данного кода в группе родственных кодов, определяемый по формуле:

$$W_j = \frac{\sum_i BMI_{ij}}{\sum_j \sum_i BMI_{ij}}, \quad (15)$$

где BMI_{ij} — Billing Minutes Indicator, число минут разговорного трафика на каждом операторе по каждому коду за заданный период (в системе — неделя).

Система уже могла бы работать, если бы у нас была полнота данных о качестве, то есть, если бы мы не имели нулевых элементов в матрице CASR. Но по большинству операторов у нас может не быть данных о качестве, а есть данные только о цене. Выход из ситуации можно найти следующим образом: заполнить недостающие элементы матрицы средневзвешенными по всем операторам. Если нет значений показателей качества для данного кода ни на одном из операторов, то он принимается равным средневзвешенному по операторам и по кодам:

$$CASR_j = \frac{\sum_{i, C_{ij} \neq 0} C_{ij} CASR_{ij}}{\sum_i C_{ij}}, \quad (16)$$

где C_{ij} — число вызовов по j -му коду через i -го оператора.

Если для данного i $C_{ij} = 0$ для любого j , то усреднение производим по всем вызовам:

$$CASR_j = \frac{\sum_j \sum_{i, C_{ij} \neq 0} C_{ij} CASR_{ij}}{\sum_j \sum_i C_{ij}}. \quad (17)$$

Далее, из ASR определяется CASR и производится расчет коэффициентов принятия решения. Выборы для направления группы кодов трафика на операторов определяются в порядке убывания коэффициентов принятия решения D_i .

На заключительном этапе работы в алгоритм был добавлен учет достоверности качественных показателей. Известно, что такие показатели доступности как ASR и NER могут быть трафико-зависимыми. То есть, если трафик на сети того оператора, который производит измерения, невелик, то показатели доступности будут высокими; как только увеличивается трафик на сети оператора, производящего измерения, показатели снижаются. В случае, если направление имеет трафико-зависимый профиль ASR, измерение, произведенное на каналах к оператору, на которого, например, направлена лишь 1/10 часть трафика, является менее достоверным, чем измерение, произведенное на каналах к оператору, куда направлено 9/10 всего трафика.

Если не учитывать этот факт в работе системы, то произойдет следующее. Система предложит направить трафик на оператора с лучшим качеством. В этот момент показатель качества на этом новом операторе упадет, в то же время увеличится показатель на операторе, куда трафик шел раньше. Система, обнаружив это, снова предложит переставить операторов по выборам, и так до бесконечности.

Таким образом, мы рекомендуем для реализации в системе нелинейный множитель – такой весовой коэффициент, который при небольших объемах (вблизи порогового значения в 50 минут) незначительно больше 1, далее растет относительно быстро, а при достоверных объемах (10 тыс. минут в сутки) выходит на насыщение к некоторому, в разы большему значению. Данный коэффициент позволит сделать работу системы более сглаженной, более стабильно поддерживать существующий план маршрутизации.

Как вариант такого коэффициента, реализующего при управлении некоторый “гистерезис”, может быть использована, например, функция:

$$f(x) = \frac{3x + c}{x + c}, \quad (18)$$

где c — некоторая постоянная, которая соответствует значению x , при котором $f(x) = 2$. При $c = 600$ мин/час (10 Эрл) функция ведет себя довольно правильно: оператор, который успешно пропустил 10 Эрл разговорной нагрузки, получает удвоение весового коэффициента по сравнению с оператором, через которого прошло лишь несколько десятков вызовов.

Таким образом, “ядро” алгоритма расчета приоритета принимает вид:

$$D = \frac{CASR f(x)}{(P - a P_{min})}. \quad (19)$$

В таком виде данный алгоритм удалось эффективно использовать в системе управления трафиком сети ОАО “МТТ”.

6. ВЫВОДЫ

1. В зависимости (2) числа повторных вызовов на один первичный вызов величина p представляет собой лишь часть наблюдаемых потерь — те потери, при которых информация, поступающая к абоненту А, заставляет его произвести повторную попытку вызова.
2. Зависимость для числа k повторных вызовов на один первичный вызов описывается формулой (5).
3. Для того, чтобы выяснить, какие значения могут принимать в выражении (5) постоянные коэффициенты H и p_c , возникла необходимость проанализировать большой объем данных.

Для этого было разработано и применено специальное программное обеспечение, позволяющее определять экспериментальные значения q и k для трафика по различным направлениям.

4. Диапазон наблюдаемых значений для настойчивости H — от 0.3 до 0.8, для постоянной составляющей потерь p_c — от 0.1 до 0.5. Некоторые выборки представляют собой результат работы сети в двух различных режимах: областях низкой и высокой настойчивости.
5. Для направлений с высокой настойчивостью постоянная составляющая потерь тоже довольно высока (не менее $p_c = 0.2$). Даже для выборок с самой низкой доступностью направления настойчивость не превышала $H = 0.8$. Это позволило рекомендовать предельные параметры $H = 0.8$ и $p_c = 0.2$ в качестве оценки для постоянных коэффициентов для (5).
6. Учесть влияние повторных вызовов предложено путем использования “чистых” показателей ASR — CASR (Clean ASR). Стоимость терминации трафика на каждого из операторов предлагается учитывать в регулируемом коэффициенте переплаты. Влияние объемов трафика на достоверность качественных показателей учтено в специальном нелинейном множителе.
7. Ядро (19) алгоритма обобщено для работы с родственными кодами, а также для работы в условиях недостатка информации (большого числа нулевых данных).

СПИСОК ЛИТЕРАТУРЫ

1. Андрианов Г.А., Порязов С., Цитович И.И. Обслуживание транзитного трафика на сети оператора связи. Информационные процессы, 2010, № 1.
2. William C. Hardy. QoS Measurement and Evaluation of Telecommunications Quality of Service – John Wiley & Sons, Ltd. 2001.
3. ITU-T Rec. E.425 (03/2002) Internal automatic observations.
4. ITU-T Rec. E.420 (11/1988) Checking the quality of the international telephone service – general considerations.
5. ITU-T Rec. E.426 (10/1992) General guide to the percentage of effective attempts which should be observed for international telephone calls.
6. Корнышев Ю.Н., Фань Г.Л. Теория распределения информации. М.: Радио и связь, 1985.