УДК 004.4'242

АДМИНИСТРИРОВАНИЕ ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ НА ОСНОВЕ СТАТИСТИЧЕСКОГО АНАЛИЗА ТРАФИКА

А.К. Скуратов [1], Д.С. Безрукавный [2]

Государственный НИИ информационных технологий и телекоммуникаций (ГНИИ ИТТ «Информика») (1);
Московский государственный университет леса (2), г. Москва

Представлена директором ТамбовЦНИТ В.Е. Подольским и членом редколлегии профессором В.И. Коноваловым

Ключевые слова и фразы: анализ трафика; временной ряд; сетевое администрирование; спектральная характеристика; статистический анализ; телекоммуникационная сеть.

Аннотация: Рассматриваются вопросы оперативной диагностики характеристик трафика в IP-сетях, имеющей важное значение для повышения качества передачи информации. Анализируется состояние сети для решения задач сетевого администрирования, мониторинга маршрутизаторов и другого оборудования магистральных сетей с целью выявления аномального поведения системы или сбоев в работе сети. Используется сбор и анализ различной статистической информации по IP-трафику. Определяются характер сбоя в сети, необоснованные рост или падение интенсивности трафика, изменения в стационарном характере трафика, чрезмерное повышение интенсивности использования отдельных частей сети и т.д. Решения таких задач часто основываются на субъективных методах, появившихся в процессе долговременной работы. В статье описано программное средство, позволяющее автоматизировать этот процесс.

Совершенствование технологической базы анализа и мониторинга телекоммуникационных систем и компьютерных сетей, во многом определяющих развитие страны, имеет важное народно-хозяйственное значение.

Оперативная диагностика характеристик трафика в IP-сетях имеет важное значение для повышения качества передачи информации. Вследствие влияния различных факторов на режим работы информационной сети, она, как правило, включает случайную составляющую, и количественный анализ характеристик исследуемой системы возможен на основе использования статистического подхода и аппарата теории вероятности. На сегодняшний день наибольший интерес и значимость представляет задача исследования характеристик сетевого трафика с целью анализа состояния сети для решения одной из важнейших задач сетевого администрирования, которой является мониторинг маршрутизаторов и другого оборудования магистральных сетей с целью выявления аномального поведения системы, или сбоев в работе сети. Данную задачу можно решать с помощью сбора и анализа различной статистической информации по IP-трафику, проходящему через тот или иной интерфейс сетевого устройства. Аномалии в поведении трафика определяют характер сбоя в сети и могут представлять собой, например, необоснованные рост или падение интенсивности трафика, изменения в стационарном характере трафика, чрезмерное повышение интенсивности использования отдельных частей сети и т.д. Выявление и распознавание аномального поведения

сети администраторами очень часто основывается на субъективных методах, появившихся в процессе долговременной работы в области управления сетью.

Развитие телекоммуникационных систем и компьютерных сетей обуславливает необходимость создания и надежного функционирования большого набора инфокоммуникационных сервисов, обеспечивающих эффективную работу пользователя с разнородной информацией в гетерогенной телекоммуникационной сети. Вместе с тем, исторически сложившаяся неоднородность, как телекоммуникационных систем, компьютерных сетей, сетевых информационных ресурсов, так и аудитории пользователей, которой данная информация адресована, осложняет объективный анализ и мониторинг телекоммуникационных архитектур и ресурсов. Поэтому представляется актуальным, что при эксплуатации телекоммуникационных систем и компьютерных сетей должен быть использован достаточно широкий спектр современных и научно обоснованных технических и технологических решений их анализа и мониторинга. Практика использования и эксплуатации гетерогенных телекоммуникационных систем и компьютерных сетей, связанная с недостаточной их прозрачностью, сложностью, организационными ограничениями и спецификой определяет необходимость более широкого и научно-обоснованного внедрения статистических методов их анализа и мониторинга на основе открытой потоковой информации, которую можно получить, используя предлагаемые методы и средства [1].

Для облегчения задачи сетевого администрирования с целью статистического анализа и мониторинга телекоммуникационной системы (сети) были разработаны теоретические подходы и реализована программа «Анализатор трафика». В функции этой программы входят:

– непрерывный мониторинг трафика сети;

– сравнение текущих параметров с нормальными;

– выдача системному администратору предупреждений и рекомендаций, в случае возникновения отклонений.

У программы есть два режима работы: обучения и анализа. После установки программа запускается в режиме обучения. В этом режиме программа собирает информацию о трафике, воспринимая его как нормальный режим работы сети. Длительность обучения, которая может настраиваться пользователем, равна по умолчанию двум неделям. В течение срока программа раз в пять минут фиксирует количество информации, прошедшей через канал, и запоминает его в файле в текстовом формате. По истечении времени обучения программа обрабатывает накопленные данные и переключается в режим анализа. Обработка накопленной информации заключается во-первых, в спектральном анализе, во вторых, в вычислении тренда, и, в третьих, в определении частоты появления выбросов (выбросом является кратковременное резкое повышение загрузки канала).

Для спектрального анализа используется представление данных в виде ряда Фурье [2]

$$\sum_{i=1}^{n} A_i \cos(\omega_i t_i) + B_i \sin(\omega_i t_i) \, . \qquad (1)$$

Коэффициенты A_i и B_i вычисляются по формулам

$$A_i = \frac{2}{n} \sum_{t=1}^{n} u_t \cos(\omega_i t),$$

$$B_i = \frac{2}{n} \sum_{t=1}^{n} u_t \sin(\omega_i t), \qquad (2)$$

$$\omega_i = 2\pi \frac{i}{N}.$$

Здесь $n = N/2$ (N – количество точек ряда).

Далее из полученных гармоник отделяются те, которые являются значимыми по критерию Фишера

$$\frac{I(\omega_i)/\nu_1}{\sigma^2/\nu_2} > F(0.05, \nu_1, \nu_2),$$
$$I(\omega_i) = A_i^2 + B_i^2, \tag{3}$$

где σ^2 – оценка дисперсии ряда; ν_1 и ν_2 – количество степеней свободы (2 и N).

Полученные гармоники покрывают 95 % дисперсии ряда, и они запоминаются в отдельном файле как спектральная характеристика нормального режима функционирования сети.

Для вычисления тренда используется метод скользящего среднего. Простое сглаживание основывается на составлении нового ряда из простых средних арифметических, вычисленных для промежутков времени длиной q:

$$\bar{x}(k) = \sum_{t=k}^{q+k} x(t)/q, \quad (k = 1,2,..., n-q+1), \tag{4}$$

где длина периода сглаживания q зависит от характера временного ряда, а также от цели сглаживания и выбирается исследователем экспертным путем; k – порядковый номер средней точки окна.

Обработка выбросов заключается в том, что вычисляется среднее количество их появления в течение одной недели.

Все эти параметры сохраняются в отдельном файле и программа начинает работать в режиме анализа. В этом режиме программа раз в 5 минут фиксирует загрузку сети и сравнивает это значение с модельным значением, вычисленным с помощью спектра нормального режима. Помимо этого, программа хранит в памяти данные за последние две недели работы. В случае отклонения полученного значения от эталона более чем на $\sqrt{3}\sigma$ (где σ – дисперсия эталонного ряда) на протяжении 6 измерений (т.е. в течение часа) программа фиксирует серьезное отклонение и начинает определять его причину. Для этого производится спектральный анализ, вычисление тренда и количества выбросов за последние две недели.

Соответственно, может возникнуть одна или несколько из следующих ситуаций:

- изменился тренд;
- изменилась частота появления выбросов;
- изменилась спектральная характеристика ряда;
- изменилось среднее значение ряда.

Самый низкий уровень опасности имеет увеличение количества выбросов. Если это зафиксировано, то администратору будет выдано предупреждение и рекомендация проверить сеть на появление вирусов, т.к. велика вероятность того, что выбросы связаны с деятельностью сетевых «червей».

Более высокий уровень опасности представляет изменение спектральной характеристики ряда. Очевидно, что в нормальном режиме работы сеть подчиняется суточным и недельным колебаниям загрузки. Если же произошли отклонения, значит один или несколько пользователей перешли на круглосуточный (и/или без выходных) режим работы и администратору будет предложено уделить этим пользователям внимание.

Еще более высокий уровень опасности представляет изменение тренда. Оно означает, что происходит медленный, но верный рост загрузки сети и в не очень отдаленном будущем возможно переполнение ее пропускной способности. Адми-

нистратору будет рекомендовано рассмотреть возможность увеличения пропускной способности каналов.

Самым опасным вариантом является скачкообразное увеличение среднего значения ряда. Это означает, что сеть работает в режиме, близком к предельному, и может не справляться с загрузкой. Администратору выдается предупреждение и рекомендация срочно обратить внимание на загрузку сети для определения причины ее увеличения и исправления ситуации.

Помимо случаев отклонения от эталона, программа проводит регулярные проверки параметров сети (по умолчанию – раз в неделю). Это необходимо, так как изменения могут происходить настолько медленно, что не будут фиксироваться описанным выше критерием, но реакция на них все равно необходима. При регулярной проверке проводится та же самая процедура, что и при фиксировании отклонений, как описано выше.

В случае проведения работ по реконструкции или модернизации сети, ее нормальные параметры функционирования изменятся. Поэтому в программе предусмотрена возможность переобучения по новым данным. При ее активации программа переходит в режим обучения и фиксации данных для создания нового эталона, отвечающего текущему состоянию сети.

Программа «Анализатор трафика» написана на языке Borland Delphi 6.0 и предназначена для использования с операционными системами семейства Windows.

В результате обработки статистической информации о функционировании телекоммуникационной сети можно определить нормальный профиль сети (этап анализа). Выявление и предсказание отклонений от нормального профиля сети (этап мониторинга) проводится системным администратором с целью определения возникновения нештатной ситуации и принятия соответствующего решения об изменении конфигурации сети.

Таким образом, является актуальным разработка методов и средств статистического анализа и мониторинга телекоммуникационной сети, обработки первичной информации с использованием выбранных статистических методов анализа и выработка рекомендаций по реконфигурации сети.

В статье представлены отдельные результаты научно-исследовательской работы в этой области, выполняемой в рамках гранта РФФИ 02-07-90026.

Список литературы

1. Енюков И.С. Статистический анализ и мониторинг научно-образовательных интернет-сетей / И.С. Енюков, И.В. Ретинская, А.К. Скуратов. Под. ред. А.Н. Тихонова. – М.: Финансы и статистика, 2004. – 320 с.

2. Кендалл М.Дж. Многомерный статистический анализ и временные ряды / М. Дж. Кендалл, А.М. Стьюарт: М. Наука, 1976.

Administration of Telecommunication Network on the Basis of Statistic Analysis of Traffic

A.K. Skuratov[1], D.S. Bezrukavny[2]

Informika IT and Telecommunication National Research Institute, Moscow (1); Moscow State University of Woods (2), Moscow

Key words and phrases: statistic analysis; telecommunication network; traffic analysis; network administration; spectrum characteristics; temporal row.

Abstract: Being important for improving the quality of information transfer, matters of quick diagnosis of traffic features in IP-networks are studied. To solve the tasks of network administration, monitoring of routers and other equipment of trunk networks, and reveal system abnormal behavior or network disturbance the network state is analyzed. The type of network disturbance, unreasonable growth or drop in traffic intensity, changes in stationary character of traffic, excessive increase in intensive use of single network parts, etc., are revealed. The solution of such problems is often based on subjective methods used in the long course of work. Software means enabling to automate this process is described in the paper.

Administrieren des Telekommunikationsnetzes auf Grund der statistischen Analyse des Verkehres

Zusammenfassung: Es werden die Fragen der operativen Diagnostik der Verkehrscharakteristiken in den IP-Netzen, die die wichtige Bedeutung für die Erhöhung der Qualität der Informationsübergabe haben, betrachtet. Es wird der Zustand des Netzes für die Lösung der Aufgaben des Netzadministrierens, des Monitorings der Einrichtung der Hauptleitungsnetze in der Absicht der Entdeckung des anomalen Verhaltens des Systemes oder der Störungen in der Netzarbeit analysiert. Es wird das Sammeln und die Analyse der verschiedenen statistischen Information nach dem IP-Verkehr verwendet. Es werden der Charakter der Störung ins Netz, das unbegründete Steigen oder das Fallen der Intensität des Verkehrs, die Veränderungen im stationären Charakter des Verkehrs, die übermäßige Erhöhung der Intensität der Benutzung der einzelnen Teile des Netzes u.s.w. bestimmt. Es ist die Software, die diesen Prozeß zu automatisieren zuläßt, beschrieben.

Administration du réseau des télécommunications à la base de l'analyse statistique du trafic

Résumé: Sont envisagés les problèmes du diagnostic opératif des caractéristiques du trafic dans les réseaux IP ayant une grande importance pour l'augmentation de la qualité de la transmission de l'information. Est analysé l'état du réseau pour la solution des problèmes de l'administartion de réseau, du monitoring de l'équipement des réseaux dans le but de l'analyse du conditionnement anomal du système ou bien des décalages dans le fonctionnement du réseau. Sont utilisés le stockage et l'analyse de la différente information statistique d'après le trafic IP. Sont analysés le caractère du décalage, l'augmentation ou bien la chute de l'intensité du trafic, les changements dans le caractère stationnaire du trafic, l'augmentation de l'intensivité de l'emploi des parties du réseau, etc. Les solutions de tels problèmes sont souvent fondées sur les méthodes subjectives qui sont apparues au cours du long travail. Dans l'article est décrit le logitiel qui permet d'automatiser ce processus.