

УДК 004.78

В.И. Саенко, А.С. Голубев

Харьковский национальный университет радиоэлектроники, Харьков

МЕТОД ОЦЕНИВАНИЯ СОСТОЯНИЯ СЕРВИСА В КОМПЬЮТЕРНОЙ СЕТИ ДЛЯ СИСТЕМ НЕПРЕРЫВНОГО МОНИТОРИНГА

В статье описан метод оценивания состояния сервиса в компьютерной сети. Сеть рассматривается как динамическая система, основными функциональными объектами которой являются сервисы. Предложено оценку состояния сервисов свести к формированию качественных оценок состояния и вычислению тренда значений этих оценок. Правильность полученного решения подтверждается на примере.

компьютерные сети, QoS, качество сервиса, непрерывный мониторинг

Описание проблемы и анализ известных результатов исследований

Компьютерные сети сегодня становятся неотъемлемой частью инфраструктуры любого предприятия. Возрастает уровень сложности компьютерной сети, при этом актуальными становятся вопросы повышения эффективности использования имеющихся информационных ресурсов. Информационные ресурсы неотделимы от информационных сервисов, посредством которых и обеспечивается доступ к ресурсам. С ростом сложности компьютерной сети увели-

чивается число поддерживаемых сервисов и возрастает нагрузка на имеющиеся ресурсы. Большинство компьютерных сетей функционируют в режиме дефицита ресурсов, поэтому особую актуальность приобретают вопросы оценивания качества поддерживаемых сервисов. В литературе эти проблемы известны как оценивание QoS (Quality of Service).

Для любой компьютерной сети задача выборочного (разового) оценивания показателя QoS может быть хорошо формализована и решена, например, [1]. В этом случае мы получаем статическую оценку, а объект, компьютерная сеть, также рас-

смачивается как статический (динамические характеристики не принимаются во внимание).

В случае, когда компьютерная сеть рассматривается как динамическая система, задача оценивания показателя QoS становится на порядок сложнее. Дело в том, что состояние динамической системы описывается некоторой непрерывной функцией и, следовательно, показатель QoS, должен быть описан непрерывной функцией или оцениваться через некоторые дискретные интервалы измерения.

Данная задача не нова и для разных классов систем имеются различные решения. Например, в [2] рассматривается метод определения качества сервиса Java сообщений в условиях динамических изменений состояния компьютерной сети. В [3] рассмотрены методы формирования показателя качества для динамического оценивания, при этом предлагается рассматривать задержки и отказы в сети как единственные параметры, определяющие качество. Интересным также является решение, представленное в [4] посвященное оцениванию отдельных сервисов для обеспечения наилучших локальных условий использования ресурсов в сети. Однако при этом не изученными остаются вопросы, связанные с обобщенной формализацией методов оценивания показателей качества.

Данная работа является продолжением статьи [1], в которой описан метод оценивания качества информационных сервисов в компьютерной сети, определенного как статический объект. Предложенный в [1] метод позволяет получить обобщенную оценку при сравнении нескольких показателей качества.

Цель работы: сводится к поиску путей оценивания сервисов компьютерной сети в условиях динамического изменения её основных характеристик.

Структура статьи. Постановка задачи описывает сеть, как объект исследования. Затем описана концепция оценивания качества сервиса. В пункте "Вычисление показателя качества сервиса" дано краткое изложение метода [1]. Далее описывается выбор весовых коэффициентов и формирование качественных оценок, нахождение и оценка тренда. Затем дано описание метода оценивания состояния сервиса как обобщение и систематизация предыдущих пунктов. После краткого анализа метода дается пример. В завершении формализованы основные научные и практические результаты.

Постановка задачи и описание объекта исследования

Пусть имеется некоторая компьютерная сеть Net. В сети присутствуют пользователи $U = \{u_i\}, i = 1..m$. Основными функциональными объектами сети являются сервисы, которые она поддерживает $SType = \{stype_j\}, j = 1..p$. Пользователи сети формируют запросы на предмет предоставления сетью сервиса определенного типа. При удовле-

творении запроса пользователя в обслуживании создается определенный экземпляр сервиса $s_{i,j}$, где i – пользователь сервиса, j – тип сервиса. Сервис $s_{i,j}$ может быть охарактеризован некоторым показателем качества $QoS(s_{i,j})$. В дальнейшем предлагается опускать индексы для экземпляра сервиса, предполагая, что любой экземпляр сервиса определяется типом сервиса и пользователем: для $\forall s_{i,j} \exists u_i \in U$ и $\exists stype_j \in SType$, поэтому $s_{i,j}$ обозначим как s .

В большинстве случаев качество обслуживания пользователя оговаривается в специальном соглашении обеспечения уровня сервиса (SLA – Service Level Agreement) и задается в виде ограничения наблюдаемого параметра, например, $d \leq d^0$ – задержка d доступа к ресурсу не должна превышать некоторой пороговой величины d^0 . В процессе функционирования сети показатель задержки доступа есть некоторая функция (рис. 1.). Следовательно, значение показателя качества также будет изменяться во времени.

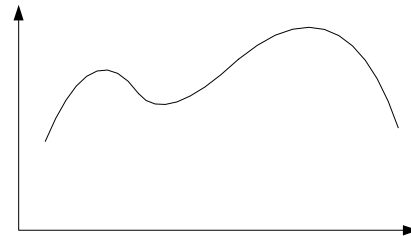


Рис. 1. Изменение задержки во времени при использовании некоторого сервиса

Таким образом, для вычисления характеристик качества сервиса необходимо получить ответы на вопросы, которые и определяют метод оценивания. К ним относятся:

1. Как оценить показатель качества? Очевидно, что в реальных условиях это должно быть дискретное оценивание.
2. С каким шагом по времени осуществлять оценивание этих показателей?
3. Что рассматривать в качестве точечного показателя? Множественное значение или усредненное?
4. Что делать с накопившимися со временем значениями?
5. Что делать с полученными численными результатами, как их оценить?

Современные технологии для наблюдения за состоянием систем предполагают построение специальных систем непрерывного мониторинга компьютерной сети. В рамках построения таких систем предложено решение по оцениванию состояния сервиса.

Следовательно, *постановка задачи* сводится к развитию методов оценивания состояния сервисов в

системах непрерывного мониторинга состояния компьютерной сети. В основе предлагается использовать метод, предложенный в [1].

Концепция оценивания качества сервиса в компьютерной сети

А. Показатель качества $QoS(s)$ рассматривается на основе определенных параметров состояния $[q_i]$. Данные параметры в каждый момент времени принимают определенные значения. Каждый показатель отражает некоторое свойство сети.

$$QoS(s) = f(q_1, q_2, \dots, q_n) = f(q_i), \quad (1)$$

где i – порядковый номер показателя состояния сервиса s , n – число показателей состояния, $QoS(s)$ – показатель качества, безразмерная количественная величина.

В. Для каждого типа сервиса $styp_e_i$ (а в общем случае и для пользователя, то есть для элемента s) задаются ограничения, например, пороговые оценки $QoS^a(s) = (q_1^a, q_2^a, \dots, q_n^a)$, $QoS^b(s) = (q_1^b, \dots, q_n^b)$, $i = 1..n$ – нижняя и верхняя граница показателей состояния. Требуемые значения показателей состояния $QoS^o(s) = (q_i^o)$, $i = 1..n$ – задаются в соответствии с требованиями соглашения SLA для каждого пользователя. При этом выполняются соотношения $QoS^a(s) \leq QoS^o(s) \leq QoS^b(s)$. Допустимые отклонения требуемых показателей задаются как $\Delta QoS(s) = (\Delta q_i)$.

С/ Соотношение вычисленной оценки $QoS(s, t)$ и пороговых оценок представлено на рис. 2.

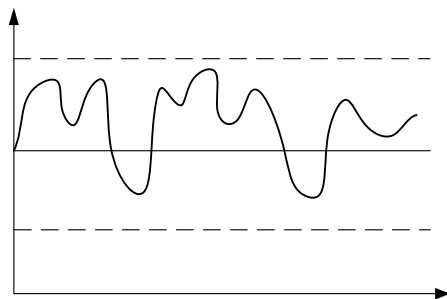


Рис. 2. Соотношение показателя качества сервиса и пороговых оценок

Д. Определение $QoS(s, t)$ осуществляется дискретно, с некоторым заданным интервалом дискретизации и последующей интерполяцией (предпочтительно линейной).

Е. Выбор шага дискретизации осуществляется в зависимости от типа задачи мониторинга (мониторинг при пиковых нагрузках, штатный мониторинг, мониторинг результата изменения конфигурации).

Ф. В качестве точечных значений примем $QoS(s, t) = QoS(s, [t_k])$, где $[t_k]$ – дискретная точка проведения измерений, выбирается некоторое усредненное значение $QoS(s, [t_k])$. При этом: $QoS(s, [t_k]) = f(q_1[t_k], q_2[t_k], \dots, q_n[t_k])$. (2)

Г. Накапливаемый в процессе измерения значений параметров состояния показатель $QoS(s, [t_k])$ сохраняется только на протяжении некоторой истории, определяемой типом решаемой задачи и с учетом эффекта старения информации.

Н. Принимая во внимание, что в системах непрерывного мониторинга актуальной является задача контроля значения показателя, предлагается вместо количественных значений показателя качества использовать качественные значения. При этом вводится новая шкала оценивания показателя качества.

И. Численные значения (количественных) показателей качества предлагается использовать для формирования оценки тренда поведения процесса $QoS(s, t_k)$, с учетом факта старения информации для выборки предыстории дискретных значений $QoS(s, [t_k])$.

Ж. Основными требованиями к методу оценивания качества сервиса являются:

- нересурсоемкость, для выполнения оперативных расчетов,
- интерпретируемость, наглядность.

Формирование и выбор весовых коэффициентов

Весовые коэффициенты позволяют усилить либо уменьшить влияние любого показателя на величину обобщенной оценки. Обычно весовой коэффициент w_i определен на интервале $w_i \in [0, 1]$, так что он интерпретируется как коэффициент ослабления для какого либо показателя. В этом случае используется формула (3).

В более общем случае коэффициент можно задавать целыми числами в любой бальной шкале. Тогда используем формулу для преобразования

$$w_i = \frac{w'_i - w'_{\min}}{w'_{\max} - w'_{\min}}, \quad \text{где } w'_i \text{ - любая бальная}$$

оценка, w'_{\min}, w'_{\max} - граничные значения шкалы оценивания для бальной оценки.

Такой подход позволяет администратору более гибко использовать правила выделения значимости любого показателя при оценивании состояния сервиса.

Формирование качественных оценок

В соответствии с (3) получаем значения показателя качества для $QoS(s)$ текущего момента времени. В системе непрерывного мониторинга эта величина сама по себе ни о чем не говорит, кроме того, накапливаемые значения оказываются неисполь-

зованными. Целесообразно вместо них использовать оценку $Q_1(s)$, отражающую сравнительное состояние сервиса в текущий момент времени, по отношению к пороговым значениям. Для вычисления сравнительной оценки используем пороговые значения в соответствии с требованиями для показателя $q_i - \{q_i^a, q_i^b, q_i^o, \Delta q_i\}$.

В общем случае заранее задать требования к QoS невозможно из-за невысокой интерпретируемости этого параметра. В то же время для показателей q_i такие требования могут быть сформированы достаточно точно. Но для получения качественных оценок для показателя $QoS(s)$ необходимо вместо абсолютных значений q_i иметь относительные значения $QoS(s)$.

Рассматривая ограничения для q_i вычислим соответствующие значения для $QoS(s)$, а именно $QoS^a(s) = 0$ - нижний рабочий порог, $QoS^b(s) = 1$ - верхний рабочий порог. Термин рабочий означает, что $QoS^a(s)$, $QoS^b(s)$ определены для условий, когда любое q_i изменяется в пределах допустимого интервала $[q_i^a, q_i^b]$.

Следует отметить, что при реализации метода непрерывного мониторинга величина $QoS(s)$ может быть и больше 1 и меньше 0. В процессе функционирования наблюдаемой системы значения q_i могут выйти за пределы интервала $[q_i^a, q_i^b]$ и тогда $QoS(s)$ может быть либо $QoS(s) < 0$, либо $QoS(s) > 1$.

Таким образом считаем, что система условно находится в состоянии нормального функционирования, если выполняются условия:

$$0 < QoS(s, [t_k]) < 1, \quad (5)$$

$$\forall i \quad 0 < \tilde{q} < 1. \quad (6)$$

Условие (6) более строгое, чем (5).

Система должна функционировать так, чтобы выполнялись требования SLA, тогда можно вычислить еще две оценки, соответствующие коридору требуемого качества. Так как известны (q_i^o) и (Δq_i) , то $QoS^c(s)$ и $QoS^d(s)$ - соответственно верхние и нижние границы будут получены как:

$$q_i^c = q_i^o - (g_i) * \Delta q_i, \quad (7)$$

$$q_i^d = q_i^o + (g_i) * \Delta q_i, \quad (8)$$

где $G = (g_i)$; $g_i = \begin{cases} 1, & \text{если } q_i \rightarrow \text{opt} \sim q_i \rightarrow \text{max}; \\ -1, & \text{если } q_i \rightarrow \text{opt} \sim q_i \rightarrow \text{min}. \end{cases}$

$$QoS^c(s) = \frac{\sum_{i=1}^n w_i \tilde{q}_i^c}{\sum_{i=1}^n w_i}, \quad (9) \quad QoS^d(s) = \frac{\sum_{i=1}^n w_i \tilde{q}_i^d}{\sum_{i=1}^n w_i}. \quad (10)$$

Соотношение пороговых показателей состояния с меняющимся со временем показателем состояния показано на рис. 3.

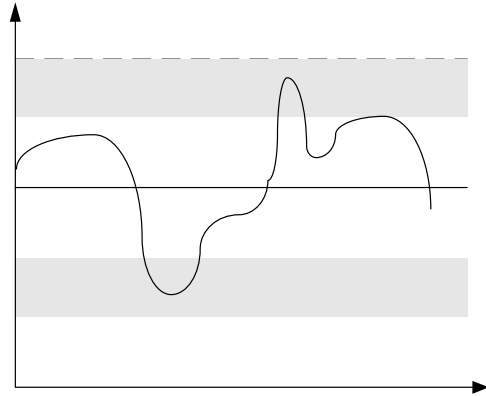


Рис. 3. Соотношение ограничений и показателя состояния сервиса

Качественные оценки предлагается формировать на основании некоторой шкалы. Допустимы следующие значения оценки:

$$Q_1(s) \in \{ < \text{отлично} >, < \text{хорошо} >, < \text{удовлетворительно} >, < \text{плохо} >, < \text{очень плохо} > \}$$

Таблица 1

Качественные оценки состояния сервиса

Оценка $Q_1(s)$	Условие
Очень плохо	$QoS(s) < 0$
Плохо	$0 < QoS(s) < QoS^c(s)$
Нормально	$QoS^c(s) < QoS(s) < QoS^d(s)$
Хорошо	$QoS^d(s) < QoS(s)$
Очень хорошо	$QoS(s) > 1$

Оценивание тренда

Все предыдущие результаты относились к получению точечных оценок и не учитывали динамики процесса. Этого вполне достаточно для решения задач текущего мониторинга для $QoS(s)$. Использование накопленных результатов $QoS(s, [t_k])$ за некоторый интервал $[T_o, T_n]$ позволяет спрогнозировать текущие значения, а также оценить тренд процесса. Основное назначение такой операции состоит в более точном оценивании текущей ситуации. При этом предлагается прогнозировать на два шага вперед. Для вычисления прогнозирующих оценок используем линейную регрессию, которую строим на основе

метода наименьших квадратов, на выборке (истории) длиной $r = 5$. Величина $r = 5$ принимается аксиоматически достоверной для процессов вычисления QoS. Дело в том, что выборка длиной $r \leq 3$ изначально не представительна, а $r = 10$ - избыточно велика, т.к. для значений QoS актуальными являются значения в интервале $[t_{k-2}, t_{k+1}]$ на множестве $\{t_{k-2}, t_{k-1}, t_k, t_{k+1}\}$.

Таким образом, для оценки состояния сервиса $Q(s)$ добавляем оценку состояния прогнозируемого показателя качества $Q_2(s)$ (табл. 2). Текущий момент времени – $QoS'(s) = QoS(s, [t_{k+2}])$ t_k . Оценка имеет три состояния:

$$Q_2(s) \in \{ < \text{очень плохо} >, < \text{плохо} >, < \text{нормально} > \}.$$

Таблица 2
Качественные оценки прогнозируемого состояния

Оценка $Q_2(s)$	Условие
Очень плохо	$QoS'(s) < 0$
Плохо	$0 < QoS'(s) < QoS^c(s)$
Нормально	$QoS'(s) > QoS^c(s)$

Метод оценивания

Метод оценивания будет представлен в виде перечня действий:

1. Выбираемый наблюдаемый сервис s_{ij} определенного типа $styp_{e_j}$, для некоторого пользователя u_i .
2. Для выбранного экземпляра сервиса s_{ij} определяем вектор показателей состояния (q_1, q_2, \dots, q_n) .
3. Для каждого показателя задаем весовой коэффициент значимости w_i .
4. Выполняем операции нормирования и нормализации в соответствии с (4).
5. Рассчитываем показатель качества $QoS(s)$ (3).
6. Определяем показатели качества для порогов (7)-(10) и формируем условия для получения оценки качества $Q_1(s)$ с учетом пороговых требований.
7. Рассчитываем оценку прогнозируемого состояния $Q_2(s)$.
8. Определяем окончательное состояние системы в виде $Q(s) = (Q_1(s, [t_k]), Q_2(s, [t_k]))$.

Анализ метода

Описанный метод обладает следующими характеристиками:

- оперативность расчета, не ресурсоемкость – достигается малым числом анализируемых значений;
- возможность задания порога – в данном методе учитываются минимально и максимально допустимые показатели качества, заданные множеством показателей состояния;
- возможность работы с разными типами показателей для формирования комплексной оценки –

показатель качества вычисляется из множества параметров, определяющих качество сервиса (по [1]);

- возможность изменения влияния различных параметров сервиса на конечную оценку – при изменении приоритетов, вес показателей может меняться;
- учет требований заданных в юридических документах – SLA;
- возможность прогнозирования показателя качества на два шага вперед;
- благодаря тренду есть возможность определить – ухудшается или улучшается состояние сервиса;
- хорошая интерпретируемость – возможно представление результатов работы метода в виде графиков, а также на оценочной шкале.

Пример

1. Сервис s является сервисом типа TCP. Программным приложением является приложение IP телефонии $styp_e = \langle TCP, IP\ phone \rangle$.

2. Показатели состояния определены как $QoS(s) = (q_i) = (b, d)$, где b – скорость передачи данных, измеряется Kbps; d – задержка (временной интервал от отправления данных до их получения, измеряется в миллисекундах).

Измерены показатели состояния для моментов времени t_k , где $k=1..5$. Текущим моментом времени является $t_k, k=5$.

Таблица 3
Значения скорости передачи данных и задержки сервиса s_i для моментов времени t_k

k	$b[t_k]$ (Kbps)	$d[t_k]$ (ms)
1	300	30
2	320	36
3	325	40
4	381	60
5	390	90

Заданы пороговые значения показателей состояния, показатели состояния оптимального показателя качества и допустимые изменения показателей состояния:

$$QoS^a(s) = (100\text{kbps}, 150\text{ms}), \quad QoS^b(s) = (390\text{kbps}, 25\text{ms}),$$

$$QoS^o(s) = (150\text{kbps}, 60\text{ms}), \quad \Delta QoS(s) = (10\text{kbps}, 10\text{ms}).$$

Необходимо оценить состояние сервиса путем определения качества в текущий момент времени $k=5$, а также определить тренд и оценить предсказанное значение.

3. Введем весовые коэффициенты на основе экспертного оценивания на интервале $[0,1]$: $W=(0,2;0,8)$.

4. Представим (2) в виде матрицы имеющих значения, нормируем и нормализуем показатели состояния (4).

$$QoS(s, [t_k]) = \begin{pmatrix} 300 & 30 \\ 320 & 36 \\ 325 & 40 \\ 381 & 60 \\ 390 & 90 \end{pmatrix}, \quad Q\tilde{d}S(s, [t_k]) = \begin{pmatrix} 0,69 & 0,96 \\ 0,75 & 0,912 \\ 0,78 & 0,88 \\ 0,97 & 0,72 \\ 1,00 & 0,48 \end{pmatrix}.$$

При этом количество столбцов будет равно количеству показателей состояния – 2, а количество строк – длине истории, в течении которой сохранялись измеренные значения – 5:

5. Рассчитываем показатель качества по (3): $Q\tilde{d}S(s) = Q\tilde{d}S(s, [t_5]) = 0,58$.

6. Определяем показатели качества для порогов: $QoS^a(s) = 0$, $QoS^b(s) = 1$. $q_1^c = 140\text{Kbps}$; $q_2^c = 70\text{ms}$ по (7); $q_1^d = 160\text{Kbps}$; $q_2^d = 40\text{ms}$ по (8) при этом $G=(1,-1)$. Вычислим показатели качества для порогов требуемого качества: $QoS^c(s) = 0,56$ по формуле (9), $QoS^c(s) = 0,7$ по формуле (10).

Определим оценку качества, учитывая пороговые требования: согласно табл. 1 выполняется условие $0 < QoS(s) < QoS^c(s)$ – что соответствует оценке $Q_1(s) = \langle \text{нормально} \rangle$.

7. Рассчитаем оценку прогнозируемого состояния, для этого сначала определим численные значения показателей качества для моментов времени t_k , где $k=1..5$:

$$Q\tilde{d}S(s, [t_k]) = (0,9; 0,88; 0,85; 0,76; 0,58).$$

Определим тренд процесса для предложенной методики оценивания. Для получения значений $QoS(s, [t_k])$ для точек $[k+1]$ и $[k+2]$ используем линейную регрессию вида $y=mx+b$. Коэффициенты m и b определяем на основании выборки длиной $r=5$ по методу наименьших квадратов: $m=-0,08$; $b=1,03$. При этом линейная функция принимает вид $y = -0,8x+1,03$. Соотношение значений $QoS(s, [t_k])$, значений функции регрессии и порогов представлено на рис. 4.

В соответствии с табл. 2. получаем $Q_2(s) = \langle \text{плохо} \rangle$, так как выполняется условие

$$0 < QoS'(s) < QoS^c(s).$$

Таким образом, состояние системы для точки текущего мониторинга оценивается как

$$Q(s) = (Q_1(s, [t_k]), Q_2(s, [t_k])) = \langle \text{нормально} \rangle, \langle \text{плохо} \rangle.$$

Делаем вывод, что система требует применения управляющих воздействий, для того, чтобы избежать дальнейшего ухудшения качества.

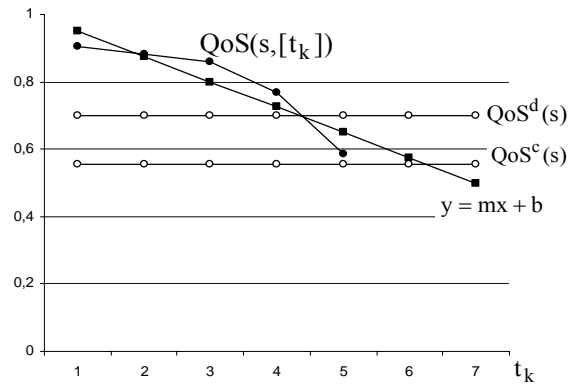


Рис. 4. Соотношение порогов, аппроксимирующей функции и значений показателей качества

Выводы

В работе предложен метод оценивания состояние сервиса в рамках создания и развития систем непрерывного мониторинга. Сервис рассматривается, как функциональный объект сети, характеризующийся множеством показателей состояния сервиса, меняющихся во времени. Качество сервиса определяется множеством этих показателей и также меняется во времени. Предложенный метод, учитывая заданные ограничения в виде порогов, позволяет оценить сервис по его качеству в данный момент времени и по предсказанному качеству для последующих (двух) моментов времени.

Научная новизна состоит в том что: получил дальнейшее развитие метод оценивания состояния сервиса. Предложено оценку состояния сервисов свести к формированию качественных оценок состояния и вычислению тренда значений этих оценок.

Практическая значимость состоит в повышении управляемости сети и снижении расходов на её эксплуатацию за счет возможности анализа качества сервиса во времени, предсказания его изменений, что позволяет своевременно вырабатывать соответствующие управляющие решения.

Сравнение с лучшими аналогами. Данную работу можно рассматривать как дальнейшее развитие идей, высказанных в [1], позволяющих определить динамические изменения качества сервиса. Предложенный метод предлагает обобщенный формальный аппарат для оценивания состояния сервисов в отличии от работ, где оценивание производится только для определенного типа сервиса [2], либо на основании заранее заданных параметров состояния сервиса [3]. Результаты могут быть использованы вместе с методами [4] для оценки и контроля предоставляемого качества.

Направления дальнейших исследований. Дальнейшая работа будет направлена на создание методов определения влияния правил политики на качество сервисов в сети при наличии требований к качеству.

Также будет создано программное приложение для мониторинга состояния заданного процесса на основе описанного метода.

Список литературы

1. Саенко В.И., Голубев А.С. Метод оценивания качества информационных сервисов в корпоративной сети // Вестник НТУ «ХПИ». – Х.: НТУ «ХПИ», 2007. – №18. – С. 124-132.

2. Chen S., Greenfield P. QoS Evaluation of JMS: An Empirical Approach // Proc. of the 37th Annual Hawaii International Conference on System Sciences (HICSS' 04) – 2004. – Track 9. – P. 10-18.

3. Саенко В.И., Кривонос Т.И. Метод оценивания качества доступа к ресурсам для пользователей компьютерной сети // Радиоэлектроника и информатика. – Х.: ХНУРЭ, 2006. – № 4. – С. 66-41.

4. Chen Lee, Lehozky J., Rajkumar R., Siewiorek D. On quality of service optimization with discrete QoS options // Real Time Technology and Applications Symposium. – Vancouver. – 1999. – P. 276-286.

Поступила в редколлегию 4.10.2007

Рецензент: д-р техн. наук, проф. Н.И. Самойленко, Харьковский национальный университет радиоэлектроники, Харьков.