

УДК 681.325

ОПРЕДЕЛЕНИЕ ОПТИМАЛЬНОГО РАЗМЕРА БУФЕРА ОНЛАЙНОВОГО ПОТОКА ДАННЫХ

П. Е. Пустовойтов

Кандидат технических наук, доцент
 Кафедра системы информации
 Национальный технический университет «Харьковский политехнический институт»
 ул. Фрунзе, 22, г. Харьков, Украина
 Контактный тел.: 097-737-44-91
 E-mail: pustovoitov@kpi.kharkov.ua

У роботі пропонується розрахунок раціонального обсягу буфера пакетів даних, при якому імовірність відмови системи не перевершить задану величину. Отримано аналітичне вираження й побудований графік залежності ймовірності відмови від кількості пакетів у буфері

Ключові слова: комп'ютерні мережі, математичне моделювання, відмовостроможність

В работе предлагается расчет рационального объема буфера пакетов данных, при котором вероятность отказа системы не превзойдет заданную величину. Получено аналитическое выражение и построен график зависимости вероятности отказа от количества пакетов в буфере

Ключевые слова: компьютерные сети, математическое моделирование, отказоустойчивость

It was suggested a calculation of rational value of package buffering, that grants the probability of system failure do not exceed the set value. It was received the analytic equation and graph of system refuse dependence of package quantity was built

Key words: networks, mathematical modeling, reliability

1. Введение

При построении компьютерной сети или при разработке сетевого программного обеспечения не всегда бывает достаточно для корректной работы программ гарантировать надежность или скорость передачи пакетов данных. Для таких приложений как аудио- и видеопередача не так уж важно какое время занимает передача пакетов пока время доставки постоянно. В данном случае необходимо учитывать флуктуацию пакетов – колебание времени доставки (среднеквадратическое отклонение).

Если одни пакеты будут доставляться за время t_1 , а другие за t_2 , то изображение или звук начнут дрожать.

Выходом из сложившейся ситуации является сохранение пакетов в буфере на принимающей стороне перед тем, как доставлять потребителю [1,2]. Выдача пакетов осуществляется с некоторой задержкой, но зато пакеты можно выбирать из буфера через равные промежутки времени.

При этом, буферизация не сказывается на пропускной способности и надежности канала. Данный алгоритм, обеспечивающий выборку пакетов с постоянной скоростью, называется «алгоритм дырявого ведра» [3].

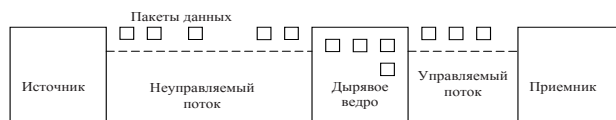


Рис. 1. Алгоритм «дырявого ведра»

Его принцип состоит в следующем. Независимо от скорости поступления пакетов некоторого потока буфер ведра накапливает пришедшие пакеты и отдает с постоянной скоростью приемнику потока, таким образом, к примеру, картинка проигрываемого фильма не дергается, а кадры обрабатываются с одинаковой скоростью. Другими словами, получается так, что пакеты, доставленные быстрее среднего времени ожидают в буфере больше, а если отстают по времени, то ждут обработки меньше. Данная технология решается как аппаратно, так и программно операционной системой.

Если входящий поток данных превосходит интенсивность выходного потока ведра, то пакеты в нем накапливаются, если временно интенсивность входа снижается, то из ведра вынимается больше пакетов, чем в него поступает, и ведро начинает пустеть. Когда в ведре не останется ни одного пакета, то, к примеру, при просмотре видео картинка зависнет или при прослушивании mp3 остановится песня. Постоянное ожидание каждой новой буферизации сделает просмотр

мультимедиа очень не комфортным. С другой стороны, долгое ожидание первой большой буферизации может привести к тому, что пользователь посчитает, что канал потерян и откажется от просмотра.

Таким образом, естественно поставить задачу определения оптимального уровня буферизации онлайн-потока данных такого, чтобы одновременно снизить вероятность простоя проигрывателя мультимедиа, но и чтобы время первичной буферизации было не велико.

2. Постановка задачи исследования.

Рассмотрим возможную приемлемую для анализа модель процесса функционирования в описанных условиях. На вход узла поступает нестационарный пуассоновский поток с интенсивностью $\lambda(t)$, закон изменения которого статистически оценен. Если вариации значений интенсивности входного потока пакетов происходит не слишком быстро, то на интервале управления обслуживанием $[T_0, T_{max}]$ допустима пуассонно-линейная аппроксимация $\lambda(t)$ [4, 5, 6] в форме

$$\lambda(t) = \begin{cases} \lambda_1, & t \in [T_0, t_1), \\ \lambda_2, & t \in [t_1, t_2), \\ \dots & \dots \\ \lambda_k, & t \in [t_{k-1}, t_k), \\ \dots & \dots \\ \lambda_s, & t \in [t_{s-1}, T_{max}). \end{cases}$$

При этом на каждом интервале система обслуживания ведет себя как одноканальная марковская цепь с пуассоновским входящим потоком интенсивностью λ , очередью неограниченной длины и детерминированным обслуживанием с заданным числом пакетов μ , отбираемых из очереди один цикл обслуживания. В соответствии с теоремой Севастьянова для описания процесса функционирования такой системы может быть использована Эрланговская теорема. Будем считать, что накопление пакетов в очереди продолжается до тех пор, пока их число достигает некоторого значения n , после чего начинается порционный отбор пакетов из очереди. Поставим задачу отыскания такого значения n , при котором вероятность того, что в системе не окажется ни одного пакета, не будет превышать заданную P_3 .

3. Основные результаты

Как известно [7,8,9], для одноканальной марковской системы с входящим потоком интенсивности λ и заданным числом n мест для ожидания в очереди вероятность того, что система окажется свободной равна

$$P_0(n, \lambda) = \frac{1}{1 + \alpha + \alpha \sum_{s=1}^n \alpha^s} = \frac{1}{\sum_{s=0}^{n+1} \alpha^s} = \frac{1 - \alpha}{1 - \alpha^{n+2}}, \quad (1)$$

где $\alpha = \frac{\lambda}{\mu}$ – приведенная плотность входящего потока.

Для отыскания требуемого значения n сформируем уравнение

$$P_0(n) = \frac{1 - \alpha}{1 - \alpha^{n+2}} \leq P_3. \quad (2)$$

Отсюда

$$1 - \alpha^{n+2} = \frac{1 - \alpha}{P_3}, \quad \alpha^{n+2} = 1 - \frac{1 - \alpha}{P_3},$$

$$n = \frac{\ln\left(1 - \frac{1 - \alpha}{P_3}\right)}{\ln \alpha} - 2.$$

Из (2) следует, что с увеличением n значение вероятности $P_0(n)$ снижается, причем предельное значение этой вероятности равно $1 - \alpha$.

Это означает, что рациональная интенсивность отбора пакетов из очереди должна быть максимально близка к интенсивности их поступления.

Не более сложным является решение задачи для случая, когда интенсивность входящего потока - случайная величина с заданной плотностью распределения $\phi(\lambda)$. Тогда соотношение

$$P_0(n) = \int_{\lambda_{min}}^{\lambda_{max}} P_0(n, \lambda) \phi(\lambda) d\lambda \quad (3)$$

определяет среднее значение вероятности того, что система не содержит ни одного пакета в произвольный момент времени. Зададим плотность $\phi(t)$, например, нормальным распределением

$$\phi(\lambda) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{(\lambda - \lambda_0)^2}{2\sigma^2}\right\}, \quad \sigma < \frac{\lambda_0}{3}.$$

Пусть интенсивность отбора равна μ_0 . Тогда состояние (3) принимает вид

$$P_0(n) \cong \int_{\alpha_{min}}^{\alpha_{max}} P_0(n, \alpha) \phi(\alpha) d\alpha = \int_{\alpha_{min}}^{\alpha_{max}} \frac{1 - \alpha}{1 - \alpha^{n+2}} \times \frac{\mu_0}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{(\alpha - \alpha_0)^2 \mu_0^2}{2\sigma^2}\right\} d\alpha. \quad (4)$$

Теперь для заданных значений n проведем численное интегрирование (4) и затем построим график зависимости $P_0(n)$, представленный на рис. 2.

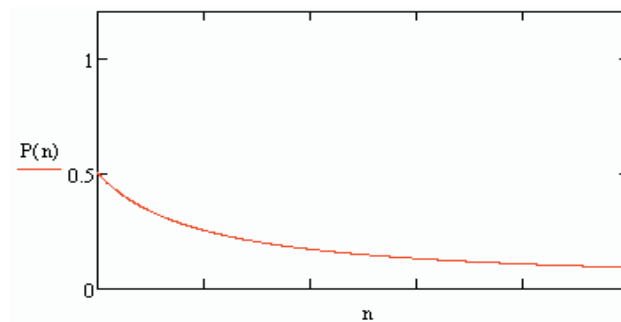


Рис. 2. Зависимость $P_0(n)$, $\mu = \lambda_0$.

Требуемое значение n отыскивается из графика. Таким образом, предложена технология отыскания рациональной продолжительности накопления пакетов до момента начала их использования по назначению, обеспечивающая с заданной вероятностью непрерывный режим работы системы.

4. Выводы

Таким образом, была получена методика определения оптимального размера буферизации пакетов перед

началом использования данных, такого, который бы с заданной вероятностью гарантировал безотказность постоянного поступления пакетов. Данная методика может быть использована в программном обеспечении для проигрывания мультимедиа, такого как просмотр онлайн видео фильмов и музыки, а также онлайн телевидения и радио. Внедрение разработанной технологии не требует особого нового оборудования или переписывания протоколов, а может быть легко добавлено в виде модуля проигрывателя мультимедиа на компьютере клиента.

Литература

1. Таненбаум Э. Компьютерные сети. 5-е издание: пер. с англ. / Э. Таненбаум. – СПб. – ПИТЕР, 2010. – 844с.
2. Столлингс В. Современные компьютерные сети. 2-е изд.: пер. с англ. / В. Столлингс. – СПб.: ПИТЕР, 2003. – 783с.

3. Олифер В. Г. Компьютерные сети. Принципы, технологии, протоколы: Учебник для вузов. 4-е изд. / Олифер В. Г., Олифер Н. А. – СПб.: Питер, 2010. – 944 с.
4. Гмурман В.Е. Теория вероятностей и математическая статистика / В.Е. Гмурман. – М.: Высш. шк., 1972. – 368с.
5. Старков И.И. Статистическая обработка наблюдений / И.И. Старков. – М.: БИНОМ, 2003 – 312с.
6. Крамер Т. Математические методы статистики: пер. с англ. / Крамер Т. – М.: МИР, 1975. – 648с.
7. Хемди А. Таха Введение в исследование операций, 7-е издание.: пер. с англ. / Хемди А. Таха – М.: Издательский дом «Вильямс», 2005. – 912с.
8. Кельберт М. Я. Вероятность и статистика в примерах и задачах. Т. II: Марковские цепи как отправная точка теории случайных процессов и их приложения / М. Я. Кельберт, Ю. М. Сухов – М.: МЦНМО, 2009. – 295 с.
9. Кемени Дж. Конечные цепи Маркова: пер. с англ. / Дж. Кемени, Дж. Снелл – М.: Наука, 1970. – 271с.

Розглянуто проблему вирішення неоднозначності слів природної мови. Запропоновано метод і алгоритм вирішення неоднозначності слів на основі застосування компонентного аналізу із використанням тезаурусу семантичних полів для побудови семантичних обсягів слів

Ключові слова: неоднозначність слів природної мови, компонентний аналіз, тезаурус

Рассмотрена проблема разрешения неоднозначности слов естественного языка. Предложены метод и алгоритм разрешения неоднозначности слов на основе применения компонентного анализа с использованием тезауруса семантических полей для построения семантических объемов слов

Ключевые слова: неоднозначность слов естественного языка, компонентный анализ, тезаурус

The problem of resolution of ambiguity words of natural language is considered. The method and algorithm disambiguate words based on the use of component analysis using a thesaurus of semantic fields for the construction of semantic volumes of words are proposed

Key words: ambiguity words of natural language, component analysis, thesaurus

УДК 681.3:519.76

РАЗРЕШЕНИЕ НЕОДНОЗНАЧНОСТИ СЛОВ ЕСТЕСТВЕННОГО ЯЗЫКА

Ю. Ю. Черепанова

Ассистент

Кафедра «Программное обеспечение ЭВМ»

Харьковский национальный университет

радиоэлектроники

пр. Ленина, 14, г. Харьков, Украина, 61166

Контактный тел.: (057) 402-14-46

E-mail: cher_y@list.ru

1. Введение

Одной из проблем, усложняющих автоматическую обработку информации на естественном языке, является неоднозначность языковых знаков. Она может проявляться в нескольких формах:

- простые омонимы (слова одной части речи, одинаковые по написанию, но разные по лексическому значению);

- слова, обозначающие разные лексические значения многозначного слова (имеющие в своих значениях что-то общее);