

Зиятдинов А.Ю.,
Толстикова Е.В.,
Холявкина Т.В.

СИСТЕМНЫЙ АНАЛИЗ ПРИКЛАДНЫХ ЗАДАЧ РАСПРЕДЕЛЕННОЙ ПЕРЕДАЧИ ДАННЫХ

Факультет компьютерных наук
Национального авиационного университета

Разработан метод оценки потенциальной эффективности распределенной передачи данных в сетях с гетерогенной структурой. Рассчитаны количественные характеристики выигрыша для широкого диапазона условий применения методов передачи с фрагментацией пакетов данных. Даны асимптотические оценки оптимального числа фрагментов пакета данных по критерию эффективность/стоимость.

Введение

Современные компьютерные сети представляют собой набор фрагментов с разными физическими средами передачи данных, эксплуатационными и техническими характеристиками, различными топологиями [1]. По существу, они представляют собой составные сети с высокой степенью гетерогенности.

Благодаря компьютеризации сетевого оборудования, совершенствованию средств мониторинга, расширению функциональности появляется возможность приблизиться к потенциальным характеристикам сети, в первую очередь, по пропускной способности и качеству сервиса в целом. Кроме того, появляется возможность управлять логической топологией в широких пределах при фиксированной физической топологии сети [2]. Эту задачу можно автоматизировать и решать в реальном времени. Поэтому проблема фрагментации пакетов при передаче как по сети в целом, так и для отдельных, слабо связанных между собой сегментов является актуальной.

Путем фрагментации пакетов и отправки по N параллельным маршрутам можно добиться таких результатов:

– ускорение доставки и повышение надежности доставки (короткие пакеты быстрее проходят через транзитные узлы, а в беспроводных сетях и сетях радиодатчиков уменьшаются вероятности колли-

зий, потерь пакетов при занятости буфера и т.д.);

– повышение оперативности передачи, сохранности, целостности информации, снижение вероятности перехвата, расшифровки и модификации информации по перехваченным коротким фрагментам.

При этом параметры маршрутов (в первую очередь, средняя задержка) не должны сильно отличаться друг от друга, например, коэффициент вариации задержки σ_x/m_x должен быть заметно меньше единицы: $\frac{\sigma_x}{m_x} < 0,1$. В против-

ном случае задача оптимизации числа параллельно передаваемых фрагментов будет иметь множество локальных экстремумов, скаляризовать задачу не удастся, а глобальный экстремум может распасться на множество близких по величине мелких локальных экстремумов. Вследствие этого достоверность, а, следовательно, ценность результата оптимизации будет низкой.

Одним из показателей качества сервиса является отношение числа ошибочно переданных бит к общему числу переданных бит (BER). Чем хуже BER , тем больше искажений пакетов и повторных передач. Поэтому задача количественной оценки BER в процессе передачи с фрагментацией представляет не только теоретический, но и практический интерес.

Постановка задачи

Следуя работе [3], при организации процесса передачи разобьем его на последовательность отдельных этапов (фаз) по следующим причинам:

1. Некоторые фрагменты данных, приложений, файлов могут быть слабо связанными друг с другом, поэтому они могут выполняться или обрабатываться параллельно. Более того, они могут распределяться по обслуживающим приборам (ОП), которые находятся в различных сегментах одной сети или даже в сетях разных операторов. При этом сокращается среднее время передачи.

2. Вероятность безошибочной передачи данных и успешного завершения короткого фрагмента пакета выше, чем длинного фрагмента и, тем более, чем пакета целиком.

На рис. 1 показана условная схема процесса разбиения пакета данных и направления отдельных фрагментов на разные ОП. Отметим, что поток заявок интенсивностью λ делится на N потоков с средними интенсивностями λ/N , поэтому задача обслуживания упрощается.

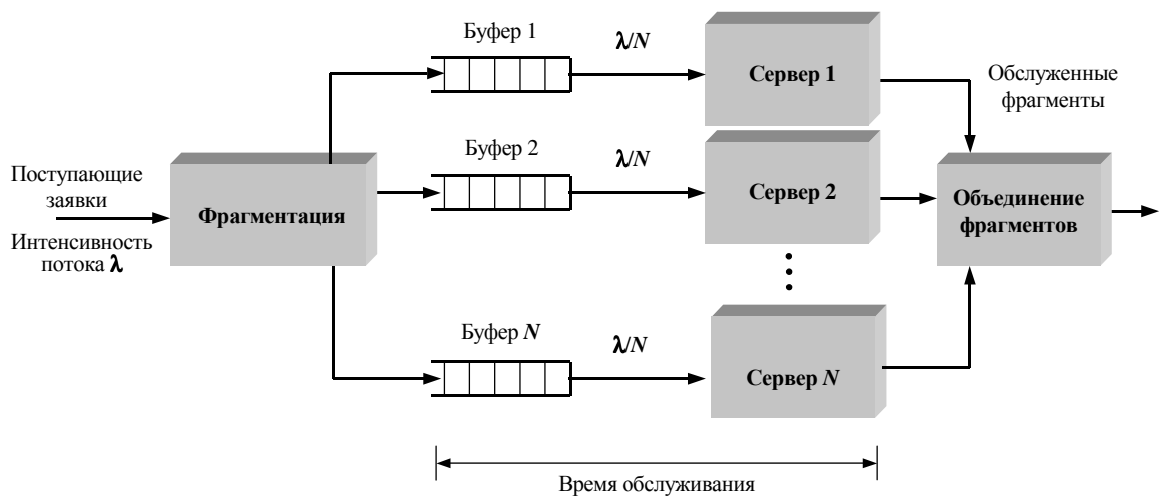


Рис. 1. Процессы фрагментации, обслуживания и объединения фрагментов пакета

Допустим, что пакет состоит из элементарных (неделимых, атомарных)

фрагментов общим числом N . Безусловная вероятность передачи n фрагментов

$$p_n(\alpha) = p(\alpha_n) p(\alpha_1, \alpha_2, \dots, \alpha_{n-1} | \alpha_n), \quad (1)$$

где $p(\alpha_n)$ – априорная вероятность передачи n -го фрагмента;

ненны передачи предыдущих $(n - 1)$ фрагментов.

$p(\alpha_1, \alpha_2, \dots, \alpha_{n-1} | \alpha_n)$ – условная вероятность выполнения передачи n -го фрагмента при условии, что успешно выпол-

Конкретизируем выражения для безусловных вероятностей для фрагментов от первого до N -го.

I этап: $p_1(\alpha) = p(\alpha_1);$

II этап: $p_2(\alpha) = p(\alpha_2) p(\alpha_1 | \alpha_2) = p(\alpha_2) p(\alpha_1);$

III этап: $p_3(\alpha) = p(\alpha_3) p(\alpha_1, \alpha_2 | \alpha_3) = p(\alpha_3) p(\alpha_2) p(\alpha_1).$

Без потери общности можно предположить, что

$$p(\alpha_1) = p(\alpha_2) = \dots = p(\alpha_k) = \dots = p(\alpha_N).$$

Тогда накапливаемая вероятность $p_s(\alpha)$ выполнения передачи есть пока-

зательная функция числа атомарных фрагментов:

$$p_s(\alpha, n) = \left[p(\alpha) \right]^n, \quad (2)$$

$$n = \overline{1, N}$$

При разбиении пакета на короткие фрагменты, состоящие из малого числа n атомарных фрагментов, очевидно, вероятность успешного выполнения передачи повышается.

В свою очередь, для реализации J -фазной передачи необходимо подавать на ОП пакет данных длиной K символов. При этом возникают проблемы искажения символов в пакете и вытекающей из этого необходимости повторных передач.

Если p_k – вероятность искажения k -го символа в пакете, то максимальная вероятность повторной передачи определяется вероятностью искажения хотя бы одного из K символов в пакете:

$$p_{DL} = 1 - (1 - p_k)^K. \quad (3)$$

Пусть среднее число символов в каждом из фрагментов одинаково и равно

$$p_{DN\Sigma} = (p_{DM})^M = \left[1 - (1 - p_k)^{\lfloor K/M \rfloor} \right]^M. \quad (5)$$

Очевидно, при увеличении числа M (теоретически – до величины L) вероятность (5) стремится к величине

$$p_{DN\Sigma} = \left[1 - (1 - p_1) \right]^L.$$

Теоретически пакет длиной L символов пользовательской информации можно разбить на L пакетов. В каждом пакете длиной K символов $K = K_{\text{сл}} + K_{\text{польз}} + K_0$ будет $K_{\text{сл}}$ символов служебной информации (например, в ячейке АТМ 5 байт), один символ $K_{\text{польз}} = 1 = K_1$ пользовательской информации и K_0 пустых (не используемых) символов: $K_0 = K - K_{\text{сл}} - 1$.

Конечно, ясно, что крайние случаи максимальной длины пакета $K_{\text{max}} = L$ и минимальной длины $K_{\text{min}} = K_{\text{сл}} + K_1 + K_0$ неприемлемы по частным критериям оптимальности. Если считать, что реально вероятность потери пакета (и повторной передачи из-за потери), начиная с некоторой длительности K_A , не зависит от длины па-

$L = \lceil K/J \rceil$, где $\lceil x \rceil$ – оператор округления x до ближайшего целого числа. Разобьем весь пакет длиной K символов на M коротких пакетов, в каждом из которых будет содержаться $\lceil J/M \rceil$ атомарных фрагментов и, соответственно, $\lceil K/M \rceil = \lceil LJ/M \rceil$ символов (без учета служебной информации – заголовка и концевика пакета).

Тогда вероятность искажения хотя бы одного символа и повторной передачи короткого пакета

$$p_{DN} = 1 - (1 - p_k)^{\lfloor K/N \rfloor} < p_{DL}. \quad (4)$$

Соответственно, вероятность искажения m пакетов из общего числа M пакетов равна $p_{Dm} = (p_{DM})^m$, а вероятность искажения всей информационной посылки, состоящей из M коротких пакетов:

кета, то размерность задачи векторной оптимизации уменьшается на единицу. Кроме того, при этом дальнейшее уменьшение длины фрагмента теряет смысл.

Таким образом, в данной работе, в отличие от [3], предлагается более точная модель процесса фрагментации и, соответственно, повышается точность определения диапазона оптимальных размеров фрагментов.

Результаты численного анализа

С использованием выражений (1–5) выполнены расчеты величин выигрыша в вероятности успешной передачи фрагментированных пакетов при следующих исходных данных:

– длина исходного пакета – от 576 байт – применяемое по умолчанию ограничение максимальной длины IP-пакетов – до 1500 байт – широко применяемый

стандарт длины пакетов в сетях доступа на сегментах *Ethernet*;

– исходные пакеты фрагментируются на короткие пакеты длиной от 53 байт (стандарт *ATM*) до 250 байт;

– вероятность искажений хотя бы одного символа в пакете выбрана равной 0,001. Это весьма мягкое требование. Например, в каналах связи без дополнительной защиты она составляет, как правило, $10^{-4} \dots 10^{-6}$, а в оптоволоконных линиях – до 10^{-9} [4].

Приведены результаты расчетов вероятности повторных передач вследствие искажения хотя бы одного символа в ис-

ходном пакете длиной 576 байт и в фрагментированных пакетах (рис. 2, 4, 6) и выигрыша в вероятности искажения при фрагментации (рис. 3, 5, 7). Здесь P_{DL} – вероятность искажения не фрагментированного пакета; P_{DN} – вероятность искажения фрагментированного пакета; $P_{DN\Sigma}$ – вероятность искажения нескольких фрагментированных пакетов. Отметим, что при фрагментации на более короткие пакеты выигрыш, как и следовало ожидать, растет.

1. Фрагментация на пакеты по 53 байта

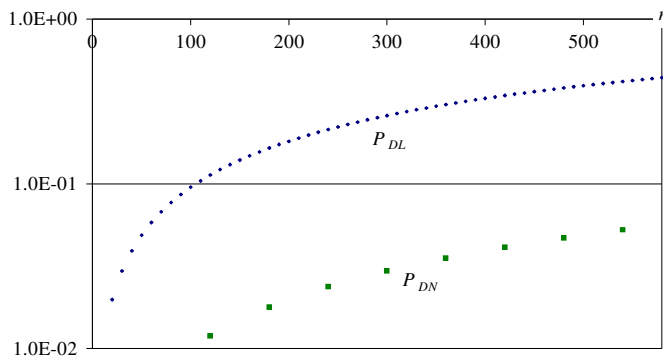


Рис. 2. Зависимости вероятности искажения пакетов от их длины n

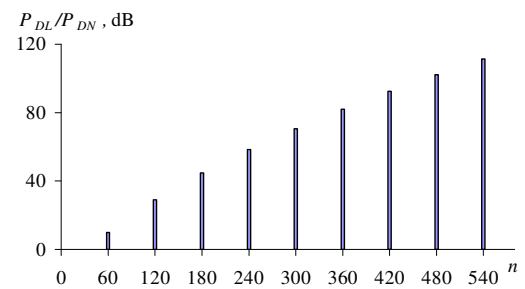


Рис. 3. Выигрыш в вероятности искажения при фрагментации

2. Фрагментация на пакеты по 110 байт.

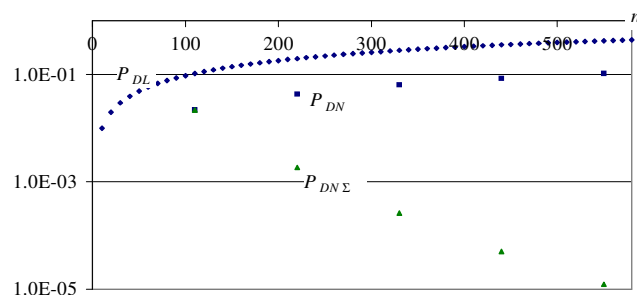


Рис. 4. Зависимости вероятности искажения пакетов от их длины n

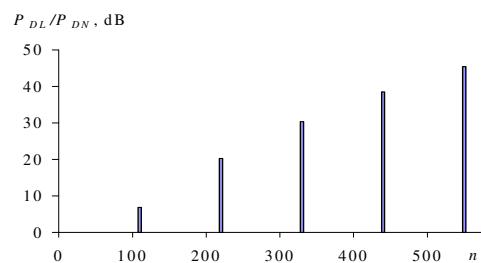


Рис. 5. Выигрыш в вероятности искажения при фрагментации

3. Фрагментация на пакеты по 270 байт.

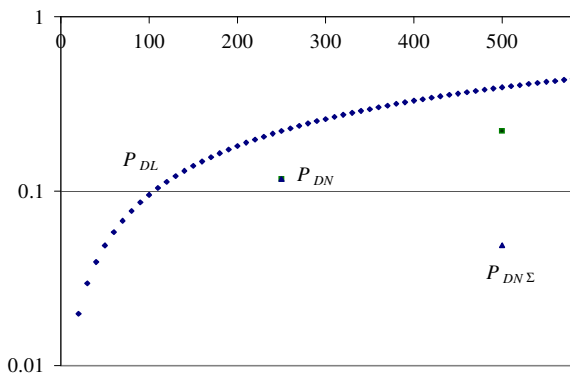


Рис. 6. Зависимости вероятности искажения пакетов от их длины n

На рис. 8 приведен графики зависимости вероятности искажения символов в пакетах в зависимости от коэффициента

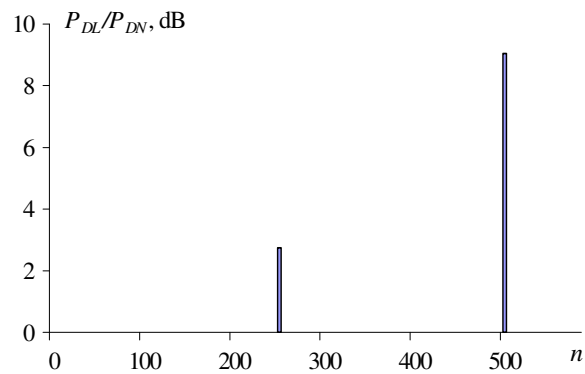


Рис. 7. Выигрыш в вероятности искажения при фрагментации

фрагментации $k_{fr} =]K/M[$ при разных априорных вероятностях искажений отдельного символа.

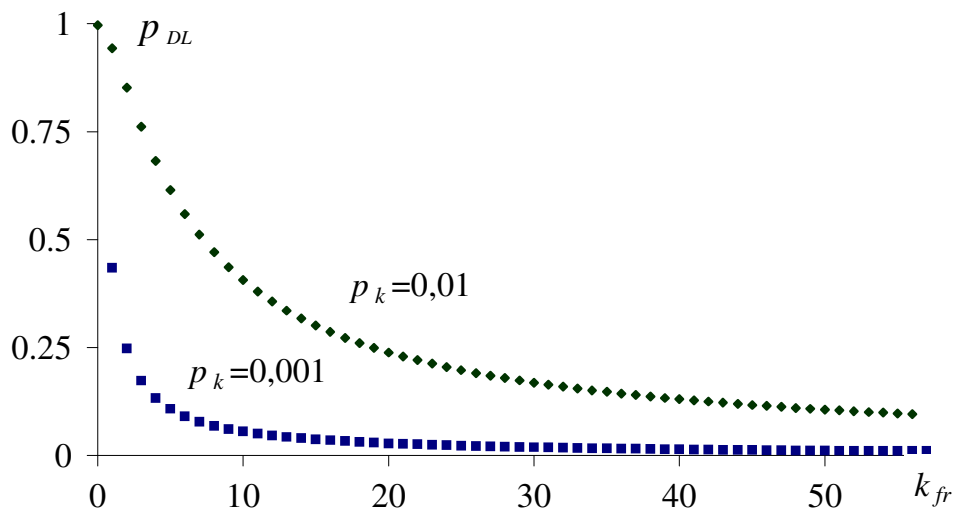


Рис. 8. Зависимость вероятности искажения пакета от коэффициента фрагментации

При выполнении расчетов по методике, принятой в [3], не учитывались такие факторы, как задержка обработки в ОП, расходы на фрагментацию и плата за использование нескольких ОП (возможно, принадлежащих разным операторам). Кроме того, была выбрана недостаточно реалистичная модель (длина фрагмента стремится к нулю, что невозможно в принципе). Отсутствует учет потерь вре-

мени $t_{ож}$ на ожидание фрагментов и $t_{об}$ – объединение фрагментов в устройстве дефрагментации.

Для уточнения модели примем максимальное значение $t_{ож} = t_{ож\ ср} + 3\sigma_{дост}$, где $\sigma_{дост}$ – среднеквадратическое отклонение (СКО) времени доставки.

На рис. 9 изображена типичная зависимость $\sigma_{дост}$ от длины фрагмента $l_{фр}$.

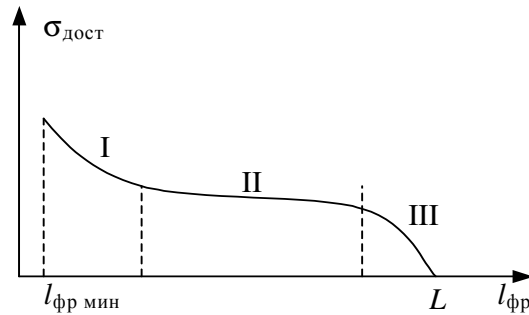


Рис. 9. Розброс времени доставки фрагментов разной длины

На первом участке (короткие фрагменты) разброс обусловлен разной политикой обработки пакетов разной длины в транзитных узлах. Для фрагментов средней длины (второй участок) разброс обусловлен в основном параметрами сетевых узлов коммутации и маршрутизации. Уменьшение разброса на третьем участке вполне логично обусловлено уменьшением числа фрагментов вплоть до передачи без фрагментации (падение до нуля).

Задержка обработки в ОП складывается из задержки коммутации, которая может составлять от долей до тысяч миллисекунд, времени буферизации и ожидания пакета в очереди и времени перемещения пакета в выходной порт [4]. Основной вклад в задержку обработки вносит время буферизации. При уменьшении длины пакета задержка буферизации убывает, и величина суммарной задержки асимптотически приближается к величине времени перемещения пакета. Для описания зависимости задержки от длины пакета можно взять рациональную функцию вида $\tau_d = \frac{a}{b +]K/M[}$ или экспоненту $\tau_d = \exp(-c]K/M[) + d$, где постоянные a, b, c, d выбираются с учетом технических и эксплуатационных характеристик оборудования [5].

Расходы на фрагментацию возникают из-за увеличения общей длительности фрагментированных пакетов, поскольку к каждому такому пакету необходимо присоединить заголовок и концевик. К ним, как отмечалось выше, добавляются рас-

ходы на использование оборудования других операторов.

Для конкретизации функции стоимости фрагментации используем общие подходы системотехники [5,6]: расходы пропорциональны квадрату увеличения ресурса – мощности механизма, машины, генератора, расхода топлива, числа применяемых приборов, устройств, машин и т.д. Поэтому в качестве функции стоимости возьмем квадратичную функцию вида $C_{\Sigma}(]K/M[) = C_0(]K/M[)^2 + C_p$, где C_0 – константа, выбираемая, с учетом стоимостных показателей оборудования, C_p – условно-постоянные начальные затраты.

С учетом приведенных соображений результирующий критерий эффективности фрагментации можно представить в виде линейной свертки частных показателей Q_i :

$$k_{eff} = \sum \beta_i Q_i, \quad (6)$$

где β_i – весовые коэффициенты, выбираемые опытным путем.

Для расчета были выбраны исследованный ранее показатель – вероятность искажений и повторных передач, а также функции задержки и стоимости. На рис. 10 представлен графики частных показателей и результирующая функция эффективности фрагментации (для наглядности выбрана функция, обратная коэффициенту эффективности). Как видно, минимум достигается при коэффициентах фрагментации от 10 до 15, что соответствует диапазону разбиения исходного пакета длиной 576 байт на короткие пакеты длиной от 40 до 60 байт.

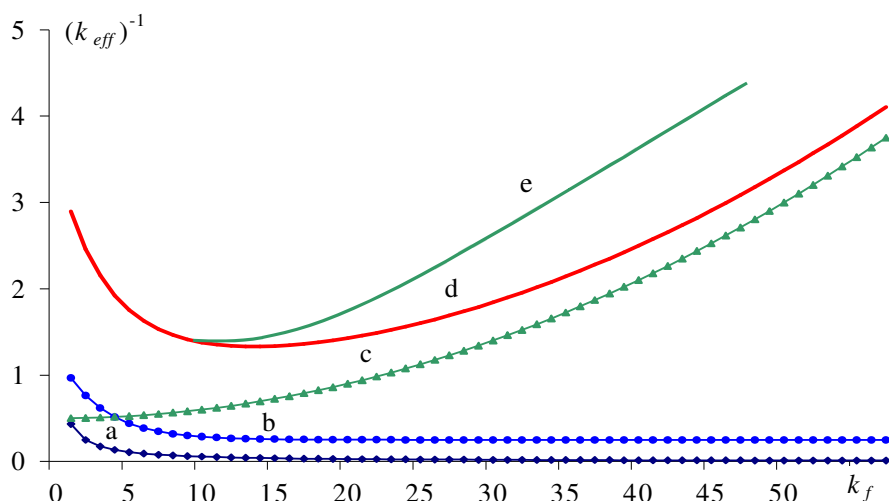


Рис. 10. Зависимость параметра суммарной эффективности $(k_{eff})^{-1}$

от коэффициента фрагментации: а – вероятность искажений символов; b – задержка; с – стоимость; d – сумма показателей по методике [3]; e – сумма показателей по уточненной методике

Выводы

1. При разбиении пакета данных на отдельные фрагменты, и распределении этих фрагментов по разным обслуживающим приборам (маршрутам), достигается существенный выигрыш в эффективности обслуживания – от 10 до 100 дБ в зависимости от величины коэффициента фрагментации. Кроме того, смягчаются требования к качеству линий передачи данных – даже при достаточно высоких вероятностях искажений отдельных символов в пакете результирующая вероятность искажений и повторных передач остается в допустимых пределах.

2. С учетом затрат на распределенное обслуживание – фрагментацию и сбор пакетов – определен диапазон эффективной фрагментации – разбиения пакета на 10...12 коротких фрагментов. При этом выигрыш в эффективности обслуживания будет достигать 35...40 дБ.

3. Исследования проблемы распределенной передачи целесообразно продолжить в направлении учета сравнительной важности выбранных частных показателей эффективности и вероятностных характеристик сохранения целостности информации в процессе передачи фрагментированных данных.

Список литературы

1. Олифер В.Г., Олифер Н.А. Компьютерные сети. Принципы, технологии, протоколы: Учебник для вузов. 3-е изд. – СПб.: Питер, 2006. – 958 с.
2. Armitage G. Quality of Services in IP Networks. – Indianapolis. – IN.: Macmillan Technical Publishing. – 2000. – 309 p.
3. Баховский П.Ф., Виноградов Н.А. Асимптотические оценки эффективности распределенной реализации виртуальных технических функций в сетях мобильной связи новых поколений // Наукові записки УНДІЗ, №4(12), 2009. – С. 40 – 46.
4. Вишневский В.М., Ляхов А.И., Портной С.Л., Шахнович И.В. Широкополосные беспроводные сети передачи информации. – М.: Техносфера. – 2005. – 592 с.
5. Дружинин В. В., Конторов Д. С. Конфликтная радиолокация (опыт системного исследования). – М.: Радио и связь. – 1982. – 124 с.
6. Моисеев Н.Н. Математические задачи системного анализа. – М.: Наука, 1981. – 488 с.

Подано до редакції 01.10.2010