

УДК 004.2

Скулиш М. А., канд. техн. наук (Тел.: +380 05 607 42 29. E-mail: mb_s@ukr.net)
(Національний технічний університет України «КПІ», Інститут телекомунікаційних систем, м. Київ)

ОРГАНІЗАЦІЯ РОБОТИ ГРУПИ СЕРВЕРІВ ДЛЯ ЗАБЕЗПЕЧЕННЯ ПОТРЕБ РОЗПОДІЛЕНОЇ СИСТЕМИ ТАРИФІКАЦІЇ ПОСЛУГ

Скулиш М. А. Організація роботи групи серверів для забезпечення потреб розподіленої системи тарифікації послуг. Запропоновано новий підхід до процесу керування роботою систем онлайн тарифікації послуг. Система розглядається як мережа, кількість вузлів мережі змінюється залежно від навантаження, що потребує обслуговування (заявки на тарифікацію). Керування потоками в мережі здійснюється за принципами балансування навантаження на вузли мережі. Розроблено метод регулювання кількості обслуговуючих вузлів, який застосовується для поточної перевірки достатності ресурсів і дозволяє проводити оцінку динаміки вхідного навантаження та поточного стану технічних засобів.

Ключові слова: телекомунікаційна система, розподілена система, тарифікація послуг, динаміка навантаження, якість послуг

Скулиш М. А. Организация работы группы серверов для обеспечения потребностей распределенной системы тарификации услуг. Предложен новый подход к процессу управления работой систем онлайн тарификации услуг. Система рассматривается как сеть, количество узлов сети изменяется в зависимости от нагрузки, которая требует обслуживания (заявки на тарификацию). Управление потоками в сети осуществляется по принципам балансировки нагрузки на узлы сети. Разработан метод регулирования количества обслуживающих узлов, который применяется для текущей проверки достаточности ресурсов и позволяет проводить оценку динамики входной нагрузки и текущего состояния технических средств.

Ключевые слова: телекоммуникационная система, распределенная система, распределенная система, тарификация услуг, динамика нагрузки, качество услуг

Skulysh M. A. The organization of work of group of servers for ensuring requirements of the distributed system of tariffing of services. New approach to process of management of work of systems online of tariffing of services is offered. The system is considered as a network, the number of nodes of a network changes depending on loading which demands service. Control of streams in a network is exercised by the principles of balancing of load of network nodes. The method of regulation of quantity of the serving nodes which is applied to the current check of sufficiency of resources is developed and allows to carry out an assessment of dynamics of entrance loading and current state of technical means.

Key words: telecommunication system, the distributed system, the distributed system, tariffing of services, the loudspeaker of loading, quality of services

Вступ. Забезпечення роботи телекомунікаційних систем невід'ємно пов'язано з роботою систем тарифікації послуг. Проблема тарифікації гостро постає як перед операторами зв'язку, так і перед провайдерами різноманітних послуг. Головною складністю таких системи є значна кількість абонентів, які мають одночасно отримати розрахунок. Крім того процес ускладнюється через різні процедури тарифікації, які передбачені для різних типів послуг.

Існуюча тенденція реалізації надання послуг на базі IP-протоколу приводить до того, що мобільні користувачі 3G/4G мають безпосередній бездротовий високошвидкісний доступ до Інтернету, де мобільні термінали - це IP хости, підключені до Інтернету [1, 2]. Спрощення інтерфейсів взаємодії мобільних терміналів із Інтернет відкриває можливість абонентам користуватися значним спектром різноманітних послуг, кількість яких все зростає. Послуги потребують все більше обчислень для визначення їх вартості. В існуючих мобільних мережах це призводить до підвищення навантаження на системи тарифікації, як наслідок:

- зниження ефективності обслуговування викликів;
- зниження гнучкості та ефективності тарифікаційних механізмів;
- неможливості гарантування кризової якості надання послуг.

Для вирішення цих проблеми були створені проекти 3G Partnership Project (3GPP).

Значна кількість робіт та підходів присвячена контролю якості обслуговування абонентів. Розробляються методи забезпечення показників якості в процесі передачі інформації на різних етапах обслуговування (доступ до мережі, організація транспорту інформаційного потоку) [3, 4], частина робіт провячено забезпеченню ефективної тарифікації, при цьому розглядаються питання розробки правил, в яких задається політика обслуговування індивідуально для кожного типу сервісу залежно від тарифного плану [1, 5]. Роботи [6, 7] присвячені оптимізації білінгових систем пов'язані з дисципліною обслуговування направлених на диференціацію вхідного потоку та створення нових правил послідовності тарифікації абонентів.

Однак, для операторів зв'язку достатньо гостро постає проблема перевищення часу обслуговування заявки саме на сервері тарифікації, оскільки не враховується принцип розподілу ресурсів технічних засобів, що у періоди пікових навантажень є критичними для якості обслуговування. Вирішення цієї проблеми може полягати у залученні до обслуговування додаткових орендованих технічних ресурсів. Які у моменти пікових навантажень будуть динамічно долучатися до обробки заявок на тарифікацію. Це дозволить уникнути втрат, які пов'язані з перевищенням часу очікування заявки на тарифікацію через дефіцит технічних ресурсів системи. Перевищення допустимої тривалості обслуговування призводить до відхилення виклику, відповідно до економічних втрат та зниження репутації компанії.

Внаслідок цього є актуальною науково-технічна задача планування процесу динамічного залучення ресурсів для забезпечення безперебійного процесу тарифікації підсистемами системи OCS (Online charging system), яка б, *по перше*, враховувала системи тарифікації та навантаження, яке створюється різними типами послуг, і, *по-друге*, включала відповідні механізми розрахунку необхідної кількості ресурсів для статистичних даних за тривалий період часу, а також поточну динаміку вхідного потоку, що дозволило б подолати описані вище недоліки.

Робота системи тарифікації. Обслуговування абонентів на сервері тарифікації є необхідною складовою частиною процесу обслуговування абонентів та надання їм телекомунікаційних послуг. Оскільки кожна послуга є платною, оператор зв'язку залежно від типу тарифного плану здійснює тарифікацію заявки або перед наданням послуги в режимі реального часу, або в режимі off-line, необхідно перевірити наявність коштів на рахунку абонента, зробити перерахунок, тобто здійснити ряд стандартних операцій.

Час виконання операцій є обмеженим, у разі його перевищення заявка вважається втраченою і абонент отримує повідомлення про неможливість отримання послуги. У такій ситуації оператор несе збитки, а у разі систематичних відмов у наданні послуг псується репутація компанії. Тому важливо підібрати ресурси серверу таким чином, щоб зменшити ймовірність перевищення часу обслуговування абонентських заявок.

Складність процесу розрахунку кількості ресурсів необхідної для тарифікації різних типів послуг пояснюється чотирма основними передумовами:

1. Процес тарифікації передбачає послідовне виконання операцій, які потребують різної кількості ресурсів.
2. Кожний тип послуг, незважаючи на стандартність операцій, які виконуються в процесі тарифікації, потребує для здійснення розрахунків різної кількості ресурсів.
3. Одночасно надходить велика кількість заявок на тарифікацію різних типів ресурсів.
4. Інтенсивність надходження заявок залежить від дня тижня та часу доби.

Схему обслуговування заявок на сервері тарифікації показано на Рис. 1.

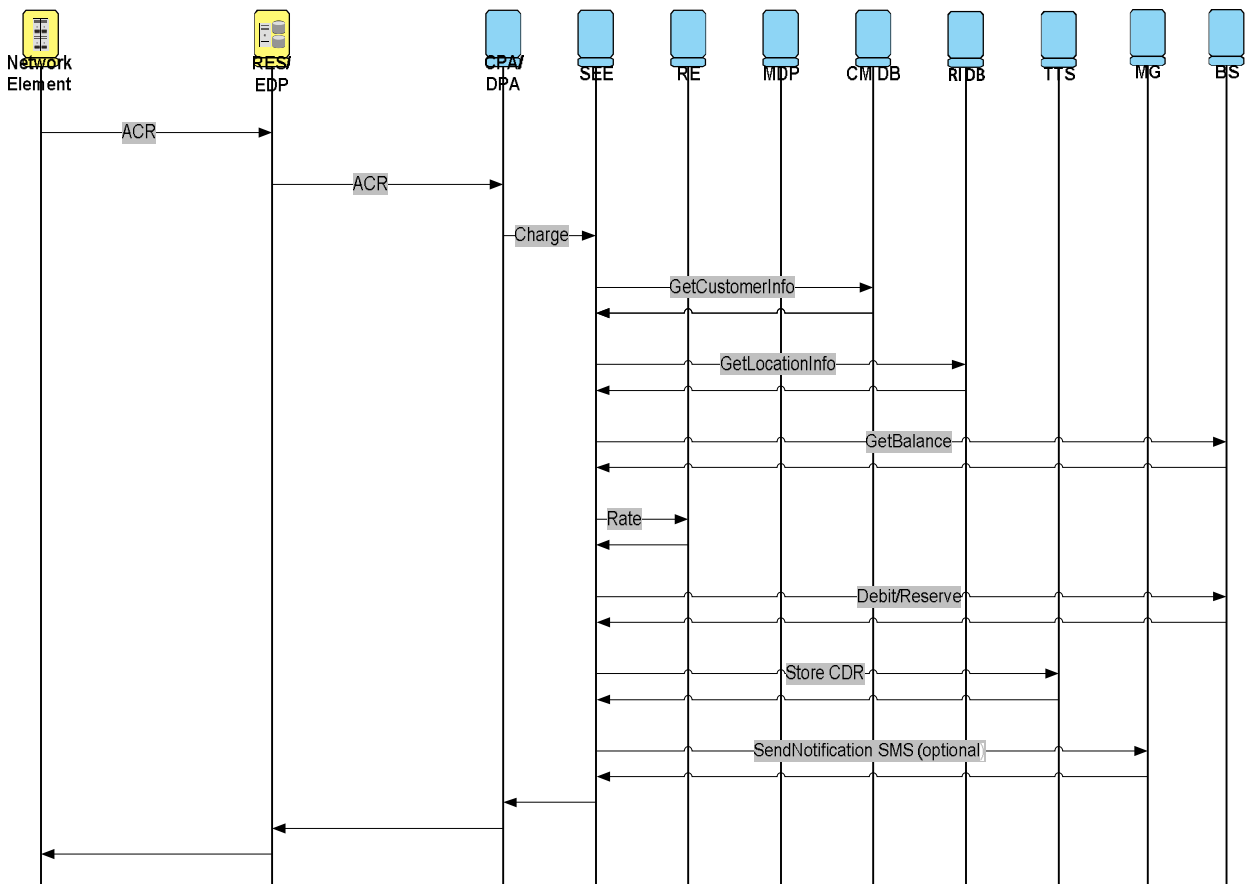


Рис. 1. Загальна схема online тарифікації послуг на сервері тарифікації

Модулі EDP (Enhanced Diameter Proxy) та CPA (Cap protocol adapter) декодують потоки, що надходять за протоколами cap та diameter.

Розкодовану справу модуль передає на сервер бізнес логіки SEE (Service Execution Environment) через модуль маршрутизації. Сервер бізнес логіки SEE являється ядром системи тарифікації та забезпечує середовище для виконання послідовності операцій, які передбачає процес обслуговування заявок.

Послідовність операцій які повинні бути виконані для успішного обслуговування заявки включає в себе наступні дії:

1. Вилучення інформації про абонента. Для цього модуль SEE звертається до бази даних керування абонентами (CM DB Customer management data base).
2. Вилучення інформації про місце розташування абонента. Для цього модуль SEE звертається до бази даних, що зберігає структуру мережі (RE DB Resource Inventory data base).
3. Вилучення інформації про стан рахунку абонента. Для цього модуль SEE звертається до бази даних рахунків абонентів (Balance Storage).
4. Здійснюється розрахунок вартості послуги на основі тарифної книги та поточної тарифної моделі абонента. Для цього модуль SEE звертається до модулю розрахунків (RE – Rating Engine).
5. З рахунку абонента знімається плата за послугу, залежно від типу сервісу паралельно може здійснюватися резервування заданої суми, тільки для сервісів з резервацією (SCUR - Session Charging with Unit Reservation та ECUR - Event Charging with Unit Reservation), коли тривалість надання послуги є не відомою та не можна підвести остаточний

розрахунок. Для здійснення операцій дебету та резервації коштів модуль SEE звертається до бази даних рахунків абонентів (Balance Storage).

6. Формується звіт про надання послуги CDR (call data record). Для цього модуль SEE звертається до програмного модулю формування CDR (TTS – Toll Ticket Server).

7. Якщо це передбачено послугою, відправляється повідомлення абоненту про результати наданих послуг. Для цього модуль SEE звертається до модулю для посилки повідомлень абоненту (MG- message GateWay).

Як видно процес тарифікації є багатоетапним, при цьому операції, які послідовно виконуються у ядрі бізнес-логіки SEE із залученням різних підсистем – різноманітні, відповідно потребують різної кількості оперативної пам'яті, процесорного часу та дискового простору. Етапи обслуговування на сервері тарифікації будемо називати функціональними блоками. На якість обслуговування заявок на тарифікацію значний вплив має швидкість обробки заявок, яка залежить від технічних засобів, що здійснюють обслуговування. Тому необхідно знати, які потужності технічних ресурсів потрібні для обслуговування вхідного потоку заявок на тарифікацію в той чи інший період часу.

Необхідно врахувати як загальну кількість ресурсів, що обслуговує сервер в цілому, так і потреби у потужностях кожної з підсистем. Оскільки з метою забезпечення успішного функціонування системи може виникнути потреба перегрупування технічних ресурсів для забезпечення виконання завдань логічного рівня на кожному етапі обслуговування.

Проблеми організації роботи білінгової системи на множенні технічних засобів.

Однією з перешкод на шляху до масового використання в системах онлайн тарифікації зовнішніх технічних ресурсів є відсутність відповідних методів та алгоритмів організації злагодженої роботи групи серверів, які забезпечують роботу систем онлайн тарифікації як єдиного цілого.

В даній статті розглядається два алгоритми, які забезпечать безперебійну та економічну роботу системи тарифікації. Економічність полягає в тому, що відповідно до статистичних даних про вхідне навантаження, відбувається включення потрібної кількості серверів, які забезпечують роботу системи тарифікації або його підсистем.

Розробка розкладів включення обладнання. Метод визначення моменту включення додаткового технічного засобу. В умовах необхідності заощаджувати витрати на утримання серверів, будь то орендна плата за використання ресурсів, які розташовані у хмарах, або енергетичні ресурси для забезпечення роботи власних серверів. Компанії оператору необхідно розробити план включення серверів, який би задовольняв всі потреби користувачів у поточний період часу.

Метод визначення моменту включення додаткового технічного засобу (серверу) полягає в тому, що із заданим інтервалом часу оцінюється поточна статистика, будується тренд, або лінійна апроксимація, зібраних даних про кількість заявок, що обслуговується, після чого робиться прогноз ймовірності того, що кількість заявок перевищить допустиму межу. Межею в даному випадку виступає допустима кількість заявок, яка може бути обслужена вже залученим увімкненим ресурсом.

За час для якого робиться оцінка добре вибрати час завантаження ще одного ресурсу, це дозволить, завчасно залучити необхідну кількість ресурсів.

Метод застосовується в процесі моніторингу системи, з метою виявлення моментів коли необхідно залучити додаткові ресурси.

Дана задача може вирішуватися як для системи в цілому, тоді оцінюється статистика кількості заявок на тарифікацію які надходять до системи онлайн тарифікації, так і для підсистем, коли оцінюється статистика заявок-звернень до підсистем, системи тарифікації. Вибір точки застосування методу залежить від вибраної архітектури залучення серверів.

Отже, на основі поточної статистики навантаження, що створюється абонентськими заявками на тарифікацію; оцінки верхньої границі кількості заявок, які одночасно можуть оброблятися на потужностях доступних серверів буде визначена ймовірність того, що протягом заданого часу кількість заявок, що надійде не перевищить допустимого значення.

Алгоритм методу контролю достатності ресурсів системи для обробки заявок на тарифікацію.

Вхідні дані:

- 1) інтервал $T1$ – інтервал часу, протягом якого буде проводитися аналіз статистики;
- 2) інтервал dt – інтервал дискретизації часу (малий інтервал);
- 3) дані системи моніторингу про кількість заявок на тарифікацію, які надійшли за малі інтервали часу, загалом $(T1/dt)$ результатів вимірів (див. Табл. 1);

Табл. 1.

t	$t0-T1$	$t0-T1+ T1/dt$	$t0-T1+ 2T1/dt$	$t0-T1+ 3T1/dt$...
X	$x(t)$	$x(t)$	$x(t)$	$x(t)$	

- 4) час протягом якого може бути завантажено сервера – T ;
- 5) таблиця значень Функції Лапласа (розподілу функції $N(0,1)$);
- 6) допустима ймовірність помилки, наприклад 10^{-4} ;
- 7) M – максимально допустима кількість заявок, які можуть обслуговуватися при поточній кількості ввімкнених серверів (ресурсів).

Включення серверу відбувається, якщо ймовірність недостачі ресурсів вище допустимого значення (див. Рис. 2).

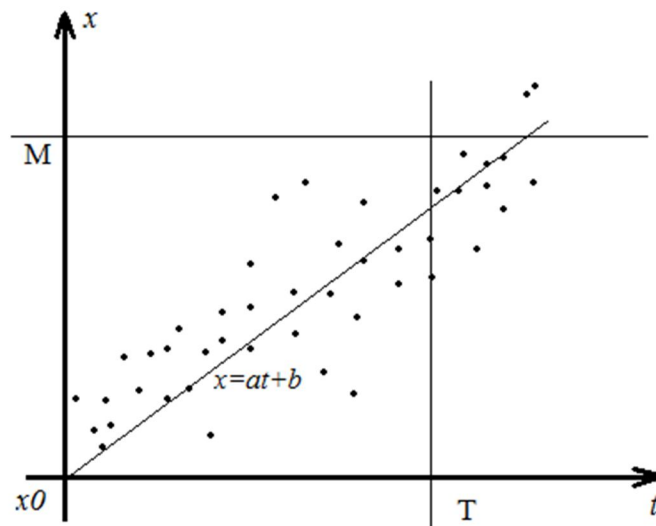


Рис. 2. Аналіз динаміки навантаження на сервер онлайн тарифікації.

Алгоритм методу включає в себе три основні кроки:

1. Аналіз статистичних даних за час $(t0-T1)$, де $t0$ – поточний час, я якого виконується розрахунок. На основі статистичних даних системи моніторингу для пари значень (t,x) за

методом найменших квадратів розраховується оцінка коефіцієнту a для прямої, яка апроксимує значення вхідного навантаження з Табл.1 $x=at+b$.

2. Оцінка ймовірності помилки протягом заданого часу розраховується за формулою:

$$P(\text{помилка за час } T) \leq 2P(N(0,1) \geq y),$$

де $y = ((M-x_0) - \hat{a} \cdot T) / \sqrt{T}$; x_0 – поточне навантаження на систему, M – допустиме навантаження на систему.

Ймовірність $P(N(0,1) \geq y)$, того що випадковий процес, який має нормальний розподіл, перевищить значення y знаходиться з таблиці значень Функції Лапласа.

3. Якщо ймовірність помилки протягом допустимого часу T ($P(\text{помилка за час } T)$) перевищує допустимий поріг, тоді відбувається включення додаткового серверу.

Застосування даного алгоритму дозволить контролювати динаміку зростання навантаження на сервер онлайн тарифікації та визначити момент коли необхідно залучити для обслуговування додаткові ресурси. Час T обирається виходячи з міркувань скільки потрібно часу від запуску до повної працездатності додаткового серверу.

Як видно з Рис. 2 лінія $x=at+b$ проходить нижче від точки з координатами (T, M) , оскільки даний метод враховує структуру випадкового процесу та середньоквадратичне відхилення яким характеризується випадковий процес надходження заявок на тарифікацію.

Дана функція перевірки є частиною системи моніторингу, запускається із заданою періодичністю для забезпечення надійності та безперебійної роботи системи, забезпечує контроль за достатністю ресурсів системи.

Оцінка ефективності. Спектр проблем, які можуть вирішуватися за допомогою запропонованого методу та алгоритмів досить широкий. Основною ознакою систем, для яких може бути застосований запропонований метод, є виконання великої кількості процедур, ініціатором яких є люди або інші програми. Виконання процедур здійснюється за допомогою програмного забезпечення серверу, при цьому використовуються технічні ресурси системи.

На сьогоднішній день проблема перевантаження вирішується за рахунок не допуску в систему надлишкової кількості запитів на тарифікацію, тобто якщо сервер є перевантаженим, то поступає сигнал на керуючий пристрій і тимчасово заявки на тарифікацію не приймаються. При цьому одночасно задіяні (знаходяться у режимі очікування) всі ресурси серверу, завантаження обладнання на 20-30% є нормальним в процесі обслуговування.

Був проведений експеримент роботи імітаційної моделі в двох режимах:

1. На обслуговування виділявся обмежений технічний ресурс, у разі перевантаження заявки відкидалися (Режим без застосування запропонованих методів).

2. Обслуговування заявок може проводитися від одного до трьох аналогічних серверів з обмеженими технічними ресурсами, на основі статистичних вибірки від оператора зв'язку був сформований розклад включення серверів відповідно до алгоритму пошуку моментів переключення, далі було згенеровано вхідний потік максимально наближений до реального. Включення серверів відбувалося за розкладом, а за методом включення додаткового технічного обладнання кожену три хвилини відбувалась перевірка достатності ресурсів для обслуговування вхідного потоку.

За результатами роботи імітаційної моделі було проведено дослідження використання електричної енергії, яке показало скорочення споживання енергії на 60% (Рис. 3,а). В той же час кількість втрачених заявок скоротилася в 5 разів від п'яти відсотків до одного.

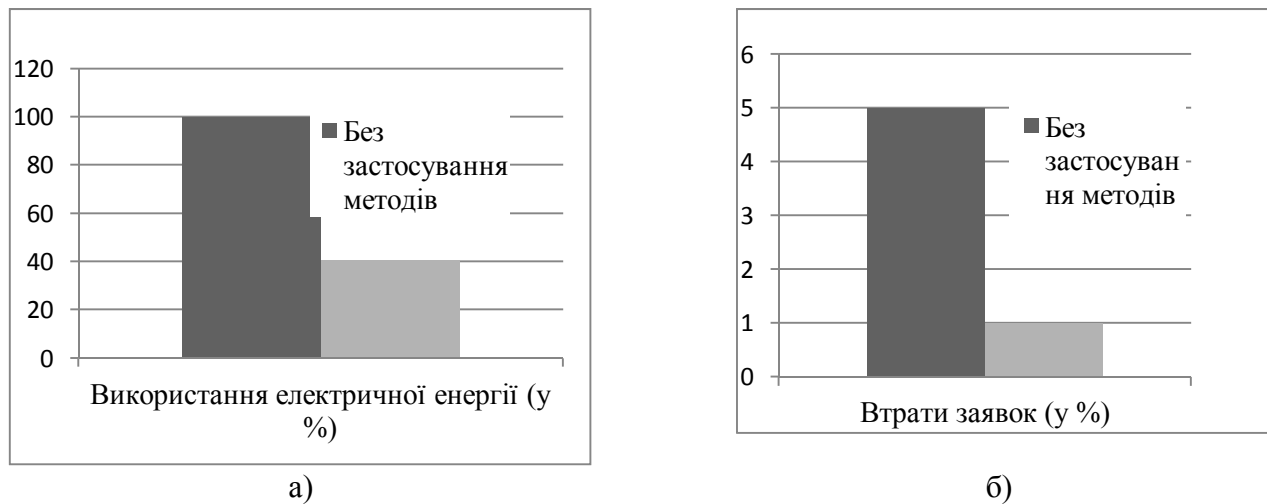


Рис. 3. а) Порівняння використання електричної енергії
б) Порівняння кількості втрачених заявок на тарифікацію

Висновки. В статті розглянуті проблеми організації роботи системи онлайн тарифікації. Запропоновано підхід до системи керування роботою системи онлайн тарифікації, який полягає в тому, що система розглядається як мережа, керування потоками в мережі здійснюється із застосуванням системами балансування, де кількість доступних вузлів варіюється залежно від навантаження, яке потребує обслуговування (заявки на тарифікацію).

Запропоновано метод визначення моменту включення додаткового технічного засобу, який дозволяє проводити оцінку динаміки вхідного навантаження та поточного стану технічних засобів, та завчасно увімкнути додатковий ресурс, та запобігли перевантаженню існуючих ресурсів.

Проведене моделювання показало, що скорочується споживання електричної енергії на 60%, а втрати заявок через зайнятість серверу зменшуються у 5 разів від 5% до 1%.

Література

1. Chaitanya T. V. K. Improving 3GPP-LTE uplink control signaling performance using complex-field coding / T. V. K. Chaitanya, E. G. Larsson // IEEE Transactions on Vehicular Technology. – 2013. – Vol. 62, No. 1. – P. 161-171.
2. Cuevas A. IMS Service Platform: A Solution for Next-Generation Network Operators to Be More than Bit Pipes / Antonio Cuevas, Jose Ignacio Moreno, Hans Einsiedler. // IEEE Communication Magazine. – 2006, Aug. – P. 75-81.
3. Ghosh A. LTE-advanced: next-generation wireless broadband technology / A. Ghosh, R. Ratasuk, B. Mondal, N. Mangalvedhe, T. Thomas // IEEE Wireless Communications. – 2010. – Vol. 17, No. 3. – P. 10-22.
4. Globa L. Nodal routing with traffic classification / Larisa Globa, Mariia Slukysh // Polish association for knowlage management Series: Studies&Proceedings. – 2011. – №42. – P. 37-46.
5. Costa-Requena J. SDN integration in LTE mobile backhaul networks / J. Costa-Requena // Information Networking (ICOIN), 2014 International Conference. – 10-12 Feb. 2014. – P. 264-269.
6. Sok-Ian Sou. Signaling overhead of Policy and online Charging Control for bearer sessions in LTE network / Sok-Ian Sou, Jeu-Yih Jeng, Yinman Lee // Consumer Electronics, 2009. ISCE '09. IEEE 13th International Symposium. – 2009. – P. 593-597.
7. Malandrino F. LTE offloading: When 3GPP policies are just enough / Francesco Malandrino, Claudio Casetti, Carla-Fabiana Chiasserini // Wireless On-demand Network Systems and Services (WONS), 2014 11th Annual Conference. – P. 1-8.