

ВДОСКОНАЛЕНА МОДЕЛЬ ДЛЯ ВИЗНАЧЕННЯ ТОПОЛОГІЧНОЇ ЛОКАЛЬНОСТІ В МЕРЕЖІ ІНТЕРНЕТ

Порєв Г.В.

Анотація: Методи визначення топологічної відстані є ключовими для підвищення ефективності однорангових та розподілених мереж. На відміну від багатьох існуючих або запропонованих методів, які зазвичай використовують зовнішню інфраструктуру, ми розробили і вдосконалили суто обчислювальну клієнтську метрику CARMA, яка базується на попередньо обчислених топологічних структурах Інтернет. Як показала експериментальна перевірка, ця метрика досить добре відповідає дійсній вимірній топологічній відстані.

Abstract: The topological distance estimation methods are keys to efficiency of peer-to-peer and distributed networks. Contrary to many existing or proposed methods which usually employ external infrastructure, we have developed and improved purely computational client-side CARMA metric based on the precomputed topological structure of the Internet. As the experimental validation results show, this metric fairly well corresponds to the actual measured distance.

Анотация: Методы определения топологического расстояния являются ключевыми для повышения эффективности одноранговых и распределённых сетей. В отличие от многих существующих и предложенных методов, которые обычно используют внешнюю инфраструктуру, мы разработали и усовершенствовали чисто вычислительную клиентскую метрику CARMA, которая базируется на предвычисленных топологических структурах Интернет. Как показала экспериментальная проверка, эта метрика достаточно хорошо соответствует реальному измеренному топологическому расстоянию.

Ключові слова: Інтернет, топологія, метрика відстані.

Вступ

В порівнянні з початком 20-го сторіччя, коли однорангові мережі тільки з'явилися, їх розвиток і різноманітність вирішуваних задач значно розширено. На сьогоднішній день існує ціла множина похідних від однорангових мереж технологій, зокрема файловий обмін, миттєві текстові повідомлення, VoIP, мультимедіа на вимогу, доставка програмного забезпечення тощо.

Різноманітні дослідження структури трафіку Інтернет вказують, що від 30% до 50% згенерованого користувачами трафіку має відношення до однорангових технологій, включаючи також USENET та нефайлообмінні служби. Хоча ідеологія однорангових мереж полягає, зокрема, в послабленні навантаження на централізовані серверні ферми, притаманна асиметрія каналів зв'язку у кінцевих користувачів змусила провайдерів Інтернет збільшити витрати на обслуговування та оновлення обладнання «останньої милі» для збереження якості надання послуги на належному рівні. Деякі оператори зв'язку також ввели непопулярні рішення, які виявляють та примусово знижують якість або швидкість каналу для трафіку, який визнано належним до однорангових мереж, що, в першу чергу, впливає на файловий обмін.

З цих причин дослідники шукають шляхи для оптимізації функціонування однорангових мереж в контексті регуляції трафіку, вважаючи потенціалом для оптимізації притаманну Інтернет кластерну структуру. Загальна ідея полягає в збільшенні швидкості обміну даними в середині певного сегменту мережі при мінімізації трафіку між такими сегментами. В даному випадку, масштаб поняття «сегмент» не визначено, оскільки звичайно можна виділити більш ніж один рівень сегментації Інтернет.

Деякі дослідники запропонували низку рішень, які передбачають побудову зовнішньої (по відношенню до шару однорангової мережі) інфраструктури, що спеціалізується на відстеженні умов внутрішньо-мережевих та між-мережевих каналів зв'язку, запам'ятовує явно знайдені шляхи маршрутизації та на основі цих даних видає прогнози маршрутизації та іншу інформацію. Такими рішеннями є, наприклад, P4P [1] та iPlane[2]. Інші пропозиції, такі як [3] передбачають пряме втручання в трафік однорангової мережі для приведення його у відповідність до політики оператора зв'язку.

В даній роботі ми пропонуємо інший підхід, який не потребує зовнішньої негарантованої інфраструктури для обчислення топологічних даних. Такий підхід був окреслений в попередній роботі, зокрема [4], де топологічна відстань була розглянута в контексті метричної дистанції для оптимізації побудови мінімальних покриваючих дерев.

Таким чином, метою даної роботи є вдосконалення як структурної моделі мережі Інтернет так і алгоритму знаходження метричного показника локальності для підвищення ефективності організації та функціонування однорангових мереж.

Огляд попередніх доробок

В роботі [4] метрична модель CARMA (Combined Affinity Reconnaissance Metric Architecture) надавала можливість оцінювати локальність між парою довільних адрес IPv4. Основною відмінністю цієї моделі від аналогів було те, що оцінка виконувалась суто внутрішніми обчисленнями, без прямих вимірювань або звертань до зовнішніх служб. З огляду на це, CARMA можна вважати наближеною математичною моделлю топології мережі Інтернет в цьому конкретному контексті відносної локальності мережних вузлів.

CARMA розпочинає роботу з попереднього завантаження структурної інформації з публічно доступних служб — регіональних Інтернет-реєстрів (PIP) та перетворення її на внутрішню графо подібну

структуру. На відміну від рішень, що базуються на інфраструктурі PlanetLab або таких, що використовують функції RouteViews, служби RIR є обов'язковими, публічними та суттєвими для функціонування Інтернет, і, таким чином, значно надійнішими. Файли, важливі для функціонування CARMA включають множину *delegated-latest-** з переліком зареєстрованих адрес та автономних систем та інші файли з бази даних WHOIS-сервера, які містять інформацію про під-діапазони та множини автономних систем (MAC).

Оскільки швидкість приросту бази даних RIR порівняно низька, системі CARMA буде достатньо оновлювати актуальні дані лише один раз на кілька днів. Як тільки дані завантажені, CARMA будує модель наближення топології Інтернет з деякими спрощеннями, яка містить наступні 4 структурні рівні: а) діапазони IPv4 розділені на б) під-діапазони і одночасно оголошені як належні до в) автономних систем (АС), які, в свою чергу, об'єднані в г) MAC. В рамках побудованої моделі вважається, що сутності нижчого рівня є явно з'єднаними між собою через сутність вищого рівня, в якій вони оголошені.

На основі такої структури та отримавши на вході пару довільних IPv4 адрес CARMA може обчислити їх взаємну локальність, яка може бути виражена в чотирьох класах:

- *Під-діапазон (subrange)* — визначає присутність другого вузла в тому самому під-діапазоні IPv4, що і перший, у відповідності до його визначення в базі даних під-діапазонів WHOIS сервера. Здебільшого це означає присутність обох вузлів в сегменті мережі, за яку відповідає один і той самий маршрутизатор. Наприклад, це можуть бути кінцеві користувачі, підключені до точки присутності оператора зв'язку або вузли всередині мережі університету, які зазвичай мають підключення лише через одного оператора зв'язку.

- *Діапазон (range)* — визначає присутність другого вузла в тому самому діапазоні IPv4, що і перший, у відповідності до його визначення в базі даних делегацій (*delegated-latest-**). Здебільшого це означає, що обидва вузли знаходяться у внутрішній мережі відділу або невеликої організації.

- *АС (as)* — визначає присутність другого вузла в адресному просторі, який анонсовано тією ж самою АС, що і для першого вузла. Хоча це не гарантує такої ж однозначної та безпосередньої зв'язності, як два попередніх класи, пакети між вузлами навряд чи вийдуть за межі корпоративної мережі або АС, оскільки АС є основною одиницею маршрутизації в Інтернет [5], і тому вони будуть оброблені та передані всередині мережі оператора зв'язку (пакети, адресовані на ті самі АС, керуються тими самими правилами маршрутизації впродовж всього шляху).

- *MAC (asset)* — визначає що вхідні вузли належать до різних автономних систем, об'єднаних в одну множину. Така множина може виявитися точкою обміну трафіком (IX), якщо кількість учасників точки обміну досить велика. Перевага такого визначення не одразу очевидна, але потрібно згадати, що в країнах де інфраструктура Інтернет лише розвивається, різниця в якості каналу та видимій швидкості передачі може прямо залежати від такого класу локальності та сягати кількох порядків.

- *Горизонт (distant)* — показує, що відносна локальність двох мережних вузлів не може бути однозначно обчислена і тому вони вважаються такими, що розташовані топологічно далеко.

Перевага, яку може надавати CARMA за допомогою цих класів, полягає, наприклад, в перепорядкуванні списку вузлів, за допомогою яких даний вузол вносить себе в множину обміну або з яких починає запитувати потрібну інформацію. В принципі, вибираючи вузли з найближчим класом локальності, програмне забезпечення вузла може зменшити час відклику та збільшити швидкість обміну.

Аналіз швидкодії CARMA

Для оцінки швидкодії моделі CARMA було виконано ряд тестових запусків з вузлів, що мають IP адреси користувацького простору ВАТ «УкрТелеком», при цьому другі вузли пари бралися з 6 випадкових зграйок BitTorrent з одного з найбільш популярних публічних трекерів — RuTracker (www.rutracker.org). Розподіл результатів наведено на рис. 1.

Очевидно висока кількість вузлів що попали під клас MAC пояснюється тим, що історично RuTracker має досить значну користувацьку базу з України. Модель CARMA змогла виявити до 19% таких вузлів, оскільки більшість їх з топологічної точки зору «належить» двом основним точкам обміну даних — UA-IX та DataGroup-IX [6]. Це дає на виході моделі CARMA клас MAC для всіх таких адрес, якщо перший вузол також «належить» до такої точки обміну.

Завдяки усталеній технологічній та бізнесовій практиці провайдерів-учасників українських точок обміну, полоса пропускання та ціна трафіку зовні та всередині UA-IX можуть значно відрізнятися, до кількох порядків. З огляду на це, ми вважаємо, що модель CARMA має значний потенціал оптимізації. Якщо правила відбору віддалених вузлів будуть формуватися у відповідності до класів локальності моделі CARMA, наприклад, в мережах типу BitTorrent, для вузлів що працюють в адресному просторі UA-IX або аналогічній національній конфігурації, значно менше вузлів повинні будуть з'єднуватися зовні своїх точок обміну і значно більше вузлів зможуть вибрати партнерів для обміну даними з тих, у кого більше шансів мати швидший канал зв'язку.

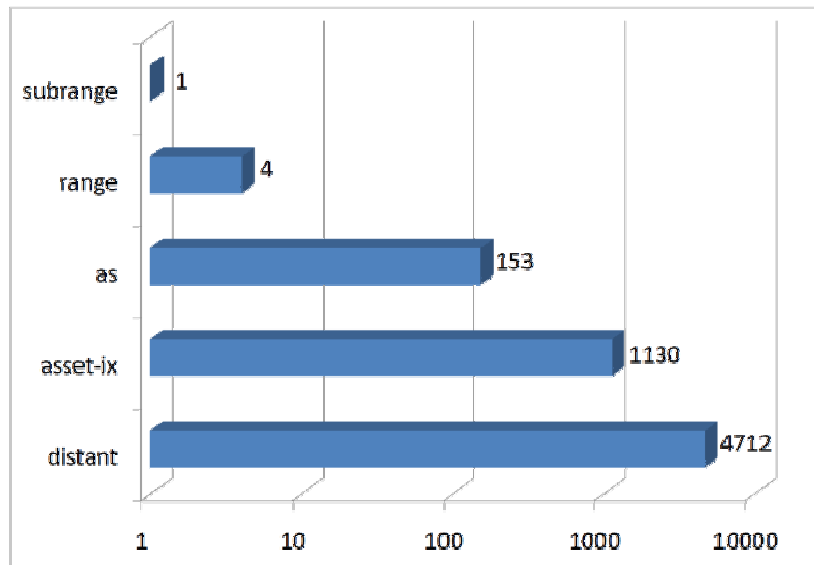


Рисунок 1 — Розподіл класів локальності по RuTracker

Вдосконалення моделі CARMA

Результати, отримані в серії тестів лише щодо трекера RuTracker не можуть бути визначальними. В подальших тестах ми вирішили взяти множини адрес зі зграйок BitTorrent щонайменше з двох трекерів — RuTracker та TorrentsNetUa (www.torrents.net.ua). Ці трекери відрізняються в одному ключовому аспекті, важливому для перевірки та підкреслення переваг моделі CARMA — відсоток вузлів в адресному просторі тієї самої точки обміну трафіком, що і вузол спостереження. За нашими приблизними оцінками, RuTracker має біля 20% а TorrentsNetUa має біля 95% українських користувачів он-лайн в будь-який момент. Якщо CARMA зможе експериментально підтвердити таку переважну кількість, це буде першим підтвердженням коректності моделі.

Однак подальші тести виконані з адресного простору UA-IX на зграйках TorrentsNetUa показали, що кількість вузлів, що були віднесені до класу «горизонт» значно менша, ніж очікувалося (рис.2).

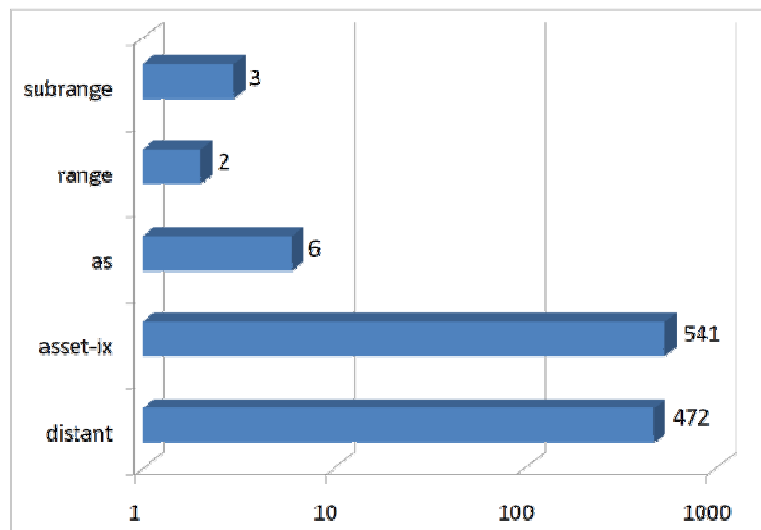


Рисунок 2 — Розподіл класів локальності по TorrentsNetUa

Вихідними даними цього графіку є випадково вибрані 1024 вузли зі зграйки на трекері TorrentsNetUa. Припускається, що переважна більшість з них працює в адресному просторі UA-IX, але тільки 54% були визначені як такі моделлю CARMA.

Така очевидна невідповідність зумовила пошук шляхів підвищення точності визначення класів. Оскільки модель CARMA демонструє досить високу швидкість (приблизно 10 пар адрес за секунду) обчислення класів навіть на комп’ютерах середнього рівня, нами було запроваджено додаткові перевірки

локальності в рамках згенерованої топологічної моделі, зокрема для двох додаткових різновидів взаємодії MAC.

Додаткові два класи локальності, що знаходяться вище «АС» та нижче, ніж «горизонт», такі:

- *MAC-Зв'язок (asset-link)* — показує, що вхідні адреси належать різним MAC, але принаймні один з цих MAC включає в себе інший.

- *Магістраль (backbone)* — визначає, що автономні системи, до яких належать задані діапазони IP-адрес, складають різні MAC, існує третя MAC, яка включає в себе ці дві.

Ці два додаткові класи були введені також і на основі детального аналізу структури та властивостей точки обміну UA-IX, який наведено в роботі [6].

Експериментальна перевірка

Тепер, коли модель CARMA може віднести пару адрес до одного з семи класів локальності, її ефективність не може бути оцінена просто з видимої відповідності між прогнозом та розподілом класів в експерименті.

З цієї причини ми вирішили використати модифікований метод ICMP *traceroute* у вигляді модуля в програмі пакетної обробки CARMA. Модуль вважає що перша адреса на вході є власною і виконує серію запитів ICMP *traceroute* щодо другої адреси. Модифікований метод ICMP *traceroute* відрізняється від стандартного таким чином:

- Протокол ICMP — аналогічно Windows-версії засобу, метод використовує ICMP замість UDP в аналогічному засобі середовища GNU. Здебільшого це зумовлено тим, що середовищами для запуску CARMA є Windows x64 та x86, де частиною IP Helper API є легкі в використанні та з програмової точки зору інтерфейси ICMP.

- Не проводиться перетворення DNS імен — адреси проміжних вузлів шляху не перетворюються в імена вузлів, і самі адреси не виводяться, оскільки в даному випадку інтерес представляє лише їх кількість.

- Адаптивна перевірка — на відміну від засобу командного рядку, час чекання відклику задано одною секундою. Якщо останній вузол, що відповів на запит не є другим вузлом вихідної пари і перевищення часу очікування сталося будь-де на протязі маршруту, весь процес вимірювання кількості вузлів запускається спочатку з першого проміжного вузла. Однак це не може трапитися більш ніж тричі. Якщо цільового вузла не досягнуто на другу і третю спробу, за кількість вузлів переходу приймається найбільше число переходів з трьох спроб. Це ефективно прибирає наслідки випадкових мережних затримок, які можуть викликати тимчасове перевищення часу очікування.

Модифікований метод ICMP *traceroute* було перевірено з користувачького підключення типу ADSL всередині UA-IX, яке характеризується середнім часом відклику біля 50 мілісекунд до найближчого оточення з точки обміну та біля 300 мілісекунд з інших вузлів. Експерименти показали, що середня тривалість циклу вимірювання кількості переходів для вузла становить не більше 0.5 секунди, якщо вузол відповідає на запит; не більше 5 секунд, якщо вузол не відповідає на запити до часу очікування; не більше 3 секунд, якщо маршрутизатори повертають помилку досяжності. Однак, в середньому, вимірювання кожної тисячі пар вузлів займає близько 1 години.

Модуль було інтегровано в програму пакетної обробки CARMA таким чином, щоб для кожної пари після обчислення класу локальності було проведено вимірювання реальної кількості переходів з подальшим записом у спільний протокол експерименту. Результат вимірювання кількості переходів не заноситься в протокол тільки у випадках, коли виявлено петлю маршрутизації (приблизно 2 випадки на 1000 пар адрес).

Щоб отримати різноманітний розподіл класів локальності CARMA, як і розподіл кількості переходів, нами було використано обидва публічних трекери, як показано вище. Однак, обсяг та зграйок серйозно відрізнявся — RuTracker зміг надати зграйку в 3610 адрес, тоді як TorrentsNetUa мав певні труднощі зі звітуваннями зграйок обсягом вище 1000 адрес. Причиною цього є те, що на RuTracker дозволяється (а з недавнього часу є обов'язковим) включення механізмів розповсюдження DHT та PEX. Ці механізми протоколу BitTorrent [7] дають можливість отримувати повний список адрес безпосередньо від вузлів, які вже приймають участь в зграйці замість часткових (обмежених в кількості вузлів у відповіді) запитів до трекеру, які можуть суттєво навантажувати його і спричинити збої в обслуговуванні.

Тестова зграйка з RuTracker складається з 3610 вузлів, з яких, як вважається, біля 20% має належати до простору UA-IX. TorrentsNetUa звітував зграйку обсягом 891 вузол, з яких практично всі належать до адресного простору UA-IX. Кожний вузол був оброблений вдосконаленою моделлю CARMA з адресою локального вузла (також всередині простору UA-IX). Потім було вимряно умовну відстань між цими вузлами в термінах кількості перехідних вузлів.

Експеримент тривав 5 годин. Щоб надати уявлення про характер результатів, їх значна частина наведена та нормалізований розподіл класів локальності наведені на рис.3.

Аналіз результатів та висновки

Аналіз отриманих результатів показав два суттєві покращення в функціонуванні моделі CARMA.

По-перше, розподіл класів по TorrentsNetUA показав, що кількість вузлів в класі «горизонт» зменшилася до 2.5% в порівнянні з попередніми експериментами, що підтверджує зроблений нами раніше прогноз.

Однак, потрібно відзначити, що в деяких випадках цей клас може означати вузол, який не оперує в адресному просторі UA-IX, але його адресний блок зареєстровано в Україні — якщо, наприклад, інформація в базах даних RIR є некоректною. В той же час можлива і обернена ситуація — в даному випадку не всі вузли, що отримали клас «магістраль» належать адресному простору українських точок обміну завдяки наявності кількох швидкісних каналів зв'язку між ними та російськими точками обміну через інфраструктуру TeliaSonera AB.

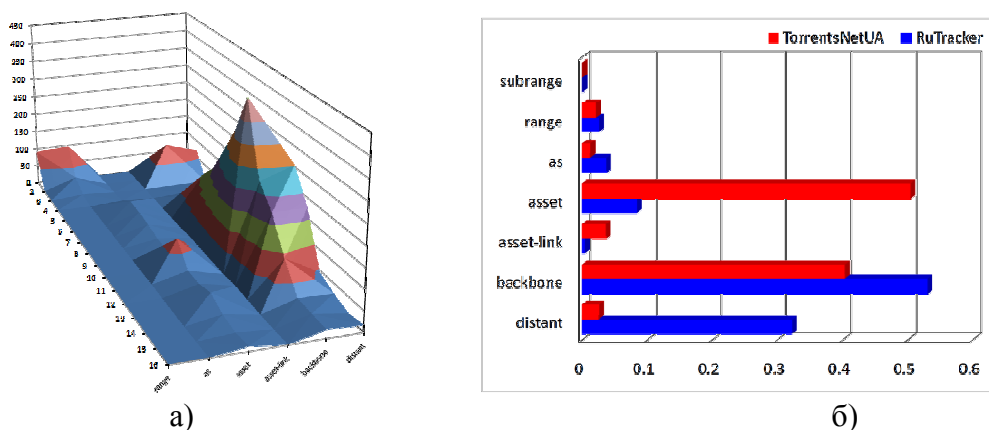


Рисунок 3 — а) Характер відповідності обчислених класів локальності вимірюваній кількості перехідних вузлів; б) — нормалізований розподіл класів локальності в експерименті.

По-друге, помітний малий відсоток вузлів, що отримали клас «МАС-зв'язок» для обох трекерів може свідчити про порівняно низьку кількість кросс-АС зв'язків, які не пов'язані спільною МАС більш високого рівня, або загальну тенденцію до побудови ієрархічних політик маршрутизації в національних конфігураціях мережі Інтернет, що відповідає рекомендаціям, даним в роботі [6].

Досить несподіваним став факт відсутності вузлів з класом «під-діапазон». Це могло статися або через невеликий обсяг виборки, або завдяки тому факту, що в базі даних RIR існує перекриваючий діапазон з тими ж параметрами — в цьому випадку модель CARMA повертає більш високий клас локальності «діапазон».

Із застереженнями, що числове представлення обчисленого класу локальності, на відміну від кількості вузлів переходу, не має ніякого фізичного значення та з розумінням, що природа цих двох параметрів є повністю відмінною, з рис.3а можна зробити висновок про високий рівень відповідності обчисленого класу локальності в моделі CARMA та топологічної відстані в мережі.

Дана робота є логічним розвитком попередніх доробок авторів [4] з частковим залученням механізмів 2-го рівня загальної моделі CARMA. В даній роботі, зокрема, було виконане вдосконалення як структурної моделі мережі Інтернет так і алгоритму знаходження метричного показника локальності шляхом введення двох додаткових класів локальності та операцій для їх знаходження, що дає підґрунтя для подальшої розробки та ефективного використання комбінованої метрики локальності в рамках загальної моделі CARMA.

Впровадження повної реалізації загальної моделі CARMA надає механізми для оптимізації як існуючих розподілених і однорангових мереж так і розробки оптимальної архітектури нових мереж.

Робота виконана в рамках гранту президента України для підтримки наукових досліджень молодих вчених, НДР GP/F27-0040 «Дослідження регуляційних та рейтингових методів в однорангових мережах типу BitTorrent» договір № Ф27/8-2010 від 12 лютого 2010.

Список літератури

1. H. Xie, A. Krishnamurthy, A. Silberschatz, and R. Y. Yang, "P4P: Explicit communications for cooperative control between P2P and network providers", 2007: http://www.dcia.info/documents/P4P_Overview.pdf
2. H. Madhyastha, T. Isdal, M. Piatek, and C. Dixon, "iPlane: An information plane for distributed services" In Proceedings of the 7th USENIX Symposium on Operating Systems Design and Implementation. USENIX, 2006, pp. 367–380.

3. T. Karagiannis, P. Rodriguez, and K. Papagiannaki, "Should internet service providers fear peer-assisted content distribution?" in IMC'05: Proceedings of the 5th ACM SIGCOMM conference on Internet measurement. New York, NY, USA: ACM, 2005, pp. 63–76.
4. G. Poryev, H. Schloss, and R. Oechsle, "CARMA based MST approximation for multicast provision in P2P networks" In Proceedings of the Sixth International Conference on Networking and Services (ICNS 2010). Cancun, Mexico: IEEE, 2010, pp. 123–128.
5. J. Hawkinson and T. Bates, "Guidelines for creation, selection, and registration of an Autonomous System (AS)". RFC 1930 (Best Current Practice), Internet Engineering Task Force, mar 1996. <http://www.ietf.org/rfc/rfc1930.txt>
6. Фурашев В., Зубок В., Ланде Д. «Параметры украинского сегмента интернет как сложной сети» // Открытые информационные и компьютерные технологии: Сб. научн. трудов. Вып. 40.—Харьков: Нац. аэрокосм. ун-т. "ХАИ", 2008.—С.235-242.
7. B. Cohen, "Incentives build robustness in BitTorrent" in the First Workshop on the Economics of Peer-to-Peer Systems, 2003.—pp.140-144.

Відомості про авторів

Порев Геннадій Володимирович — к.т.н., старший науковий співробітник, докторант Національного технічного університету України «КПІ», пр.Перемоги, 37, м.Київ, 03056, тел.. 258-52-09, +380-68-321-345-1, core@barvinok.net