

Г.В. Порєв

Національний технічний університет України «КПІ», Київ

ЕФЕКТИВНІСТЬ ОВЕРЛЕЙНОЇ СТРУКТУРИ ОДНОРАНГОВИХ ФАЙЛООБМІННИХ МЕРЕЖ



Запропоновані показники ефективності оверлейної структури однорангової мережі з врахуванням специфіки функціонування файлообмінних мереж та використанням коефіцієнтів балансування поведінки вузла та відносної топологічної локальності.

Ключові слова: Інтернет, однорангові мережі, оверлейні структури, локальність.

Серед вагомих надбань початку ХХІ ст. можна назвати низку технологій, розподілених та однорангових комп'ютерних мереж. Їх стрімкий розвиток обумовлений, в першу чергу, потребою відходу мереж доставки значних обсягів мультимедійного контенту від архітектури «клієнт—сервер», яка в сценаріях масового використання демонструє суттєві недоліки і загальне зниження ефективності.

Автори сучасних досліджень, спрямованих на пошук шляхів оптимізації функціонування розподілених та однорангових мереж, не приділяють достатньої уваги питанням побудови внутрішніх оверлейних структур однорангової мережі взагалі та питанням топологічної диференціації вузлів таких мереж зокрема.

Більшість робіт з цієї тематики [1–4] присвячені питанню вдосконалення засобів оцінки топологічної локальності активними методами або за допомогою виділеної наглядової інфраструктури. Так, в роботі [1] продемонстрована недосконалість засобу мережної діагностики *traceroute* на прикладі повернення цією програмою неіснуючих маршрутів та запропоновано відповідний модифікований засіб. Автори робіт [3, 4] як інструмент оптиміза-

ції роботи розподілених однорангових мереж пропонують використовувати виділену наглядову інфраструктуру у вигляді розподіленої мережі серверів, які здійснюють накопичення та обробку маршрутизаційної інформації від учасників мережі, та надають автоматичні рекомендації щодо топологічного розташування вузлів мережі. В роботах [5, 6] детально проаналізовано всі основні недоліки таких методів та запропоновано новий метод оцінки взаємної топологічної локальності на основі попередньо обчисленої локальної структурної моделі національних сегментів мережі Інтернет. Також було запропоновано використовувати показник локальності для оцінки взаємної топологічної віддаленості будь-якої пари вузлів (в тому числі нелокальних) як метрику для побудови оверлейних структур в однорангових мережах.

Дана робота є логічним розвитком попередніх доробок та присвячена питанню оцінки ефективності методів оптимізації оверлейних структур за допомогою розробленого в [5, 6] показника локальності.

ЗАДАЧА ОЦІНКИ ЕФЕКТИВНОСТІ

У цій роботі буде проаналізовано формальний критерій, за яким пропонується оцінювати ефективність оверлейної структури одно-

рангової мережі на основі даних про клас локальності та прогрес завантаження, і розглянуто частковий випадок однорангової мережі файлообмінного призначення. Спільною характеристикою таких мереж є те, що вихідний файл для поширення розбивається на декілька блоків однакового розміру. У випадках, коли в складі контенту наявні декілька файлів, вони розглядаються як єдиний неперервний бінарний потік даних. При початковому розповсюдженні контенту різні частини такого потоку розподіляються серед підключених вузлів у рої. Перед тим як перейти до розробки показника ефективності, необхідно зробити певні застереження щодо області його застосування та природи показників, які враховуються при його обчисленні.

В даній роботі розглядається, в основному, мережа BitTorrent. *По-перше*, необхідно враховувати важливу архітектурну особливість цієї технології, яка полягає в тому, що мережа BitTorrent, на відміну від багатьох інших, не формує єдиного простору імен або адрес та не формує повнозв'язного графа з вузлів-учасників мережі. Це означає, що для вузлів, які зайняті в різних роях, немає необхідності обмінюватися повідомленнями і протоколом не передбачена така можливість. Взагалі не формуються також і повнозв'язні графи з вузлів, що складають один рій. Хоча в останніх версіях найбільш популярних клієнтів мережі BitTorrent, таких, як μ Torrent, BitComet, Vuze та ін., реалізовані розподілені хеш-таблиці DHT (Distributed Hash Tables), які формують спільний простір імен, вони використовуються лише для того, щоб знаходити додаткові вузли з того самого рою (в якості ідентифікатора рою прийнято infohash торрент-файла), якщо з якихось причин трекер не доступний або повертає недостатню кількість адрес.

Всі сучасні клієнти мережі BitTorrent можуть приймати участь у більш ніж одному рою одночасно, але взаємодія між різними потоками даних в межах одного клієнта не виконується на протокольному рівні, а проявляється

лише як правила надання пріоритетів швидкості та розподіл пропускної здатності каналу. Навіть тоді, коли вузли мережі BitTorrent зайняті в кількох спільних роях одночасно, службовий трафік щодо кожного з них генерується та опрацьовується окремо. Відповідно незалежними є і відносини з вузлами інших роїв і ситуація з наявністю або відсутністю блоків потоку даних. За таких обставин доцільно ввести показник ефективності файлообмінної однорангової мережі як суб'єктивний, тобто такий, який характеризує стан рою з точки зору даного вузла. Похідні критерії загальної ефективності мережі можна буде згодом ввести на його основі як такі, що враховують всі суб'єктивні оцінки з кожного вузла рою.

По-друге, всі показники які можуть бути обчислені в контексті певного вузла рою, мають парну природу, тобто обчислюються і мають зміст лише по відношенню до іншого вузла того ж рою.

ПОКАЗНИК ЕФЕКТИВНОСТІ НА ОСНОВІ РЕЛЕВАНТНОСТІ ВУЗЛІВ

Зробивши вищезазначені застереження, можна ввести позначення характеристик i -го вузла рою по відношенню до даного, в контексті якого вони обчислюються:

- ✦ N_s — загальна кількість блоків в роздачі даного рою;
- ✦ N_o — кількість блоків, завантажена даним вузлом;
- ✦ N_i — кількість блоків, завантажена i -м вузлом;
- ✦ S_{ei} — кількість блоків, завантажених i -м вузлом і яких немає у даного вузла;
- ✦ S_{ai} — кількість блоків, яких немає у i -го вузла і які завантажені даним вузлом;
- ✦ f_i — показник локальності i -го вузла по відношенню до даного вузла у вигляді номера класу від 0 до максимального f_{max} .
- ✦ T_k — коефіцієнт, який враховує зовнішні по відношенню до рою обмеження на швидкість передачі даних між i -м та даним вузлом;
- ✦ K_g — максимальна дозволена кількість з'єднань для даного вузла з усіма іншими вузла-

ми з усіх інших роїв, де присутній даний вузол;

✦ K_a — максимальна дозволена кількість з'єднань для даного вузла з іншими вузлами одного рою, де присутній даний вузол;

У популярному BitTorrent-клієнті — μ Torrent — обчислюється загальний показник релевантності

$$R_{\mu} = \frac{S_{ei}}{N_s}. \quad (1)$$

Цей показник має суто інформаційне призначення, тобто алгоритм прийняття рішення щодо черговості спроб з'єднання з вузлами рою не враховує цей показник. Навпаки, автори μ Torrent вважають, що випадкове опитування вузлів зі списку пірів у випадку, коли цей список більший, ніж локальна кількість одночасно дозволених з'єднань на рій, дозволяє рівномірно розподілити навантаження серед вузлів учасників рою.

Очевидно, що в спеціальному випадку $N_i = N_s$ (i -й вузол є seeder) релевантність вузла R повинна оцінюватися лише в термінах кількості його вільних слотів для з'єднань та правилами обмеження швидкості передачі. Однак у спеціальному випадку $N_o = N_s$ (даний вузол є seeder) релевантність вузлів має оцінюватися з врахуванням показника S_{ai} :

$$R(N_s, N_i, N_o) = \begin{cases} N_i = N_s, & \bar{R}(f_i, T_k, N_o, K_g, K_a) \\ N_o = N_s, & \bar{R}(f_i, T_k, N_i, S_{ai}, K_g, K_a). \end{cases} \quad (2)$$

У запропонованій схемі використовуються два проміжні показники релевантності, які, зокрема, не залежать від загальної кількості блоків.

Перший, «егоїстичний», показник R_e відображає цінність i -го вузла для даного, яка зростає зі збільшенням кількості потрібних і відсутніх на даному вузлі блоків S_{ei} . Цінність також є меншою для вузлів, у яких частина потрібних, але відсутніх блоків менша відносно загальної кількості завантажених i -м вузлом, оскільки у такого вузла внаслідок механізмів адаптивного керування відкриттям слотів (smart

unchoking) менша імовірність встановити з'єднання з даним за один і той же проміжок часу. Показник R_e запишеться як

$$R_{ei} = \frac{S_{ei}}{N_i}. \quad (3)$$

Другий, «альтруїстичний», показник R_a означає цінність даного вузла для i -го вузла, яка зростає зі збільшенням кількості завантажених блоків. Ці блоки потрібні i -му вузлу і аналогічним чином зменшується для вузлів, які мають більшу кількість інших (не присутніх на даному) завантажених блоків, оскільки це також знижує імовірність виклику даного вузла i -м вузлом за один і той же проміжок часу. Показник R_a для цього випадку записується як

$$R_{ai} = \frac{S_{ai}}{N_i}. \quad (4)$$

У обох випадках введення у відношення N_i замість N_s дозволяє диференціювати вузли за абсолютною кількістю завантажених ними блоків, тобто дає можливість встановлювати з'єднання, розраховані на тривалий зв'язок. Легко довести, що у випадку

$$\forall i \neq n \in N_i = N_n : R_{ai} = R_{en}, R_{ei} = R_{an}. \quad (5)$$

Потрібно враховувати і той факт, що на імовірність з'єднання між i -м вузлом і даним впливає також політика обмеження швидкості передачі даних та користувальницькі настройки, зокрема глобальної кількості одночасних зв'язків K_g та кількості одночасних зв'язків K_a на кожен активний рій на вузлі. Важливим є те, що в поточній реалізації протоколу таку інформацію про віддалений вузол отримати неможливо. Але при подальшому розвитку протоколів передачі даних у файлообмінних мережах, зокрема BitTorrent, видається доцільним доповнити специфікацію протоколу службовими повідомленнями про відповідний стан вузлів.

Ще одним важливим елементом запропонованого в [5, 6] способу обчислення показника ефективності є можливість встановлення оператором вузла гнучкої політики розподілу пріоритетів між завантаженням та відвантажен-

ням блоків. Такий коефіцієнт балансування $k = [0...1]$ позначається так, що при $k = 0$ вузол працює в повністю «альтруїстичному» режимі, а при $k = 1$ – в повністю «егоїстичному».

Зрозуміло, що коефіцієнт балансування у виразі для показника ефективності повинен впливати на відповідні величини релевантності вузла таким чином:

$$R_{xi}(k) = kR_{ei} + (1-k)R_{ai}, \quad (6)$$

де $R_{xi}(k)$ – узагальнений показник релевантності i -го вузла для даного.

Також необхідно враховувати той факт, що у виразах (3) та (4) кількість потрібних вузлу блоків нормується до загальної кількості завантажених на віддаленому вузлі, а тому показник релевантності може виявитися практично однаковим для тих вузлів, які завантажили лише один блок, відсутній на даному, і для тих, які працюють в режимі «seeder» та яким залишилося завантажити один останній блок. У обох випадках показник релевантності буде дорівнювати або наближатися до одиниці відповідно:

$$\begin{aligned} \exists k \neq n: N_k = 1, S_{ek} = 1, N_n = N_s, S_{en} = \\ = N_s - 1: R_{en} \rightarrow R_{ek} = 1; \end{aligned} \quad (7)$$

аналогічно для «альтруїстичного» показника:

$$\begin{aligned} \exists k \neq n: N_k = 1, S_{ak} = 1, N_n = N_s, \\ S_{an} = N_s - 1: R_{an} \rightarrow R_{ak} = 1. \end{aligned} \quad (8)$$

З огляду на вказану причину доречно ввести додатковий множник для абсолютної кількості завантажених вузлом блоків N_i та N_o при обчисленні релевантності. Необхідно враховувати те, що значення цих параметрів на практиці можуть сягати кількох тисяч і з їх збільшенням релевантність зростає нелінійно, оскільки збільшується імовірність того, що спробі з'єднання буде відмовлено через зайнятість усіх слотів. У такому випадку доречно використовувати логарифмічну залежність, яку можна записати у загальному вигляді:

$$R_i = k \frac{S_i}{N_i} \ln(S_i + 1). \quad (9)$$

Обчислений показник локальності f_i відіграє в релевантності i -го вузла не меншу роль, ніж усі попередні характеристики. До того ж показник локальності є симетричним, тобто щодо нього виконується комутативний закон

$$\forall i, n: f_i(n) = f_n(i). \quad (10)$$

З урахуванням вищесказаного вираз для показника ефективності файлообмінної однорангової мережі можна записати як

$$C = \sum_i (f_{max} - f_i + 1) R_{xi}(k), \quad (11)$$

а з урахуванням виразів (3), (4) та (9) як

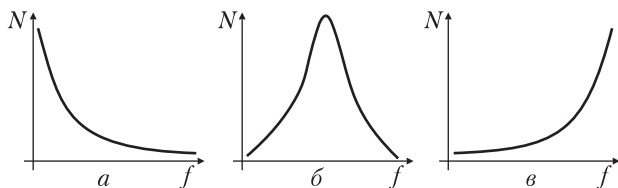
$$\begin{aligned} C = \sum_i (f_{max} - f_i + 1) \left(k \frac{S_{ei}}{N_i} \ln(S_{ei} + 1) + \right. \\ \left. + (1-k) \frac{S_{ai}}{N_i} \ln(S_{ai} + 1) \right). \end{aligned} \quad (12)$$

ПОКАЗНИК ЕФЕКТИВНОСТІ НА ОСНОВІ ПРОГРЕСУ ЗАВАНТАЖЕННЯ

Як відомо, одним з рушійних факторів виникнення однорангових мереж, зокрема файлообмінних, є необхідність розвантажити вузол мережі, що здійснює публікацію, від численних прямих запитів на завантаження однієї і тієї самої частини. У цьому полягає принциповий недолік клієнт-серверної архітектури комп'ютерних мереж, який проявляється в тих сценаріях використання, коли апаратно-програмні можливості клієнтських вузлів практично не відрізняються від серверних вузлів (в тому числі, за можливістю встановлення прямого з'єднання).

До недавнього часу вирішенням проблем перевантаження серверних ферм вважалося створення розподілених мереж доставки, які склалися з регіональних дзеркальних серверів з автоматичним (або напівавтоматичним) пере-направленням клієнтських запитів.

У сучасному інформаційному просторі технології однорангових мереж набувають дедалі більшого поширення як найбільш ефективний засіб розвантаження виділеної інфраструктури від клієнтських запитів.



Вплив кластеризації на розподіл вузлів по класах локальності: *a* – високий рівень кластеризації (*бажаний*), *б* – середній (*нейтральний*), *в* – низький (*небажаний*)

Розглянемо структуру оверлейної мережі одного рою з точки зору параметрів, важливих для інтегральної оцінки ефективності такої мережі.

Нехай $i \in [0, C]$ – номер та кількість частин публікації в поточному рої, $n \in [0, M]$ – номер та загальна кількість вузлів рою. Тоді двійкову функцію, яка показує наявність i -ї частини публікації \bar{C} на n -му вузлі рою, можна записати у такому вигляді:

$$p(n, i) = \begin{cases} 1: C_i \in \bar{C} \\ 0: C_i = \emptyset \end{cases} \quad (13)$$

При запуску початкового розповсюдження публікації з 0-го вузла ситуація з розподілами частин в рою визначена як

$$p(0, i) \equiv 1, \quad \forall n > 0: p(n, i) \equiv 0. \quad (14)$$

У процесі розповсюдження частини публікації передаються від *seeder*-вузла іншим вузлам оверлейної мережі. Можна сформулювати принцип побудови критерію оптимальності рою, виходячи з розподілу частин на вузлах. У найгіршому випадку, коли наявний на *seeder*-вузлі канал використовується для дублюючих пересилань, завантаження відбувається послідовно, тобто по одній частині публікації одночасно на всі вузли мережі. В ідеальному випадку кожна частина публікації відвантажується з кожного вузла, починаючи з *seeder*-вузла, іншим вузлам мережі тільки один раз.

Як проміжний показник Q_i виберемо кількість вузлів, на яких завантажено i -у частину публікації в мережі:

$$Q_i = \sum_{n=0}^N p(n, i) k_n. \quad (15)$$

Очевидно, що в наведеному вище гіршому випадку (на відміну від ідеального) значення Q_i матимуть значний розкид, який можна оцінити таким чином:

$$E = \sqrt{\frac{\sum_{i=0}^C \left(\frac{\sum_{l=0}^C Q_l}{C} - Q_i \right)^2}{C(C-1)}}. \quad (16)$$

Зауважимо, що ефективність оверлейної структури однорангової файлообмінної мережі визначається не тільки і не стільки ефективністю використання каналу зв'язку *seeder*-вузла, оскільки початкові умови, за яких це грає визначальну роль, тривають тільки протягом часу, потрібного на передачу перших частин. Надалі починають відігравати роль індивідуальні швидкості обміну даними між вузлами рою. А на це, в свою чергу, впливає рівень кластеризації рою на групи вузлів за ознакою локальності.

У загальному випадку, якщо вузли-учасники рою не використовують сортування списку з'єднаних вузлів за будь-якою ознакою, можна вважати, що розподіл по класах локальності серед них буде або приблизно рівномірним, або відповідати загальному розподілу по всьому рою. Однак у випадку, коли метрика локальності використовується для автоматичної кластеризації оверлейної мережі з вузлів рою, можна оцінити рівень кластеризації за допомогою нижчезазначених показників.

Нехай функція

$$D(n_1, n_2) = \begin{cases} 1: n_1 \leftrightarrow n_2 \\ 0: n_1 \nrightarrow n_2 \vee n_1 = n_2 \end{cases} \quad (17)$$

визначає наявність з'єднання між вузлами рою n_1 та n_2 . Тоді функцією

$$L(n_1, n_2, f) = \begin{cases} 1: D(n_1, n_2) \neq 0 \wedge n_1 \cup n_2 \in \bar{L}_f \\ 0: D(n_1, n_2) = 0 \end{cases} \quad (18)$$

можна визначити належність вузлів рою n_1 та n_2 до класу локальності f .

Для кожного вузла індивідуальний розподіл прямо зв'язаних вузлів по класах локальності визначиться виразом

$$L'(n, f) = \sum_{i=0}^N L(n, i, f). \quad (19)$$

Загальний розподіл вузлів по класах локальності в контексті одного рою записується як

$$E'(f) = \frac{\sum_{n=0}^N L'(n, f)}{N}. \quad (20)$$

Поведінка функції $E'(f)$ визначає рівень класифікації оверлейної мережі за групами локальності, як показано на рисунку.

З використанням виразу (20) контролююче програмне забезпечення вузлів однорангової мережі може обчислювати показник ефективності оверлейної структури та при необхідності приймати рішення про перепланування з'єднань.

ВИСНОВКИ

Використовуючи запропоновані показники ефективності оверлейної структури, розробники програмного забезпечення однорангових і розподілених мереж мають можливість контролювати поведінку оверлейних структур в реальному часі. Це, в свою чергу, надає підґрунтя для підвищення ефективності функціонування однорангових і розподілених мереж в цілому.

ЛІТЕРАТУРА

1. *Avoiding traceroute anomalies with Paris Traceroute* / B. Augustin, X. Cuvellier, B. Orgogozo et al. // IMC'06: Proceedings of the 6th ACM SIGCOMM conference on Internet measurement. — New York, NY, USA: ACM, 2006. — С. 153–158.
2. *Li J., Sollins K.* Exploiting autonomous system information in structured peer-to-peer networks // In ICCCN. — IEEE CS Press, 2004. — С. 403–408.
3. *Xie H., Krishnamurthy A., Silberschatz A., Yang R. Y.* P4P: Explicit communications for cooperative control between P2P and network providers. — 2007. www.dcia.info/documents/P4P_Overview.pdf.
4. *iPlane: An information plane for distributed services* / H. Madhyastha, T. Isdal, M. Piatek, C. Dixon // In Proceedings of the 7th USENIX Symposium on Operating Systems Design and Implementation. — USENIX, 2006. — С. 367–380.
5. *Poryev G., Schloss H., Oechsle R.* CARMA based MST approximation for multicast provision in P2P networks // In Proceedings of the Sixth International Conference on Networking and Services (ICNS 2010) / IEEE. — Cancun, Mexico: 2010. — С. 123–128.
6. *Poryev G.V., Schloss H., Oechsle R.* CARMA: A distance estimation method for internet nodes and its usage in P2P networks // International Journal on Advances in Networks and Services. — 2010. — Vol. 3, № 3, 4. — С. 114–128.

Г.В. Порев

ЭФФЕКТИВНОСТЬ ОВЕРЛЕЙНОЙ СТРУКТУРЫ ОДНОРАНГОВЫХ ФАЙЛООБМЕННЫХ СЕТЕЙ

Предложены показатели эффективности оверлейной структуры одноранговой сети с учётом специфики функционирования файлообменных сетей и использованием коэффициента балансирования поведения узла и относительной топологической локальности.

Ключевые слова: Интернет, одноранговые сети, оверлейные структуры, локальность.

G. Poryev

OVERLAY STRUCTURE EFFICIENCY IN PEER-TO-PEER FILE-SHARING NETWORKS

The efficiency indicator for the overlay structure of the peer-to-peer network, taking into account the specific aspects of the file-sharing networks, the behavioral balancing rate and the mutual topological locality have been proposed.

Keywords: Internet, peer-to-peer networks, overlay structures, locality.

Стаття надійшла до редакції 22.03.11