

## **ЄВРОПЕЙСЬКІ МЕРЕЖІ ОБМІНУ ІНТЕРНЕТ-ТРАФІКОМ ТА ЇХНІЙ ВПЛИВ НА ЗВ'ЯЗНІСТЬ МІЖ АВТОНОМНИМИ СИСТЕМАМИ**

The article dedicated to research of Europe's largest network of Internet exchange DE-CIX using the routing tables of its participants. Also proposed a calculation methods for whole network and single node parameters such as degree distribution, clustering coefficient, average geodesic path, the global effectiveness and vulnerability.

Refs: 6 titles.

*Keywords:* Internet Exchange, power law, degree distribution, global efficiency

### **Вступ**

Перші мережі обміну трафіком – “network access points” – з'явилися в ARPANET ще до того, як термін Internet набув популярності. Ідея їхнього заснування – додержуватись чітко визначених процедур підключення та правил взаємодії між мережами учасників.

Сьогодні в Європі функціонує багато мереж обміну трафіком. Вони мають різну кількість учасників та обсяги трафіку, різні процедури підключення, а головне - різну політику маршрутизації та взаємодії між учасниками. Але кожна з них має свій вплив на топологію зв'язків між автономними системами в Інтернеті.

### **Історія створення мереж обміну трафіком**

У 1983 році протоколи TCP/IP було інтегровано в „серійну” операційну систему BSD UNIX. Це суттєво спростило побудову взаємодії між серверами. У 1985 році до існуючої мережі ARPANET приєдналась наукова мережа Національної наукової фундації (NSF) – NSFnet. З того часу термін “Інтернет” став загальноновживаним.

Велика кількість вузлів Інтернет стала проблемою для керування мережею. Тому з 1986 р. в NSFNet застосовувалась трирівнева мережна архітектура: місцеві (локальні мережі, мережі кампусів) підключались до регіональних мереж, які в свою чергу підключались до опорної мережі (backbone), що підтримувалась 6 загальнонаціональними суперкомп'ютерними центрами NSF, як наведено на рис. 1.

Велика кількість вузлів Інтернет стала проблемою для керування мережею. Тому з 1986 р. в NSFNet застосовувалась трирівнева мережна архітектура: місцеві (локальні мережі, мережі кампусів) підключались до регіональних мереж, які в свою чергу підключались до опорної мережі (backbone), що підтримувалась 6 загальнонаціональними суперкомп'ютерними центрами NSF, як наведено на рис. 1 [1].



Рис. 1. Мережа NSFNet в 1986 – 1988 роках.

До 1988 року основні канали передачі даних в мереж працювали на швидкості 56 кбіт/с. В 1987 році консорціум за участі Merit (міжуніверситетської комп'ютерної дослідницької мережі), IBM, MCI виграв грант NSF на реінжиніринг та обслуговування NSFNet. В ході робіт канали були змінені на канали T1 (1,5 Мбіт/с). Канали T1 об'єднали 6 суперкомп'ютерних центрів та ще 13 вузлів мереж по всій території США.

За словами Елізи Геріх, віце-президента ICANN, нинішнього керівного органу Інтернету, яка особисто приймала участь в цих роботах, така інфраструктура мала б слугувати становим хребтом мережі протягом 5 років. Але вже за 18 місяців, у 1991 році, обсяги трафіку змусили переводити канали T1 на T3 - 45 мбіт/с (з виступу на форумі RIPE NCC Regional Meeting, Москва, 29.09–01.10.2010). Схему нової мережі наведено на рис.2. Точки доступу (network access point – NAP) спочатку були розділені на державні (федеральні - FIX) та комерційні (CIX). Приблизно з 1994 року такий розподіл зник. Разом з цим, зник розподіл користувачів на державні, академічні та комерційні організації. Поняття NAP та вимоги до них були формалізовані, і це відкрило шлях до побудови нових точок.

NAP, за термінологією NSFNET, це комплекс з 1 чи більше швидкісних комутаторів, до яких під'єднується певна кількість маршрутизаторів для обміну трафіком. Всі мережі при підключенні до NAP автоматично отримують дозвіл обміну трафіком з іншими мережами. Пропускна спроможність мережі, що підключається, має бути соразмірною з вже підключеними мережами.

Для кожної NAP призначалась посада – менеджер NAP. До його обов'язків входило розгортання, технічне, організаційне, комерційне обслуговування NAP та планування її розвитку.



Рис. 2. Т3-мережа NSFNet в 1992 – 1995 роках.

Але завжди існувала можливість взаємного підключення вузлів поза межами NAP, тому топологія Інтернет в результаті розвитку отримала так звану безмасштабну структуру.

В результаті розробки та публікації «NSFNet Acceptable Use Policy» та відкритості точок доступу Інтернет набула стрімкого розвитку. Тому у 1995 році близько 44% префіксів в таблиці маршрутизації були з поза меж США.

В топології сучасного Інтернету NAP, чи як їх зараз називають – мережі обміну трафіком чи біржі трафіку (Internet Exchanges) відіграють дуже важливу роль. Лише деякі з них керуються політикою та процедурами того часу. Одною з „класичних” мереж обміну трафіком є Українська мережа обміну трафіком (UA-IX).

### **Сучасні приклади мереж обміну трафіком в Європі**

Сьогодні в Європі функціонує багато мереж обміну трафіком. Найвизначнішими з них за кількістю учасників та обсягами трафіку є DE-CIX у Франкфурті-на-Майні, AMS-IX в Амстердамі, LINX в Лондоні та інші. Вони мають різну кількість учасників та обсяги трафіку, різні процедури підключення, а головне - різну політику маршрутизації та взаємодії між учасниками. Але кожна з них має свій вплив на топологію зв'язків між автономними системами в Інтернеті.

За обсягами трафіку, що передається, Українська мережа обміну трафіком посіла в 2010 році шосте місце в Європі (рис.3). А перше місце належить мережі DE-CIX, із понад 1.3 Гбіт сумарного трафіку (рис.4).

Взаємодія між учасниками UA-IX здійснюється за протоколом BGP-4 шляхом обміну таблицями маршрутизації між маршрутизатором учасника та серверами маршрутизації (routing servers, RS) UA-IX. При цьому трафік між учасниками проходить безпосередньо через порти комутатора, минаючи сервери маршрутизації.

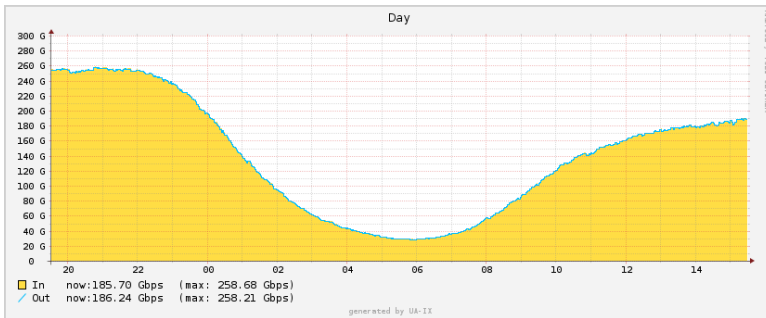


Рис.3. Графік навантаження UA-IX протягом доби в грудні 2010 року (в Гбіт/с).

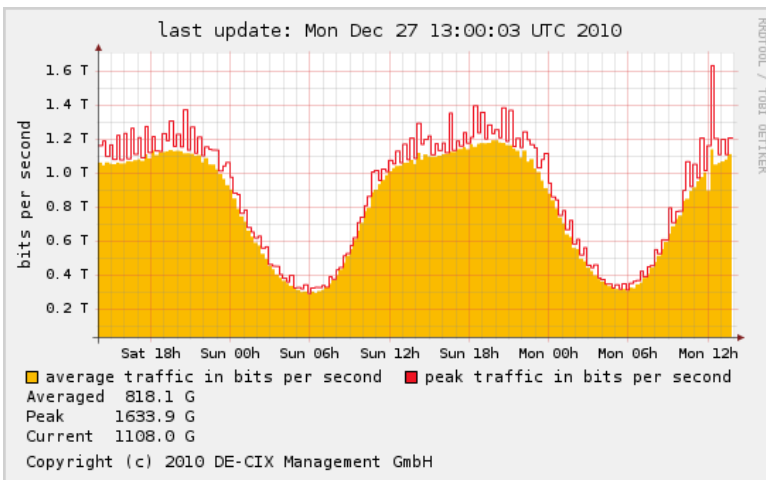


Рис.4. Графік навантаження DE-CIX в грудні 2010 року (Тбіт/с).

У відповідності до політики маршрутизації UA-IX, кожен учасник зобов'язаний віддавати на RS таблиці маршрутизації своєї автономної системи і має право анонсувати таблиці маршрутизації автономних систем своїх клієнтів за умови обов'язкового прийому ними таблиць маршрутизації від UA-IX. Таким чином досягається симетрія трафіку в мережі, а така політика має назву «відкритої».

Технічна політика DE-CIX є більш широкою, ніж UA-IX. Обмін маршрутами через RS є лише одним з варіантів. На відміну від UA-IX, правила DE-CIX передбачають можливість побудови двосторонньої взаємодії між автономними системами, оминаючи RS (так званий private peering) Також можлива побудова приватних мереж обміну трафіком на технічних майданчиках DE-CIX. Це означає, що DE-CIX пропонує оренду місця для комунікаційного обладнання, а також активну та пасивну мережеву інфраструктуру. З точки зору дослідження взаємодії автономних систем, ці

мережі обміну мають досліджуватись окремо, безвідносно до автономної системи DE-CIX.

### **Характеристики сегментів глобальних інформаційних мереж**

У прикладних дослідженнях звичайно застосовують такі типові для мережного аналізу характеристики, як розмір мережі, мережна щільність, розподіл ступіню вузлів, ступінь центральності тощо.

При аналізі мереж на базі теорії графів виділяють:

- розрахунок параметрів окремих вузлів;
- розрахунок параметрів мережі в цілому;
- виділення мережних підструктур.

#### *Параметри вузлів мережі*

Для окремих вузлів виділяють наступні параметри:

- вхідний ступінь вузла - кількість ребер графа, що входять у вузол;
- вихідний ступінь вузла - кількість ребер графа, що виходять з вузла;
- відстань від даного вузла до кожного з інших;
- середня відстань від даного вузла до інших;
- ексцентричність (eccentricity) — найбільше з геодезичних відстаней від даного вузла до інших;
- проміжність (betweenness) — число шляхів, що включають даний вузол;
- центральність (загальна кількість зв'язків даного вузла стосовно інших).

#### *Загальні параметри мережі*

Для розрахунку індексів для мережі в цілому використовують такі параметри:

- число вузлів;
- число ребер;
- середня геодезична відстань  $l$  (найкоротший шлях) між вузлами;
- середня відстань від одного вузла до інших;
- розподіл ступіню вузлів  $P(k)$ ;
- щільність (відношення кількості ребер у мережі до можливої максимальної кількості ребер з даною кількістю вузлів);
- кількість симетричних, транзитивних і циклічних тріад;
- діаметр мережі (найбільша геодезична відстань у мережі).

#### *Виявлення мережних підструктур*

Існує декілька актуальних завдань дослідження складних мереж:

- виділення компонентів (частин мережі), які зв'язані усередині та не зв'язані між собою;
- знаходження блоків і перемичок; вузол називається перемичкою (cutpoint), якщо при його вилученні мережа розпадається на незв'язані частини;

- виділення групувань - груп еквівалентних вузлів (які мають максимально схожі профілі зв'язків).

#### *Шлях між вузлами*

Шлях між вузлами - це послідовність вершин і ребер, що з'єднують дві вершини. Відстань між вузлами - кількість кроків, які потрібно зробити, щоб добратися від однієї вершини до іншої. Вершини в графі можуть бути зв'язані прямо або опосередковано. Для всієї мережі введемо поняття середньої відстані між вершинами:

$$l = \frac{1}{\frac{1}{2}n(n+1)} \sum_{i \geq j} d_{ij}, \quad (1)$$

де  $d_{ij}$  - найкоротша відстань між вузлами  $i$  та  $j$ .

#### *Коефіцієнт кластеризації*

Д. Уаттс і С. Страттс у 1998 році визначили такий параметр соціальних мереж, як коефіцієнт кластеризації, що відповідає рівню зв'язності вузлів у мережі.

Коефіцієнт кластеризації для окремого вузла мережі визначається в такий спосіб. Нехай з вузла виходить  $k$  ребер, які, мабуть, з'єднують його з  $k$  іншими вузлами, найближчими сусідами. Якщо припустити, що всі найближчі сусіди з'єднані безпосередньо один з одним, то кількість ребер між ними складало би  $\frac{1}{2}k(k-1)$ . Тобто це число, що відповідає максимально можливій кількості ребер, якими могли б з'єднуватися найближчі сусіди обраного вузла. Відношення реальної кількості ребер, що з'єднують найближчих сусідів до максимально можливого й називається коефіцієнтом кластеризації вузла. Природно, ця величина не перевищує одиниці.

Коефіцієнт кластеризації може визначатися як для кожного вузла, так і для всієї мережі. Відповідно, рівень кластеризації мережі визначається як нормована по кількості вузлів сума відповідних коефіцієнтів окремих вузлів. Розглянутий нижче феномен «малих світів» безпосередньо пов'язаний з рівнем кластеризації.

#### *Розподіл ступіню*

Розподіл ступіню  $P(k)$  визначається як вірогідність того, що вузол  $i$  мережі має ступінь  $k_i=k$ . Відомо, що мережі із різним розподілом ступіню демонструють різну „поведінку” в разі змін в мережі.  $P(k)$  може бути експонентним розподілом, розподілом Пуассона чи степеневим розподілом [2].

#### **Попередні дослідження**

У рамках роботи [3] досліджувалися параметри українського сегменту Інтернет як складної мережі. В ході дослідження було встановлено, що

сукупність транзитних вузлів українського сегменту Інтернет є складною мережею із степеневим розподілом ступіню, коефіцієнтом кластерності 0,04 та середнім найкоротшим шляхом 3,2. Це підтвердило теоретичні висновки щодо відсутності суттєвої кластеризації мережі, наявності дуже великої кількості вузлів (майже 90%), які взагалі не виконують транзитних функцій з точки зору взаємодії автономних систем, та малої кількості вузлів з високим ступенем посередництва, які насправді утворюють опорну мережу. Виявлення та обрання двох чи більше саме таких вузлів для підключення нової автономної системи має суттєві переваги з точки зору зберігання зв'язності підчас часткової відмови мережі.

З цієї точки зору є дуже цікавим дослідження частини європейського сегменту «навколо» DE-CIX та отримання аналогічних характеристик стосовно найбільшої в Європі мережі.

### **Вхідні дані для дослідження**

Протокол BGP-4 має засоби фільтрації маршрутної інформації, встановлення пріоритетів у відповідності із політикою, але не має засобів публікації цієї політики. Для публікації правил маршрутизації автономних систем використовуються централізовані бази даних – реєстри маршрутизації (Internet Routing Registry, IRR). Але треба зважувати на те, що дані IRR відображають правила (політику) взаємодії, а не реальний стан такої взаємодії на певний момент часу.

Найбільш повну і актуальну інформацію про зв'язки між вузлами мережі на рівні автономних систем можна отримати, дослідивши вхідні таблиці маршрутизації з маршрутизаторів, які взаємодіють з найбільшою кількістю автономних систем. Для реалізації цієї методики дослідження необхідно мати безпосередній доступ до такого маршрутизатора, чи через так звані сервери-«дзеркала» (looking glass servers). Ця інформація, що знаходиться в цій таблиці, відображає актуальний стан взаємодій «прикордонних» маршрутизаторів автономних систем.

У рамках роботи [3] Вихідні дані для аналізу були отримані від учасників UA-IX. Аналіз проводився на основі BGP-таблиці маршрутизації, отриманої безпосередньо з маршрутизатора, що взаємодіє з UA-IX.

Для дослідження DE-CIX ми використали базу даних учасників DE-CIX (<https://www.peeringdb.com/private/index.php>), DE-CIX Looking Glass ([http://www.de-cix.net/content/network/looking\\_glass.html](http://www.de-cix.net/content/network/looking_glass.html)), таблиці маршрутизації BGP, а також інформацію з Internet Routing Registry Європейського реєстру (<http://www.db.ripe.net/whois>).

Для аналізу мережі було обрано таку множину вузлів та ребер:

- у якості вузлів мережі було обрано автономні системи, які мають транзитну функцію, тобто виконують „реєкспорт анонсів” (анонсують не лише префікси, які походять від власної AS, але й ретранслюють анонси сусідніх AS);

- у якості ребер (неспрямованих) було обрано BGP-піринг, тобто факт наявності обміну анонсами між автономними системами, які є вузлами мережі, що досліджується.

BGP-взаємодія та анонси префіксів відображаються маршрутизаторами у вигляді, наведеному на рис. 5.

```
> sh ip bgp neighbor 80.81.192.145 route
BGP table
Status codes: * valid, > best
Origin codes: i - IGP, e - EGP, ? - incomplete

  Network                Next Hop          Metric LocPrf Weight Path
*> 193.107.109.0/24      80.81.192.145    100         21219 21354 50027 i
*> 213.5.192.0/21       80.81.192.145    100         21219 50130 50130 50130 50130 i
*> 193.34.72.0/22       80.81.192.145    100         21219 39189 39189 i
*> 193.107.108.0/24     80.81.192.145    100         21219 21354 50027 i
*> 213.111.104.0/21     80.81.192.145    100         21219 48683 i
*> 195.95.189.0/24     80.81.192.145    100         21219 41105 i
*> 193.107.111.0/24    80.81.192.145    100         21219 21354 50027 i
```

Рис. 5. Фрагмент вхідної таблиці маршрутизації BGP-4, отриманої з route-сервера DE-CIX.

## Результати дослідження

### *Кількісні показники DE-CIX*

За наведеними вище даними було підраховано, що в рамках відкритої політики маршрутизації за допомогою route servers DE-CIX взаємодіють 410 автономних систем, але приблизно 10% з них є неактивними (тобто BGP-взаємодія відсутня). Загалом вони анонсують приблизно 43000 префіксів, які або походять з їхніх AS, або є реекспортом анонсів, отриманих від AS їхніх клієнтів.

Номер автономної системи DE-CIX – 6695. Кількість автономних систем, які зустрічаються в DE-CIX в BGP-атрибутах шляху (origin, as\_path) перевищує 9000. За кількістю анонсованих префіксів найбільшим з учасників DE-CIX є оператор Telecom Italia Sparkle (AS6762), який анонсує майже 15000 префіксів. Серед українських учасників DE-CIX найбільшим є оператор «Датагруп» (AS21219), який анонсує понад 500 префіксів.

### *Методики обрахунку розподілу ступіню та інших показників*

Для Інтернету, як і багатьох мереж, які самоутворилися в результаті людської діяльності, є характерним степеневий розподіл ступіню [4], тобто вірогідність того, що вузол  $i$  має ступінь  $k$  зменшується степеневим законом із зменшенням  $k$ :

$$P(k) \sim 1/k^\gamma, \quad k \neq 0, \quad \gamma > 0 \quad (2)$$

при цьому для Інтернету є характерним  $2 < \gamma < 3$  [5].

Досліджувана мережа не має бути виключенням. До неї входить відносно дуже мала кількість вузлів, які мають ступінь вище 100, тобто не



менше 100 зв'язків з іншими AS. Вузли з високим ступенем мають назву габів [2]. Також вони мають низький коефіцієнт кластеризації, і, відповідно, високий показник посередництва (betweenness).

Середній геодезичний шлях мережі має бути характерним для „малого світу” [4,5]. Зворотна величина від середнього геодезичного шляху - дуже тісно пов'язана характеристика мережі – *глобальна ефективність* [6].

$$E = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{d_{ij}}. \quad (3)$$

З глобальної ефективності розраховується показник *вразливості*  $V$  для певного вузла  $i$ , який визначає, наскільки зміниться середній геодезичний шлях (знизиться ефективність мережі) в разі вилучення певного вузла  $i$ , відповідно, всіх його зв'язків:

$$V_i = \frac{E - E_i}{E} \quad (4)$$

Таким чином виконується моделювання характеристик нової мережі, яка не містить певного вузла. Слід зазначити, що, якщо під  $i$  розуміти сукупність з двох чи більше вузлів, можна отримати справжню модель вразливості мережі до атак на габи.

Результати дослідження параметрів вузлів з найвільшим ступенем зв'язності в мережі можна оформити у вигляді таблиці, в якій буде наведено про кожний вузол:

- номер AS;
- назва оператора;
- кількість зв'язків  $k$  з транзитними AS;
- коефіцієнт посередництва;
- вразливість.

### **Проблеми обрахунку показників мережі та напрямки оптимізації**

Проблемою в ході даного дослідження є розмір мережі, точніше – розмір матриці суміжності, яка є основною структурою даних, що використовується для представлення графів в комп'ютерних програмах.

У випадку дослідження взаємодії автономних систем розмір матриці суміжності складає  $N_{as}^2$ , де  $N_{as}$  - кількість автономних систем мережі. І хоча ми розглядаємо ненаправлений граф, та головна діагональ матриці складається з нулів, „інформативний” розмір матриці складає  $N_{as}/2 - N_{as}$ , що для досліджуваної мережі дає приблизно 81 000 000 значень. При цьому результуюча матриця є сильно розрядженою та більшість значень буде дорівнювати нулю.

В результаті аналізу таблиць маршрутизації (рис.5) було визначено 10470 атрибутів `as_path`, що є унікальними (не повторюються).

Збір первинної інформації та отримання даних для обчислень займає

лічені хвилини. Але пошук зв'язаних пар AS1-AS2 для кожного з 81 млн варіантів по переліку з понад 10 тис унікальних атрибутів `as_path` є або досить витратною задачею з точки зору обчислювальних ресурсів (потребує розподілених обчислень), або займатиме неприйнятний час (мільйони секунд, тобто десятки діб). Подальшими напрямками оптимізації є:

- обґрунтування зменшення діаметру досліджуваної мережі та, як наслідок, розмірності матриці, за рахунок видалення автономних систем із ступенем 1 та 2, тобто таких, що не мають значного впливу на транзити;
- спрощення `as_path` за рахунок виключення штучного подовження шляху (`prepend`);
- побудова матриці суміжності на основі пошуку сумісних пар AS замість перевірки всіх можливих пар AS на сумісність.

### **Висновки**

Викладені методики дозволяють застосувати математичний апарат для вивчення особливостей як групи Інтернет-вузлів, так і властивостей певного вузла та його ролі в мережі. Отримана інформація допоможе зробити обрання оператора для підключення до мережі більш технічно обґрунтованим. Це особливо стосується випадків обрання другого провайдера та побудови взаємодії між автономними системами по протоколу BGP-4.

Аналіз зниження ефективності мережі при вилученні вузлів, груп вузлів та їхніх зв'язків дозволяє моделювати вразливість мережі до вилучення (від'єднання) ключових учасників мережі.

Оскільки поставлена задача має не лише суто дослідницький, але й прикладний характер, необхідно продовжувати пошук шляхів оптимізації методів обрахунку параметрів складних мереж та реалізувати їх практично.

1. *National Science Foundation. NSF and the Birth of the Internet.* - [www.nsf.gov](http://www.nsf.gov).
2. Ю. Головач, К. фон Фербер, О. Олемської, Т. Головач, О. Мриглод, І. Олемської, В. Пальчиков. Складні мережі // Журнал фізичних досліджень, 2006. –Т. 10. - С. 247–291.
3. Ланде Д.В., Зубок В. Ю., Фурашев В.Н. Исследование сетевых параметров украинского сегмента Интернет. // Открытые информационные и компьютерные технологии: Сб. науч. трудов. Вып. 40. – Харьков: Нац. аэрокосм. ун-т „ХАИ”: 2008.
4. A.-L. Barabasi, E. Bonabeau. Scale-Free Networks // Scientific American. – May 2003.-pages 50-59.
5. Faloutsos M., Faloutsos P., & Faloutsos C. On Power Law Relationships of the Internet Topology // Comput. Commun. Rev. 29, (1999) 251-263.
6. V.Latora, M.Marchiori // Efficient behavior of small-world networks. Physical Review Letters, vol. 87, no. 19, 5 Nov 2001.

*Поступила 27.01.2011р.*