

Зубок В.Ю.

АНАЛІЗ ВПЛИВУ ТОПОЛОГІЇ НА ЕФЕКТИВНІСТЬ ТА ВРАЗЛИВІСТЬ В ГЛОБАЛЬНИХ КОМП'ЮТЕРНИХ МЕРЕЖАХ

Аннотация:

В статье продолжено изучение возможностей применения теории сложных сетей для исследования обширных сегментов Интернет на основе таблиц маршрутизации ее участников. На примере Украинской сети обмена трафиком (UA-IX) рассчитаны такие показатели как распределение степени узлов, средний геодезический путь, глобальная эффективность и уязвимость. В работе приведена методика моделирования топологических изменений определенного сегмента Интернет и сравнения параметров сети до и после изменений. Эмпирические данные показывают, что чувствительность сети к удалению определенного узла не зависит от количества связей данного узла.

Библиогр.: 10 наим.

Ключевые слова: Интернет, сложные сети, распределение степени, средний путь, глобальная эффективность, уязвимость.

Анотація:

У статті продовжено вивчення можливостей застосування теорії складних мереж для дослідження великих сегментів Интернет на основі таблиць маршрутизації її учасників. На прикладі Української мережі обміну трафіком (UA-IX) обраховано такі показники мережі як розподіл ступеню вузлів, середній геодезичний шлях, глобальна ефективність та вразливість. В роботі наведена методика моделювання топологічних змін певного сегменту мережі Интернет та порівняння параметрів мережі до та після змін. Емпіричні дані вказують, що чутливість мережі до вилучення певного вузла не залежить від кількості зв'язків цього вузла.

Бібліогр.: 10 найм.

Ключові слова: Интернет, складні мережі, розподіл ступеню, середній шлях, глобальна ефективність, вразливість.

Abstract:

The article continues to research the ways of exploring wide segments of the Internet based on the routing tables of its participants. Provided calculation for such network parameters as degree distribution, average geodesic path, the global effectiveness and vulnerability. This review shows a technique for modeling topological changes of a certain segment of the Internet and compare the network parameters before and after the change. Empirical evidence suggests that the sensitivity of the network to the removal of a particular node does not depend on the number of links of this node.

Refs: 10 titles.

Keywords: Internet, complex networks, degree distribution, average path, global efficiency, vulnerability.

Вступ

Людство розбудовує Інтернет понад 30 років, та понад 20 років активно користується його можливостями. В той же час і Інтернет вплинув на людство – сьогодні саме Інтернет є унікальним та всебічним інформаційним, обчислювальним, навчальним, підприємницьким та навіть виховним доккіллям.

Глобальна інтеграція стала можливою багато в чому завдяки прискоренню інформаційного обміну. Територіальна розгалуженість призвела до активного впровадження Інтернет у бізнес-процеси (насамперед в документообіг). Тому отримало величезне значення питання загальнодоступності (по співвідношенню вартості до якості) глобальних мереж передачі документальної, аудіо- та візуальної інформації із забезпеченням цілісності, захищеності та доступності на прийнятному для суб'єктів телекомунікацій рівні.

Цікаво, що саме сімейство мережевих протоколів TCP/IP є найбільш вдалим експериментом з побудови потужної, надійної, незалежної від фізичної середовища передачі телекомунікаційної мережі [1]. Але переваги Інтернет у певній мірі є її недоліками. Ці недоліки відчутні кожному користувачу, насамперед, в разі втрати зв'язку з мережею чи окремими її частками, чи в незадовільній пропускну здатності між користувачем інформації та підмережею, в якій розташований необхідний інформаційний ресурс. Зазвичай, пересічний користувач чи навіть чимале підприємство не може вплинути на розподіл ресурсів мережі поза межами зони своєї відповідальності. Але надійність взаємодії з Інтернет має суттєву залежність від обрання точок взємоз'єднання в глобальною мережею.

Основні знання щодо маршрутизації в Інтернеті

Важливою особливістю Інтернет була та є можливість динамічної зміни маршрутів передачі пакетів між вузлами, які фізично не поєднані. Для такої маршрутизації розроблені та використовуються у якості стандартів єдині „міжвузлові” протоколи маршрутизації та правила оформлення політик маршрутизації.

Наявність зв'язків між вузлами Інтернет визначається наявністю прикордонної взаємодії між групами мережевого обладнання. З точки зору цієї прикордонної взаємодії, вузлом мережі ми будемо називати групу маршрутизаторів, які спільна політика маршрутизації об'єднує в автономну систему. Під автономною системою (autonomous system, AS) ми розуміємо групу IP-мереж, які належать одному чи декільком операторам, та мають єдину чітко визначену політику маршрутизації [2]. Автономні системи між собою обмінюються інформацією про маршрути з використанням дистанційно-векторного протоколу BGP-4 [3].

Роздивимось взаємодію чотирьох AS на практичному прикладі (рис. 1).

Кожна AS надає інші AS, з якою в неї є прикордонна взаємодія, адреси підмереж, або „префікси”, трафік до яких вона готова прийняти. AS4 має взаємодію з AS3 по каналу d та анонсує свої префікси до AS3. Це означає, що AS3 відтепер знає щонайменше один шлях, по якому можна доправити трафік, адресований підмережам, що їх анонсувала AS4. В той же час, AS3 анонсує до AS4 свої префікси.

Як показано на схемі, AS3 має також прикордонну взаємодію з AS1 та AS2 по каналах b та c. Тому, в залежності від політики маршрутизації, AS3 може анонсувати до

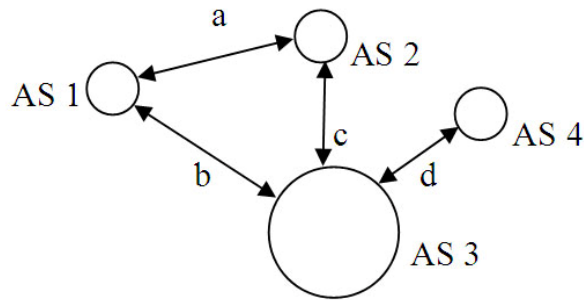


Рис. 1. Прикордонна взаємодія автономних систем.

AS 1,2,3,4 – автономні системи; a, b, c, d – канали зв'язку.

AS3 ще й префікси, отримані від AS1 та AS2, а префікси, які отримані від AS4, анонсувати до AS1 та(або) AS2. Таким чином AS3 стане транзитною для всіх інших AS.

Крім того, ми бачимо на прикладі, що AS1 та AS2 мають між собою прикордонну взаємодію (a), тому в трикутнику AS1-AS2-AS3 вони можуть бути транзитними одна для одної. Ці дані ми можемо використати, аби представити певний сегмент мережі у вигляді графа.

Інтернет - складна мережа

Багато явищ в природі та в людських співтовариствах відбуваються за участю цілої мережі схожих взаємодіючих об'єктів чи тіл, що складаються в мережу. Виявляється, загальні характеристики цих явищ (наприклад, стійкість, здібність до адаптації і т. д.) залежать, як правило, не від конкретних об'єктів, а від математичних властивостей мережі в цілому: зв'язаності, однорідності, кластеризації, ієрархії.

Теорія мереж знайшла численні застосування в житті — у вивченні транспортних потоків, шляхів розповсюдження інформації та знань, біологічних схем, тощо. На цей час з метою дослідження мережі розподіляють на соціальні, інформаційні, технологічні та біологічні [4].

Інтернет як телекомунікаційна мережа не має центру планування та керування. Централізація стосується лише деяких операцій, таких як розподіл адресного простору та керування системою доменних імен верхнього рівня. Топологія мережі не є визначеною чи навіть типовою, кількість та структура вузлів також не є константами. Постійно змінюючись, Інтернет набув характеристик складної мережі.

Одною з найважливіших характеристик складних мереж є розподіл ступіню (кількості зв'язків) вузлів $P(k)$, який визначається як вирогідність того, що вузол i має ступінь $k(i)$. Для орієнтованих мереж окремо розраховується вхідний та вихідний ступінь вузла. Мережі з різними $P(k)$ демонструють різну поведінку [5].

Емпіричне вивчення Інтернету як складної мережі проводиться з початку 1990-х років. Їх проводять на рівні маршрутизаторів або на міждоменному рівні. У першому випадку вузли мережі представлені маршрутизаторами, спеціальними комп'ютерами, що контролюють рух даних у мережі, а в другому — вузлами є так звані автономні системи (АС). АС є підмережею, яка складається з багатьох маршрутизаторів, що реалізують єдину та чітку політику маршрутизації [ripe-181].

В комплексному дослідженні „On Power-Law Relationships of the Internet Topology” [6] Інтернет проаналізований і на рівні АС (на основі трьох карт, створених від листопада 1997 р. до грудня 1998 р.), і на рівні маршрутизаторів (на основі даних 1995 року). Автори

дослідили, що розподіл багатьох кількісних характеристик інтернету має степеневий характер.

Емпіричні дослідження тривають й досі. Найвідоміші з них — Cooperative Association for Internet Data Analysis (<http://www.caida.org>), який у 2006 році об'єднав зусилля з проектом Active Measurement Project (National Laboratory for Applied Network Research), а також RIPE NCC Test Traffic Measurements (<http://www.ripe.net/projects/ttm/>).

Результати досліджень топології Інтернету

Автори всіх досліджень підтвердили, що Інтернет, типово до „штучних” мереж, утворених людством, має безмасштабну (scale-free) топологію. Основна властивість такої побудови — степеневий (power-law) розподіл ступеню вузлів:

$$P(k) \sim 1/k^\gamma \quad (1)$$

де k — це кількість зв'язків (ступінь) випадково обраного вузла,

γ — масштабуюча експонента ($2 < \gamma < 3$),

$P(k)$ — розподіл ступеню вузла, тобто вірогідність того, що випадково обраний вузол має k зв'язків.

Безмасштабним мережам властива наявність вузлів з величезною кількістю зв'язків (їх називають габами — hubs), в той час як середня кількість зв'язків є відносно невеликою. Саме в цьому сенсі такі мережі отримали назву безмасштабних.

Досліди показали, що і природним, і утвореним людством мережам властивий розвиток, і що з ростом кількості вузлів вони перетворюються в безмасштабну. Причина в тому, що „старі” вузли в мережі мають більше можливостей щодо отримання зв'язків з новими вузлами при їхньому утворенні [7]. Стосовно Інтернет, цей феномен добре знайомий пересічному користувачеві на прикладі всесвітнього павутиння World Wide Web: чим довше знаходиться ресурс в WWW без зміни адреси, тим більше на нього посилань нараховується пошукових системах. Два фактори — ріст мережі та переваги її нових членів — пояснюють існування габів.

Параметри, що характеризують складні мережі

Зупинимось ще на деяких найважливіших поняттях з теорії складних мереж, які, крім розподілу ступеню, також будуть використані в даній праці.

Відстань між вузлами можна визначити як кількість кроків, які необхідно зробити, щоб дістатись від одного вузла до іншого. Природно, що вузли можуть бути з'єднані безпосередньо (за допомогою єдиного ребра) або опосередковано, декількома ребрами через інші вузли. *Шляхом* між вузлами називається найкоротша відстань між ними. В зарубіжній літературі зустрічається визначення „геодезичний шлях” (geodesic path) [4].

Коефіцієнт кластерності [8] показує, скільки найближчих сусідів заданого вузла № також найближчими сусідами один для одного. Він характеризує тенденцію до утворення груп взаємопов'язаних вузлів — так званих клік (clique). Для окремого вузла мережі, який має ступінь k , тобто з якого виходить k ребер, що з'єднують його з k іншими вузлами (так званими найближчими сусідами), коефіцієнт кластерності визначається як відношення реальної кількості ребер E_m , якими з'єднані найближчі сусіди вузла, до максимально можливого, яке дорівнює, як відомо, $k(k-1)/2$:

$$C_m = \frac{E_m}{k(k-1)/2} \quad (2)$$

Рівень кластерності для всієї мережі визначається як арифметичне середнє значення коефіцієнтів кластерності по всіх вузлах мережі та вказує на ймовірність існування зв'язку між двома випадково взятими найближчими сусідами вузла, а також містить інформацію про наявність у мережі „трикутників” (циклів довжиною три).

Однією з важливих характеристик вузлів є посередництво (betweenness, або betweenness centrality). Посередництво $\sigma(m)$ відображає роль вузла в установленні зв'язків у мережі й показує, скільки найкоротших шляхів проходить через цей вузол:

$$\sigma(m) = \sum_{i \neq j} \frac{B(i, m, j)}{B(i, j)}, \quad (3)$$

де $B(i, j)$ — загальна кількість найкоротших шляхів між вузлами i та j , а $B(i, m, j)$ кількість найкоротших шляхів між i та j , таких, що проходять через вузол m .

Для всієї мережі можна визначити поняття *середнього шляху*, як середньої по всіх парах вузлів найкоротшої відстані між ними:

$$l = \frac{2}{n(n-1)} \sum_{i>j} d_{ij}, \quad (4)$$

де d_{ij} — шлях (найкоротша відстань) між вузлами i і j .

Однак деякі мережі можуть бути непов'язаними, тобто в них є вузли, відстань між якими є нескінченною. Відповідно, середній шлях може також дорівнювати нескінченності. З метою урахування таких випадків існує визначення *середнього інверсного шляху* між вузлами, який обчислюється за формулою:

$$il = \frac{2}{n(n-1)} \sum_{i>j} \frac{1}{d_{ij}}, \quad (5)$$

де d_{ij} — шлях (найкоротша відстань) між вузлами i і j .

Враховуючи, що, природно, "швидкість" передачі інформації мережею зворотно пропорційна середньому найкоротшому шляху, цей показник часто називають також *глобальною ефективністю E*.

З глобальної ефективності розраховується показник *вразливості* V_i 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:

$$V_i = \frac{E - E_i}{E} \quad (6)$$

Таким чином виконується моделювання характеристик нової мережі, яка не містить певного вузла.

Без перебільшення можна вважати, що вузли з більшим посередництвом (габи) грають головну роль у встановленні зв'язків між іншими вузлами мережі. Синонімом посередництва є термін "навантаження" (load), який відіграватиме дуже важливу роль в аналізі надійності Інтернет. Слід зазначити, що, якщо у формулі (6) під i розуміти

сукупність з двох чи більше вузлів, можна отримати справжню модель вразливості мережі до атак на габи.

Як було сказано раніше, взаємодію Інтернет-вузлів можна представити у вигляді графа. А мережу в цілому можна представити у вигляді безмасштабної моделі із степеневим розподілом ступіню.

Мережі зі степеневим розподілом ступеню мають ще декілька важливих властивостей, які витікають з їхньої топології. Вони „стійкі” до випадкового вилучення вузлів, але вразливі в разі цілеспрямованого вилучення навіть невеликої кількості габів. У випадку з телекомунікаційною основою Інтернет ця вразливість проявляється в тому, що вихід з ладу кожного вузла, який має велику кількість зв’язків, може різко уповільнити швидкість перетікання інформації по мережі в цілому [7, 9], і ось чому.

В теоретичній випадковій мережі (з розподілом ступеня за Пуассоном) всі вузли приблизно рівні за ступенем, а ребра з’єднують їх „від сусіда до сусіда”, і вилучення вузла чи навіть групи вузлів не зробить ніяку частину мережі а ні ізольованою, а ні кластеризованою. Але висока стійкість до пошкоджень досягається ціною вкрай низької ефективності передачі інформації на далекі відстані [5] через дуже довгий середній шлях.

В реальних мережах — із степеневим розподілом — вилучення габа призведе до перерозподілу великої кількості шляхів між іншими вузлами, рівень їхнього посередництва (або „навантаження”) значно виросте?

В роботі 2008 року [10] наведено аналіз Української мережі обміну трафіком UA-IX, вузлами якої виступають автономні системи (AS) українського сегменту Інтернет. За допомогою методології складних мереж досліджено розподіл ступеню вузлів, середній інверсний шлях, коефіцієнт кластерності, розподіл ступенів вузлів.

В ході дослідження виявлено, що з 1034 автономні системи, анонси яких фігурують в UA-IX, більше половини, а саме 777 AS, мають ступінь 1, тобто не є посередниками і не виконують транзитних функцій. Високий рівень посередництва властивий лише декільком вузлам із ступенем від 5 до 43. В разі вилучення цих одиниць габів, сотні вузлів опиняться в ізоляції від інших вузлів цієї умовної мережі. На практиці зв’язки не зникнуть, але найкоротші шляхи між вузлами, які наразі з’єднує UA-IX, простягнуться через вузли, які є за межами української частини Інтернет, збільшуючи навантаження на так звані „зовнішні”, міжнародні Інтернет-канали.

Моделювання вилучення вузла

Станом на листопад 2011 р. в UA-IX присутні анонси 115 учасників, що безпосередньо підключені до цієї мережі. Вони обмінюються через UA-IX анонсами, які походять від 1540 автономних систем. Розподіл ступеню в UA-IX зберігає типовий степеневий характер (рис. 2).

На рис. 3 зображено зв’язки 150 вузлів UA-IX з найвищим ступенем. Обрано метод візуалізації сегменту Інтернет в двовимірних полярних координатах із змінним радіусом в залежності від ступіню вузла, а зміна кута є постійним інкрементом. В подальшому планується розробити методи використання кута вузла для відображення однієї з його характеристик.

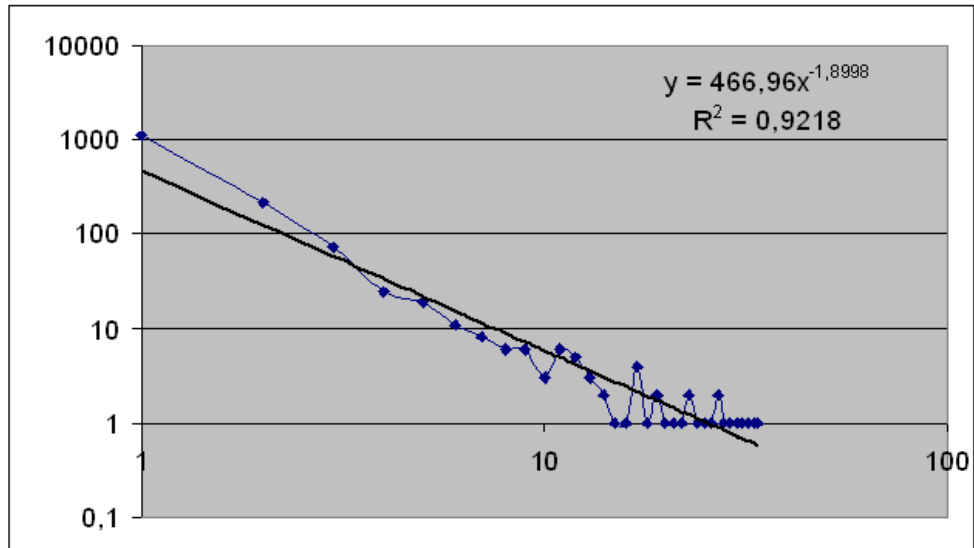


Рис.2. Розподіл ступеню вузлів в UA-IX та результат апроксимації степеневою функцією

Максимальну кількість зв'язків з іншими AS – 172 зв'язки – демонструє AS 21219 (ПрАТ “Датагруп”), яка знаходиться на рис. 3 приблизно в центрі координат. Спробуємо змодельовати, як зміниться середній шлях мережі (4), глобальна ефективність (5) та вразливість (6) в разі вилучення з мережі вузла AS 21219.

Для цього в матриці інцидентності обнулимо всі елементи матриці в k -стовбці та в k -рядку. Це еквівалентно вилученню вузла k та всіх його зв'язків, та рівнозначно до видалення з таблиці маршрутизації всіх анонсів, що походять від вузла k чи реекспортуються через нього. Результати розрахунків наведено в таблиці 1.

Таблиця 1

Зміна параметрів мережі при вилученні вузла з найвищим ступенем

№	Параметр	Значення до вилучення AS21219	Значення після вилучення AS21219	Різниця, %
1.	Середній шлях (l)	4.274	втрата зв'язності	-
2.	Глобальна ефективність (E)	0.25	0.201	-19.6
3.	Вразливість (V)	-	0,196	-

Тепер обрахуємо зміну цих параметрів в разі вилучення вузла із значно меншою кількістю зв'язків (115) - AS15645. Результати наведено в таблиці 2.

Таблиця 2

Зміна параметрів мережі при вилученні другого за ступенем вузла

№	Параметр	Значення до вилучення AS15645	Значення після вилучення AS15645	Різниця, %
1.	Середній шлях (l)	4.274	втрата зв'язності	
2.	Глобальна ефективність (E)	0.25	0.202	-19,2%
3.	Вразливість (V)	-		-

Література:

1. *Craig Hunt*. TCP/IP Network Administration, 3rd Edition // O'Reilly Media, 2002. - 752 стор.
2. *T. Bates, E. Gerich etc.* Representation of IP Routing Policies in a Routing Registry (ripe-181). - 1994. <ftp://ftp.ripe.net/ripe/docs/ripe-181.txt>
3. *Y. Rekhter, P. Gross*. RFC 1772. Application of the Border Gateway Protocol in the Internet. – <http://tools.ietf.org/html/rfc1772>
4. *M.E.J. Newman*. The structure and function of complex networks // SIAM Review. - 2003. - Vol. 45. pp. 167–256.
5. *Ю. Головач, К. фон Фербер, О. Олемської, Т. Головач, О. Мриглод, І. Олемської, В. Пальчи́ков*. Складні мережі // Журнал фізичних досліджень, 2006. –Т. 10. - С. 247–291.
6. *Faloutsos M., Faloutsos P., & Faloutsos C.* On Power Law Relationships of the Internet Topology // Comput. Commun. Rev. 29, (1999) 251-263.
7. *A.-L. Barabasi, E. Bonabeau*. Scale-Free Networks // Scientific American. – May 2003.- pages 50-59. *Ланде Д.В.*,
8. *D.J. Watts, S.H. Strogatz*. Collective dynamics of “small-world” networks. // Nature. - 1998. - Vol. 393. pp. 440-442.
9. *V.Latora, M.Marchiori* // Efficient behavior of small-world networks. Physical Review Letters, vol. 87, no. 19, 5 Nov 2001.
10. *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.