

Вимірювання і візуалізація Інтернету

Вимірювання в Мережі можуть бути найрізноманітнішими. Обсяг відомої анотованої бібліографії по вебометрії складає 25 сторінок. Найпростіше вимірювання полягає в тому, щоб намагатися визначити розміри і темпи росту Інтернету — наприклад, у числі сторінок чи числі сайтів. Брайан Мюррей і Олвин Мур з компанії «Cyveillance» в одній зі своїх статей подали графік росту Інтернету [24].

Структура зв'язків у Мережі досить нетривіальна [1-5]. Як зараз прийнято вважати, вона відповідає моделі «краватки-метелика» (Bow Tie). Ця краватка складається з наступних компонентів:

1. Сильно зв'язаний компонент (SCC) чи «центральне ядро» чи «вузол краватки». Це така безліч веб-сторінок, що кожна з них може бути досягнута з будь-якої іншої по відповідних орієнтованих гіперпосиланнях. Це ядро включає приблизно 30% веб-сторінок.
2. Стартові веб-сторінки. З центрального компонента SCC на ці сторінки потрапити не можна, але самі вони містять гіперпосилання, що дозволяють в остаточному підсумку потрапити в SCC.
3. Фінальні веб-сторінки. З цих сторінок не можна потрапити в SCC, але в центральному SCC містяться орієнтовані гіперпосилання, що відправляють в остаточному підсумку на фінальні сторінки.
4. Ізольовані ресурси. Це веб-сторінки цілком відділені від SCC. Вони, у свою чергу, можуть бути виділені в менші по розмірі сильно зв'язані компоненти. Серед них є також «веб-острова», узагалі не взаємодіючі з іншими ресурсами Інтернет. Ці острови не можуть бути знайдені ніякими пошуковими машинами.

Останні три категорії ресурсів: стартові, фінальні й ізольовані — приблизно однакові по розмірах (22-24% веб-сторінок).

Вимірювання компактності соціуму

Багато біологічних і соціальних систем містять у собі мережі «тісного світу» («Small World») [6 – 14]. У цих «тісних світах» виявляється дуже коротким відстань між двома будь-якими індивідуумами. Докладний опис концепції і математики «тісних світів» мається в монографії молодого американського математика Дункана Уоттса (Duncan J. Watts) «Малі Світи: динаміка Мереж між порядком і випадковістю». Дункан Уоттс одержав свою Ph D. по теоретичній і прикладній механіці в 1997 році в Корнельському університеті. Його книга присвячена систематичному погляду на «малі світи» як цікавий об'єкт у теорії графів. Фахівці з теорії графів відкрили «тісні світи» ще в 60-і роки, однак, ці об'єкти не були як слід зрозумілі. Ці математичні об'єкти мають застосування в широкому діапазоні наукових дисциплін, включаючи нейробіологію, соціологію, екологію, економіку й епідеміологію.

Кожен знайом з явищем «тісного світу»: зустрівши з новою людиною, ми нерідко виявляємо, що маємо загальних друзів, чи зв'язані через короткий ланцюжок знайомих. У розмовному англійському це описується вираженням «Six degrees of separation»: очевидно, важко знайти в будь-якім співтоваристві двох людей, яких би не дозволил з'єднати ланцюжок знайомих, що не перевищує по довжині, скажемо, шість чоловік. Той факт, що між індивідуумами утворюються досить сильні зв'язки, використовується в декількох популярних іграх. Наприклад, відома в Америці гра «шість рівнів Кевина Бэкона» полягає в тому, щоб спробувати знайти найкоротший шлях від будь-якого актора до Кевину Бэкону. Оскільки акторський світ досить тісний (у тому числі у формальному змісті *теорії графів*), важко знайти якого-небудь американського актора, що відстоїть більш, ніж на чотири кроки від актора Кевина Бэкона. У цій грі метрика визначається спільною грою двох акторів в одному фільмі. Вважається, що відстань між акторами дорівнює одиниці, якщо вони знімалися спільно. Якщо ж вони не знімалися спільно, але мається хоча б один актор, що знімався з кожним з них у якому-небудь фільмі, то відстань у цій парі вважається рівним двом і т.д. Ця відстань називається в даній грі числом Бэкона і затверджується, що в Інтернет-базі даних по кіно дуже мало акторів з числом Бэкона

вище чотирьох. Мається трохи сайтів, що містять гру Кевина Бэкона. Один з них розроблений кафедрою комп'ютерних наук Школи інженерних і прикладних наук університету Вірджинії (www.cs.virginia.edu/oracle/).

Існує також відома гра для математиків (чи навіть більше, ніж гра) «Число Ердеша». Ця гра полягає в тому, щоб знайти відстань між будь-яким вченим і найвідомішим угорським математиком Ердешом (Erdos). Метрика в цьому науковому світі визначається аналогічно метриці у світі акторів. Якщо вчений був співавтором Пауля Ердеша, то його число Ердеша дорівнює одиниці, якщо ж він має спільну публікацію з ким-небудь, хто коли-небудь був співавтором Пауля Ердеша, його число Ердеша дорівнюють двом і т.д. Учені досить серйозно сприймають гру в числа Ердеша – і особливо, математики.

Ердеш був одним з найбільших математиків двадцятого століття, — а, можливо, усіх часів. У всякому разі, по числу опублікованих їм статей (більш півтори тисяч) він перевершив усіх, навіть великого Леонарда Ейлера. Він формулював і вирішував найважливіші проблеми теорії чисел і інших областей математики. Він один із засновників *дискретної математики*, що лежить у фундаменті комп'ютерних наук. Він був мандрівником у цьому світі, кочуючи з країни в країну, з університету в університет. Протягом багатьох літ у нього не було власного будинку. Наприкінці 30-х років у Принстоні йому відмовили в продовженні стипендії, знайшовши його «не вихованим і не зручним». Він не тільки писав багато сам, але і спонукав до творчості інших людей: у нього було більш п'ятисот співавторів протягом життя. Зовні він був схожий на іншого великого мандрівника — Андрія Сахарова.

Числа Ердеша були частиною математичного фольклору по усьому світі. Потім вони перетворилися у формальний Проект «Числа Ердеша». Вэб-сайт цього Проекту підтримується Джерри Гроссманом в університеті Окленда в співробітництві з поруч інших математиків. Для обчислення чисел Ердеша використовується база даних журналу «Mathematical Reviews» Американського Математичного Суспільства. Ця база даних містить зведення про більш ніж півтора мільйони статей, написаних 337 тисячами різних авторів. Приблизно дві третини статей написані одним автором, чверть – двома авторами, біля

семи відсотків – трьома авторами, інші мають більше число співавторів. Частка статей, написаних одним автором безупинно зменшується починаючи з 90% у сорокові роки, до менш 50% у наші дні. Число Ердеша, рівне двом, має 5 713 чоловік, середнє значення числа Ердеша – 4,69. Найбільше число Ердеша, рівне 15, має Р.Камалов. Розподіл чисел Ердеша таке, що майже кожен математик з кінцевим числом Ердеша має його значення менше, ніж 8 – і тільки в 2% воно вище. Вчені в інших науках, включаючи соціальні дисципліни, також можуть мати малі числа Ердеша. Джерри Гроссман помітив якось з похмурим гумором, що він навіть чув про коняку, що оголосила, що її число Ердеша дорівнює трьом. Багато лауреатів Нобелівської премії в самих різних науках мають не дуже велике число Ердеша. Наприклад, в Альберта Ейнштейна воно дорівнює двом. В інших великих фізиків воно теж мало: у Вольфганга Паулі, Макса Борна, Мюрея Гелмана, Абдусса Салама дорівнює трьом, Вернера Гайзенберга, Поля Дирака, Ричарда Феймана, Юліана Швінгера, Гансу Беті – чотирьом. Найвідоміша людина комп'ютерного світу Вільям (Білл) Гейтс має число Ердеша, рівне чотирьом. У 1979 році Білл Гейтс опублікував статтю з Кростосом Пападимитриоу, що публікувався раніше в співавторстві з Ксайо Тань Денг, що публікувався зі співавтором Ердеша Павол Хелл — це воістину тісний світ!

У генеалогічну гру, подібну до чисел Ердеша почали недавно грати і люди зі сфери комп'ютерних наук. Одна з професійних асоціацій цієї сфери підтримує базу даних по привласнених докторських ступенях (випускник, науковий керівник, рік і місце захисту) по теоретичних комп'ютерних науках по усьому світі (<http://sigact.acm.org/genealogy>). Коли-небудь, ця база даних дозволить визначати відстань між фахівцями, по ланцюжку вчителів.

Вимірювання компактності Інтернету

Лада Адамик з Каліфорнійського дослідницького центра фірми «Хегох», очевидно, першої помітила, що Всесвітня Мережа WWW також має топологію тісного (чи маленького) світу. Мережа має тенденцію розпадатися на кластери. Однак у теж час тільки кілька гіперпосилань розділяє будь-який сайт від будь-якого іншого.

Мовою *теорії графів* сайти⁷ відіграють роль вершин графа, а гіперпосилання є орієнтованими ребрами графа. Відповідно до визначення Дункана Уотса граф тісного світу має наступні властивості:

- коефіцієнт утворення кластерів набагато більше, ніж у випадковому графі з тим же самим числом вершин і тим же самим середнім числом ребер у розрахунку на одну вершину (іншими словами, інтенсивність зв'язків вище, ніж у світі випадкових зв'язків);
- середня відстань між двома вершинами майже також мала, як для відповідного випадкового графа.

На противагу графам тісного світу випадкові графи не так схильні утворювати кластери, але мають короткі середні відстані, у той час як регулярні ґрати мають довгі відстані і тенденцію до утворення кластерів.

Лада Адамік обробила деяку випадкову вибірку сайтів, зібрану Джімом Пітковим з компанії «Хегох» (вибірка містила більш чверті мільйона сайтів і п'ятдесят мільйонів сторінок). Найбільша SCC у цій вибірці містила близько 65 тисяч сайтів. Розподіл сайтів у SCC по довжині найкоротшого шляху між ними має чітко виражений максимум при довжині шляху, рівним чотирьом гіперпосиланням. Коефіцієнт кластировання в цій вибірці майже в сто разів перевершує відповідний коефіцієнт для графа з випадковими зв'язками. Максимальна довжина найкоротшого шляху, рівна ланцюжку в 15 гіперпосилань, зустрілася в цій безлічі в дев'ятох випадках.

Лада Адамік виконала аналогічне дослідження для безлічі освітніх сайтів освітнього домена .edu. З 11000 сайтів домена .edu, включених у список Ладой Адамік, у ядро – тобто найбільшу SCC, увійшло 3456 сайтів. Середня відстань між сайтами в цьому ядрі склала 4,062. Це дуже близько до середнього шляху випадкового графа тих же розмірів: 4,048. У теж самий час коефіцієнт кластирування сайтів домена .edu на два порядку перевершує відповідний коефіцієнт для випадкового графа. Розходження між випадковим графом і мережею виявляється в «хвостах» розподілів найкоротших шляхів. Для SCC домена .edu досить часто зустрічаються довгі шляхи, аж до шляху довжиною в 13 гіперпосилань. А для випадкового графа такого ж розміру самий довгий з

найкоротших шляхів має довжину 8 посилань, і більш довгі шляхи малоймовірні.

Лада Адамик зауважує, що це спостереження повинне вплинути на розробки пошукових машин Інтернету. Гарна пошукова машина повинна використовувати «тісноту» Мережі (Лада Адамик використовує трохи незграбний словотвір *small-worldness*, що можна передати не менш незграбним українським словом «малосвітність»).

Розумно припустити, що гарний веб-документ посилається на інші гарні документи аналогічного контенту (це можна розглядати як сховане визначення гарного документа). У такій групі взаємозалежних документів є такі, котрі є гарними центрами, у тім змісті, що відстані від них до будь-якого іншого документа в групі в середньому мінімально. Гарна пошукова машина у відповідь на конкретний запит повинна пропонувати Вам не безформні довжелезні списки, а насамперед показувати Вам центри кожної SCC. Сортуючи потім центри по розмірі відповідних SCC, можна надати користувачу максимальний простір пошуку з мінімальним числом входів у цей простір. Одержавши гарну стартову крапку, користувач може експлуатувати ці SCC за своїм розсудом.

Врахування топології тісного світу може значно спростити навігацію в Мережі і прискорити витяг корисної інформації. Врахування цієї структури дозволяє також розібратися в структурі людських співтовариств, що породжують дану структуру гіперпосилань за допомогою документів, що вони створюють. Тут можна згадати наївні перші дні Інтернету, що вже ідуть у сиву давнину тисячоріччя, що завершилося; можна згадати, що Тім Бернерс-Ли задумував свій проект Мережі тільки лише для того, щоб відобразити і так існуючі неформальні співтовариства усередині Церна, щоб полегшити збереження й обмін інформації в цих співтовариствах. У ці сиві дні Тім Бернерс-Ли для задуманого їм створення мав тільки одне ім'я *Mesh* (петля, мережа, пастка). На звичному тепер для всіх словосполученні *World Wide Web* він зупинився роком пізніше при написанні програми для автоматичного перекладу документів у стандарт HTML (www.w3.org/History/1989/proposal.html).

Алгоритми візуалізації

Які би вимірювання ми ні проводили в Інтернеті, в остаточному підсумку важливо – чи можна це представити в наочній формі [15 – 23]. Будь-які вимірювання генерують безліч чисел. Однак не усі люблять числа. І навіть люблячі математику навряд чи будуть у захваті від перегляду масивів, що містять десятки тисяч (чи навіть десятки мільйонів) чисел. Необхідність наочного представлення інформації про Інтернет була усвідомлена давно, на ранніх стадіях розвитку Всесвітньої Павутини. Геометрична наочність у представленні результатів потрібна глобальним пошуковим машинам; потрібна вона і при пошуку корисної інформації навіть у локальних мережах помірної розміру. Наприклад, звичайний університетський веб-ресурс може складатися з декількох сотень сайтів і декількох тисяч зв'язків між ними. Пошук інформації навіть у такому обмеженому просторі, що є усього лише яскравою краплею в мислячому океані Інтернету, – може стати досить обтяжним заняттям.

Найбільш очевидним математичним об'єктом, що відображає Інтернет, є граф. Вершинами графа є індивідуальні сайти, роль орієнтованих ребер грають спрямовані гіперпосилання. Однак адекватне геометричне відображення графа навіть з помірним числом вершин у три-п'ять сотень, не говорячи вже про графів з мільйонами вершин, — задача важка чи попросту нерозв'язна. Відома безліч спроб розробити відповідні математичні моделі і програмне забезпечення, що дозволили б справитися з проблемою візуалізації вебметричних даних. Деякі з цих алгоритмів — наприклад, техніка «Риб'ячого ока» були розроблені ще в доінтернетовську епоху, для адекватного і наочного представлення деяких картографічних даних. Піонерські роботи цього напрямку виконав Джордж Фурнас з лабораторії Белла на початку 80-х років. Наочна графічна інтерпретація математичної моделі Фурнаса належить Маноджит Саркар і Маркові Брауну з дослідницького центру фірми DEC (<ftp://ftp.digital.com/pub/DEC/SRC/research-reports/SRC-084a.pdf>). Ця робота була виконана в 1992 році. Наукові і прикладні дослідження з цієї тематики не припинялися і в наступні роки. Остання відома нам дослідницька робота — фундаментальна дисертація Тамари Мюнцнер — захищена в

Стенфордском університеті в червні 2000 року (http://graphics.stanford.edu/papers/munzner_thesis/).

Деякі роботи, виконані в цій області в середині 90-х років, досить часто цитуються в спеціальній літературі. Можна вважати, що це «класика» візуалізації. Сучасному програмісту корисно мати представлення про ці роботи — деякі ідеї можуть виявитися ефективними при розробці локальних пошукових машин і візуалізації даних у локальних мережах.

Співробітники Технологічного інституту штату Джорджія Сугата Мухерьяа і Джеймс Фоли розробили спеціальний засіб – побудовувач навігаційного огляду (The Navigational View Builder) — якій дозволяє користувачу інтерактивно створювати корисні візуалізації інформаційного простору. Ці автори використовували зв'язування, групування, фільтрацію і ієрархізацію в якості чотирьох основних стратегій побудови наочних зображень. Ці стратегії використовували комбінацію структурного і контент-аналіза відповідного простору для побудови візуалізації. Уже тоді було усвідомлено, що будь-які засоби візуалізації дають усього лише двох - чи тривимірні проекції багатомірних гіпермедиа-мереж. Але навіть зображення цих двох - чи тривимірних проекцій у зрозумілому і естетично привабливому виді – досить важка задача. Була усвідомлена також складність проблеми «многовузельності» відповідних зображень. Представлення всієї інформаційної структури на екрані може бути скрутним. Якщо розміри скорочені для приміщення на екрані, то деталі стають занадто маленькими. Перегляд складної схеми вроздріб за допомогою скролінга і простежування дуг гіперпосилань затінює глобальну структуру. Метою повинне бути розміщення як деталей, так і всього контексту гладко інтегрованими в одному екрані.

Ще одна ідея Сугата Мухерьяа і Джеймса Фоли дотепер виявилася не реалізованою: щоб бути дійсно корисної, візуалізація повинна давати представлення не тільки про структуру, але і про актуальній контент вузлів і зв'язків. Ці автори намагалися проводити групування мережних даних (кластерування) двома шляхами: структурне кластерування (пошук субмереж даної мережі); контекстного кластерування (кластерування по подібних контент-атрибутах, наприклад, по ключових словах чи по прізвищах авторів). Фільтрація також

вироблялася на базі контент-аналізу, чи по структурних ознаках, чи по характері гіперпосилань. Усі процедури, що базуються на аналізі контенту, досить важко автоматизувати. Наявні пошукові засоби нерідко ототожнюють контент зі словами на сторінці. А в цьому випадку виникає добре відома лінгвістам *проблема словника*: існує багато способів виразити те саме, використовуючи будь-яку природну мову. При аналізі тільки гіперпосилань (структурному аналізі) можна виділити в загальній безлічі сайтів ієрархічні підструктури – і звичайно це зображується за допомогою «конічних дерев». Навіть якщо вся система і не є ієрархією, виділення ієрархічних підструктур дає навігаційні і пізнавальні вигоди. Викладені принципи були застосовані для візуалізації структури веб-ресурсу Технологічного інституту Джорджії. Для більш великих структур застосування техніки цієї класичної роботи може бути не досить ефективним. Графічні ілюстрації цієї роботи зараз сприймаються як трохи наївні.

Автор іншої класичної роботи тієї епохи — Тім Брэй, старший віце-президент Open Text Corporation — ввів такі терміни як *видимість* сайта (число гіперпосилань, що ведуть до цього сайту) і *світність* сайта (число гіперпосилань з даного сайту на інші сайти). Ця робота була представлена на П'ятій міжнародній конференції по World Wide Web у травні 1996 року. Як розмір сайта він трактував число сторінок. Тім Брэй знайшов, що деякі сайти надзвичайно *видими*: десятки тисяч гіперпосилань ведуть до них. Значна більшість сайтів замічені деякими іншими Інтернет-ресурсами (мають від однієї до десяти вхідних посилань). Є сайти з нульовою видимістю, непомічені ніким. По статистиці Тіма Брэя майже 80% сайтів середини 90-х років не містили жодного зовнішнього URL. Очевидно, що тільки мала порція веб-сайтів тримає на собі все навантаження веб-навігації, відповідно до відомого правила 20/80: «Двадцять відсотків людей випивають вісімдесят відсотків пива». Старше покоління програмістів знає це правило по класичній книзі 60-х років «Фізики жартують». У той час програмістів і юзерів ще було мало, і жартували тільки фізики...

Тім Брэй придумав дуже наочне представлення, що дозволяє на одному малюнку побачити розмір сайта, його *світність* і його *видимість*. Він представив сайти як

вавілонські східчасті вежі – зіккурати. Ці зіккурати увінчані сферами. Діаметр вежі виражає число сторінок сайту, висота визначається видимістю. Розмір кулі, що плаває над вежею визначається світністю. Колір вежі визначався доменом сайта. Розподіл зіккуратів у просторі було засновано на силі зв'язків між ними. Подібні «зіккурат-шахи» можна сприймати при не дуже великому числі фігур, можливо до однієї-двох сотень. Цей надзвичайно наочний засіб не дуже зручний для візуалізації величезних ресурсів Інтернету.

Ще одна класична робота, також представлена на паризької п'ятої конференції World Wide Web, була виконана Стивеном Лэммом і його співавторами з університету Іллінойсу і Netscape Communications Corporation. Стивен Лэмм і його співавтори розробили систему візуалізації веб-даних, засновану на географічній інформації. Безумовно, багато користувачів Інтернету не мають потребу в зведеннях про фізичну локалізацію отриманої ними інформації. Можливо, це є одна з привабливих сторін Інтернету. Однак для вивчення Всесвітньої Мережі як єдиного об'єкта, а також з погляду багатьох застосувань і, насамперед, е-комерції – географічна прив'язка даних є необхідною. Багато Інтернет-ресурсів допускають географічну прив'язку по доменному імені при використанні баз даних типу «Who is». Подібні бази даних містять інформацію про доменах, хостах, мережах і інших Інтернет-адміністраторах. Ця інформація звичайно, хоча і не завжди, включає і поштову адресу. Якщо по цій поштової адресі можна установити довготу і широту місця, то відповідний сайт допускає географічну локалізацію. Зображуючи на моніторі глобус, можна уникнути відомих проблем картографічного перекручування. Сайти на цьому глобусі в роботі Лэмма і співавторів зображувалися у виді хмарочосів. Висота хмарочоса показувала кількісні характеристики сайту – наприклад, число байтів чи число запитів на інші сайти. Кольори елементів хмарочоса представляли розподіл типів документів чи інтервалів між послідовними запитами. Цей надзвичайно наочний спосіб візуалізації також працює при не дуже великому числі відображуваних елементів. Зручність цього способу візуалізації складається, зокрема, у можливості наочного представлення

часової динаміки: можна зіставити відповідне зображення глобуса в різні години, дні тижня, місяці і роки.

Заслугує згадування ще одна класична робота тієї епохи. Джеромі Кар'єр з канадської фірми Nortel і Рик Кацман з Пітсбургського інституту інжиніринга програмного забезпечення (Software Engineering Institute) розробили ряд засобів пошуку і візуалізації в Мережі, заснованих на аналізі звязності і змісту. Зокрема, вони використовували вже згадані конічні дерева.

Веб-сайт CAIDA

Сучасні засоби візуалізації, засновані на аналізі значних сегментів Інтернету, представлені на сайті Каліфорнійської організації CAIDA (<http://www.caida.org>). Повна назва цієї організації – Кооперативна Асоціація для Аналізу Даних Інтернету (Cooperative Association for Internet Data Analysis). Ця незалежна дослідницька група розташована в Суперкомп'ютерному Центрі Каліфорнійського університету Сан Дієго. Ця група вважає своєю місією створення нейтральної бази для забезпечення більшого співробітництва в розвитку і застосуванні засобів для вимірювання, аналізу і візуалізації Інтернет-даних. Відповідні засоби повинні забезпечити розробку і підтримку робастної, масштабно-інваріантної глобальної інфраструктури Інтернету. Діяльність CAIDA підтримана такими спонсорами як Національний Науковий Фонд США (NSF) і Агентство Перспективних Оборонних Дослідницьких Проектів (DARPA). Її спонсують також деякі Інтернет-провайдери і вендори устаткування. Деякі подробиці про групу CAIDA і про програму її досліджень можна знайти в статті головного дослідника CAIDA з невимовним ім'ям Кс Клэффи (Кс Claffy), що одержала свій докторський ступінь по комп'ютерних науках у 1994 році в Сан Дієго (<http://computer.org/internet/v5n1/caida.htm>). У 1999 році CAIDA почала трирічний проєкт «Атлас Інтернету». У рамках цього проєкту розробляються методологія, програмне забезпечення і протоколи для відображення Інтернету. Один із засобів візуалізації, створений групою CAIDA, називається Walrus (морж), цей засіб використовує неевклідовий гіперболічний простір для забезпечення огляду величезних спрямованих графів з мільйонами вершин. Walrus забезпечує безупинне

перетворення (уже згадане раніше «Риб'яче око»), що дозволяє користувачу розглянути деталі в малій області, зберігаючи загальний погляд на весь граф. Ці дані потім проєктуються на поверхню одичної сфери.

Інший проєкт групи CAIDA називається Skitter. Згідно Лонгмановському словнику сучасної англійської мови дієслово *skitter* означає «бігти легко і швидко». Як назву проєкту це слово може сприйматися як щось схоже на назву військового корабля: «Стрімкий». Програмне забезпечення скіттеру використовується для вивчення топології шляхів в Інтернеті. Технічно це здійснюється за допомогою посилання малих спробних пакетів ICMP (Internet Control Message Protocol) по широкому спектрі простору Інтернет-адресів. Потім акумулюється й аналізується час відгуку й інша інформація, отримана у відповідь на спробний пакет. Програма Скіттер акумулює також безліч інших даних. Наприклад, розподіл шляхів по числу стрибків до іншого сайту, розподіл часу повернення сигналу; корелює цей час повернення з географічними даними і відстанню від джерела сигналу (відстань виражена в числі стрибків, необхідних для досягнення даної адреси). Усі ці дані можна знайти на сайті групи CAIDA. Найбільш вражаюче зображення даних, що збираються Скіттером, виходить при використанні бібліотеки Nupviewer, призначеної для представлення великих графів. Ця бібліотека створена Тамарою Мюнцнер у Стенфордському університеті. Вона використовує те ж перетворення, що і програма Walrus.

Нам представляється, що викладений у статті матеріал може бути використаний не тільки при навчанні студентів молодших курсів спеціальності інформатика, але і навіть на уроках інформатики у старших класах загальноосвітньої школи.

ЛІТЕРАТУРА

1. Темненко В.А., Сейдаметова З.С. Нові елементи curricula для ІТ-фахівців: Вебологія, Вебометрія, ... // «Інформатика та комп'ютерно орієнтовані технології навчання» – збірник наук. праць Всеукраїнської науково-практичної конференції, м. Хмельницький, 16-18 травня 2001 року, с. 188 – 189.
2. Темненко В.А., Сейдаметова З.С. Вебометрія І. Глобус віртуального мира // «ComputerWorld/Київ» № 9 (315)

3. Темненко В.А., Сейдаметова З.С. Вебометрия II. Тесный мир Всемирной Паутины // «ComputerWorld/Киев» № 10 (316)
4. Темненко В.А., Сейдаметова З.С. Вебометрия III. Увидеть невидимое // «ComputerWorld/Киев» № 11 (317)
5. Темненко В.А., Сейдаметова З.С. Вебометрия IV. Состязания поисковых машин // «ComputerWorld/Киев» № 12 (318)
6. Albert, Réka; Jeong, Hawoong, Barabási, Albert-László (1999). Diameter of the World Wide Web. *Nature*, 401: 130-131.
7. Almind, Tomas C., Ingwersen, Peter (1997). "Informetric analyses on the World Wide Web: A methodological approach to "webometrics"". *Journal of Documentation*, September 1997, 53(4):404-426. <http://ix.db.dk/cis/texts/abstract3.htm>
8. Chen, C., Newman, J., Newman, R., Rada, R. (1998). "How did university departments interweave the Web: A study of connectivity and underlying factors". *Interacting With Computers*, 10(4), 353-373.
9. Chen, C., Carr, L. (1999). "Visualizing the evolution of a subject domain: A case study". *Proceedings of IEEE Visualization 99*. October 24-29, 1999. San Francisco, California, USA. <http://www.brunel.ac.uk/~cssrccc2/papers/IEEEVis99.pdf>
10. Chen, Hsinchun; Chung, Yi-Ming; Ramsey, Marshall, Yang, Christopher C. (1998). A Smart Itsy Bitsy Spider for the Web. *JASIS*, 49(7):609-618.
11. Ciolek, T. Matthew. (1997). "The Size, Content and Geography of Asian Cyberspace: An Initial Measurement". *The Journal of East Asian Libraries, CEAL*, 1997
12. Cronin, Blaise; Snyder, Herbert W.; Rosenbaum, Howard; Martinson, Anna, Callahan, Ewa (1998). Invoked on the Web. *JASIS*, 49(14): 1319-1328.
13. Cui, Lei (1999). "Rating Health Web sites using the principles of Citation Analysis: A Bibliometric Approach". *Journal of Medical Internet Research* 1999; 1 (1): e4.
14. Cunningham, Sally Jo (1998). Applications for bibliometric research in the emerging digital libraries. *Scientometrics*, 43(2): 161-175.
15. Kushal, Khan, Locatis, Craig (1998). Searching through Cyberspace: The Effects of Link Cues and Correspondence on

Information Retrieval from Hypertext on the World Wide Web. JASIS, 49(14):1248-1269.

16. Kushal, Khan, Locatis, Craig (1998). Searching Through Cyberspace: The Effects of Link Display and Link Density on Information Retrieval from Hypertext on the World Wide Web. JASIS, 49(2):176-190

17. Larson, Ray A. "Bibliometrics of the World Wide Web: An Exploratory Analysis of the Intellectual Structure of Cyberspace". <http://sherlock.berkeley.edu/asis96/asis96.html>

18. Lawrence, Steve, Giles, C. Lee. (1998). "Context and Page Analysis for Improved Web Search". IEEE Internet Computing, 2(4), July/August 1998:38-46.

<http://www.neci.nj.nec.com/homepages/lawrence/papers/search-ic98/>

19. Lawrence, Steve, Giles, C. Lee. (1998). "Searching the World Wide Web". Science, 280(5360):98.

20. Lawrence, Steve, Giles, C. Lee (1998). "Inquirus, The NECI Meta Search Engine". Proceedings of the Seventh International World Wide Web Conference, Brisbane, Australia, 1998.

21. Lawrence, Steve, Giles, C. Lee (1998). "Searching the Web". Letter to Science, 281(5374):175.

22. Lawrence, Steve, Giles, C. Lee (1999). "Accessibility of information on the Web". Nature, 400:107-109.

23. Leighton, H. Vernon, Srivastava, Jaideep (1997). "Precision among World Wide Web Search Services (Search Engines): Alta Vista, Excite, Hotbot, Infoseek Lycos". <http://www.winona.msus.edu/library/webind2/webind2.htm>

24. Moore, A., Murray, B. H. (2000). Sizing the Internet. Cyveillance. <http://www.cyveillance.com/newsroom/3012.asp>