DOI: 10.35681/1560-9189.2019.21.4.199409

УДК 004.7

В. Ю. Зубок
Інститут проблем моделювання в енергетиці ім. Г.Є. Пухова
вул. Генерала Наумова, 15, 03164 Київ, Україна
тел. (+38044) 4241063, e-mail: vitaly.zubok@gmail.com

Формальний опис об'єктів і процесів глобальної маршрутизації у мережі Інтернет для оцінки впливу кібератак на маршрутизацію

Атаки на глобальну маршрутизацію, зокрема викрадення та перехоплення маршрутів, є однією із масштабних проблем кібербезпеки. Атаки використовуються для маніпуляцій із трафіком з метою дестабілізації мережі Інтернет, шпигунства, крадіжок, нанесення матеріальної шкоди. Одним із найважливіших етапів на шляху моделювання впливу атак на маршрутизацію є побудова формальної моделі глобальної інтернет-маршрутизації. Запропоновано формальний опис об'єктів глобальної маршрутизації і відносин між ними, а також процесу вибору маршруту.

Ключові слова: *глобальна маршрутизація, формальна модель глобальної маршрутизації, перехоплення маршрутів, оцінка ризиків, кібербезпека.*

Вступ

Атаки на глобальну маршрутизацію, зокрема викрадення та перехоплення маршрутів (route hijack, route leak) є можливими тому, що протокол прикордонного шлюзу (BGP) був запроваджений до того, як був усвідомлений його вплив на кібербезпеку. 30 років тому зловживання його незахищеністю були відсутні, а 20 років тому — не були такими поширеними, як сьогодні [1]. В оригінальній специфікації BGP не було визначено заходів безпеки для запобігання навмисних чи ненавмисних помилок конфігурації мережі. Відсутність вбудованих заходів безпеки робить BGP вразливим як до викрадення префіксу, так і до витоків маршруту [2]. Викрадення префіксу — це явище, при якому Автономна система (AS) нелегітимно оголошує себе як джерело префіксу замість справжнього джерела. Витік маршруту означає, що AS нелегітимно, з порушенням політики маршрутизації, пропонує маршрути до чужих префіксів. Ці нелегітимні маршрути забруднюють таблиці маршрутизації BGP, спотворюють шляхи проходження мережевого трафіка та впливають на конфіденційність, цілісність і доступність IP-комунікацій.

© В. Ю. Зубок

Вплив цих атак на дестабілізацію мережі Інтернет настільки великий, що сценарії відключення країни від Інтернету, що здавалися нереалістичними, стають все більш імовірними. Для проведення атаки, здатної порушити стабільність роботи глобальної мережі в окремо взятій країні, потрібна тривала підготовка, але технічно це можливо. Реєстратори доменних імен — це частина критичної інфраструктури країни. Так як порушення їхньої роботи впливає на функціонування глобальної мережі, вони є об'єктом атак з боку проурядових атакуючих. Як показали минулі місяці, найбільш небезпечними стали зломи типу DNS (hijacking), внаслідок проведення яких атакуючі могли управляти DNS-записами для MITM-атак, у поєднанні з BGP (hijacking), у ході яких здійснюється перехоплення маршруту та перенаправлення трафіка окремих префіксів автономної системи через стороннє обладнання. Так, 24 квітня 2018 року така атака була застосована до інфраструктурного IP-префіксу широко відомого хмарного сервісу Amazon AWS, метою якого була фішингова атака на криптовалютний сервіс «MyEtherWallet» шляхом перенаправлення трафіка. 12 листопада 2018 року збій глобальної маршрутизації, що торкнувся сервісів G Suite, Google Пошук і Google Аналітика, стався завдяки невеликому нігерійському провайдеру за участю China Telecom та Ростелекому. Пізніше цей інцидент фахівці визнали умисними діями. В червні 2019 року такої самої атаки зазнав відомий сервіс мережевого захисту CloudFlare.

Сучасні напрямки протидії перехопленню маршрутів

На даний час мережі покладаються на практичні реактивні механізми, щоб спробувати захиститися від викрадення префіксу, оскільки запропоновані проактивні механізми (наприклад, RPKI [2]) є повністю ефективними лише при глобальному розгортанні, а оператори неохоче розгортають їх через пов'язані з цим технічні та фінансові витрати. Захист проти викрадення реактивно складається з двох етапів: виявлення та пом'якшення.

Аналіз механізмів проведення атаки залежно від її цілей і варіанти її реалізації детально описано в [3, 4]. Виявлення в основному забезпечується сторонніми службами, такими як BGPMon, які сповіщають адміністратора мережі про підозрілі події, пов'язані з їхніми префіксами, на основі інформації про маршрутизацію, відслідковуючи реальні маршрути шляхом трасування та спостерігаючи за оновленнями анонсів маршрутів у BGP. У разі інциденту, постраждалі мережі приступають до пом'якшення наслідків події, наприклад, анонсуючи більш специфічні префікси своїх мереж або звертаючись до інших AS для фільтрації хибних анонсів.

Однак, через поєднання технологічних і практичних проблем розгортання, існуючі реактивні підходи значною мірою недостатні. Зокрема, у найсучасніших технологіях виникають такі основні проблеми:

— різноманітність типів атак на маршрутизацію, комбінування методів призводять до того, що немає надійної методики виявлення перехоплення маршруту;

— операторам необхідно завчасно інформувати про легітимні зміни в своїй політиці маршрутизації (утворення нових взаємодій між AS, анонсування нового префіксу тощо) аби такі зміни не вважатися підозрілими подіями для систем виявлення атак в умовної третьої сторони. З іншого боку, прийняття менш суворої політики щодо компенсації відсутності оновленої інформації та скорочення кіль-

кості false positives, несе небезпеку знехтування реальними подіями і невиявленням інцидентів (false negatives);

— кілька хвилин перенаправлення трафіка можуть спричинити великі фінансові втрати через недоступність послуги або порушення безпеки. А швидкість реагування на інциденти є повільною у будь-якому разі, бо за існуючою практикою є необхідність вручну перевірити попередження, що надходять від систем моніторингу та сторонніх сервісів. Тривалість широко відомих інцидентів становила від декількох годин до місяців [3].

У процесі оцінювання ризику особливі вимоги висуваються до якості інформації (максимально можливий рівень повноти, точності і відповідності на момент її отримання) та якості її джерел [4, 5]. Результат ідентифікації ризику повинен бути структурованим та охоплювати чотири елементи: джерела виникнення ризику, безпосередні події у результаті реалізації загроз, причини цих подій та очікувані наслідки. Побудова формальної моделі інтернет-маршрутизації сприятиме рішенню задачі прогнозування наслідків подій.

Модель мережі Інтернет у вигляді графа

Телекомунікаційна мережа є невід'ємною частиною інформаційно-телекомунікаційних систем і характеризується різними за формою зв'язками, а також різними видами взаємодії. Часто математичною моделлю таких мереж може служити граф. Граф можна уявити як набір точок, званих вершинами чи вузлами, з'єднаних собою лініями, які називаються дугами чи зв'язками. Кожному зв'язку і вузлу графа може відповідати певна кількість параметрів, що характеризують природні обмеження. Наприклад, мережа може бути представлена у вигляді графа, в якому дуги відповідають каналам зв'язку, а вершини — вузлам комутації. Важливі параметри включені в модель у вигляді чисел або ваг, приписаних до дуг і вершин графа. Ці ваги можуть бути фіксованими та випадковими. Так, для мережі типова вершина, що представляє вузол комутації, може мати такі ваги: максимальну пропускну здатність, обсяг запам'ятовуючого пристрою, і т.ін. Типова дуга — лінія зв'язку — може мати такі ваги: максимальну пропускну здатність, середню затримку передачі по каналу зв'язку, надійність каналу зв'язку тощо.

Доцільність побудови моделі у вигляді графа залежить від фізичної природи досліджуваної мережі. Найбільш очевидна доцільність користування моделлю у вигляді графа при вирішенні задач зв'язності. Так, у загальному випадку, нас може цікавити завдання доставки інформації з будь-якої точки в будь-яку. Це структурна задача, в якій необхідно встановити чи існує принаймні один шлях з будь-якої вершини в будь-яку іншу. Іншим важливим завданням є пошук найкоротшого шляху між двома, кількома, або всіма вершинами графа, а також пошук найкоротшого шляху з урахуванням різних обмежень. За допомогою графів можна також вирішувати завдання синтезу топології мережі, тобто побудови оптимальної системи. Одним із можливих критеріїв оптимальності можуть бути ступінь живучості або надійності телекомунікаційної системи. Оскільки телекомунікаційній мережі властиві пошкодження і відмови (наслідком чого є порушення зв'язку), можливим завданням може бути побудова системи, в якій наслідки порушень роботи є мінімальними при заданих умовах роботи.

Математичний апарат теорії графів, а пізніше — теорії складних мереж, у застосуванні до топології глобальної телекомунікаційної мережі Інтернет дозволяє аналізувати ступінь вузлів, розподіл ступеню, шлях між парою вузлів, середній шлях у мережі, показники кластерності, посередництва тощо. На даний час у багатьох дослідженнях граф використовують для побудови моделі мережі Інтернет на рівні автономних систем [7–9]. Мережа Інтернет у них представляється графом $G := (V, E)$, де V є множиною автономних систем (AS), а E — їхні зв'язки, утворені протоколом маршрутизації BGP-4. При цьому пропускна спроможність зв'язків між вузлами Інтернет не приймається до уваги і не впливає на вагу ребер, таким чином, граф є незваженим. Було досліджено стосунки (node relationships) між автономними системами в Інтернеті та виділено декілька типів таких стосунків, зокрема «клієнт-провайдер», «партнер-партнер» та інші. У випадку взаємодії двох рівних за статусом операторів, вони анонсують один іншому власні префікси та префікси мереж своїх клієнтів. У разі взаємодії провайдера та клієнта, провайдер анонсує клієнтові всі наявні в нього префікси, а клієнт анонсує префікси власних мереж. Таким чином, зв'язки між автономними системами представляються двосторонніми, отже ребра є ненаправленими, а граф — неорієнтований.

За допомогою математичного апарату теорії графів було запропоновано метричну функцію відстані в Інтернеті для такого графа і доведено її відповідність аксіомам метрики [9].

Об'єкти глобальної маршрутизації і формальний опис відносин між ними

Отже, для виявлення атак з перехоплення маршрутів, дослідження масштабів впливу на топологію, а також подальшої оцінки ризиків необхідно мати модель мережі Інтернет на рівні глобальної маршрутизації, тобто — з використанням BGP-зв'язків. Першою відмінністю від наведених у попередньому розділі підходів є те, що сам процес маршрутизації невід'ємно пов'язаний з вибором напрямку, отже при дослідженні втручання в маршрутизацію, коли результатом є несанкціонована зміна напрямку, ми не зможемо використовувати як модель незважений граф. Для визначення необхідних якостей нової моделі пропонується формалізувати поняття маршрутизації.

Сформулюємо вихідні дані.

Існує адресний простір мережі Інтернет A — множина унікальних IP-адрес a, які згруповані в IP-префікси p (надалі — просто «префікси»):

$$A = \{a_1, a_2, a_3, ...a_{|A|} : a_i \neq a_j, \{i, j\} \leq |A|\},$$

$$a \in p \subset A.$$

Префікси, в свою чергу, групуються (інколи вживають термін «агрегуються») з більш специфічних у менш специфічні, як це визначено у [8] і наведено в таблиці. Повний перелік префіксів наведено в документації по CIDR. З ілюстрації зрозуміло, що будь-яка IP-адреса входить до 32 префіксів, які «інкапсулюються», тобто входять один в одного за допомогою зміни мережевої маски. При цьому весь адресний простір A може бути описаний одним префіксом з довжиною мере-

жевої маски 0, або об'єднанням двох префіксів з довжиною маски 1, або чотирьох з довжиною маски 2, і так далі:

$$p_3 \subset p_2 \subset p_1 \subset p_0;\ |p_0| = 2|p_1| = 4|p_2| = 8|p_3|.$$

Агрегація IPv4-префіксів згідно RFC 4632

Нотація префіксу	Довжина netmask, біт	Кількість адрес у префіксі	Загальна кількість префіксів IPv4
x.x.x.x/32	32	1	4294967296
x.x.x.x/31	31	2	2147483648
x.x.x.x/30	30	4	1073741824
x.x.x.x/29	29	8	536870912
.....			
x.x.x.0/24	24	256	16777216
x.x.x.0/23	23	512	8388608
.....			
x.x.0.0/17	17	32768	131072
x.x.0.0/16	16	65536	65536
x.x.0.0/15	15	131072	32768
.....			
x.0.0.0/9	9	33554432	512
x.0.0.0/8	8	16777216	256
x.0.0.0/7	7	8388608	128
.....			
x.0.0.0/1	1	2147483648	2
0.0.0.0/0	0	4294967296	1

У загальному випадку можемо так виразити відношення між підмножинами IP-адрес, визначених певними префіксами, в множині всіх IP-адрес:

$$\begin{cases} p_i = 2^{j-i} p_j, \\ i \le j, \\ 0 \le \{i, j\} \le \log_2 |A|. \end{cases} \quad (1)$$

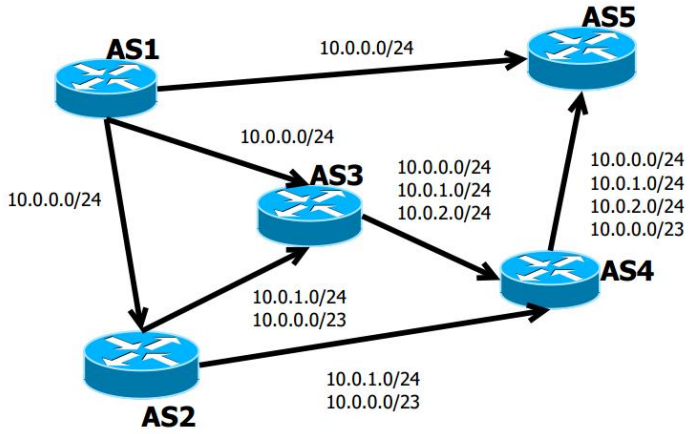
Префікси анонсуються автономними системами (в іноземній літературі використовується термін «originating», отже в кожного префіксу є свій «origin» — автономна система, що анонсує його. Приклад взаємодії AS наведено на рисунку.

У нормальному стані кожен префікс анонсує лише одна автономна система.

До кожного префіксу p існує принаймні один маршрут $m(p)$, який складається з множини направлених ребер e_p, що відповідають напрямкам анонсів префіксу між вершинами множини v_p (автономними системами):

$$m(p) := \left(v_p, e_p\right), \quad (2)$$

До префіксу, який не анонсується, маршрутів нема, і він не може вважатись учасником Інтернету.



Приклад експорту анонсів IP-префіксів у процесі взаємодії автономних систем

Сукупність маршрутів m_p до певного префіксу p може бути представлена у формі зв'язного орієнтованого графа без кілець:

$$M_p := (V_p, E_p), \tag{3}$$

де V_p — зв'язки між автономними системами, по яких надходять анонси префіксу p; E_p — автономні системи, в яких присутні маршрути до префіксу p. Вершини графа по вхідних ребрах приймають анонси префіксу і по вихідних — ретранслюють їх до інших вершин (чи не ретранслюють, залежно від політики маршрутизації чи стану каналу). Одна із вершин має тільки вихідні ребра, це — origin. Інші вершини обов'язково мають вхідні ребра (бо приймають анонси) і можуть мати вихідні (якщо ретранслюють анонси).

Поєднання (ансамбль) усіх графів $M_p := (V_p, E_p)$ — це сукупність усіх маршрутів до всіх префіксів. Вона утворює такий граф

$$G := (V, E) : V = \bigcup_p V_p \; ; \; E = \bigcup_p E_p ,$$

який і можна інтерпретувати як мережу Інтернет, представлену на рівні глобальної маршрутизації.

Поглибимось у процес маршрутизації. Якщо префікс p_j є підмножиною префіксу p_i, тобто $p_j \subset p_i$, це не означає, що в них обов'язково однаковий origin, у загальному випадку префікси мають origin незалежно від вкладеності. Приклад: частина великого провайдерського префіксу може санкціоновано анонсуватись одним із клієнтів провайдера в напрямку інших AS. З іншого боку, якщо в певній автономній системі відсутні анонси до префіксу p_j, це не означає, що через неї цей префікс недосяжний: наявність в якомусь вузлу анонсу до p_i означає також і можливість маршрутизації до p_j (це одностороннє твердження):

$$p_j \subset p_i \;\Rightarrow m(p_j) \subset m(p_i) . \tag{4}$$

Кардинальним прикладом цього є кінцевий мережевий пристрій, підключений до мережі своїм єдиним мережевим інтерфейсом. Його таблиця маршрутизації може містити виключно маршрут до загального префіксу 0.0.0.0/0 (див. рисунок), який також зветься маршрутом за замовчанням (default route). Слід зауважити, що префікс 0.0.0.0/0 анонсується в BGP виключно для учасників, які використовують обладнання, що не здатне обробити «full view» — повну таблицю маршрутизації (зазвичай це кінцеві споживачі — невеликі мережі, автономні системи яких не є транзитерами).

Формальний опис процесу вибору маршруту

Задача знаходження оптимального маршруту є складним завданням. Для вирішення задачі має бути відомою топологія мережі, пропускні спроможності ліній зв'язку, середня довжина повідомлення. Це задача з класу цілочисельного нелінійного програмування, для яких доки не існує алгоритмів навіть з поліноміальною складністю. Відомі лише евристичні алгоритми, які дозволяють отримувати лише наближене рішення завдання оптимізації. Тому в стеку TCP/IP прийнятий так званий однокроковий підхід до оптимізації маршруту просування пакета (next-hop routing) — кожний маршрутизатор і кінцевий вузол приймають участь у виборі тільки одного кроку передачі пакета. Однокроковий підхід означає розподілене обчислення задачі вибору маршруту, і це — перевага, яка лежить в основі умовно безкінечної масштабованості мережі Інтернет.

Першим етапом обрання напрямку доставки є вибір префіксу. В таблиці маршрутизації має бути обраний з урахуванням (1) найбільш специфічний префікс:

$$p(a) = \{\min_j(p_j) \ : \ a \in p \subset A, \ 0 < j \le |A|\}. \qquad (5)$$

Як це викладено в попередньому параграфі, вибір маршруту властивий не всім учасникам мережі. Суб'єктом вибору маршруту в глобальній маршрутизації є вузол графа, тобто автономна система. Передумовою для початку процесу вибору є наявність більше ніж одного маршруту до одного і того самого префіксу p. Один із доступних маршрутів може бути визначений як шлях між двома вузлами графа (2), де початковим вузлом є автономна система, в якій приймається рішення, а кінцевим вузлом — автономна система, яка є origin для префіксу p. Якщо не враховувати специфічні локальні атрибути маршруту та ті, що встановлюються адміністративно, єдине, що приймається до уваги, це топологія мережі на момент прийняття рішення. Загальним критерієм обрання шляху є його довжина (best path) — кількість транзитних вузлів між початковим і кінцевим вузлами. З урахуванням (2) та (3), шляхом до префіксу буде такий маршрут $\pi_v(p)$:

$$\pi_v(p) = \{\min_v(m_v(p)) : \pi \in M_p, v \in V_p\}, \qquad (6)$$

де v — вихідний вузол, в якому приймається рішення.

Отже, виділено та надано формальний опис двом складовим процесу маршрутизації: розрахунку префіксу (5) і вибору шляху (6).

Відносини між об'єктами глобальної маршрутизації і різні типи IP-адрес

На сьогоднішній день в Інтернеті використовується два типи IP-адрес, які відрізняються розрядністю. Традиційна IP-адреса складалася з 32 біт. Це наклало обмеження на кількість можливих адрес у $2^{32} = 4294967296$. Зростання Інтернету, попри впровадження економного розподілу адресного простору, призвело до нестачі адрес. Сучасна IP-адреса, яка має умовну назву «IPv6-адреса» (на відміну від традиційної, яку стали називати «IPv4-адреса») має довжину 128 біт. Ця різниця впливає на функціонування канального та мережевого рівнів (згідно моделі OSI). Однак принципи CIDR та маршрутизації в цілому не зазнали змін при впровадженні IPv6.

З точки зору CIDR, відмінність від IPv4 полягає в наступному:

— максимальна довжина мережевої маски складає 128 біт замість 32;

— кожна адреса входить до 128 IPv6-префіксів, включених один в одного.

Протокол BGP-4 для IPv4 та IPv6 не відрізняється — BGP-спікери використовують одні й ті самі протокольні повідомлення, атрибути шляху, критерії вибору маршруту як для адресного простору IPv4, так і для адресного простору IPv6. Тому опис моделі глобальної маршрутизації, запропонований в попередньому розділі, зокрема вирази (1), (4)–(6), є незмінним незалежно від типу IP-префіксів.

Висновки

Важливим кроком на шляху до оцінки ризику, що спричинений атаками на глобальні маршрутизацію, є прогнозування наслідків атаки, а саме — оцінка масштабу атаки (шляхи розповсюдження, зона впливу, кількість «пошкоджених» маршрутів). Запропоновано формальні описи об'єктів глобальної маршрутизації — автономних систем, IP-префіксів, BGP-анонсів і відносин між ними. Виділено та надано формальний опис двом складовим процесу маршрутизації: розрахунку IP-префіксу і вибору шляху. Ці етапи є важливим поступом до моделювання кібератак на глобальну маршрутизацію з метою формулювання задачі поводження з ризиками для певного вузла від перехоплення маршрутів як задачі пошуку для нього найбільш ефективної топології зв'язків.

1. Pavlos Sermpezis, Vasileios Kotronis, Alberto Dainotti, and Xenofontas Dimitropoulos. A survey among network operators on BGP prefx hijacking. *ACM SIGCOMM Computer Communication Review*. 2018. **48**(1). C. 64–69.

2. Andreas Reuter, Randy Bush, Italo Cunha, Ethan Katz-Bassett, Thomas C Schmidt, and Matthias Wählisch. Towards a rigorous methodology for measuring adoption of RPKI route validation and fltering. *ACM SIGCOMM Computer Communication Review*. 2018. **48**(1). P. 19–27.

3. Зубок В. Ретроспективний аналіз інцидентів кібербезпеки, пов'язаних з атаками на глобальну маршрутизацію. *Моделювання та інформаційні технології*: зб. наук. праць. 2019. Вип. 86. С. 42–49.

4. Зубок В. Метричний підхід до оцінки ризику атак на глобальну маршрутизацію. «Информационные технологии и безопасность». Материалы XVIII Междунар. науч.-практ. конф. ИТБ–2018. Київ: ООО «Инжиниринг», 2018. С. 43–47.

5. Зубок В., Мохор В. Дослідження зв'язку між топологією та ризиком внаслідок кібератак на глобальну маршрутизацію. *Моделювання та інформаційні технології*: зб. наук. праць. 2018. Вип. 85. С. 23–26.

6. Risk Management — Vocabulary (ISO Guide 73:2009, IDT): ДСТУ ISO Guide 73:2013. [Чинний від 2014–07–01]. Київ: Мінекономрозвитку України, 2014. 13 с. (Національні стандарти України).

7. IPv4 and IPv6 AS Core: Visualizing IPv4 and IPv6 Internet Topology at a Macroscopic Scale. URL: http://www.caida.org/research/topology/as_core_network/ (Last accessed: 20.10.2019).

8. Pavlos Sermpezis, Vasileios Kotronis et al. ARTEMIS: Neutralizing BGP Hijacking within a Minute. URL: https://arxiv.org/abs/1801.01085 (Last accessed: 11.06.2019).

9. Мохор В., Зубок В. Формування міжвузлових зв'язків в Інтернет з використанням методів теорії складних мереж. Київ: «Прометей», 2017. 175 с.

10. Fuller V., Li T. Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan. URL: https://tools.ietf.org/html/rfc4632 (Last accessed: 11.10.2019).

Надійшла до редакції 20.11.2019