УДК 004.052:004.738.2 (045)

Мінаєв Ю.М., д.т.н.,
Толстікова О.В.

ЛОКАЛЬНЕ СТРУКТУРУВАННЯ ТРАФІКУ КОМП'ЮТЕРНИХ МЕРЕЖ ТЕНЗОРНИМИ МОДЕЛЯМИ

Національний авіаційний університет

В основу виявлення аномальних станів (атак) на підставі тензорної моделі моніторингу параметрів трафіку покладено принцип структурування трафіку, представленого у вигляді тензора парних рангів. Це дозволило створити узагальнену цифрову модель трафіку з урахуванням всіх параметри, що характеризують трафік. При цьому відпадає необхідність враховувати семантику конкретних параметрів трафіку.

Вступ

При дослідженні процесів та об'єктів комп'ютерних систем і мереж, зокрема, мереж радіодатчиків [1] представлення об'єкту дослідження (виміру) у вигляді тензора є більш адекватним, ніж представлення у вигляді величини. Тензорна модель, яка розглядається як матрична проекція, дозволяє аналізувати об'єкт в різних системах координат, тобто експертні оцінки об'єкту розглядаються як один і той же об'єкт в різних системах координат; відповідність між цими системами координат, що характеризуються тензорами, може бути встановлена за допомогою «тензора приєднання», який дозволяє узгодити і зв'язати різні точки зору. Не виключено, що багатоманітність точок зору - це багатоманітність систем координат, якщо привести до однієї системи ці різнопланові, на перший погляд точки зору, вони виявляться адекватними [2].

Постановка проблеми

Повнота застосування властивостей тензорів залежить від типу простору, що представляє дану предметну область. Застосування простору-структури мереж дозволяє представити структуру та процеси систем з різних предметних областей єдиним методом за допомогою мережних моделей.

Властивості тензора, що залишаються незмінними при перетвореннях координат, визначаються системою його інваріантів, що є коефіцієнтами характеристичного рівняння. Інваріанти - константи, значення яких зберігаються при зміні системи координат. Можливість заміни аналізу трафіку, представленого тензором, аналізом його інваріантів, відкриває нові шляхи у визначенні аномальних станів комп'ютерних систем і мереж.

Шляхи вирішення проблеми

Виявлення аномальних станів трафіку комп'ютерних мереж (КМ) на підставі парадигми структурування трафіку (локально- структурований трафік) КМ тензорними моделями та ідентифікація аномальних станів на підставі інваріантів тензорних моделей трафіку є одним з найефективнішим шляхом вирішення цієї проблеми.

Невизначеність поняття «аномальний стан» змушує вводити певні уточнення і обмеження на характер і склад задачі, що розглядаються. Зокрема, у аномальні дії обмежені розглядом наступних ситуацій:

- непомітний вплив, відсутність відповіді;
- шкідливий вплив, фатальна реакція;
- вплив відсутній, реагують всі;
- нестандартний вплив, ідентифікуюча відповідь.

Намагання уникнути нечітких визначень призводить до того, що з нечіткими поняттями виконуються дії з використанням стандартного математичного апарату, який є абсолютно неадекватним визначенням. Це, природно, пливає на достовірність отриманих висновків.

Аналіз трафіку базується на вмісті пакетів, розгляд ведеться для стандартного та прихованого запитів, дослідження пакетів ведеться на підставі і з допомогою *TCPdump*. Досліджується:

- *IP* заголовок, зокрема, визначається місце, де він закінчується;

- інші поля з вказівкою довжини, зокрема довжина *IP*- дейтаграми, довжина заголовка *TCP*-сегмента, збільшення фіксованої довжини або досліджується весь пакет.

Зазначимо в цьому зв'язку наступне. Власне трафік КС або КМ вважається як семантичне поняття цілком дослідженим, хоча автори, використовують достатньо різний набір компонентів трафіку. При цьому зовсім не визначаються причини, з яких такий набір визначається, дуже часто ці причини є техніко-економічними і природно не дають відповіді на питання наскільки повний (необхідний та достатній) набір параметрів з точки зору ефективності вияв-лення аномального стану. Відповідно до такого підходу дослідження полів *IP*- заголовка орієнто-ване на виявлення атак зі вставкою та прихованих атак.

Аналіз трафіку широко використовує поняття «нестандартного» (незвичного) трафіку, під яким розуміють загальну картину того, що відбулося, власне трафік, незвичайне сканування, хости-відправники та хости-приймачі, розташування конкретних хостів, значення полів *TTL*, розмір вікна, параметри *TCP*, повторні запити і т.ін.

Очікуваний трафік. Намагання продемонструвати всі можливі варіанти нормального трафіку не є можливими. Можна представити різні стандартні ситуації і зразки трафіку, котрий передається частіше за все. Звернемо увагу на зразки. Особливу увагу слід звернути відповідній реакції хостів і маршрутизаторів при отриманні різної інформації при різних обставинах її отримання і при різних протоколах. Пояснити, що означає „нормальний трафік" неможливо, неможливо розглянути всі безкінечні варіанти нормального трафіку. Напевне, можна вважати, що найкращою характеристикою нормальності слід визнати відсутність нормальності, що передбачає розгляд можливо найбільшої кількості прикладів трафіку, який відрізняється від норми.

Проблема виявлення аномалій в роботі КС, в т.ч. на підставі аналізу трафіку, відноситься до проблем, що важко формалізуються, розв'язання яких сучасна наука бачить у використанні інтелектуальних технологій.

Інтелектуальні технології (системи) представляють собою такі нформаційні технології, у яких передбачені:

- наявність баз знань, що відбивають досвід конкретних людей, груп, суспільств, людства в цілому, у рішенні творчих задач у виділених сферах діяльності, що традиційно вважалися прерогативою інтелекту людини (наприклад, такі недостатньо формалізовані задачі, як прийняття рішень, проектування, витягнення змісту, пояснення, навчання і т.ін.);

- наявність моделей мислення на основі баз знань: правил і логічних висновків; аргументації і міркування; розпізнавання і класифікації ситуацій; узагальнення і розуміння і т.ін.;

- здатність формувати цілком чіткі рішення на основі нечітких, несуворих, неповних, недовизначених даних;

- здатність пояснювати висновки і рішення, тобто наявність механізму пояснень;

- здатність до навчання, перенавчанню і, отже, до розвитку.

Вони ґрунтуються на ідеології штучного інтелекту.

Параметри трафіку у загальному вигляді визначені: $x = \{x_i\}, i = 1,9$:

x_1 – *Protocol ID* протокол, зв'язаний з подією (*TCP=0, UDP=1, ICMP=2, unknown=3);*

x_2 – номер порту джерела;

x_3 – номер порту хоста призначення;

x_4 – *IP*-адреси джерела;

x_5 – *IP*-адреси приймача;

x_6 – *ICMP Type* тип *ICMP*-пакет (*Echo Request or Null);*

x_7 – *ICMP Code* кодове поле з *ICMP*-пакета (*None or Null*);

x_8 – *Raw Data Length* довжина даних у пакеті;

x_9 – *Raw Data* порція даних у пакеті.

Представлення трафіку КМ тензором, матриця котрого може бути сформована на підставі різних принципів, дозволяє, по-перше, агрегувати *всі* потрібні для дослідження параметри трафіку, по-друге, дозволяє уніфікувати аналіз, використовуючи для цієї мети властивості матричного аналізу. Поняття тензора має декілька визначень: класичне, математичне, сучасне, прикладне [3, 4]. Згідно до одного з цих визначень, тензор представляє собою частину таблиці, в нашому випадку – частину системного журналу, де ведеться запис трафіку. Структуризація трафіку, параметри якого представленого у вигляді таблиці БД (системний журнал), може йти у два способи (рис.1):

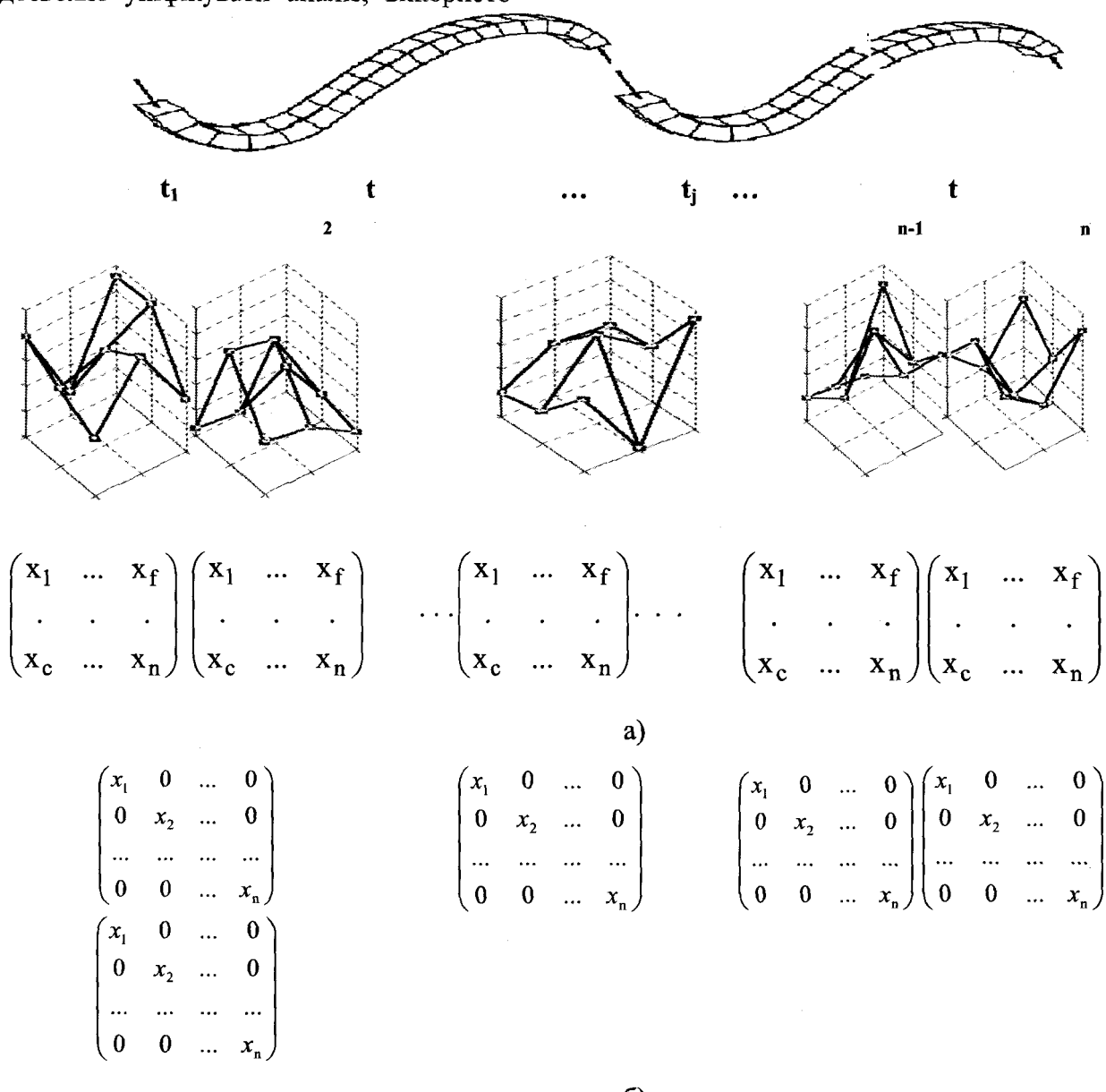


Рис. 1. Тензор-трафік КС: $t_1, t_2, ..., t_p$ - проміжки часу, коли визначаються параметри трафіку; $t_j \rightarrow [x_1,...,x_n]^{(j)}$ - параметри трафіку визначені у часі t_j; форми структурування трафіку: а) тензор парних рангів з матрицею m×m=n; б) діагональна матриця, $\mathrm{diag}(x) = \{x_1, x_2, ..., x_n\}$.

- якщо кількість параметрів трафіку може бути представлена як Зп, п=2, 4, тензор-трафік можна представити у вигляді тензора парних рангів;

- якщо кількість параметрів не відповідає умові Зп, п=1, 2, 3, 4, то тензор-трафік представлено у вигляді діагональної матриці, 6= {хі, х2, ..., хп}, хієсі - і-тий параметр трафіку [1,4, 5].

Для уніфікованого аналізу потрібно визначити функції від тензора (відповідно від матриць), що представляють тензор, який, у свою чергу, моделює трафік (рис. 1).

Нагадаємо, що найпростішими функціями від тензорів є їхні розкладення на симетричний-асиметричний, ізотропний девіатор, інші функції від тензорів можуть бути представлені як функції від інваріантів. У багатьох випадках для визначення аномальних станів КС доцільно використовувати функції від матриць, наприклад, трафік КС за час {И ї2 ... Щ} можна представити сукупністю відповідних тензор-трафіків {х(П) х(ї2)... х(їп)} [2, 5].

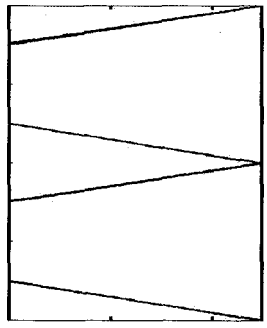
В основу виявлення аномальних станів (атак) на підставі тензорної моделі моніторингу параметрів трафіку покладено принцип структурування трафіку, представленого у вигляді тензора парних рангів [4, 5], що дозволило створити узагальнену цифрову модель трафіку, яка враховує всі параметри, що характеризують трафік, без необхідності враховувати семантику конкретних параметрів трафіку. На параметри трафіку накладено обмеження у вигляді функцій належності, яка враховує можливу роль окремого параметра у визначенні стану комп'ютерної системи, всі параметри приймаються відносними, тобто без врахування їхньої вимірності.

Створення даного тензора (рис. 2) реалізовано шляхом тензорного добутку рядка (вектора) безвимірних (приведених) параметрів трафіку -*X={x$* на колонку (вектор) функцій належності:

M={ц},]=1,2,...,9;

T=(X={x,})-(M=

де • - знак тензорного добутку, *T=[u^,ї,]=1,9; T* - символ транспонування, у зв'язку з тим, що трафік вміщує 32 параметрів, отримуємо тензор 4-го ранга.



TRAFIC T = x®(1)x , x=[xi, x2, x3, x4, x5, хб, x7, x8, x9],

0) x=[1 000 1000 I]1

1	XI	x2	x3	x4	x5	хб	x7	x8	x9
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
1	xi	x2	x3	x4	x5	хб	x7	x8	x9
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
1	xi	x2	x3	x4	x5	хб	x7	x8	x9

а) б) в) г)

Рис. 2. Тензорна модель трафіку для окремого моменту часу:
а) спосіб утворення тензорного добутку; б), в) характеристична функція; г) тензор-трафік

Зазначимо, що у вигляді тензорів може бути представлена окрема компонента трафіку *Xi є X* яка розглядається у часі. Трафік КМ наведено на рис.3.

Структурування трафіку КМ у вигляді сукупності тензорів його наведено на рис.4.

Інший шлях структурування трафіку полягає в наступному. Припустимо що параметри трафіку *X={xi}* i=1,__,9; обчи-

слені у вигляді підмножин $X(н)=\{x[к)\}$ $к=1\ K;\quad X\{')=\ =\{x[8)\}\quad 5=7\quad Г\ m)=\{x[т)\}$ $m-1\quad M;\ Г\ n)a\ X\ (x)\ aX<m)\ c\ X\quad M<\ Б\ <K.$ Для кожної підмножини збудуємо бінарне дерево і відповідну 2-адичну матрицю:

$X\ н>\ ^>B(к)$

$X^*)\qquad -Ж\ ®\ 2$

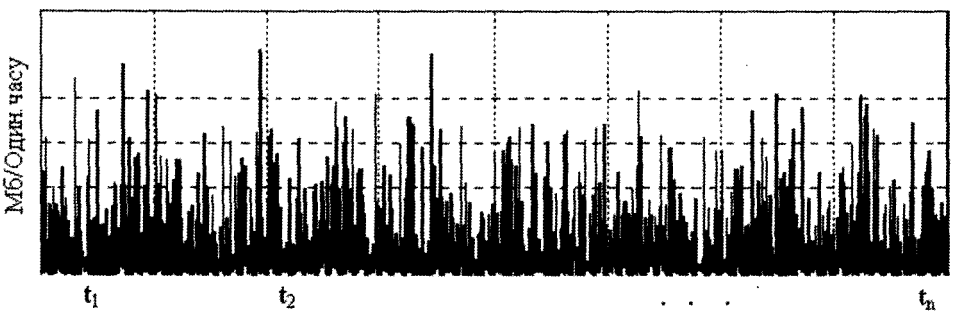
$X\ m)-ю\ (m)\ ->м<m\geq 2.$

Отримуємо множину матриць (рис. 5), наприклад, для кожної хвилини, години, доби, місяця і ін.

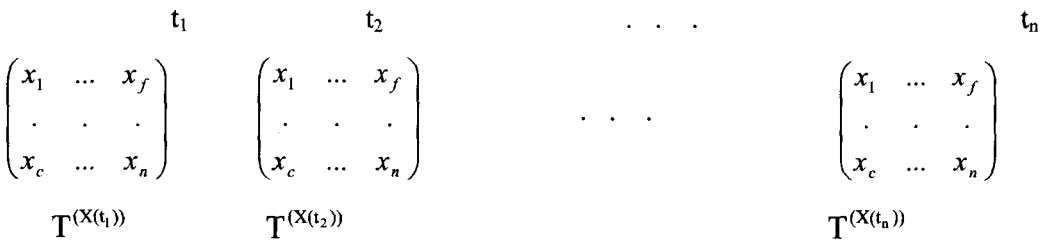
TRAFIC -p

\ \	*x2*	**x3**	**x4**	**x5**	**хб**	**x7**	**x8**	**x9** ^
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
xi	**x2**	**x3**	**x4**	**x5**	**хб**	**x7**	**x8**	**x9**
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
чк1	**x2**	**x3**	**x4**	**x5**	**хб**	**x7**	**x8**	**x9 y**

Рис. 3. Трафік КМ



а)



б)

Рис.4. Структурований трафік: а) залежність X={x(i0[])} i=1 *щ;* їо- реальний час; б) структурований трафік X(l)= {x(Id0>}<0 i=1 т; т=пхп ї= імп[/т I- тензорний час (ї= пт i0)

$$X® \ >т\ хб)—= \begin{matrix} Ч & \bullet.. & xf \end{matrix}$$

.. де

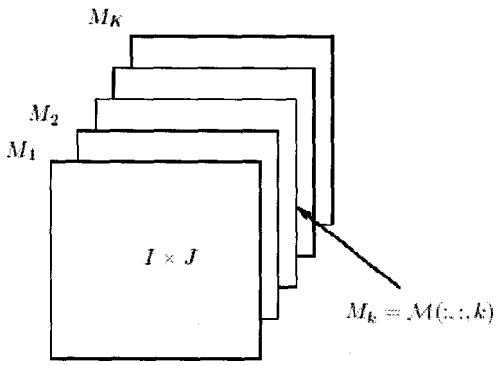


Рис. 5. Множина K матриць $M_k \in R^{I \times J}$ представлених як $I \times J \times K$ тензор M

Множини матриць K матриц $M_k \in R^{I \times J}$ можна розглядати як $I \times J \times K$ тензор M.

Висновки

В основу виявлення аномальних станів (атак) на підставі тензорної моделі моніторингу параметрів трафіку покладено принцип структурування трафіку, представленого у вигляді тензора парних рангів, що дозволило створити узагальнену цифрову модель трафіку, яка враховує всі параметри, що характеризують трафік, без необхідності враховувати семантику конкретних параметрів трафіку. На параметри трафіку накладено обмеження у вигляді функцій належності, яка враховує можливу роль окремого параметра у визначенні стану комп'ютерної системи, всі параметри приймаються відносними, тобто без врахування їхньої вимірності. Створення даного тензора реалізовано шляхом тензорного добутку рядка (вектора) безвимірних параметрів трафіку – $X=\{x_j\}$ на колонку (вектор) функцій належності M, $T=(X=\{x_j\}) \bullet (M= =\{\mu_j\}^T)$, $T=[t_{ij}]$.

Список літератури

1. Минаев Ю.Н., Толстикова Е.В. Метод эффективной передачи данных при наличии аномалий в сетях радиодатчиков // Проблеми інформатизації та управління. – К.: НАУ, 2011. – Вип. 2(34). – С. 91–95.

2. Мінаєв Ю.М., Філімонова О.Ю. Тензорна нейроматематика // Матеріали X-Міжнародної науково-технічної конференції "Автоматика-2003", Севастополь, 15-18 вересня 2003 р. – С. 117-125.

3. Минаев Ю.Н., Филимонова О.Ю. Тензорный базис в концепции нечеткости и формальных методах // Материалы 10-ой международной конференции по автоматическому управлению. г. Севастополь, 15-19 сентября 2003 г. – С. 153-156.

4. Минаев Ю.Н., Филимонова О.Ю. Тензорный базис как основа новых алгоритмов решения задач в условиях неопределенности // Материалы VI Всероссийской научно-технической конференции «Новые информационные технологии», Москва, 23-24 апреля 2003 г. Сб. трудов. Том 1. – С. 142-147.

5. Минаев Ю.Н., Филимонова О.Ю. Тензорный базис как основа мягких вычислений в условиях неопределенности // Материалы V Международной научно-технической конференции "ABIA-"2003", 23-25 квітня 2003. – С.*155-159.*