УДК 004.621.3

Єжова Л.Ф., Скачек Л.М.

ЗАСТОСУВАННЯ ТЕОРІЇ ГРАФІВ ЩОДО ОПИСУ ІНФОРМАЦІЙНИХ ПРОЦЕСІВ

В статті розглядаються інформаційні процеси з метою визначення можливості їх математичної формалізації за допомогою теорії графів та обробки інформації в паралельних обчислювальних системах на об'єкті.

Ключові слова: інформаційний процес, паралельна обробка, теорія графів.

Широкомасштабне використання обчислювальної техніки і телекомунікаційних систем, перехід до безпаперової інформаційної технології, збільшення обсягів оброблюваної інформації приводять до ускладнення інформаційних процесів, якісно нових можливостей несанкціонованого доступу до ресурсів і даних інформаційної системи, до їх високої уразливості.

Подальша розробка теоретичних та методичних основ формування механізму захисту інформації на будь-якому об'єкті неможлива без залучення нових знань стосовно структури та протікання інформаційних процесів, без застосування нових підходів і методів їх моделювання.

Будь-який інформаційний процес на об'єкті визначається зміною параметрів, які його задають, або приведенням одних параметрів (вихідних — $\Phi_1,\Phi_2,...,\Phi_m$) у відповідність з іншими (вхідними - $x_1,x_2,...,x_n$) за законом:

$$\begin{cases} \Phi_1 = \varphi_1(x_1,...,x_n), \\ \Phi_2 = \varphi_2(x_1,...,x_n), \\ \quad\vdots \\ \Phi_m = \varphi_m(x_1,...,x_n), \end{cases} \tag{1}$$

де $(x_1,x_2,...,x_n)\in D\subseteq R^n$.

Нехай $\varphi_i\in C^1(D),\ i=\overline{1,m}$. Припустимо, що значення одного з вихідних параметрів Φ_j однозначно визначається сукупністю значень інших $\Phi_1,...,\Phi_{j-1},\Phi_{j+1}...,\Phi_m$, тобто, якщо $\Omega_0\subseteq R^{m-1}$ є множина точок, що відповідають точкам $(x_1,x_2,...,x_n)\in D$, то в Ω_0 буде мати місце функціональна залежність:

$$\Phi_j = f(\Phi_1,...,\Phi_{j-1},\Phi_{j+1},...,\Phi_m), \tag{2}$$

де $f\in C^1(\Omega),\ \Omega\subseteq R^{m-1},\ \Omega\supseteq\Omega_0$, Ω - відкрита множина, а при підстановці (1) в (2) виходить тотожність відносно $(x_1,x_2,...,x_n)\in D$:

$$\varphi_j(x_1,...,x_n)\equiv f(\varphi_1(x_1,...,x_n),...,\varphi_{j-1}(x_1,...,x_n),\varphi_{j+1}(x_1,...,x_n),...,\varphi_m(x_1,...,x_n)). \tag{3}$$

У цьому випадку будемо казати, що функція φ_j залежить від функцій $\varphi_1,...,\varphi_{j-1},\varphi_{j+1}...,\varphi_m$ в області D. У загальному випадку функції $\varphi_1,\varphi_2,...,\varphi_m$ називаються залежними в області D, якщо одна з них залежить від інших. У протилежному випадку функції $\varphi_1,\varphi_2,...,\varphi_m$ називаються незалежними в D. Відповідно вихідні параметри інформаційного процесу на об'єкті будуть незалежними (залежними), якщо незалежними (залежними) в області D будуть функції (1), що їх визначають.

У випадку незалежності вихідних параметрів процеси їх визначення відповідно до (1) можуть проводитися одночасно (паралельно), що значно скорочує час аналізу та реалізації інформаційного процесу на об'єкті. Цей процес можна представити як сукупність не зв'язаних між собою «простих» процесів, результатом кожного з яких є отримання лише

одного параметра Φ_i, а дослідження поданого інформаційного процесу на об'єкті зведеться до дослідження скінченної сукупності «простих».

Для визначення залежності (незалежності) вихідних параметрів функцій Φ_j $(j=1,...,m)$ можна скористатися властивостями матриці частинних похідних функцій φ_i, $i=\overline{1,m}$, — матриці Якобі неперервного інформаційного процесу на об'єкті:

$$\begin{pmatrix} \frac{\partial\varphi_1}{\partial x_1} & \frac{\partial\varphi_1}{\partial x_2} & \dots & \frac{\partial\varphi_1}{\partial x_n} \\ \frac{\partial\varphi_2}{\partial x_1} & \frac{\partial\varphi_2}{\partial x_2} & \dots & \frac{\partial\varphi_2}{\partial x_n} \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ \frac{\partial\varphi_m}{\partial x_1} & \frac{\partial\varphi_m}{\partial x_2} & \dots & \frac{\partial\varphi_m}{\partial x_n} \end{pmatrix} \quad (4)$$

Нехай $n \geq m$. Якщо хоч один визначник m-го порядку, складений з елементів матриці (4), відмінний від нуля в області D, то в цій області функції φ_i, $i=\overline{1,m}$ є незалежними, а тому і вихідні параметри інформаційного процесу на об'єкті, незалежні.

Необхідно відзначити, що інформаційний процес на об'єкті, зокрема процес функціонування системи захисту інформації на всіх інформаційних об'єктах, включає у множину вхідних параметрів не тільки параметри об'єктів системи, але й параметри зовнішніх збурень інформаційної системи та параметри керуючої підсистеми інформаційної системи.

Під керованою підсистемою розуміємо сукупність декількох функціональних систем, об'єднаних єдністю вибраної мети функціонування інформаційної системи, під керуючою — сукупність засобів, що забезпечує досягнення керованою підсистемою поставленої мети.

В якості прикладу розглянемо структуру інформаційних потоків у підсистемі захисту інформації (рис. 1)

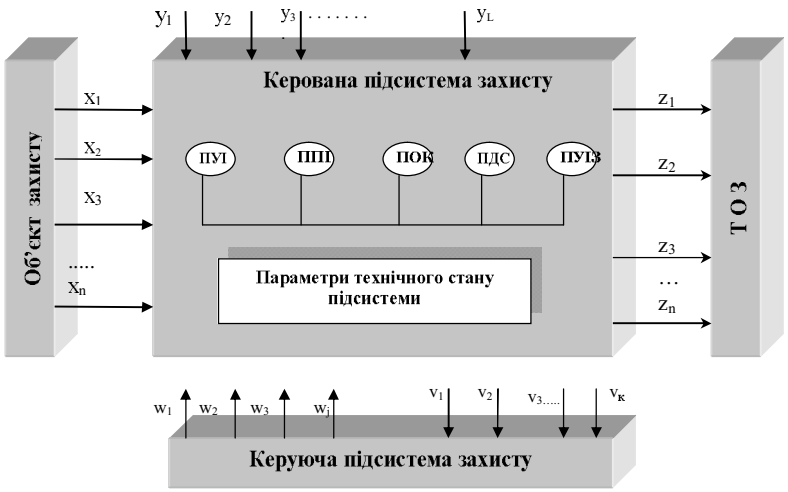

Рис. 1. Структура інформаційного процесу у підсистемі захисту інформації

Керовану підсистему будемо характеризувати наступними групами змінних:

$X\{x_1,.....x_n\}$ - множина вхідних сигналів підсистеми, обумовлених фізичними полями та випромінюваннями, властивими конкретному ОЗ;

$Z\{z_1.....z_m\}$ - множина вихідних сигналів підсистеми, що є залишковими випромінюваннями різної фізичної природи;

$Y\{y_1,.....y_i\}$ - множина невизначених зовнішніх обурень, супутніх функціонуванню підсистеми;

$V\{v_1,.....v_k\}$ - множина спостережених параметрів (ознак), що характеризують технічний стан підсистеми;

$W\{w_1,.....w_j\}$ - множина дій, що керують функціональними структурами підсистеми.

Типовий варіант узагальненої структури підсистеми захисту включає декілька функціональних підсистем, основними з яких є:

ПУІ - підсистема утаєння інформації ;

ППІ - підсистема створення помилкової інформації;

ПОК - підсистема оперативного контролю виробничого процесу;

ПДС - підсистема діагности стану;

ПУІЗ - підсистема управління, передачі інформації та зв'язку.

Довільному інформаційному процесу на об'єкті поставимо у відповідність орієнтований граф (орграф) наступним чином. Зіставимо k-й вершині графа отримання величини u_k. Перші n вершин будуть точками введення початкових даних $u_1,...,u_n$ і називатися вхідними, а інші вершини — обчислення u_k як значень функцій F_k - вихідними. Будемо вважати, що дуга йде з i-ї вершини в j-у тоді й тільки тоді, коли при обчисленні величини u_j величина u_i використовується як аргумент.

При такій побудові отримаємо дводольний орграф, оскільки множина його вершин V була розбита на дві підмножини V_1 і V_2 так, що кожне ребро графа є впорядкованою парою виду $< v_1, v_2 >$, де $v_1 \in V_1$, а $v_2 \in V_2$.

Крім того, зі способу побудови графа випливає, що граф інформаційного процесу на об'єкті є ациклічним.

У графі, як моделі інформаційного процесу на об'єкті, відображено, які окремі перетворення вхідних даних пов'язані між собою інформаційно, які з них можуть виконуватися одночасно, які потрібно виконувати пізніше або раніше, чим інші. Граф інформаційного процесу на об'єкті може бути використаний для аналізу інформаційного процесу об'єкта в цілому та його структури.

Отриманому графу на об'єкті ставиться у відповідність матриця Ф розміром $m \times N$ з елементами φ_{ij}:

$$\varphi_{ij} = \begin{cases} -1, & \text{якщо } j = i+n \\ 1, & \text{якщо } j \text{ є одним із чисел } (i+n)_1,...,(i+n)_{S_{i+n}}. \\ 0, & \text{в інших випадках} \end{cases} \qquad (5)$$

Очевидно, k-ий стовпець матриці Ф відповідає параметру u_k, а k-ий рядок - параметру u_{k+n}. В k-ому рядку елемент -1 знаходиться в тому стовпці, номер якого відповідає номеру параметра, що обчислюється - u_{k+n}. Елементи +1 знаходяться у тих стовпцях, номери яких відповідають номерам аргументів параметра, що обчислюється - u_{k+n}. Матриця Ф описує зв'язок параметрів u_k між собою і називається матрицею інформаційної зв'язності інформаційного процесу на об'єкті.

Якщо ненульові елементи матриці змінюють свої значення, то по такій оновленій матриці можна однозначно відновити граф інформаційного процесу на об'єкті, а тому може використовуватися для аналізу цього процесу.

Для орієнтованого графа, що відповідає інформаційному процесу на об'єкті, стандартно визначається $N \times N$-матриця суміжності B з елементами b_{ij}

$$b_{ij} = \begin{cases} 1, & \text{якщо з } i\text{-ї вершини в } j\text{-у йде ребро} \\ 0, & \text{в інших випадках} \end{cases}, \quad (6)$$

і матриця інцидентності A з елементами a_{ij}

$$a_{ij} = \begin{cases} 1, & \text{якщо } j\text{-е ребро виходить з } i\text{-ї вершини} \\ -1, & \text{якщо } j\text{-е ребро входить в } i\text{-ту вершину.} \\ 0, & \text{в інших випадках} \end{cases} \quad (7)$$

Матриця B тісно пов'язана з матрицею Ф інформаційної зв'язності інформаційного процесу на об'єкті: Ф - підматриця, яка складається з останніх m рядків матриці $B^T - I$, де I - одинична матриця відповідного розміру. У зв'язку з цим для приведення матриці інформаційної зв'язності інформаційного процесу на об'єкті до більш зручного виду можна перенумерувати рядки і стовпці та здійснити відповідну перенумерацію для матриці суміжності.

Виявлення й наступне використання внутрішнього паралелізму на основі відповідного графа приведе до значного зменшення витрат часу, необхідних для дослідження властивостей інформаційного процесу на об'єкті, зокрема, для вивчення паралельних форм інформаційного процесу.

Поява паралельних обчислювальних систем і впровадження їх у практику рішення великих прикладних задач привела до необхідності аналізу інформаційних процесів з метою визначення можливості їх математичної формалізації та обробки в паралельних обчислювальних системах на об'єкті. Результатом аналізу є виявлення таких частин процесу, які інформаційно між собою не пов'язані, а це свідчить, що інформаційному процесу на об'єкті властивий внутрішній паралелізм.

Будемо використовувати граф $G = (V, E)$ інформаційного процесу на об'єкті, де V - множина вершин, а E - множина впорядкованих пар вершин (ребер) для аналізу його структури, не накладаючи ніяких обмежень на вид вхідних і вихідних параметрів. Граф інформаційного процесу на об'єкті не накладає, взагалі кажучи, ніяких обмежень і на порядок виконання операцій, що входять до складу процесу, крім одного: до моменту початку реалізації будь-якої операції повинні закінчити своє виконання всі ті операції, які поставляють для неї параметри-аргументи. Таким чином граф інформаційного процесу на об'єкті визначає множину припустимих порядків виконання його операцій.

Будь-який інформаційний процес на об'єкті - це процес, що протікає в часі, будь-яка його реалізація породжує певне сортування вхідних до його складу операцій. Це сортування будує розбивку множини операцій (вершин відповідного графа) на такі групи, які виконуються послідовно, а операції всередині групи можуть виконуватися одночасно.

Ототожнюючи інформаційний процес на об'єкті з його графом, будемо припускати, що в графі відображені операції отримання всіх параметрів і зв'язки, вплив яких на реалізацію процесу підлягає вивченню.

Нехай $G = (V, E)$ - довільний орієнтований ациклічний граф з n вершинами. Тоді існує таке натуральне число $s \le n$, що всі вершини графа можна так помітити одним з індексів $1,2,...,s$, що якщо ребро йде з вершини з індексом i у вершину з індексом j, то $i < j$.

Така розмітка вершин називається топологічним сортуванням графа або паралельною формою. Очевидно, що ніякі дві вершини з однаковим індексом не є суміжними. Крім того, для будь-якого натурального $s \le n$, більшого довжини критичного шляху, існує топологічне сортування, при якому використовуються всі s індексів, тобто граф має не єдине топологічне сортування.

Якщо співвідношення $i < j$ замінити на $i \le j$, то отримаємо узагальнене топологічне сортування графа інформаційного процесу.

Результатом топологічного сортування є виявлення можливостей паралельного виконання чи аналізу операцій, що входять в інформаційний процес на об'єкті. Сукупність

всіх топологічних сортувань графа інформаційного процесу на об'єкті визначає його паралельні форми реалізації (обробки, аналізу). Операції, відповідні до вершин графа однієї паралельної форми, є інформаційно незалежними, а тому можуть виконуватися паралельно. Групи операцій, що відповідають різним топологічним рівням, виконуються послідовно в порядку зростання номерів вершин графа, що входять у них.

Зрозуміло, чим складніше граф, чим більший його розмір, тим важче побудувати його топологічне сортування. Для прискорення цього процесу можна розбити граф на підграфи меншого розміру з наступною побудовою топологічних сортувань підграфів і відновленням сортування всього графа по сортуваннях підграфів. За допомогою перенумерації вершин графа інформаційного процесу на об'єкті можна спростити його опис. Наведемо приклади простих і паралельних процесів обробки інформації (рис. 2).

Однією з характеристик протікання інформаційного процесу на об'єкті є час його виконання. При безпосередній реалізації визначені моменти виконання всіх операцій. Порушення певного часу виконання деякої операції в ході процесу є сигналом можливих збоїв, атак, спрямованих на засоби захисту, що забезпечують протікання процесу, і т.д.

Перенумеруємо вершини графа інформаційного процесу на об'єкті довільно й кожній i-й вершині поставимо в співвідношення час t_i закінчення

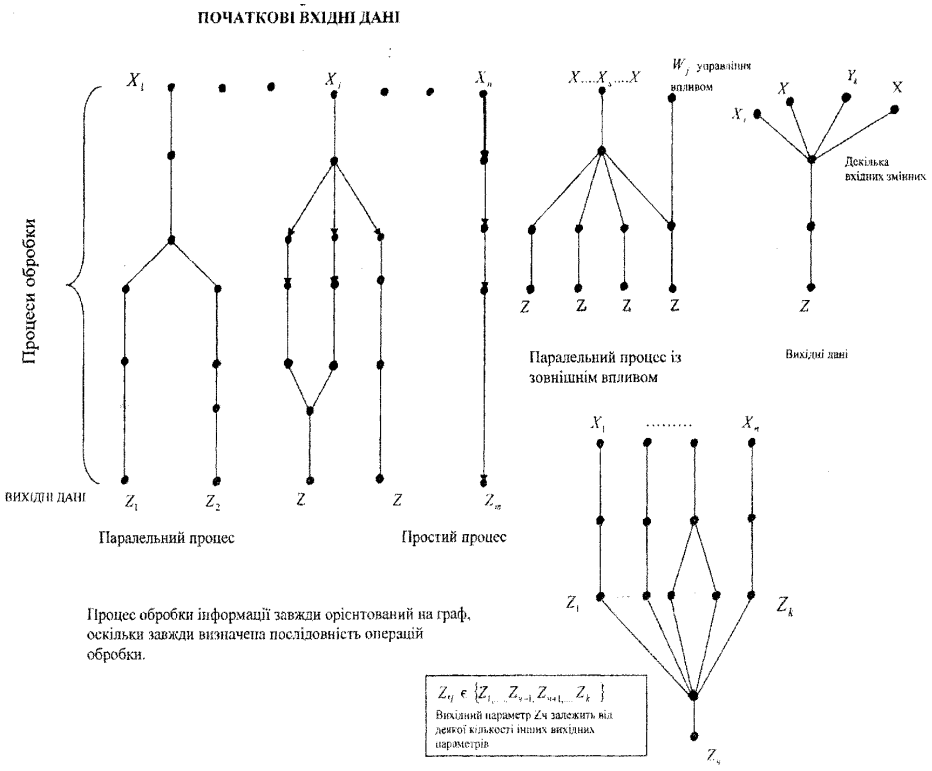

Рис. 2. Приклади простих і паралельних процесів обробки інформації

виконання відповідної операції. Таким чином, з інформаційним процесом можна зв'язати вектор часового розгортання процесу $t=(t_1,t_2,...,t_n)^T$, який показує перебіг процесу на об'єкті у часі. Для елементів вектору t можна визначити деякі обмеження, наприклад, задати час h_j для реалізації j-ї операції. Але, якщо у графі інформаційного процесу існує ребро, що йде з i-ї вершини в j-у, то повинне виконуватися співвідношення:

$$t_j - t_i \geq h_j,$$

Оскільки час, що проходить від закінчення i-ї операції до закінчення j-ї, містить у собі не тільки час виконання j-ї операції, але й час, який витрачається на передачу інформації,

необхідної для виконання j-ї операції. Невід'ємний вектор $h=(h_1,h_2,...,h_n)^T$ називають вектором реалізації інформаційного процесу на об'єкті.

З кожним ребром графа зв'язують не тільки ту інформацію, що передається від однієї вершини до іншої. Будь-яке часове розгортання однозначно визначає час появи цієї інформації t_i й час $t_j - t_i$ її існування, частково в незмінному, частково в перетвореному в j-й вершині вигляді. Таким чином в момент t_j стара інформація повністю закінчує своє існування й народжується нова. Час $t_j - t_i$ є часом затримки інформації на дузі, яка пов'язує i-у та j-у вершини. При реалізації інформаційних процесів на об'єкті на часи затримок накладаються обмеження знизу. Вони спричиняються часом передачі інформації по каналах і лініях зв'язку, часом зберігання інформації й іншими факторами на об'єкті. Ці обмеження задаються апріорі й вважається, що час затримки інформації на дузі, що зв'язує i-у і j-у вершини, не менше певного невід'ємного числа w_{ij}:

$$t_j - t_i \geq w_{ij}.$$

Вектор w з координатами w_{ij} називають вектором затримок.

Нехай інформаційний процес на об'єкті починає свою реалізацію в нульовий момент часу, і в кожний додатний цілочисловий момент виконується хоча б одна операція. Розглянемо відповідне часове розгортання. Згідно з розгортанням множина вершин графа розбивається на неперетинні підмножини, де в одну підмножину входять ті й тільки ті вершини, які відповідають операціям, що виконуються одночасно. Кожній з побудованих підмножин приписується індекс, що дорівнює моменту виконання відповідних операцій. Очевидно, що отримана розбивка вершин графа інформаційного процесу на об'єкті визначає деяку паралельну форму його протікання. Отже, очевидно, що аналіз часових розгортань є перспективним напрямком в галузі досліджень інформаційних процесів на об'єкті.

Таким чином з'ясовано, що граф інформаційного процесу на об'єкті є ациклічним, у графі, як моделі інформаційного процесу на об'єкті, наочно представлені відомості про те, як окремі перетворення при виконанні процесу пов'язані між собою інформаційно, які перетворення в ході їх моделювання можуть виконуватися одночасно, які потрібно виконувати пізніше або раніше, чим інші. Граф інформаційного процесу на об'єкті описує всю картину поширення інформації, а тому може бути використаний для аналізу інформаційного процесу об'єкта в цілому, його структури.

Графи можна використовувати не тільки для описування простих і складних потоків інформації та інформаційних процесів, але й для їх аналізу, оцінювання тощо.

ЛІТЕРАТУРА

1. Скачек Л.Н. Оптимизация системы защиты информации в условиях предприятия / Л.Н.Скачек, В.А.Хорошко // Вісник Київського Національного університету ім.Т.Шевченко. Військово-спеціальні науки, вып.27, 2010. – С.141 –148.

2. Кривицкая Н.Я. Математическое моделирование социально-экономических процессов в Украине. Часть 1. / Н.Я Кривицкая, Л.Н.Скачек, А.В.Титов, В.А Хорошко // Вісник ДУІКТ, Т.8, №4, 2010. – С.363–375.

3. Кривицкая Н.Я. Математическое моделирование социально-экономических процессов в Украине. Часть 2. / Н.Я Кривицкая, Л.Н.Скачек, А.В.Титов, В.А Хорошко // Вісник ДУІКТ, Т.9 (2), 2011.– С.170 – 180.

4. Управління інформаційною безпекою. В 2-х томах / Єжова Л.Ф., Мачалін І.О., Невойт Я.В., Хорошко В.О. – Том 1, К.: Вид ДУІКТ, 2010 - с.350.

5. Управління інформаційною безпекою. В 2-х томах / Єжова Л.Ф., Мачалін І.О., Невойт Я.В., Хорошко В.О. – Том 2, К.: Вид ДУІКТ, 2010 - с.201.

Надійшла: 01.02.2012

Рецензент: д.т.н., проф. Скрипник Л.В.