УДК 681.3.06

Л. П. Мартиненко, студ.;
О. М. Новіков д. т. н., проф.;
А. М. Родіонов , асп.

СИНТЕЗ ОПТИМАЛЬНОЇ СТРУКТУРИ СИСТЕМИ ЗАХИСТУ ІНФОРМАЦІЇ З ВИКОРИСТАННЯМ ЛОГІКО-ІМОВІРНІСНОГО ПІДХОДУ

Розглянуто проблему синтезу системи захисту інформації інформаційно-комунікаційної системи. Розроблено модель системи захисту інформації, і виходячи з логіко-імовірнісного підходу, запропоновано функцію ймовірності атаки на об'єкти (ресурси) інформаційно-комунікаційної системи. На основі цієї функції побудовано цільову функцію для розв'язання задачі синтезу оптимальної структури системи захисту інформації, та алгоритм послідовного аналізу варіантів для її розв'язання. Наведено приклад визначення оптимальної структури системи захисту інформації.

Вступ

У зв'язку з широким використання корпоративних мереж, аналіз та синтез їх системи захисту є актуальною задачею. Особливістю корпоративних мереж є те, що це структурно-складні об'єкти, які мають розподілену структуру з різними рівнями вимог з безпеки, містять багато сервісів та служб, що взаємодіють між собою, утворюючи інформаційні потоки, надаючи користувачам при цьому необхідну функціональність.

Існує багато практичних рекомендацій щодо розробки структури (топології) корпоративних мереж (від Cisco, Microsoft та ін.), у яких вказується, де саме і як слід розміщувати відповідні сервіси мережі (Веб-сервер, сервер електронної пошти, міжмережевий екран тощо) та механізми захисту проти атак. Але не пропонується формалізації цього процесу та методики, за допомогою якої можна було б оцінити стан захищеності корпоративної мережі, та підібрати оптимальну структуру системи захисту.

Серед підходів та моделей, які оцінюють захищеність корпоративної мережі та її ресурсів, поширеними є дерево атак (attack tree) [1, 2] та більш формалізована модель — граф атак (attack graph) [3—5]. В цих моделях розглядаються можливі атаки на систему, що експлуатують відомі вразливості кожного із сервісів мережі. Згідно з наведеними підходами, для пошуку можливих шляхів проникнення до системи, будується відповідний граф атак, який враховує топологію мережі. Недоліком цих підходів є необхідність виконання перебору всіх можливих шляхів атак. Також цей підхід може використовуватися для аналізу наявних мереж, але не дозволяє синтезувати систему захисту.

У роботі [6], відповідно до моделі повного перекриття, система захисту інформації подана за допомогою логіко-імовірнісного підходу. Ґрунтуючись на цій моделі можна розв'язувати задачі аналізу та синтезу системи захисту інформації, але при цьому модель не враховує топологію інформаційно-комунікаційної мережі та можливі шляхи атак.

Враховуючи наведене вище, актуальною є розробка моделей для розв'язання задач синтезу та аналізу системи захисту інформації складних інформаційно-комунікаційних систем.

Метою роботи є запропонувати модель та підхід для синтезу оптимальної системи захисту інформації, який би враховував структуру інформаційно-комунікаційної системи, шляхи атак та можливі ймовірності самих атак.

Модель системи захисту інформації

Структура корпоративної мережі являє собою набір сервісів, склад яких та взаємодія між собою визначаються відповідно до функціональних потреб організації та прийнятої політики безпеки. З кожним сервісом пов'язана певна множина загроз, які можуть діяти на різних рівнях, від приклад-

ного до мережевого. Відповідно до моделі повного перекриття для кожної загрози має існувати механізм захисту. Механізми захисту — це протидія атакам.

Таким чином, для кожної атаки на сервіси мають бути визначені відповідні механізми захисту. Повний набір механізмів захисту й буде утворювати систему захисту інформації.

Введемо такі поняття:

Ймовірність спроби атаки — визначає те, з якою ймовірністю зловмисник спробує здійснити цю атаку у порівнянні з іншими атаками. Наприклад, сканування портів буде здійснено атакуючим з ймовірністю близькою до 1. Ймовірність спроби підбору стандартних паролів теж є високою.

Ймовірність реалізації атаки — визначає з якою ймовірністю буде реалізована спроба відповідної атаки. Ймовірність реалізації атаки буде залежати від механізмів захисту. Якщо механізми захисту будуть відсутні, то ймовірність реалізації даної атаки буде дорівнювати 1.

Виходячи з цього, ймовірність атаки на відповідний сервіс мережі визначимо як добуток ймовірність спроби атаки на ймовірність реалізації атаки (з урахуванням механізмів захисту). Або у формалізованому вигляді:

$$P_i = E_i\left(1 - M_i K_i\right) \tag{1}$$

де, $E_i \in [0,1]$ — ймовірність спроби атаки; $M_i \in \{0,1\}$ — факт наявності або відсутності механізму захисту; $K_i \in [0,1]$ — міцність механізму захисту.

Міцність механізму захисту означає з якою ймовірністю механізм захисту зможе протидіяти даній атаці.

Таким чином, для кожного сервісу, що працює у інформаційно-комунікаційної мережі, можемо формально зобразити ймовірність атаки на нього.

Далі, запропонуємо модель та цільову функцію, що визначає інтегральну оцінку захищеності компонентів інформаційно-комунікаційної мережі, з урахуванням топології мережі, можливих шляхів атаки, ймовірностями атак та розташуванням механізмів захисту.

Функція ймовірності захищеності об'єктів інформаційно-комунікаційної системи

У якості цільової функцією, яка б враховувала можливі сценарії атак та топологію мережі, використаємо функцію ймовірності атаки, яка отримана за допомогою логіко-імовірнісного підходу [7] для опису структурно складних систем.

Відповідно до даного підходу, імовірність переходу об'єкта корпоративної мережі до небезпечного стану буде формулюватись таким чином [8]:

$$P\left\{y(Z_1,\ldots,Z_m)=1\right\} = P\left\{\bigvee_{l=1}^{d}\left[\bigwedge_{i\in K_{\phi l}} Z_i\right]=1\right\}, \tag{2}$$

де, Z_i — події, які необхідні виконати зловмиснику для здійснення атаки,

$Z_i = \begin{cases}1, & i\text{-та подія відбулась;}\\ 0, & i\text{-та подія не відбулась;}\end{cases}$ K_{ϕ_l} — послідовність дій зловмисника у корпоративній комп'ютерній системі, що приводить, до небезпечного стану визначеного об'єкта системи, яка відповідає l-му сценарію атаки; d — скінченний набір сценаріїв атак.

Маючи імовірність здійснення кожної з подій Z_i, можна обчислити

$$P\left\{y(Z_1,\ldots,Z_m)=1\right\}, \text{ де } P\left\{Z_i=1\right\} = P_i.$$

P_i визначають ймовірності, того, що відбудуться події захоплення відповідних сервісів мережі. Враховуючи ймовірності атак та наявність механізмів захисту, $P_i\left(E_i, M_i K_i\right) = E_i\left(1 - M_i K_i\right)$.

Таким чином, функції ймовірності захищеності об'єктів інформаційно-комунікаційної системи будуть залежати від топології мережі, сценаріїв атак та механізмів захисту

$$P\left\{y\left(Z_1,\ldots,Z_m\right)=1\right\} \to P\left(P_1,\ldots,P_N\right) \to P\left(E_1,..,E_N; M_1K_1,..,M_NK_N\right). \tag{3}$$

Використаємо отриману функцію, у якості цільової для синтезу оптимальної системи захисту

інформації.

Синтез оптимальної структури системи захисту інформації з обмеженням на її вартість

Задачу синтезу структури системи захисту інформації сформулюємо як визначення оптимального розміщенні механізмів захисту у інформаційно-комунікаційній системі, при відомих сценаріях атак на відповідні компоненти системи. Сценарії атак при цьому будуть враховувати топологію мережі.

Для вирішення задачі побудови оптимальної структури комплексної системи захисту інформації з урахуванням обмежень на її вартість, у якості цільової функції використаємо функцію ймовірності захищеності системи, яка у якості змінних враховує ймовірності атак та міцність (наявність) механізмів захисту:

$$P(E; MK), \tag{4}$$

де E — ймовірність спроби атаки; M — факт наявності або відсутності механізму захисту; K — міцність механізму захисту.

Змінна $M_i \in \{0,1\}$ буде визначати наявність чи відсутність відповідного механізму захисту у системи. Таким чином значення змінних $M_1,\ldots,M_N$ будуть визначати структуру розташування механізмів.

Оскільки метою оптимізації є пошук рішення, прийнятного по сумарній вартості засобів захисту, мінімізація цільової функції повинна виконуватися з урахуванням обмеження на вартість реалізованих засобів захисту

$$C(M) = \sum_{i=1}^{N} M_i \cdot C_i \le C_o, \tag{5}$$

де $C(M)$ — вартість реалізації множини механізмів захисту $M = \{M_i\}, i \in \{1,...,N\}$;
N — кількість атак та механізмів захисту; M_i — параметр, значення якого (0 чи 1) визначає факт наявності/відсутності механізму захисту від i-ї атаки; C_i — вартість реалізації механізму захисту від i-ї атаки; C_o — обмеження на сумарну вартість засобів захисту.

З урахуванням наведеного вище, постановка задачі нелінійного цілочислового програмування, яку необхідно вирішити для побудови оптимальної системи захисту інформації, має вигляд

$$\begin{cases} P(M) \to \min; \\ C(M) \le C_o; \\ M = \{M_i\}, \quad M_i \in \{0,1\}. \end{cases} \tag{6}$$

Алгоритм послідовного аналізу варіантів для розв'язання задачі

Приведена вище задача, відноситься до класу задач нелінійної дискретної оптимізації.

Одним із методів її розв'язання є перебір всіх можливих варіантів, які задовольняють обмеженням.

Самим простим буде алгоритм при якому послідовно аналізуються всі комбінації значень змінних $M = (M_1,\ldots,M_N)$, тобто:

(1, 0, 0, …, 0); (0, 1, 0, …, 0); … ; (1, 1, 0, …, 0); … ; (1, 1, 1, …, 1),

де 1 — якщо механізм M_i присутній, 0 — у протилежному випадку.

Складність даного алгоритму є 2^N, де N — кількість атак та механізмів захисту.

Для зменшення кількості обчислень під час розв'язання даної задачі розглянемо алгоритм послідовного аналізу варіантів.

Метод послідовного аналізу варіантів оснований на конструюванні розв'язку крок за кроком (послідовне уточнення значень компонентів розв'язку) та відкиданні у процесі такого конструювання тих розв'язків, які не можуть бути добудовані до оптимальних.

Для загального випадку задачі дискретної оптимізації схема методу послідовного аналізу варіантів буде така [9]:

з обмеженнями

$$\min f(x_1,\ldots,x_n); \tag{7}$$

$$\begin{aligned} & g_i\left(x_1,\ldots,x_n\right) \le 0, \;\; i = \overline{1,m}; \\ & x_j \in Q_j, \;\; j = \overline{1,n}, \end{aligned} \tag{8}$$

де $Q_j = \{q_{1j},..,q_{\pi_j j}\}$ — задані кінцеві множини.

Вектор $x = \left(x_1,\ldots,x_n\right)$ назвемо розв'язком, якщо його компоненти $x_j \in Q_j, \;\; j = \overline{1,n}$. Множину всіх розв'язків позначимо Ω. Розв'язок є допустимим якщо він задовольняє нерівності $g_i\left(x_1,\ldots,x_n\right) \le 0$. Множину всіх допустимих розв'язків позначимо Ω_f. Вектор $x_{(p)} = \left(x_1,\ldots,x_p\right)$, $p < n$, будемо називати частковим розв'язком, якщо $x_j \in Q_j, \;\; j = \overline{1,p}$. Якщо при цьому він може бути добудований до допустимого розв'язку $(x_1,\ldots,x_p,x_{p+1},\ldots,x_n)$, то будемо його називати допустимим частковим розв'язком.

Для того щоб задати алгоритм розв'язання конкретної задачі, необхідно вказати правило μ вибору часткових розв'язків, які будуть розвиватись крок за кроком, тобто стратегію послідовного конструювання розв'язків, та побудувати набір σ елімінувальних тестів ξ_i, що виконують відсів часткових розв'язків, які не можуть бути добудовані ні до допустимих розв'язків, ні до оптимальних. Елімінувальні тести створюються виходячи зі специфіки задачі.

Припускаємо, що до набору σ входять тести ξ_0 — аналіз допустимості рішень та ξ_1 — порівняння допустимих рішень по значенню цільової функції. Застосовуючи до довільної множини рішень h (що містить, як повні так і часткові розв'язки) тести ξ_0 та ξ_1, виявляємо в h повні рішення та відкидаємо за допомогою ξ_0 недопустимі, а за допомогою ξ_1 — ті з допустимих, для яких в h знаходяться кращі за значенням цільової функції. Але якщо застосовуємо тільки тести $\sigma = \{\xi_0, \xi_1\}$, то отримуємо алгоритм повного перебору, наведений на початку.

Застосуємо цей підхід для розв'язання системи

$$\begin{cases} P(M) \to \min; \\ C(M) \le C_o; \\ M = \{M_i\}, \;\; M_i \in \{0,1\}. \end{cases} \tag{8}$$

Цільова функція $P\left(E_1,..,E_N;M_1K_1,..,M_NK_N\right)$, і обмеження $g = \sum_{i=1..N} C\left(M\right) - C \le 0$ є монотонним по всім змінним.

Розв'язування задачі складається з наступних етапів:

A. Знаходження початкового значення рекорду

B. Сортування змінних для більш ефективного пошуку

C. Побудова дерева пошуку

Кожен з етапів детальніше.

A. Початкове значення рекорду P^* — оцінка зверху для мінімуму цільової функції. Для його визначення знаходимо деякий допустимий розв'язок $M^* = \left[M_1^*,\ldots M_N^*\right], M_i \in \{0,1\}$ і значення цільової функції (ймовірності небезпечного стану) на ньому приймаємо за значення рекорду: $P^* = P\left(M_1^*,\ldots M_N^*\right)$. M^* будуємо таким чином:

Крок 0. $M^* = \left(0,\ldots,0\right)$. $i = 1$

Крок 1. Якщо $i = N + 1$, то завершуємо пошук. Причому, отриманий розв'язок є оптимальним.

Інакше $M_i^* = 1$.

Крок 2. Якщо $g\left(M_1\ldots M_N\right) \le 0$ (тобто отриманий розв'язок є допустимим),

То $i = i + 1$ і повертаємося на Крок 1.

Інакше, $M_i^* = 0$. Завершуємо пошук.

B. Сортування виконуємо або у порядку зменшення вартості механізму захисту, або у порядку зростання «корисності» механізму захисту. У першому випадку сортування направлене на те, щоб при побудові дерева пошуку зразу ж відкинути занадто дорогі механізми захисту. У другому ви-

падку намагаємося швидко відкинути ті механізми захисту, які незначно впливають на захищеність мережі. «Корисність» механізму захисту $U(K_i)$ обчислюється за формулою

$$U(K_i) = \sum_{j=1..L} \frac{\partial P(E_1,\ldots,E_N,0,\ldots,0)}{\partial E_j} K_{ij}, \qquad (9)$$

де K_{ij} — ефективність *i*-го механізму захисту проти j — ї атаки, L — кількість атак.

Таким чином у $U(K_i)$ враховується і значимість атак, проти яких направлений механізм захисту, і міцність самого механізму захисту.

С. Згідно з методом послідовного аналізу варіантів будуємо дерева пошуку на часткових розв'язках задачі. Виконуємо обхід дерева в ширину і на кожному етапі розвиваємо тільки перспективні часткові розв'язки. Для визначення перспективності розв'язку, обчислюємо його вартість $g\left(M_1^0,\ldots,M_{depth}^0,0,\ldots 0\right)$, та оцінку значення цільової функції $P\left(M_1^0,\ldots,M_{depth}^0,1,\ldots,1\right)$. Крім того, враховуємо, що для часткових розв'язків вигляду $\left(M_1^0,\ldots M_n^0,0\right)$ оцінка вартості буде така ж, як для батьківського розв'язку $\left(M_1^0,\ldots M_n^0\right)$. Аналогічно для часткових розв'язків вигляду $\left(M_1^0,\ldots M_n^0,1\right)$ оцінка значення цільової функції буде така ж, як $\left(M_1^0,\ldots M_n^0\right)$. Це дозволяє не обчислювати їх повторно.

Крок 0 $depth = 1$. Поміщаємо в чергу часткові розв'язки $M^0{}_{(1)} = (0)$ і $M^0{}_{(1)} = (1)$.

Крок 1. Виймаємо з черги перший елемент — частковий розв'язок $M_{(depth)}^0$.

Якщо $depth == N$ (побудували повний розв'язок),

То переходимо на Крок 6.

Крок 2. Якщо $M_{depth}^0 = 0$ (останній елемент часткового розв'язку дорівнює нулю),

То переходимо на Крок 3.

Інакше переходимо на Крок 4.

Крок 3. Оцінка значення цільової функції на частковому розв'язку.

Якщо $P\left(M_1^0,\ldots,M_{depth}^0,1,\ldots 1\right) \le P^*$ (частковий розв'язок можна добудувати до оптимального),

То поміщаємо в кінець черги часткові розв'язки $M_{(depth+1)}^0 = \left(M_1^0,\ldots,M_{depth}^0,0\right)$, $M_{(depth+1)}^0 = \left(M_1^0,\ldots,M_{depth}^0,1\right)$;

Переходимо на Крок 5.

Крок 4. Оцінка значення функції обмежень на частковому розв'язку.

Якщо $g\left(M_1^0,\ldots,M_{depth}^0,0,\ldots 0\right) \le 0$ (частковий розв'язок можна добудувати до допустимого),

То поміщаємо в кінець черги часткові розв'язки $M_{(depth+1)}^0 = \left(M_1^0,\ldots,M_{depth}^0,0\right)$, $M_{(depth+1)}^0 = \left(M_1^0,\ldots,M_{depth}^0,1\right)$

Крок 5.

Якщо $g\left(M_1^0,\ldots,M_{depth}^0,0,\ldots 0\right) \le 0$,

То $P^* = \min\left\{P\left(M_1^0,\ldots,M_{depth}^0,0,\ldots,0\right),P^*\right\}$;

Переходимо на Крок 1

Крок 6. Оцінка значення функції обмежень та цільової функції на повному розв'язку.

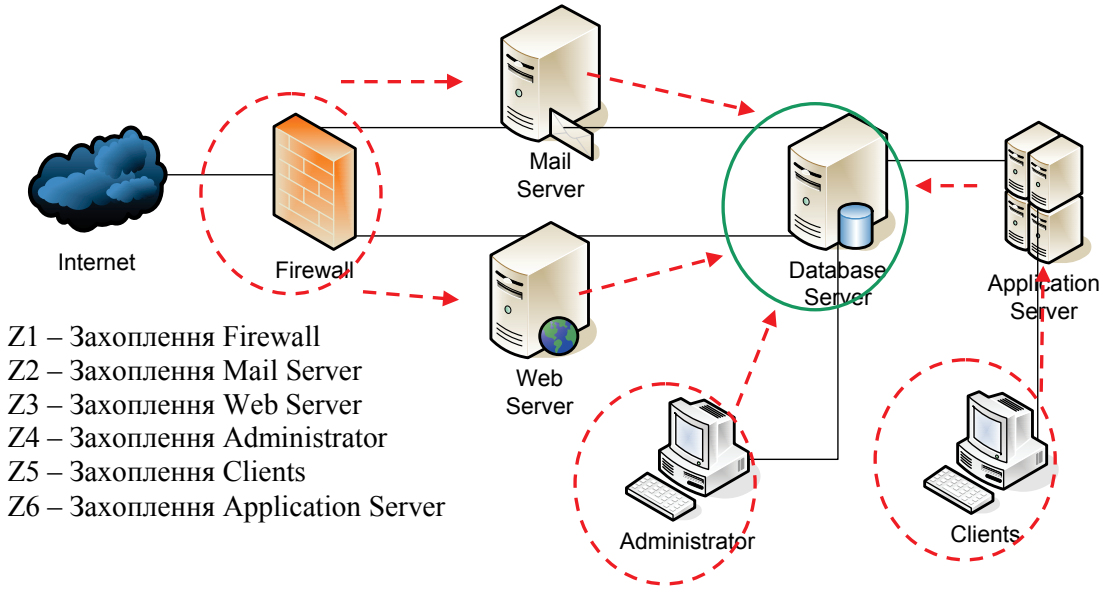
Якщо $g\left(M_1^0,\ldots,M_N^0\right) \le 0$ (розв'язок є допустимим) і $P\left(M_1^0,\ldots,M_N^0\right) \le P^*$,

То $P^* = P\left(M_1^0,\ldots,M_N^0\right)$ і $M^* = M_N^0$;

Переходимо на Крок 1.

Приклад моделювання

Розглянемо випадок визначення структури розміщення механізмів захисту для комп'ютерної мережі такого вигляду, з наведеними джерелами атак та об'єктом атаки (рисунок.



Шляхи захоплення серверу Баз даних

Функція ймовірності захищеності, для цього прикладу, була отримана у [8].

Задамо для даної мережі значення ймовірносте атак, міцність механізмів захисту та вартість їх застосування, а також максимальну вартість системи захисту. І проаналізуємо процес збіжності алгоритму.

Припустимо, що на кожен з хостів $Z_1,\ldots,Z_6$ маємо по одній атаці, з заданими значеннями ймовірностями $E_1,\ldots,E_6$. Необхідно вибрати оптимальне розташування механізмів захисту $M_1,\ldots,M_6$ проти цих атак, беручи до уваги значення їхньої міцності $K_1,\ldots,K_6$ та їх вартість $C_1,\ldots,C_6$. Сумарна вартість системи захисту не повинна перевищувати $C_0 = 5$. Процес збіжності алгоритму наведений у таблиці.

Процес збіжності алгоритму пошуку оптимального розміщення механізмів захисту

M	E	C	K	1	2	3	4	5	6	7	8	9	10	11	12	13	14
5	0,1	4	0,3	1	0	1	1	0	0	1	1	0	0	0	0	1	1
1	0,6	3	0,5			1	0	1	0	0	0	1	1	0	0	0	0
3	0,4	2	0,1							1	0	1	0	1	0	0	0
4	0,7	2	0,4													1	0
2	0,2	1	0,3														
6	0,5	1	0,6														

M	E	C	K	15	16	17	18	19	20	21	22	23	24	25	26	27	28
5	0,1	4	0,3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0,6	3	0,5	1	1	1	1	0	0	0	0	1	1	1	1	1	1
3	0,4	2	0,1	1	1	0	0	1	1	0	0	1	1	0	0	0	0
4	0,7	2	0,4	1	0	1	0	1	0	1	0	0	0	1	1	1	1
2	0,2	1	0,3									1	0	1	0	0	0
6	0,5	1	0,6													1	0

У першій колонці таблиці наведені номери механізмі в захисту які можна використати, у другій — ймовірності атак, у третій — вартість механізмів захисту, і у четвертій — коефіцієнти міцності механізмів. Далі показано по крокам робота алгоритму послідовного аналізу варіантів, з варіантами розміщення механізмів, що аналізувались. Алгоритм збігається за 28 кроків, що набагато мен-

ше ніж при застосуванні алгоритму повного перебору всіх варіантів (64 варіанти).

У результаті роботи, при обмеженні на вартість системи захисту $C_0 = 5$, отримано, що оптимальним є використання механізмів захисту від атак на міжмережевому екрані (M_1) та сервері адміністратора (M_4). При цьому ймовірність атаки зменшується з $P = 0{,}80392$ до $P = 0{,}555256$.

Таке розташування механізмів захисту також відповідає практичному досвіду побудови системи захисту.

Висновки

У статті запропоновано модель та алгоритм синтезу оптимальної структури системи захисту інформаційно-комунікаційної мережі, з врахуванням топології мережі та можливих шляхів атак.

Наведено обчислювальний приклад використання запропонованої моделі для визначення оптимального розміщення механізмів захисту з урахуванням обмеження на їхню сукупну вартість.

Отримані результати можуть бути застосовані для мереж будь-якого розміру, як локальних так і корпоративних. Для корпоративних мереж з великою кількістю сервісів можливо багато варіантів розміщення механізмів захисту, і цільова функція ймовірності захищеності буде містити багато змінних та мати складний вигляд.

Запропонована модель може бути використана при (синтезі) проектуванні системи захисту інформаційно-комунікаційних мереж, а також під час їх експлуатації.

СПИСОК ЛІТЕРАТУРИ

1. Шнайер Б. Секреты и ложь. Безопасность данных в цифровом мире / Шнайер Б. — СПб: Питер, 2003. — 368 с. — ISBN 5-318-00193-9.

2. Schneier. B. Attack Trees / Schneier. B. // Dr Dobb's Journal. — 2007. — № 24. – P. 12—18.

3. Sheyner O. Tools for Generating and Analyzing Attack Graphs // Oleg Sheyner, Jeannette Wing // Springer. — 2003. — LNSC 3188. — P. 344—371.

4. Sheyner O. Two Formal Analyses of Attack Graphs. / Sheyner O., Jha S., Wing J. // IEEE Computer Security Foundations Workshop, Cape Brenton, Nova Scotia, Canada. — 2002. — June. — P. 49—63.

5. Todd Hughes. Attack scenario graphs for computer network threat analysis and prediction. / Todd Hughes, Oleg Sheyner // Complexity. — 2003. — №9 (2). — P. 15—18.

6. Новіков О. М. Побудова логіко-імовірнісної моделі захищеної комп'ютерної системи / О. М. Новіков, А. О. Тимошенко // Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні. — 2001. — № 3. — С. 101—105.

7. Рябинин И. А. Надежность и безопасность структурно-сложных систем / Рябинин И. А. — СПб: Политехника, 2000. — 248с. — ISBN 5-7325-0549-0

8. Новіков О. М. Логіко-імовірнісна модель захищеності компонентів інформаційно-комунікаційних систем / О. М. Новіков, А. М. Родіонов // Інформаційні технології та комп'ютерна інженерія (ВНТУ). — 2008. — №1 (11). — С. 170—175.

9. Ковалев М. М. Дискретная оптимизация. Целочисленное программирование / Ковалев М. М. — М.: Едиториал УРСС, 2003. — 192с. — ISBN 5-354-00499-3

Рекомендована кафедрою обчислювальної техніки

Надійшла до редакції 8.09.08
Рекомендована до друку 20.10.08

Мартиненко Любов Петрівна — студентка, ***Новіков Олексій Миколайович*** — директор, ***Родіонов Андрій Миколайович*** — аспірант.

Фізико-технічний інститут Національного технічного університету України «Київський політехнічний інститут»