УДК 004.056.5:004.048

І.Б. Трегубенко

Черкаський державний технологічний університет, Черкаси

МЕТОД НАВЧАННЯ ІНТЕЛЕКТУАЛЬНИХ АГЕНТІВ В СИСТЕМАХ ПРЕВЕНТИВНОГО АКТИВНОГО ЗАХИСТУ ІНФОРМАЦІЇ

Запропонована ідея побудови систем превентивного активного захисту в мережевих середовищах. Обрано технологію інтелектуальних агентів для управління такими системами. Розроблено метод навчання інтелектуальних агентів для їх адаптації до оточуючого середовища.

Ключові слова: *інтелектуальний агент, навчання, захист інформації.*

Вступ

Аналіз та постановка проблеми. Сучасні інформаційні технології, зокрема в галузі інформаційної безпеки, реалізуються в середовищі складних розподілених мереж, таких як глобальна Internet мережа й корпоративні Intranet мережі, властивості та функціонування яких визначаються апаратними компонентами, програмними засобами і протоколами взаємодії. Обсяги інформації, яка зберігається та обробляється в таких системах продовжують зростати [1]. Гетерогенна структура глобальних та корпоративних мереж обумовлює особливість таких систем, а саме не структурованість інформації. Фактор неоднорідності та неструктурованості інформації значно погіршує ефективність систем захисту.

При традиційної побудови систем захисту часто методи забезпечення безпеки інформації в комп'ютерних мережах обмежуються контролем доступу в широкому сенсі. Обмежена кількість програмних засобів проводить моніторинг стану системи, що захищається. В основному це антивірусні програмні комплекси, в яких певні дії проводяться в фоновому з точки зору користувача режимі. Але ці приклади не є по суті повноцінними інтелектуальними агентами. Методи автономних програмних агентів на даний час частково застосовуються в системах пошуку текстової інформації в глобальної мережі.

Мета роботи. Необхідно запропонувати сучасні підходи до побудови систем захисту інформації в мережевих середовищах, з врахуванням великого обсягу інформації та складності й гетерогенності таких систем. З цієї точки зору методи управління такими системами є найбільш важливими. Необхідно дослідити принципи побудови систем управління та методи їх адаптації до оточуючого мережевого середовища.

Основний матеріал

Пропонується будувати системи захисту інформації на базі концепції превентивного активного захисту з застосуванням засобів інтелектуалізації. Якщо система чекає на дії порушника або на констатацію вторгнення чи порушення цілісності системи, то не можна вважати захист надійним. Значно ефективнішим представляється ситуація, коли проводиться постійний моніторинг базових параметрів конфіденційності та цілісності системи, що захищається. Зокрема в задачах ідентифікації користувачів на базі біометричних параметрів [2, 3] було запропоновано проведення постійного негласного моніторингу автентичності користувача поточному сеансу роботи.

Для вирішення цієї задачі в першу чергу необхідно побудувати сучасну технологію управління систем захисту. При традиційної побудови управління, зокрема інформаційною безпекою, строго ієрархічне. На нижньому рівні реалізується механізм управління із зворотнім зв'язком та невеликими затримками, а на верхньому рівні проходить параметрична оптимізація та програмне координування управлінських впливів. Перехід від прямого до непрямого управління підвищить ефективність систем управляння, зокрема інформаційною безпекою. Сутність концепції інтелектуального управління – інтелектуальні агенти, кожен з яких забезпечує управління частиною інформаційних ресурсів, використовуючи набір власних цільових функцій та зовнішніх управлінських впливів [1]. Інтелектуальний агент може розглядається як сукупність нейромережевих адаптивних модулів що знаходяться в постійної взаємодії з оточуючим середовищем, яке постійно змінюється. Тому є актуальною задача навчання інтелектуального агенту, яку пропонується вирішувати на базі теорії навчання з підкріпленням. Теорія навчання з підкріпленням (reinforcement learning) була розвинена в циклі робіт Р. Саттона і Е. Барто [4], та може розглядатись як розвиток теорії адаптивної поведінки, яка була розроблена в роботах М.Л. Цетліна та його послідовників [5, 6].

Побудуємо метод навчання інтелектуального агента для його адаптації до робочого середовища. Загальна схема навчання з підкріпленням, відповідно до навчання інтелектуального агента показана на рис. 1.

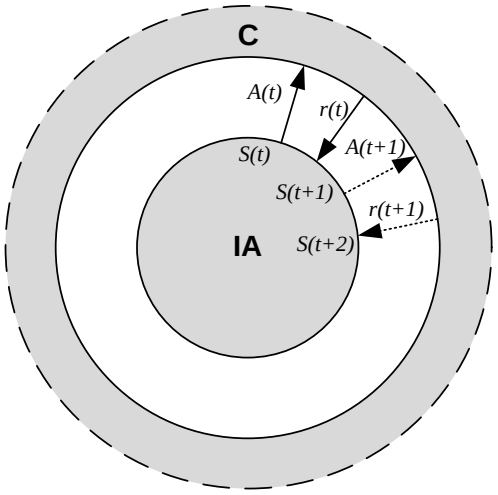

Рис. 1. Навчання з підкріпленням

Розглянемо інтелектуальний агент (ІА) у взаємодії з середовищем (С). Для проведення аналізу побудуємо згорнуту форму такої взаємодії (рис. 2).

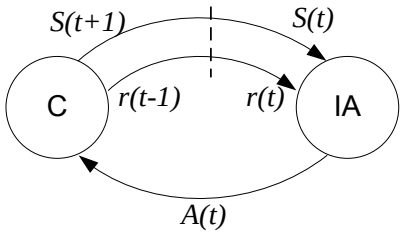

Рис. 2. Навчання з підкріпленням, згорнута форма

Приймаємо величину часу дискретною: t = 1, 2. У поточній ситуації S(t) інтелектуальний агент виконує дію a(t), отримує підкріплення r(t) і потрапляє в наступну ситуацію S(t + 1). Підкріплення може бути позитивним (стимулювання), r(t)>0, або негативним (стримання), r(t)< 0.

Задача інтелектуального агента – максимізувати сумарний стимул, який можна отримати в майбутньому протягом тривалого періоду часу. Мається на увазі, що ІА може мати свою внутрішню «суб'єктивну» оцінку сумарного стимулу і в процесі вчення постійно удосконалює цю оцінку. Ця оцінка визначається з врахуванням дисконтного чинника:

$$U(t) = \sum_{j=0}^{\infty} \gamma^j r(t+j), \quad t = 1, 2, \ldots, \qquad (1)$$

де U(t) – оцінка сумарного стимулу, очікуваного після моменту часу t; γ – дисконтний чинник, $0 < \gamma < 1$. Дисконтний чинник враховує, що чим далі ІА «заглядає» в майбутнє, тим менше у нього упевненість в оцінці стимулу. Формування дисконтного чинника γ сама по собі не тривіальна задача, яка може розглядатись окремо.

Якщо безліч можливих ситуацій $\{S_i\}$ і дій $\{A_j\}$ скінченна, то можна застосувати метод навчання SARSA, який відповідає ланцюгу подій:

$$S(t) \to A(t) \to r(t) \to S(t+1) \to A(t+1) \to r(t+1) \to S(t+2). \quad (2)$$

Необхідно сформувати оцінку значення сумарного стимулу Q(S(t),a(t)), який отримає інтелектуальний агент, якщо в ситуації S(t) він виконає дію a(t). Цей процес ітераційний. Математичне очікування сумарного стимулу дорівнює:

$$Q(S(t), a(t)) = E\{r(t) + \gamma\, r(t+1) + \gamma^2\, r(t+2) + \ldots\};$$
$$S = S(t),\ a = a(t).$$

Маємо:

$$Q(S(t), a(t)) = E[r(t) + \gamma\, Q(S(t+1), a(t+1))].$$

Похибка може визначатись:

$$\delta(t) = r(t) + \gamma\, Q(S(t+1),a(t+1)) - Q(S(t),a(t)), \qquad (3)$$

де $\delta(t)$ – різниця між оцінкою величини сумарного стимулу, який формується у інтелектуального агента для моменту часу t після обрання дії a(t+1) у наступній ситуації S(t+1) в момент часу t+1, та попередньої оцінки цієї ж величини, яка була у ІА в момент часу t. Попередня оцінка дорівнює Q(S(t),a(t)), нова оцінка дорівнює $r(t) + \gamma Q(S(t+1),a(t+1))$, що й відображає формула (3) для величини $\delta(t)$. У відповідності до цього $\delta(t)$ і обучається інтелектуальний агент.

Таким чином, у кожен такт часу виконується як вибір дії, так і навчання інтелектуального агента.

Метод навчання інтелектуального агента можна описати наступним чином:

1. Визначаються початкові умови S(t), t=1.
2. Інтелектуальний агент виконує дію a(t).
3. ІА отримує підкріплення r(t).
4. Формується оцінка сумарного стимулу Q(S(t),a(t)).
5. Визначається похибка $\delta(t)$.
6. Відповідно до похибки проводиться навчання ІА та корегування, вибір дій
7. Переходимо до наступного кроку, тобто t = t + 1 та переходимо до виконання пункту 2.

В зв'язку з безперервністю характеру загроз та швидкому зміненню оточуючого середовища казаний цикл необмежений та буде перерваний з закінченням життєвого циклу агента.

Доцільно розглянути як може відбуватись вибір дії інтелектуального агенту в пункті 6 вказаного методу. Пропонується застосувати так зване «ε – жадібне правило». Тоді вибір дії буде виконуватись наступним чином: − в момент t з вірогідністю 1 - ε обирається дія, яка відповідає максимальному значенню $Q(S(t),a_i)$: $a(t) = a_k$, $k = \text{argmax}_i\{Q(S(t),a_i)\}$ − з вірогідністю ε обирається довільна дія випадковим чином, $0 < \varepsilon < 1$.

Навчання, тобто переоцінка значень Q(S, a) проходить у відповідності з оцінкою помилки $\delta(t)$ – до значення Q(S(t),a(t)) додається значення пропорційне помилки часової різниці $\delta(t)$:

$$\Delta Q(S(t),a(t)) = \alpha\,\delta(t) = \alpha\,[r(t) + \gamma\,Q(S(t+1), a(t+1)) - Q(S(t),a(t))], \quad (4)$$

де α – параметр швидкості навчання.

Тому, що кількість ситуацій та дій обмежені, проходить формування матриці $Q(S_j,a_i)$, яка відповідає всім можливим ситуаціям S_j та всім можливим діям a_i. Оцінка сумарного стимулу Q(S(t),a(t)) може розглядатись як оцінка якості дії a(t) в поточної ситуації S(t).

Вказана технологія корелюється з методом динамічного програмування. В обох випадках загальна оптимізація багатокрокового процесу ухвалення рішення відбувається шляхом впорядкованої процедури однокрокових оптимізуючих ітерацій, причому оцінки ефективності тих або інших рішень, відповідні попереднім крокам процесу, переоцінюються з врахуванням знань про можливі майбутні кроки.

Важлива особливість метода з підкріпленням – його простота. Інтелектуальний агент отримує від зовнішнього середовища тільки сигнали підкріплення r(t). Тобто оточуюче середовище не надає пояснень об'єкту, що навчається, яким саме чином необхідно діяти. Саме це значно відрізняє даний метод від традиційного в теорії нейронних мереж методу зворотного розповсюдження помилки [7], для якого точно визначається, що саме повинно бути на виході нейронної мережі при заданому вході.

Висновки

Побудова систем захисту інформації на базі концепції превентивного активного захисту з застосуванням засобів інтелектуалізації можуть підвищити надійність захисту інформації в мережевих середовищах. При такому підході засоби захисту проводять постійний моніторинг базових параметрів конфіденційності на предмет наявності вторгнення чи порушення цілісності системи.

Для управління такими системами обрано технологію інтелектуальних агентів, зокрема для організації непрямого інтелектуального управління.

Побудовано метод навчання інтелектуального агента, розроблена загальна схема та згорнута форма навчання з підкріпленням з використанням сумарного стимулу. Запропоновано розвивати метод навчання з підкріпленням для навчання автономних інтелектуальних агентів.

Список літератури

1. Трегубенко І.Б. Концепція інтелектуального управління в складних розподілених системах / І.Б. Трегубенко // Материалы IX міжнародної наукової конференції ім. Т.А. Таран „Інтелектуальний аналіз інформації ІАІ-2009" (Київ, 19-22 травня 2009р.); ред. кол.:С.В. Сирота (гол. ред.) и др. – К. : ПРОСВІТА, 2009. – С. 391-393.

2. Трегубенко І.Б. Підвищення ефективності систем безпеки методами біометричної ідентифікації та аутентифікації користувачів / І.Б. Трегубенко, І.В. Степанушко, І.І.. Багрєєв // Вісник ЧДТУ. – 2008. – № 3. – С. 113-118.

3. Трегубенко І.Б. Побудова класифікаторів в задачах біометричної ідентифікації та аутентифікації користувачів / І.Б. Трегубенко, І.В. Степанушко, Г.Т. Олійник // Вісник ЧДТУ. – 2009. – № 1. – С. 37-40 с.

4. Sutton R. Reinforcement Learning: An Introduction [Електронний ресурс] / R. Sutton, A. Barto. – Cambridge: MIT Press, 1998. – Режим доступу до док.: http://www.cs.ualberta.ca/~sutton/book/the-book.html.

5. Цетлин М.Л. Исследования по теории автоматов и моделирование биологических систем / М.Л. Цетлин. – М.: Наука, 1969. – 316 с.

6. Варшавский В.И. Оркестр играет без дирижера / В.И. Варшавский, Д.А. Поспелов. – М.: Наука, 1984. – 420 с.

7. Rumelhart D.E. Learning representation by back-propagating error / D.E. Rumelhart, G.E. Hinton, R.G. Williams // Nature. – 1986. – 323 (6088). – P. 533-536.

Надійшла до редколегії 10.03.2010

Рецензент: д-р техн. наук, проф. А.Ю. Кочкарьов, Черкаський державний технологічний університет, Черкаси.

МЕТОД ОБУЧЕНИЯ ИНТЕЛЛЕКТУАЛЬНЫХ АГЕНТОВ В СИСТЕМАХ ПРЕВЕНТИВНОЙ АКТИВНОЙ ЗАЩИТЫ ИНФОРМАЦИИ

И.Б. Трегубенко

Предложена идея построения систем превентивной активной защиты в сетевых средах. Избрана технология интеллектуальных агентов для управления такими системами. Разработан метод учебы интеллектуальных агентов для их адаптации к окружающему среды.

Ключевые слова: *интеллектуальный агент, учеба, защита информации.*

METHODS OF EDUCATION INTELLECTUAL AGENT IN SYSTEM OF PREVENTIVE ACTIVE PROTECTION TO INFORMATION

I.B. Tregubenko

The idea of construction of the systems of preventive active defence is offered in the environments of networks. Technology of intellectual agents is select for a management such systems. The method of studies of intellectual agents is developed for their adaptation to surrounding environments.

Keywords: *intellectual agent, studies, defence of information.*