УДК 683.03

О.Ю.Ю.Афанасьєва

МЕТОДИ СЕМАНТИЧНИХ ПЕРЕТВОРЕНЬ В СТЕГАНОСИСТЕМАХ

Однією з особливостей стеганосистем (SS) є те, що вона може розміщати в цифровому середовищі (CS) інформаційний образ (IO), який може представляти інформацію в різних формах. Це обумовлено тим, що SS в існуючих і найбільш поширених алгоритмах представляє IO у вигляді модифікацій елементів, які описують CS і їх форма залежить, або тип елементів середовища залежить від способу цифрового представлення CS. В більшості випадків, інформація, яку передбачається кодувати, кодується числами, наприклад, якщо така інформація представляється текстами, а окремі букви кодуються числами, останні переводяться у бітову форму і, тоді, формулюється чергова задача модифікації цифрового представлення CS таким чином, щоб модифіковані елементи запам'ятовували відповідну бітову інформацію [1]. Можна використовувати і інші способи представлення IO, який передбачається укривати в CS, наприклад, IO може представляти собою графічний образ G. У цьому випадку, виникає задача кодування окремих елементів графічного образу, які приймаються в якості базових графічних елементів. На відміну від випадку кодування текстових IO, при використанні графічних примітивів, необхідно передбачити можливість кодування взаємного розміщення відповідниxx примітивів у повному IO. З одного боку, це приводить до збільшення кількості кодів, які необхідно вбудувати у CS, а з другого боку, це дозволяє суттєво зменшти код необхідний для опису в текстовій формі окремого примітиву.

Розглянемо в рамках даного підходу суть перетворень різних виразів та описів, що використовуються в рамках тих, чи інших теорій. Будь яке перетворення здійснюється у відповідності з наступними цілями:

- отримати більш компактну форму опису деякої сутності, що описується окремими співвідношеннями,
- отримати нову інформацію про об'єкти та закономірності, що описуюся компонентами, які використовуються, при здійсненні перетворень,
- отримати форму опису певного об'єкту, або процесу, яка придатна для її використання в процесі розвязку певної задачі, що пов'язана з предметною областю відповідних процесів.

Всі перетворення, що обумовлюються приведеними вище цілями, реалізуються з формальними формами опису відповідних процесів, оскільки формальні описи допускають достатньо широкі способи інтерпретації відповідних компонент. На відміну від формалізованих, або формальних перетворень, семантичні перетворення полягають у збереженні семантики обєктів, чи процесів, що приймають участь у перетвореннях. В цьому сенсі,

такі перетворення в більшій мірі полягають у реалізації наступних цілей:

- отриманні повних форм опису певної сутності, що являється процесом, або сукупністю об'єктів,
- перетворення описів об'єктів з одних способів опису у інші форми опису,
- перетворення, що пов'язані з реалізацією синтезу різних об'єктів та процесів у деякий новий об'єкт, або процес.

Спільною для всіх семантичних перетворень особливістю є зміна повноти опису інтерпретації об'єктів та процесів, що перетворюються. Прикладом таких перетворень можуть служити перетворення текстових описів у графічну форму їх представлення і навпаки.

В рамках задач стеганографії, семантичні перетворення розглядаються як такі основною ціллю яких є отримання нових форм опису однієї і тієї ж сутності з максимально можливим збереженням семантики об'єктів та процесів, що описуються. Для реалізації тих, чи інших перетворень, незалежно від характеру їх представлення, необхідно мати наступні компоненти, які такі перетворення роблять можливими та, з точки зору семантики, допустимими:

- способи опису процесів, чи об'єктів певного класу,
- правила перетворень відповідних способів опису в межах однієї форми їх представлення,
- правила перетворень способів опису об'єктів в різних формах їх представлень,
- методи перевірки семантичної коректності результатів, що отримані в результаті реалізації тих, чи інших перетворень,
- базові описи інтерпретацій об'єктів та процесів, що відносяться до вибраної предметної області.

Способи опису процесів, чи обєктів певної предметної області W_i грунтуються на введенні алфавіту, який, у випадку графічних способів опису, представляє собою сукупність графічних примітивів, з яких формуються слова, що представляють собою нероздільну компоненту опису обєкту, чи процесу. Сукупність окремих слів, по аналогії з текстовими структурами, будемо називати фразами φ_i опису, які можуть обєднуватися в більші фрагменти, які будемо називати реченнями опису ψ_i інформаційних образів. Основою для дослідження задач, що повязані з семантикою, є семантичні словники, які вміщають семантичний опис предметної області. Словники S^C є вихідними даними про наявність тієї, чи іншої семантики в певних IO. Структура S^C може мати різну міру складності, яка визначається особливостями задач, які досліджуються, але їх базовими елементами є інтерпретаційні описи всіх позначень, що використовуються в рамках опису предметної області та відповідних досліджень.

Приймаючи до уваги приведене вище, введемо уявлення про семантичне

перетворення у вигляді наступного визначення.

Визначення 1. Семантичним є таке перетворення, яке здійснюється над семантичними компонентами, котрі складаються з формальних ідентифікаторів і семантичних складових, що представлені у вигляді інтерпретаційних описів.

Формально, семантична компонента представляється у вигляді пари складових з, які описуються співвідношенням:

$$s_i = [x_i, j(x_i)] = [x_i, <\alpha_{i1}, ..., \alpha_{im}>] ,$$

де α_{ij} - елемент опису текстової інтерпретації, яким може бути слово ω_i, фраза φ_i, чи речення ψ_i. Використання S^C дозволяє в рамках SS використовувати останній, як засіб кодування, оскільки запис відображення x_i, по визначенню, потребує значно менше символів у порівнянні з записом відображення $j(x_i)$. Тому, в рамках SS актуальною є задача, в якій розглядається проблема скорочення форми відображення IO таким чином, щоб кодові розміри IO були мінімальними, а семантичне значення залишалось повним. Мінімізація кодового представлення IO в межах одного типу відображень може досягатися за рахунок таких перетворень, які дозволяють зберегти семантичні значення всіх компонент після реалізації відповідного перетворення. У випадку форми представлення IO у вигляді деякого тексту на мові споживача, такі перетворення можуть представляти собою перетворення $j(x_i)$ в нормалізовану форму, в якій відсутні всі можливі типи надмірностей, що характерні для природної мови споживача.

Більш ефективним перетворенням IO в код, який є меншим від початкового коду, є перетворення, яке приводить до сумісного використання різних типів відображення IO, наприклад, відображення у вигляді слів природного тексту, та елементів графічного образу. Очевидно, що S^C повинні вміщати не тільки $j(x_i)$, а і $j(g_i)$, де g_i - графічний елемент, який представляє собою графічний примітив. У випадку, коли $x_i = g_i$, то використовується інтерпретація $j(x_i, g_i)$.

Маючи описи семантичних елементів в W_i, необхідно мати змогу їх перетворювати. Для цього, необхідно сформувати правила перетворень. У відповідності з визначенням 1, такі правила повинні стосуватися не тільки формального представленого ідентифікатора x_i, а і його текстового опису $j(x_i)$. Частина s_i, яка представляє собою x_i, перетворюється на основі загальних правил перетворень, що сформульовані в математичній логіці [1], що приймається в рамках даного підходу . Частина s_i, яка представляє собою $j(x_i)$, перетворюється на основі використання модифікованих синтаксичних та семантичних правил, що відомі з граматики відповідної

мови, γ_i - окреме правило граматики Γ. Модифікація цих правил полягає у їх синтезі з загальними правилами, що запозичаються, в першу чергу, з математичної логіки з описом обмежень, що враховуються, при реалізації відповідних перетворень.

Крім модифікованих правил перетворення формалізованої складової елемента s_i, в рамках семантичних перетворень використовуються правила перетворень текстових описів відповідних компонент, які представляють собою їх інтерпретацію. Оскільки, будь які перетворення повинні грунтуватися, по можливості, на однозначній інтерпретації, критеріїв, що обумовлюють можливість їх здійснення, то текстові описи $j(x_i)$ представляються в нормалізованій формі, а параметрами, які допускають їх числову оцінку, являються параметри семантичної значимості окремих слів ω_i, фраз φ_i, чи речень ψ_i, параметри семантичної суперечності та семантичного конфлікту [2]. Оскільки правила перетворень $s_i = [x_i, j(x_i)]$ стосуються двох компонент x_i і $j(x_i)$, то між відповідними перетвореннями повинен існувати функціональний зв'язок. Серед семантичних параметрів, якими є семантична значимість ξ_i слова α_i та семантичні суперечність σ і семантичним протиріччям π, існує ієрархічна залежність, яка полягає у наступному. Базовим семантичним параметром є ξ_i, який визначається на основі S^C, а σ і π визначаються співвідношенням, що записується у вигляді:

$$\{[\mathrm{I}\xi(\alpha_i) - \xi(\alpha_j)\mathrm{I} = \varepsilon_i] \& [v^{\sigma}_{\min} < \varepsilon_i < v^{\sigma}_{\max}]\} \to [\varepsilon_i = \sigma(\alpha_i, \alpha_j)],$$

де $v^{\sigma}_{\min}$ - мінімальне значення величини ε_i, $v^{\sigma}_{\max}$ - максимальне значення величини ε_i. Семантичний конфлікт визначається співвідношенням:

$$\{[\mathrm{I}\xi(\alpha_i) - \xi(\alpha_j)\mathrm{I} = \varepsilon_i] \& [0 \le \varepsilon_i < v^{\sigma}_{\min}]\} \to [\varepsilon_i = \pi(\alpha_i, \alpha_j)].$$

Якщо в процесі перетворень компонент x_i і x_j отримуємо x_k і має місце $L(x_i, x_j) \to x_k$, то у відповідності із схемою перетворень, яку описує $L(x_i, x_j)$, то реалізуємо певну композицію $j(x_k) = M[j(x_i), j(x_j)]$. Правила формування $j(x_k)$ грунтуються на інтерпретації елементарних логічних функцій $\{\&, \vee, \to, \neg\}$, яка формується в рамках W_i і пов'язана з системою обмежень, що визначає в більшості випадків області визначення змінних x_i.

При реалізації перетворень описів V_i і CS, що відповідають певному W_i, досить часто виявляється доцільним реалізувати перетворення цих описів, у звязку з переходом від однієї форми їх відображення у іншу форму. Найбільш відомим прикладом такого типу перетворень є перетворення

логічних формул у графову структуру, перетворення процесів виводу, що реалізуються в математичній логіці, в процеси функціонування абстрактних автоматів і т.д. [5]. Різні форми відображень визначаються наступними ознаками:

- мірою загальності засобів, що дозволяють формувати описи тих, чи інших об'єктів у відповідній формі,
- мірою загальності засобів відображення, яка тісно пов'язана з мірою абстракції отриманих описів,
- різними функціональними можливостями, що можуть використовуватися при дослідженні певним способом описаних процесів, чи об'єктів,
- різними цілями, для використовуються різні форми опису, що обумовлюються вимогами відповідних задач, які необхідно розв'язати,
- різною мірою відповідності тих, чи інших засобів відображення або опису типу предметної області, в рамках якої передбачається проводити дослідження описаних процесів.

Міра загальності відображення процесів, які передбачається досліджувати, тісно пов'язана з мірою загальності результатів, які вдається отримати завдяки використанню відповідних засобів. Це дозволяє розширювати, або звужувати коло застосувань, для яких можна використовувати відповідні результати. Наприклад, якщо задача визначення пропускної здатності стеганоканалу розглядається на рівні загальності, який не передбачає враховування типу CS, в якому розміщається вбудована V_i, то таким чином більш загальний метод розв'язку задачі визначення цього параметру є більш доцільний. Особливо це актуально для розв'язку задач, що пов'язані з забезпеченням стійкості стеганограм проти атак різних типів, а також для ряду інших задач, що є актуальними для стеганографії [4].

Міра абстракції опису тих, чи інших обєктів, по своїй суті, в певному сенсі, забезпечує ту чи іншу міру загальності. Принциповою різницею між мірою абстракції і мірою загальності є те, що не залежно від міри загальності опису процесу, останній має повністю адекватну інтерпретацію в предметній області, в якій відповідна задача розглядається, в той час, як міра абстракції може приводити до неможливості враховування тих чи інших аспектів предметної області, до якої відноситься задача, що абстрагується. Міра абстракції з її збільшенням приводить до зменшення адекватності відповідного формального опису предметної області. Ця адекватність визначається виключно зменшенням повноти інтерпретаційного опису відповідних форм представлення досліджуваних процесів. Прикладом, що ілюструє відмінність міри загальності опису від міри абстракції такого опису, можуть служити теоретичні дослідження в різних областях диференціальної топології, чи задач, що полягають у дослідженнях об'єктів абстрактної алгебри і т.д. [5].

Різні функціональні можливості, при використанні різних форм опису

досліджуваних процесів, є основою для впровадження відповідних форм опису задач, які необхідно розв'язувати. Майже у всіх випадках впровадження нової форми опису задач, які необхідно розв'язувати, відповідні задачі задачі виявилося можливим розв'язати, або представити у формі, яка дозволяє відповідну задачу інтерпретувати таким чином, що остання може бути використана в певному конкретному випадку. Прикладом використання різних форм опису окремих задач, використання яких дозволило відповідні задачі розв'язувати, може служити впровадження уявлень про моделі машини Тюрінга та інші [5].

В багатьох випадках розвязки задач одного і того ж типу використовуються для досягнення різних цілей. Відмінність цілей може обумовлюватися різними типами прикладних проблем, що можуть розв'язуватися певним класом методів. Типи прикладних задач визначаються різними предметними областями, в яких передбачається розв'язувати сформовану проблему. В подальшому, в якості такої предметної області будемо розглядати предметну область, що визначається задачами стеганографії.

Завдяки тому, що елементи описуються не тільки змінними x_i, а й інтерпретаційними описами $j(x_i)$, існує можливість, з точністю до відповідного текстового опису кожної компоненти, враховувати всю інформацію, яка є доступною про відповідну компоненту. Приймається, що текстовий опис є найбільш повним по відношенню до всіх інших можливих описів. Завдяки цьому, виникає наступна можливість більш повної ідентифікації отриманих внаслідок перетворень результатів. Під ідентифікацією, даному випадку, розуміється перевірка адекватності отриманих результатів процесам, стосовно яких проводиться дослідження, частиною яких є відповідні перетворення. Суть такої перевірки грунтується на наступних факторах:

- текстовий опис інтерпретації окремих компонент та процесів, що являються вихідними даними для досліджень являється найбільш адекватним предметній області, по визначенню, в якій відповідна задача розв'язується,
- перетворення інтерпретаційних описів змінних, що приймають участь у відповідному аналізі, здійснюється у відповідності з перетвореннями, що описуються у вигляді логічних формул $L_i(x_{i1},...,x_{in})$, що формально описують відповідні компоненти,
- якщо перетворення $L_i(x_{i1},...,x_{in}) \to L_i^*(x_{i1}^*,...,x_{ik}^*)$ реалізовані коректно, то отриманий результат відповідає цілі перетворень, або $L_i^*(x_{i1}^*,...,x_{ik}^*) \to L_i^C(x_{i1}^*,...,x_{ik}^*)$,
- реалізації перетворень $L_i^F[j(x_{i1}),...,j(x_{ik})]$, які є зв'язаними з перетвореннями $L_i(x_{i1},...,x_{in})$, що записуються у вигляді наступного співвідношення:

$$L_i^F[j(x_{i1},...,x_{ik})] \infty L_i(x_{i1},...,x_{ik}), \quad (1)$$

- на перевірці семантичних параметрів результату перетворення, що здійснювалось з текстовими описами інтерпретації відповідних змінних, яке описується наступним співвідношенням:

$$L_i^F[j(x_{i1}),...,j(x_{ik})] \to L_i^{F*}[j(x_{i1}^*),...,j(x_{ik}^*)].$$

Що стосується першого фактору, то в ньому приймається сформульована умова, яка виконується при формуванні текстових описів інтерпретації всіх компонент і процесів, що описують предметну область, в якій передбачається розв'язувати задачу.

В якості базових перетворень, що використовуються в даному випадку, приймаються перетворення, що реалізуються засобами математичної логіки. Це обумовлюється наступними причинами.

1. Об'єктом перетворень можуть бути образи графічні, звукові, символічні та інші, структури яких важко, або не можливо описати формально математичними засобами, що вимагають досить точного представлення і відповідають ряду обов'язкових умов та обмежень. Наприклад, певний музичний твір, що в цифровому середовищі представляє собою досить складний сигнал, не може бути достатньо адекватно описаний системою диференціальних рівнянь, як деякий порівняно простий динамічний процес, що міняється в часі [7].

2. Текстові описи складаються з окремих елементів, якими є фрази. Фрази можуть представляти собою певну сукупність слів, яка характеризується цілим рядом семантичних параметрів, що їх характеризують. У зв'язку з цим, така фраза, як окремий елемент представляє собою досить складний об'єкт і, тому, фрази φ_i можуть знаходитися в досить складних залежностях між собою. Такі залежності, не зважаючи на те, що окремі параметри оцінюються чисельно, можна апроксимувати логічними функціями, оскільки останні надають найбільш широкі можливості у формуванні складних взаємозалежностей між різними φ_i і φ_j.

3. В рамках засобів математичної логіки існує досить широкий асортимент можливостей аналізу відповідних логічних структур, результати яких допускають інтерпретацію не тільки в рамках формалізму математичної логіки, а й в рамках предметної області, яка описується відповідними логічними формулами та засобами. Наприклад, таке уявлення в теорії абстрактної алгебри, як ідеал, значно складніше інтерпретувати в тій, чи іншій предметній області, яку можна було би описувати на рівні загальності, який вимагає формальний апарат абстрактної алгебри. Такі уявлення, як суперечність, конфлікт та ряд інших, що характерні для формалізму математичної логіки, значно простіше інтерпретувати в предметній області, для опису якої використовується логіка.

Основною особливістю представлення текстових описів є представлення фраз в послідовній формі та у використанні обмеженої кількості розділових

знаків, які несуть додаткову семантичну інформацію для користувача, що планує такою інформацією скористатися. В даному випадку, обмежимося двома розділовими знаками – точкою і комою. В рамках цієї особливості приймемо, що перетворення, які будуть здійснюватися в рамках $j(x_i)$, φ_i, чи ψ_i полягають у наступному:

- у перестановці фраз та слів місцями в межах текстового образу, що формально записується у вигляді:

$$(\varphi_{i1} * \varphi_{i2} * ... * \varphi_{in}) \to (\varphi_{i1} * \varphi_{ij} * ... * \varphi_{ij-1} * \varphi_{ij+1} * ... * \varphi_{in}),$$

- у включенні, чи виключенні окремих фраз у структуру тексту, що формально описується у вигляді:

$$\{[(\varphi_{i1} * \varphi_{i2} * ... * \varphi_{in}) \to (\varphi_{i1} * ... * \varphi_{ij-1} * \varphi_{ij+1} * ... * \varphi_{in})] \vee [(\varphi_{i1} * ... * \varphi_{in}) \to (\varphi_{i1} * ... * \varphi_{in+1} * ... * \varphi_{in})]\},$$

- у заміні окремих фраз в текстовому образі:

$$(\varphi_{i1} * \varphi_{i2} * ... * \varphi_{in}) \to (\varphi_{i1} * ... * \varphi_{ij-1} * \varphi_{in+1} * \varphi_{ij+1} * ... * \varphi_{in}).$$

Якщо обмежитися правилами виводу, що використовуються в математичній логіці, які не стосуються перетворень, що пов'язані з кванторами, то відповідні правила виводу обмежуються операціями введення, або елімінації змінних, що логічно зв'язані з формулою кон'юнкцією, диз'юнкцією та імплікацією, яка використовується в рамках правила виводу "modus ponens". Одномісний предикат заперечення $\neg$ має власну інтерпретацію в більшості натуральних мов у вигляді заперечення, наприклад, в українській мові це є слово "ні".

Всі перетворення з $L_i^F[j(x_{i1}),..., j(x_{ik})]$ реалізуються на основі аналізу вибраних семантичних параметрів, які відображаються в множину $\{0,1\}$ у відповідності з предикатами, що формуються на основі аналізу W_i. В цьому випадку, забезпечується необхідний рівень аналогії між $L_i(x_{i1},...,x_{in})$ та $L_i^F[j(x_{i1}),..., j(x_{ik})]$, що описується у співвідношенні (1).

Розглянутий семантичний підхід використовується для реалізації процесів укриття V_i в CS системою SS наступним чином. Відомо, що основою сприйняття довільної інформації користувачем, особливо, інформації з графічних образів, має семантичну основу. Тому, в рамках семантичного укриття V_i і створення відповідного SG_i, доцільно V_i перетворювати з допомогою семантичних перетворень таким чином, що семантика V_i повністю поглинулася семантикою стеганоконтейнера завдяки чому сформувалася б SG_i, в якій була б укрита V_i.

1. *Генне О.В.* Основные положения стеганографии // Защита информации. Конфидент, N3, 2000.

2. *Афанасьєва О.Ю., Дурняк Б.В.* Дослідження семантичних параметрів, що використовуються в графіці // Зб. наук. праць. ІПМЕ НАН України, N3, 2007.
3. Такеути Г. Теория доказательств. - М.: Мир, 1978.
4. *Борсуков В.С.* Стеганографические технологии защиты документов, авторских прав и информации // Обзор специальной техники. N 2, -2000.
5. *Акимов О.Е.* дискретная математика: логика, группы, графы, фракталы. - М.: Издатель АКИМОВА, 2005.
6. *Капітонова Ю.В., Кривий С.Л., Летичевский О.А., Луцький Г.М., Печурин М.К.* Основи дискретної математики. - Київ: Наукова думка, 2002.
7. *Стрижалюк Т.Г., Коновалова Н.Р.* Диференціальні рівняння. - Київ: Світ, 1997.

Поступила 11.02.2010р.

УДК 683.05

Б.Дурняк, К.Павелек

ОРГАНИЗАЦИЯ И ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИИ, ДЛЯ ФОРМИРОВАНИЯ СИСТЕМЫ ЗАЩИТЫ И ИССЛЕДОВАНИЕ ОТДЕЛЬНЫХ АСПЕКТОВ ФУНКЦИОНИРОВАНИЯ СИСТЕМЫ

Информационая технология, использующая семантику информационных компонент, позволяет создавать достаточно гибкие системы защиты авторских прав благодаря следующим факторам:

- семантика естественного языка, который используется, для описания информационных компонент системы, позволяет достаточно легко согласовывать информационные потоки между отдельными носителями системы защиты, например, между системой связанной с введением, считыванием и интерпретацией *CWZ* с подсистемой реализующей процедуры реагирования,
- поскольку в рамках системы *SZ* неизбежно участие социальных подсистем, то не требуется преобразования интерпретации выходных данных компьютерных подсистем и входных данных социальных подсистем, что не требует адаптации последней к используемой информационной технологии,
- поскольку продуктами интеллектуальной собственности, которые представлены в цифровой форме, могут быть не только продукты касающиеся информационных технологий, но и касающиеся художественных произведений, то использование текстовых представлений информационных компонент, позволяет достаточно легко отображать все аспекты функционирования *SZ* в форме доступной любому учаснику процесса защиты авторских прав, включая распространителей цифровых продуктов интеллектуальной деятельности,

