КОМП`ЮТЕРНІ СИСТЕМИ ТА КОМПОНЕНТИ

УДК 681.3:621.375

Структурна організація АЦП з прогресуючими тривалостями тактів порозрядного наближення

Азаров О.Д., Решетнік О.О., Муращеноко О.Г., Теплицький М.Ю.

Анотація: Проаналізовано принципи структурної організації пристроїв прискореного аналого–цифрового наближення для СЧВН {0, 1} та {1, -1}. Запропоновано структури АЦП порозрядного наближення із прогресуючими тривалостями тактів врівноваження з ваговою надлишковістю. Показано, що для побудови ЦАП із ваговою надлишковістю для таких АЦП можна використовувати серійні двійкові ЦАП. Запропоновано підхід щодо побудови АЦП з ваговою надлишковістю.

Аннотация: Проанализированы принципы структурной организации устройств ускореного аналого-цифрового приближения для СИВИ {0, 1} и {1, -1}. Предложены структуры АЦП поразрядного приближения с прогресирующими длительностями тактов уравновешивания с весовой избыточностью. Показано, что для создания ЦАП с весовой избыточностью для таких АЦП можно использовать существующие двоичные ЦАП. Предложен подход к постороению АЦП с весовой избыточностью.

Abstract: Sturctures of devices for fast analog to digital conversion for NSWR {0, 1} and {1, -1} are analized. Structures of SAR ADC with variable tact durations with weight redundancy are shown. Serial binary DACs can be used for uilding of DACs with weight redundancy. The method of building ADC with weight redundancy are prposed.

Ключові слова: АЦП, ЦАП, вагова надлишковість, прогресуючі тривалості тактів

Вступ

Переважна більшість сучасних порозрядних АЦП врівноваження будується на основі двійкової системи числення, а процес послідовного наближення організовано з постійною тривалістю тактів врівноваження. Багаторозрядні (14 – 18 розрядів) порозрядні АЦП відносять до пристроїв середньої швидкодії [1]. Водночас побудова АЦП порозрядного наближення з прогресуючими тривалостями тактів врівноваження на основі систем числення з ваговою надлишковістю дозволяє значно збільшити швидкодію цього класу пристроїв без значного збільшення апаратних затрат [2]. Чим відрізняється від відомого підходу на основі вагової надлишковості з використанням однакових тактів порозрядного врівноваження.

Актуальність

У теперішній час серед розробників перетворювачів форми інформації (ПФІ) зростає інтерес до використання вагової надлишковості в АЦП порозрядного наближення, з метою підвищення точності пристроїв, побудованих на неточних елементах. Перші публікації в цьому напрямку в колишньому СРСР і за кордоном з`явилися наприкінці 70-х на початку 80-х років минулого століття [2, 7]. Останнім часом кількість публікацій у США [7-9] у цьому напрямку значно зросла. Водночас, в Україні у Вінницькому національному технічному університеті дослідження у цьому напрямку продовжується з 90-х років двадцятого століття [2-6] і з`явились нові вагомі результати [10, 11]. Зокрема, реалізується комплексний підхід до одночасного підвищення точності та швидкодії ПФІ на неточних елементах. Особливо слід відзначити порозрядні АЦП з прогресуючими тривалостями тактів врівноваження, але такий підхід є новим і ще не відпрацьованим, тому аналіз структурних схем АЦП на основі систем числення з ваговою надлишковістю (СЧВН) є актуальним.

Мета

Метою досліджень є аналіз запропонованого методу структурно-функціональної організації АЦП із прогресуючим набором тривалостей тактів прискореного порозрядного наближення з ваговою надлишковістю.

Задачі

1. Аналіз підходів щодо структурно-функціональної організації схем АЦП для СЧВН {0, 1} та {1, -1}.
2. Аналіз можливостей реалізації компонентів АЦП.
3. Розглянути можливість побудови аналогових вузлів, зокрема, ЦАП на серійній елементній базі.

Розв`язання задач

Завдяки застосуванню прогресуючого набору тривалостей тактів врівноваження можна досягти значного скорочення загального часу врівноваження порівняно з порозрядним врівноваженням з однаковими тривалостями тактів врівноваженн. Наявність вагової надлишковості дозволяє застосовувати під час порозрядного врівноваження операцію порівняння з регульованою чутливістю, що дає

можливість різко (на 1-2 порядки) зменшити тривалість старших тактів та зменшити тривалість врівноваження у цілому.

Суть вагової надлишковості проявляється в тому, що сума ваг молодших розрядів більша або в крайньому випадку дорівнює вазі старшого розряду (вагова надлишковість проявляється таким чином як у системах із природнім базисом, так і в системах із штучним базисом):

$$\sum_{j=0}^{j=i-1} Q_j \geq Q_i \ ,$$

де Q_i - значення фізичної величини ваг розрядів (струм або напруга).

При цьому абсолютна вагова надлишковість визначається як:

$$\Delta Q_i = \sum_{j=0}^{j=i-1} Q_j - Q_i \ ,$$

Поняття відносної вагової надлишковості визначає можливий сумарний рівень відхилень ваг розрядів, при якому характеристика перетворення залишається нерозривною:

$$\delta Q_i = \frac{\sum_{j=0}^{j=i-1} Q_j - Q_i}{\sum_{j=0}^{j=i} Q_j} \ .$$

За умови постійності α для систем числення із природнім базисом справедливе таке співвідношення:

$$\delta Q_i \approx \frac{2-\alpha}{\alpha} \ .$$

Коли мають місце короткочасні зміни ваги розряду в процесі врівноваження, використовується поняття приведеної відносної вагової надлишковості [2]:

$$\tilde{\delta} Q_i = \frac{\sum_{j=0}^{j=i-1} Q_j - Q_i}{Q_i} \ .$$

Динамічна похибка першого роду існує короткий відрізок часу (протягом одного або декількох тактів) [2].

На рис. 1 наведено діаграму порозрядного аналого-цифрового перетворення з ваговою надлишковістю та звичайного двійкового порозрядного врівноваження з постійними тривалостями тактів. При цьому навіть для малого числа двійкових розрядів виграш по швидкодії є істотним.

Додатково підвищити швидкодію, тобто зменшити час перетворення можна за рахунок організації аналого-цифрового перетворення із прогресуючим набором тривалостей тактів врівноваження. На рис. 2. наведено діаграму врівноваження для прискореного [2] аналого-цифрового перетворення із прогресуючим набором тривалостей тактів з основою системи числення $\alpha = 1{,}618$ і числом розрядів $n = 7$. У цьому випадку на старших тактах врівноваження різниця між вхідним $A_{вх}(t)$ та компенсуючим $A_к(t)$ сигналами $\Delta A_{вх}(t) = A_{вх}(t) - A_к(t)$ потенційно може бути великою. На молодших же тактах значення цієї різниці зменшується. Відповідно чутливість схеми порівняння повинна поступово зростати від старших тактів ((n-1), (n-2), . . .) до молодших (. . ., 1, 0). Такий принцип регулювання чутливості схеми порівняння в АЦП дозволяє на старших тактах перетворення мати низьку чутливість, а високу швидкодію. Водночас на молодших тактах перетворення чутливість є високою, а швидкодія низькою

Структурна реалізація АЦП з прогресуючими тривалостями тактів порозрядного наближення з ваговою надлишковістю залежить від того, чи покращується точність чи швидкодія, від алгоритму за яким працює АЦП та від елементної бази. Узагальнивши увесь попередній досвід побудови АЦП із ваговою надлишковістю можна запропонувати такий підхід щодо структурно-функціональної організації таких АЦП [2, 14].

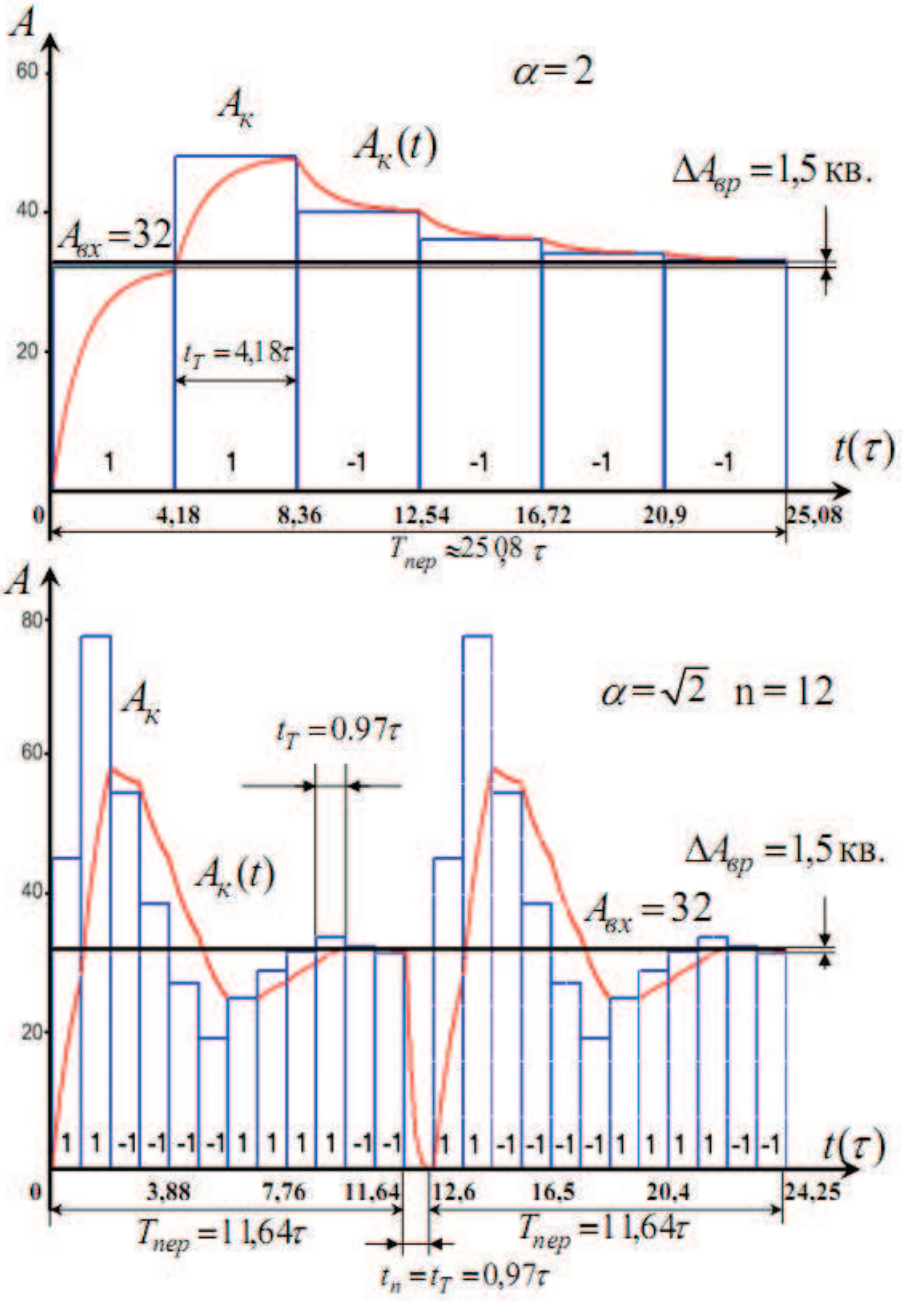


Рисунок 1 – Часова діаграма А-Ц перетворення: а) класичне $\alpha = 2$, б) прискорене $\alpha = \sqrt{2}$

Підхід щодо структурно-функціональної організації схем високопродуктивних порозрядних АЦП з ваговою надлишковістю.

1. Обираємо базис робочої системи числення [15]:
   а) якщо обрано систему числення з природним базисом, то необхідно використовувати спеціальний надлишковій ЦАП для цієї системи числення;
   б) якщо обрано систему числення із штучним базисом на основі двійкових рядів, то надлишковий ЦАП можна побудувати на базі декількох двійкових ЦАП.
2. Обираємо розрядні коефіцієнти системи числення:
   а) якщо обрано систему {1, 0}, то необхідно використовувати лише один надлишковий ЦАП, але обов’язково обирати алгоритм врівноваження з форсуючими сигналами та увести до структури блок форсуючих сигналів (БФС).
   б) якщо обрано систему {1, -1}, то необхідно використовувати два надлишкових ЦАП, що мають вихідні сигнали протилежної полярності.

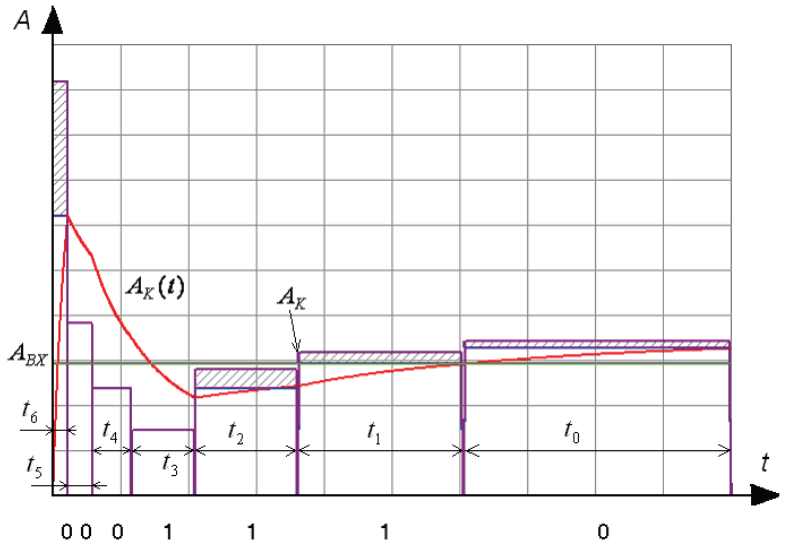


Рисунок 2 – Часова діаграма форсованого А-Ц перетворення із змінною тривалістю тактів врівноваження

3. Обираємо алгоритм врівноваження за яким працює пристрій, причому алгоритм врівноваження з форсуючими сигналами потребує спеціального блоку форсуючих сигналів:

а) для системи числення {1, -1} можна використовувати алгоритми:

1) тільки вмикання (антиглітчеве кодування).

2) прискорений з форсуючими сигналами.

б) для системи числення {0, 1} можна використовувати алгоритми:

1) прискорений з форсуючими сигналами.

2) адаптивний з форсуючими сигналами

4. Обираємо режим формування тривалостей тактів врівноваження.

а) якщо обрано врівноваження з постійними тривалостями тактів, то використовуються звичайний компаратор з постійною чутливістю та тактовий генератор для керування регістром послідовного наближеня.

б) якщо обрано прогресуючий набір тривалостей тактів врівноваження, то необхідно використовувати схему порівняння з регульованою чутливістю та спеціальний тактовий генератор для генерації прогресуючого набору тривалостей тактів врівноваження.

5. Обираємо варіанти застосування СЧВН

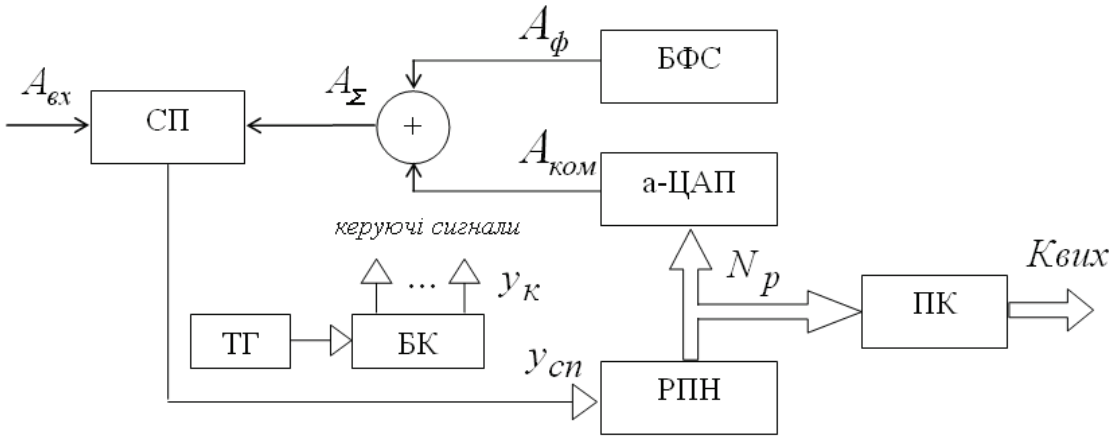
а) якщо необхідно підвищити точність перетворення пристрою побудованого на неточних елементах, то необхідно використовувати спеціальну процедуру самокалібрування ваг розрядів.

б) якщо необхідно підвищити швидкодію, то самокалібрування не використовується.

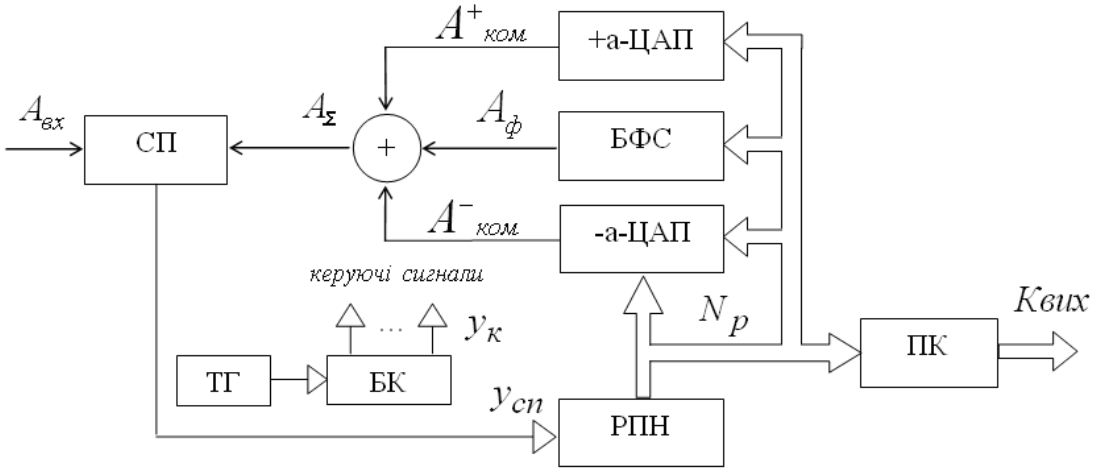
Незалежно від того які варіанти було обрано під час проектування, пристрій повинен містити стандартний набір елементів, притаманний порозрядним АЦП: регістр послідовного наближення, блок керування, перетворювач коду для перетворення робочого коду у двійковий код.

Для прикладу розглянемо структурну схему швидкодіючого АЦП порозрядного врівноваження з ваговою надлишковістю з прогресуючим набором тривалостей тактів врівноваження, який функціонує на базі СЧВН {0, 1} (рис. 3 а)). Пристрій містить: ТГ – тактовий генератор для генерування тактових імпульсів прогресуючої тривалості, БК – блок керування, а-ЦАП – ЦАП на основі системи числення з ваговою надлишковістю, БФС – блок форсуючих сигналів , РПН – регістр послідовного наближення, ”+” – суматор аналогових сигналів, СП – схема порівняння з регульованою чутливістю (змінними порогами порівняння) для організації процесу перетворення з прогресуючим набором тривалостей тактів врівноваження, ПК – перетворювач надлишкового коду N_p в двійковий код N_2. Більшість сучасних аналого-цифрових систем використовує двійкову систему числення. Таким чином АЦП на базі вагової надлишковості здійснює процес перетворення з використанням надлишкових систем числення та спеціальних алгоритмів перетворення, а інформація про результат перетворення передається на інші підсистеми як двійковий код.

АЦП для СЧВН {1, -1} відрізняється тим, що замість одного а-ЦАП використовуються +а-ЦАП та -а-ЦАП (рис. 3 б)).



а)



б)

Рисунок 3 – Структурна схема АЦП порозрядного врівноваження: а) СЧВН {0, 1}; б) СЧВН {1, -1}.

Такий пристрій може працювати навіть без БФС, оскільки використання СЧВН {1, -1} призводить до самокомпенсації похибок встановлення. Зазвичай АЦП на базі СЧВН {1, -1} працює за алгоритмом «тільки вмикання» [6], в результаті перетворення отримуємо код в якому немає нулів а є лише цифри {1, -1}. Проте варто зазначити, що пристрій може працювати за алгоритмом врівноваження в якому допускається залишати нулі у вихідному коді (це буде по суті трійкова СЧВН {1, 0, -1}), що дозволить покращити процес врівноваження та підвищити роздільну здатність.

Для реалізації запропонованої методики проектування структурних схем високопродуктивних порозрядних АЦП з ваговою надлишковістю, доцільно використовувати таблицю, що містить набір вузлів АЦП залежно від його виду (Табл. 1).

Водночас, слід відзначити, що уведення вагової надлишковості може бути реалізовано не тільки на основі відомих СЧВН [2], а й з використанням традиційних двійкових ЦАП [13]. Реалізація пристрою на основі запропонованого підходу може бути досить простою, оскільки відомі виробники електронних компонентів випускають моделі мікросхем, що містять в собі декілька двійкових ЦАП: AD5428, AD5433 та інші (Analog Devices) [12], DAC5652, DAC7558 (Texas Instruments), ISL5927, ISL5929 (Intesil). Проте, запропонований підхід є новим і недостатньо дослідженим.

Можна відзначити принаймні два підходи щодо побудови надлишкових ЦАП з нерозривною передатною характеристикою:

1) Побудова ЦАП на основі природного базису з ваговою надлишковістю з використанням спеціалізованої елементної бази;
2) Побудова ЦАП на основі штучного базису з ваговою надлишковістю з використанням серійних двійкових ЦАП.

Таблиця 1

Структурний склад АЦП з ваговою надлишковістю

Вид АЦП	СП	ТГ	РПН	а-ЦАП	-а-ЦАП	БФС	ПК	∑	БК
На базі СЧВН {0, 1}	+	-	+	+	-	+	+	+	+
На базі СЧВН {1, -1}	+	-	+	+	+	-	+	+	+
На базі СЧВН {1, -1} прискорений	+	-	+	+	+	+	+	+	+
На базі СЧВН {0, 1} з прогресуючим набором тривалостей тактів	+	+	+	+	-	+	+	+	+
На базі СЧВН {1, -1} з прогресуючим набором тривалостей тактів	+	+	+	+	+	-	+	+	+
На базі СЧВН {1, -1} прискорений з прогресуючим набором тривалостей тактів	+	+	+	+	+	+	+	+	+

Слід відзначити, що двійковий ЦАП може мати розриви на характеристиці перетворення [2], якщо ваги розрядів мають відхилення. ЦАП із ваговою надлишковістю, як правило, має нерозривну передатну характеристику (рис. 4).

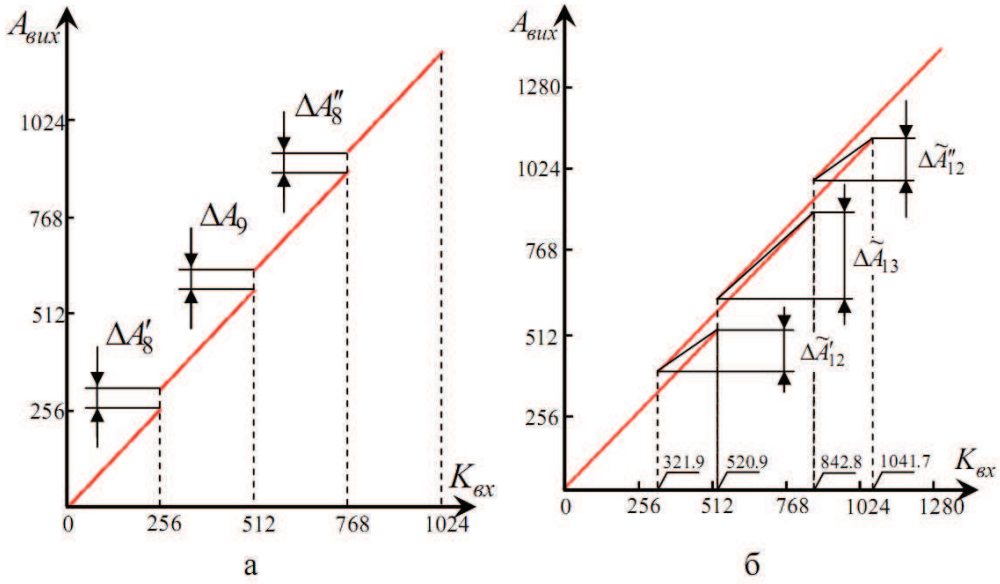


Рисунок 4 – Характеристики перетворення ЦАП за наявності статичних відхилень ваг розрядів у двох старших розрядах: а) $\alpha = 2$ і $n = 10$; б) $\alpha = 1.618$ і $n = 14$

Водночас методи побудови надлишкових ЦАП на базі двійкових ЦАП розділяються на чотири категорії:

1) ЦАП на базі двох однакових двійкових ЦАП з використанням суматора [13];
2) ЦАП на базі двох однакових двійкових ЦАП з використанням суматора і блока масштабування;
3) комбінований ЦАП на базі m однакових двійкових ЦАП з використанням суматора і блоків масштабування;
4) комбінований ЦАП на базі кількох однакових (або лише одного) двійкових ЦАП з використанням комутаційної схеми.

Для остаточного прийняття рішення про вибір якогось з них треба оцінити їх ефективність. При цьому виходимо з таких міркувань: порівняно з двійковою в надлишковій системі числення подовжується розрядна сітка [2]. Для рядів із сталим відношенням між сусідніми членами ступінь подовження визначається коефіцієнтом подовження:

$$\gamma_n = \frac{\ln 2}{\ln \alpha}.$$

Проте дана формула має обмеження. Її недоцільно використовувати при наближенні α до 1.0, а при $\alpha = 1.0$ її використання буде некоректним оскільки $\ln 1 = 0$. При цьому треба використовувати співвідношення:

$$\gamma_n = \frac{2^n}{n}.$$

Якщо ж відношення між сусідніми членами ряду не є сталим, то коефіцієнт подовження розрядної сітки можна визначити як відношення відповідної кількості надлишкових розрядів до кількості двійкових розрядів за умови однаковості діапазонів зображення чисел [2].

Ефективність застосування СЧВН можна оцінити через питому вагову надлишковість, яка враховує збільшення кількості обладнання, тобто відношення рівня вагової надлишковості до коефіцієнту подовження розрядної сітки:

$$E = \frac{\ln \alpha \left(\sum_0^{n-2} Q_i - Q_{n-1} \right)}{\ln 2 \sum_0^{n-1} Q_i}.$$

За результатами досліджень, найкраще значення ефективності використання вагової надлишковості E=0,206 для надлишкового ряду ваг розрядів з постійною основою $\alpha = \sqrt{2}$. Для інших варіантів значення E є меншими. Залежність відносної вагової надлишковості, коефіцієнту подовження розрядної сітки та питомої вагової надлишковості від основи системи числення α наведено на рис. 5. Причому криву коефіцієнту розрядної сітки нормовано по 0,1 по осі абсцис, тобто значення на графіку 0,2 відповідає коефіцієнту 2.

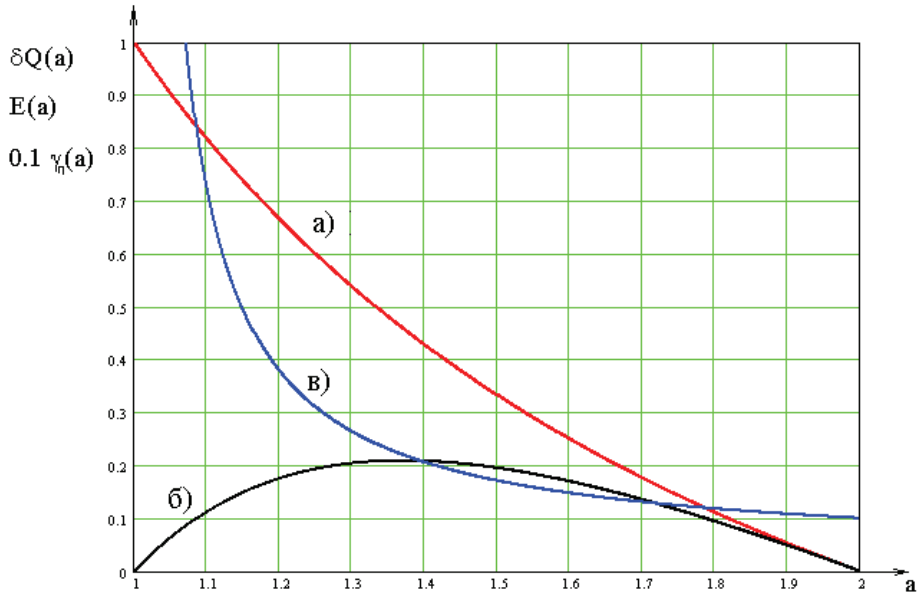


Рисунок 5 – Функціональні залежності: а) $\delta Q(\alpha)$, б) $E(\alpha)$, в) $\gamma_n(\alpha)$

Висновки

1. Запропоновано узагальнений підхід щодо структурно-функціональної організації схем АЦП порозрядного врівноваження для СЧВН {0, 1} та {1, -1}. В рамках запропонованого підходу можна синтезувати структурні схеми АЦП для організації процесу порозрядного врівноваження з прогресуючими тривалостями тактів зокрема.

2. Показано, що порозрядні АЦП з ваговою надлишковістю можна будувати на основі серійних мікросхем (блоків). Уцілому аналогова та цифрова частини АЦП з ваговою надлишковістю простіші порівняно з аналогічними високопродуктивними двійковими АЦП. Це пов’язано з тим, що зменшуються вимоги до точності формування ваг розрядів ЦАП.

3. Для отримання максимальної ефективності від застосування вагової надлишковості при побудові АЦП і ЦАП доцільно створити спеціалізовану елементну базу. Водночас навіть при

використанні для реалізації пристроїв серійної елементної бази можна отримати значний виграш швидкодії та підвищення точності порівняно з сучасними двійковими ПФІ та ПФІ з ваговою надлишковістю та постійною тривалістю тактів.

Список літератури

1. Analog-digital conversion / Edited by Walt Kester / Analog Devices Inc. 2004. 1230 pages.
2. Азаров О.Д. Аналого-цифрове перетворення на основі надлишкових систем числення з ваговою надлишковістю. Монографія. – Вінниця: ВНТУ, 2010. – 232 с.: ил/
3. Кадук О. В. Стратегії самокалібрування характеристики перетворення АЦП порозрядного кодування з ваговою надлишковістю / Азаров О. Д., Кадук О. В. // Вісник Вінницького політехнічного інституту. – 2008. – № 1 (76). – С. 102–110
4. Кадук О. В. Методи побудови АЦП порозрядного наближення, що самокалібруються / Азаров О. Д., Кадук О. В. // Оптико-електронні інформаційно-енергетичні технології. – 2008. – № 1 (15). – С. 56–66.
5. Азаров О. Д., Решетнік О. О., Гарнага В. А., Ратнюк В. В. Моделі форсуючих сигналів для прискореного порозрядного аналого-цифрового перетворен-ня з ваговою надлишковістю // Оптико-електронні інформаційно-енергетичні технології: 2006, №2.- С. 3-39с
6. Азаров О. Д., Решетнік О.О. Математична модель компаратора з регульованою чутливістю для швидкодіючого багаторозрядного АЦП із ваговою надлишковістю // Наукові праці Вінницького національного технічного університету.– 2008.– №1. Режим доступу: http://www.nbuv.gov.ua/e-journals/VNTU/2008-1/uk.htm
7. Patent 4336526, USA, H 03 K 13/05. Successive approximation analog-to-digital converter using non-binary series / Basil Wair. – Published Jun. 22, 1982.
8. Patent 7528761, USA, H 03 m 1/12. Analog/digital conversion using successive approximation and redundant weighting / Dieter Draxelmayr. – Published May. 5, 2009.
9. Patent 7046178, USA, H 03 m 1/10. Method and device for the calibration of a weighted network / Dieter Draxelmayr. – Published May. 16, 2006.
10. Азаров О. Д., Решетнік О. О., Крупельницький Л.В. Математичні моделі динамічних похибок І-го роду для швидкодіючих порозрядних АЦП із ваговою надлишковістю // Інформаційні технології та комп’ютерна інженерія: 2009, №2.- С. 8-14с
11. Азаров О. Д., Решетнік О. О., Шабатура М. Ю. Похибки усталення при аналого цифровому перетворенні з прогресуючими тривалостями тактів порозрядного врівноваження з ваговою надлишковістю // Інформаційні технології та комп’ютерна інженерія: 2009, №3.- С. 5-12 с.
12. Analog Devices inc. Dual, 12-/14-/16-bit, 1.0 GSPS D/A Converter AD9776/AD9778/AD9779. One Technology Way, P.O. Box, 9106, Norwood, MA 02062-9106, USA. 2005. – 56 p
13. Азаров О. Д., Решетнік О.О., Захарченко С.М., Лукащук О.О., Харьков О.М. Формування нерозривних передатних характеристик ЦАП і АЦП на основі вагової надлишковості // Інформаційні технології та комп`ютерна інженерія.– 2006.– N3(7).– С. 7-15с
14. Азаров О. Д., Кадук О. В., Дудник О. В., Муращенко О. Г. Пряме і зворотне перетворення «робочий код – цифровий еквівалент» в АЦП і ЦАП, що самокалібруються, з ваговою надлишковістю // Проблеми інформатизації та управління. – 2010. – №2. – С. 6-13.

Відомості про авторів

Азаров Олексій Дмитрович – д. т. н., професор, завідувач кафедри обчислювальної техніки; ВНТУ, Хмельницьке шосе, 95, м. Вінниця, 21021: тел. 58-02-25; e-mail: azarov1@vstu.vinnica.ua

Решетнік Олександр Олександрович – аспірант кафедри обчислювальної техніки, ВНТУ, Хмельницьке шосе, 95, м. Вінниця, 21021; тел. +380979693316: e-mail: de_gratnik@rambler.ru

Муращенко Олександр Геннадійович - аспірант кафедри обчислювальної техніки; ВНТУ, Хмельницьке шосе, 95, м. Вінниця, 21021: тел. 58-02-25; e-mail: oleksandr.g.m@gmail.com

Теплицький Михайло Юхимович - магістрант кафедри обчислювальної техніки; ВНТУ, Хмельницьке шосе, 95, м. Вінниця, 21021: тел. 0966555199; e-mail: misha_pirate@mail.ru

Вінницький національний технічний університет.

УДК 62-50

УПРАВЛЕНИЕ ДВИЖЕНИЕМ КОЛЕСНОЙ ПЛАТФОРМЫ С ТРЕМЯ СТЕПЕНЯМИ СВОБОДЫ

Краснодубец Л.А., д.т.н., проф., Морозова Е.Ф.

Аннотация: Ставится и решается задача аналитического конструирования закона адаптивного управления движением колесной платформы мобильного робота вдоль заданной траектории. Приводятся результаты моделирования системы автоматического управления. Предложена методика расчета параметров регулятора, реализующего сконструированный закон управления.

Анотація: Ставиться й вирішується задача аналітичного конструювання закону адаптивного керування рухом колісної платформи мобільного робота уздовж заданої траєкторії. Наводяться результати моделювання системи автоматичного керування. Запропонована методика розрахунків параметрів регулятора, який реалізує сконструйований закон керування.

Abstract: The problem of analytical designing the adaptive control law of the mobile robot wheeled platform along the given trajectory is put and solved. The results of the modeling of the automatic control system are included. The design procedure of regulator parametres realising the designed law of management is offered.

Актуальность проблемы

За последние 30 лет проблеме автоматизации управления движением технических объектов по заданной траектории посвящено множество статей. Наличие относительно недавних публикаций [1], [2], [3] и др. свидетельствует, что интерес исследователей к этой проблеме не ослабевает, и задача траекторного управления по-прежнему остается актуальной. В настоящей работе исследуется новый метод конструирования закона траекторного управления, который базируется на концепциях обратных задач динамики в сочетании с минимизацией функционалов, характеризующих энергию ускорения движения управляемых объектов в окрестности фазовых траекторий эталонных моделей [4]. При этом минимизация энергетического критерия осуществляется в процессе функционирования системы управления (локальная оптимизация). Существенно, что автоматические системы с такими алгоритмами можно отнести к классу адаптивных систем.

Целью работы является аналитическое конструирование адаптивного закона управления движением колесной платформы (КП) вдоль заданной траектории и исследование методом моделирования замкнутой системы автоматического управления.

Объект исследования

Объектом исследования является робототехническая система Robotino, которая предназначена для использования в учебном процессе при изучении инженерных дисциплин в области управления и автоматики. В состав этой системы входит мобильный робот, состоящий из приборной части и колесной платформы, имеющей три степени свободы. В состав приборной части входит процессорный модуль, комплекс сенсоров, включающий видеокамеру, разъемы для подключения дополнительных устройств и модуля доступа к беспроводной сети. Система поставляется в комплекте с RobotinoView - программным обеспечением (ПО) с графическим интерфейсом, который даёт возможность оператору контролировать текущее состояние колесной платформы в процессе ее движения и выполнять функции управления при помощи двусторонней связи с внешним компьютером.

Математическая модель колесной платформы

Математическая модель КП как объекта управления описывает движение центра масс мобильного робота, которое осуществляют три электропривода, обеспечивая всей механической системе три степени свободы. На вход модуля управления приводами КП поступают два управляющих сигнала в виде скоростей v_x, v_y движения платформы вдоль координатных осей OX и OY в горизонтальной плоскости и управляющий сигнал в виде угловой скорости ω вращательного движения мобильного робота вокруг вертикальной оси, проходящей через его центр масс. На рисунке 1 изображена схема, поясняющая вывод кинематических уравнений движения КП.

Рассматривая проекции скоростей центра масс КП на оси неподвижной системы координат, а также угол ψ, характеризующий взаимное расположение координатных систем, можно получить уравнения кинематики в виде

$$\left.\begin{aligned}\dot{x} &= u\cos\psi - v\sin\psi\\ \dot{y} &= u\sin\psi - v\cos\psi\\ \dot{\psi} &= \omega\end{aligned}\right\} \text{ или } \dot{q} = P(\psi)\vartheta, \text{ где } P(\psi) = \begin{bmatrix}\cos\psi & -\sin\psi & 0\\ \sin\psi & \cos\psi & 0\\ 0 & 0 & 1\end{bmatrix};$$

x и y – координаты местоположения платформы; ψ - угол, характеризующий угловое положение КП на плоскости; ω - угловая скорость вращательного движения КП вокруг вертикальной оси, проходящей через центр масс; $q = [x\ y\ \psi]^T$; $\vartheta = [u\ v\ \psi]^T$.

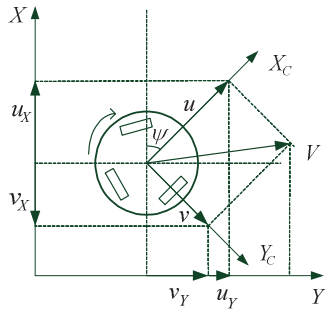


Рисунок 1. Схема, поясняющая вывод кинематических уравнений движения КП

Уравнения динамики получены в предположении, что перекрестные связи отсутствуют, на основе второго закона Ньютона и представлены в векторно-матричной форме

$$M\dot{\vartheta} = -D\vartheta + f_\vartheta , \tag{1}$$

$$M = \begin{bmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & J \end{bmatrix}, \; D = \begin{bmatrix} \mu & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \eta \end{bmatrix}, \; f_\vartheta = \begin{bmatrix} f_u \\ f_v \\ h \end{bmatrix};$$

где m – масса; μ - коэффициент трения при продольном движении; η - коэффициент сухого трения; f_u и f_v - проекции силы, приложенной к центру масс; J – момент инерции; h - вращательный момент.

Важно отметить, что при $\psi \approx 0$ матрица $P(\psi)$ вырождается в единичную и модель (1) преобразуется к виду

$$M\ddot{q} = -D\dot{q} + f , \tag{2}$$

где $f = [f_x \; f_y \; h]^T$.

Математическая постановка задачи управления движением КП

Пусть начальное состояние системы характеризуется значениями

$$t = 0 : q(0) = q_0, \dot{q}(0) = \dot{q}_0 . \tag{3}$$

Поставим задачу найти вектор управляющих функций $f = f(q, \dot{q})$ в форме обратных связей, который обеспечит перевод объекта управления (2) из начального состояния (3) в заданное стационарное состояние равновесия

$$q = \overline{q} = const, \dot{q} = 0 . \tag{4}$$

При этом потребуем, чтобы переходные процессы $q(t) \to \overline{q}$, $\dot{q}(t) \to 0$ проходили в малой окрестности эталонных процессов $q^*(t) \to \overline{q}$, $\dot{q}^*(t) \to 0$ описание которых в векторно-матричной форме имеет вид

$$\ddot{q}^* + A_1\dot{q}^* + A_0 q^* = A_0\overline{q} , \tag{5}$$

$$A_1 = \begin{bmatrix} \alpha_1 & 0 & 0 \\ 0 & \alpha_1 & 0 \\ 0 & 0 & \beta_1 \end{bmatrix}; \; A_0 = \begin{bmatrix} \alpha_0 & 0 & 0 \\ 0 & \alpha_0 & 0 \\ 0 & 0 & \beta_0 \end{bmatrix}; \; q^* = \begin{bmatrix} x^* \\ y^* \\ \psi^* \end{bmatrix}; \; \overline{q} = \begin{bmatrix} \overline{x} \\ \overline{y} \\ \overline{\psi} \end{bmatrix}.$$

Уравнение (5) описывает динамические свойства эталонной модели. При этом будем полагать, что каналы управления по координатам x и y идентичны. Коэффициенты $\alpha_0, \alpha_1, \beta_0, \beta_1$ назначаются такими, чтобы динамика эталонных моделей соответствовала требованиям к динамике проектируемой системы – по быстродействию, характеру переходных процессов и другим показателям.

Степень приближения процессов в системе и в эталонной модели будем оценивать величиной целевой функции

$$G(f) = \frac{1}{2}[\ddot{q}^* - \ddot{q}(t, f)]^2 , \tag{6}$$

где, $G(f) = [G(f_x) \; G(f_y) \; G(h)]^T$.

Компоненты вектора $G(f)$ представляют собой нормированные по массе или моменту значения энергии ускорения, вычисляемой в окрестности траектории движения эталонной модели. Чем

меньше величины компонентов вектора $G(f)$, тем в большей степени процессы в системе приближаются к процессам в эталонной модели.

Конструирование закона управления

Решение поставленной задачи будем выполнять в соответствии с рекомендациями [4] и [5] путем минимизации целевой функции методом скоростного градиента, которому соответствует дифференциальное уравнение

$$\frac{df}{dt} = -L\frac{dG}{df}, \tag{7}$$

где L - диагональная матрица, элементы λ_x, λ_y, λ_ψ которой характеризуют скорость стремления значений критериальной функции $G(f)$ к $G_{\min}(f)$.

Дифференциальный закон управления для регулятора можно получить, если в (7) подставить значение производной $\frac{dG}{df}$, найденное с учетом уравнений (2) и (5) в виде

$$\frac{dG}{df} = 2\frac{1}{2}(\ddot{q}^* - \ddot{q})\frac{d}{df}(-\ddot{q}) = -A^{-1}(\ddot{q}^* - \ddot{q}).$$

Далее имеем

$$\frac{df}{dt} = LA^{-1}[A_0(\overline{q} - q^*) - A_1\dot{q}^* - \ddot{q}]. \tag{8}$$

Сделав замену $q^* = q$, $\dot{q}^* = \dot{q}$ (что означает замену информационных связей регулятора с эталонной моделью соответствующими обратными связями, обеспечивающими регулятор информацией о текущих значениях компонент вектора управления q и его производной $\dot{q}$) и, проинтегрировав обе части уравнения (8) при нулевых начальных условиях, приводим искомый закон управления к виду

$$f = LA^{-1}[A_0\int_0^t(\overline{q} - q)dt - A_1 q - \dot{q}]. \tag{9}$$

В соответствии с (9) в структуру регулятора вводится диагональная матрица $K = LA^{-1}$, элементы которой - коэффициенты усиления $k_x = \frac{\lambda_x}{m}, k_y = \frac{\lambda_y}{m}, k_\psi = \frac{\lambda_\psi}{J}$ в соответствующих каналах управления.

Структурная интерпретация полученного закона управления (9) представлена на рисунке 2.

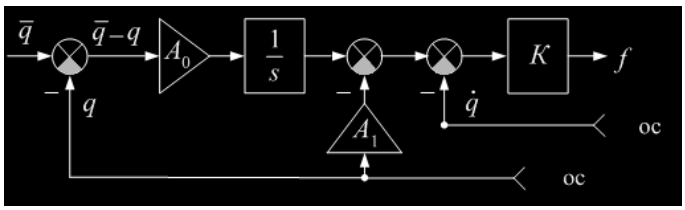


Рисунок 2. Структурная интерпретация закона управления

Расчёт параметров регулятора выполнен по формулам, которые получены при помощи канонической формы записи колебательного звена

$$\alpha_0 = \frac{3...5}{t^2_{p_{x,y}}},\ \alpha_1 = \frac{3...5}{t_{p_{x,y}}},\ \beta_0 = \frac{3...5}{t^2_{p_\psi}},\ \beta_1 = \frac{3...5}{t_{p\psi}}, \tag{10}$$

где t_p - время желаемой переходной характеристики проектируемой системы управления;

$\varsigma = 0.707$ - соответствует желаемому перерегулированию, которое не превышает 5%.

Элементы входящей в закон управления (9) матрицы K (коэффициенты усиления в каналах управления), как показано в [4], существенно влияют на степень приближения динамических свойств проектируемой системы управления к соответствующим свойствам эталонной модели. При стремлении коэффициентов усиления к бесконечности модель системы управления совпадает с эталонной моделью. По этой причине элементы матрицы K подбираются при моделировании.

Моделирование синтезированной системы управления

В качестве средства моделирования выбрана среда Simulink. Вначале моделировалось управляемое движение КП из начального состояния (3) в заданное стационарное состояние равновесия (4). Результатом моделирования является траектория в виде прямой.

Графики переходных процессов в каналах управления и их эталонные модели при коэффициентах усиления $k_x = k_y = k_\psi = 1$ и $k_x = k_y = k_\psi = 10$ приведены на рисунках 3-8, где ***сс*** – спроектированная система, ***эм*** – эталонная модель. Из приведенных графиков видно, что при увеличении значений коэффициентов усиления в каналах управления (элементов диагональной матрицы K) процессы в системе приближаются к эталонным моделям. Дальнейшее увеличение этих коэффициентов на картину процессов не влияет. При этом следует отметить, что коэффициенты k_x и k_y выбираются одинаковыми, так как соответствующие каналы управления по предположению являются идентичными. Необходимая величина коэффициента усиления k_ψ в канале вращательного движения может существенно отличаться от значений усиления в каналах продольного движения.

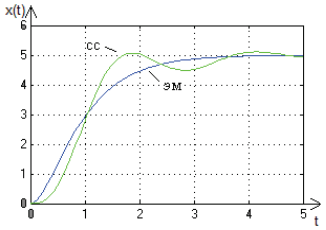


Рисунок 3 –Переходный процесс в канале управления координатой $x(t)$ и его эталонная модель ($k_x = 1$).

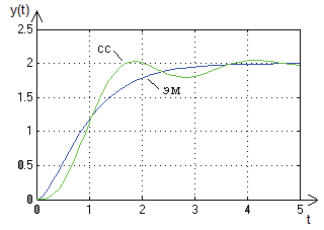


Рисунок 4 –Переходный процесс в канале управления координатой $y(t)$ и его эталонная модель ($k_y = 1$).

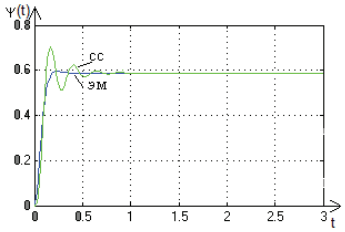


Рисунок 5 - Переходный процесс в канале управления вращательным движением и его эталонная модель ($k_\psi = 1$).

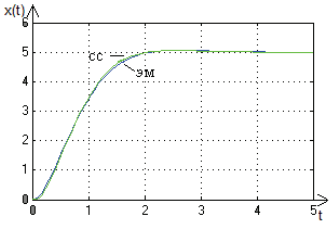


Рисунок 6 –Переходный процесс в канале управления координатой $x(t)$ и его эталонная модель ($k_x = 10$).

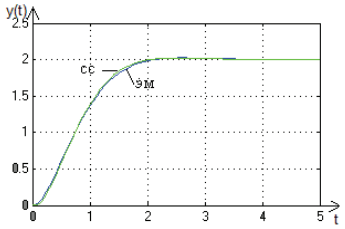


Рисунок 7 - Переходный процесс в канале управления координатой $y(t)$ и его эталонная модель ($k_y = 10$).

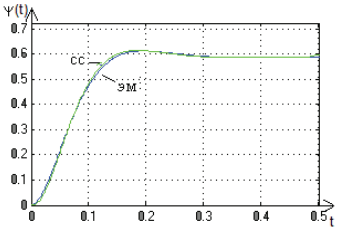


Рисунок 8 - Переходный процесс в канале управления вращательным движением и его эталонная модель ($k_\psi = 10$).

Далее моделировалось управляемое движение КП при действии постоянных параметрических возмущений в виде десятикратного увеличения массы и момента инерции. Из рисунков 9, 10 и 11, где изображены графики процессов управления в системе с координатными (по координате $x(t)$) и параметрическими (по массе и моменту инерции) возмущениями при неизменных параметрах регулятора, следует, что существенное изменение параметров объекта управления мало влияет на динамические свойства синтезированной системы.

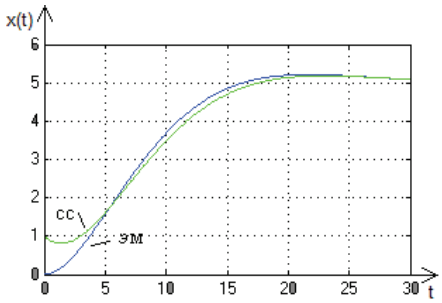


Рисунок 9 - Переходный процесс в канале управления координатой $x(t)$ и его эталонная модель ($k_x = 10$).

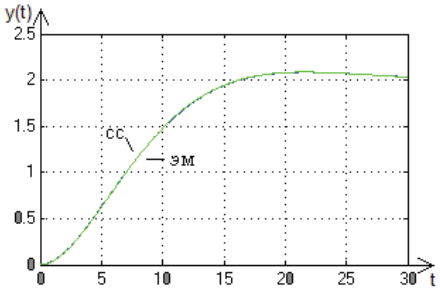


Рисунок 10 – Переходный процесс в канале управления координатой $y(t)$ и его эталонная модель ($k_y = 10$).

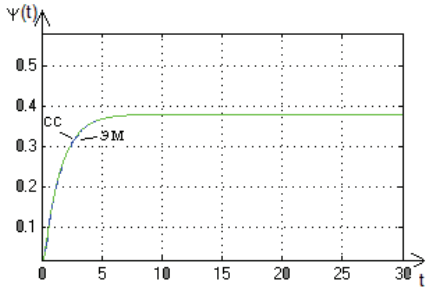


Рисунок 11 - Переходный процесс в канале управления вращательным движением и его эталонная модель ($k_\psi = 10$).

Синтезированный регулятор может так же обеспечить движение КП по траектории, отличной от прямой. При этом требуемые значения состояний $\bar{q} = [\bar{x} \quad \bar{y} \quad \bar{\psi}]^T$ в законе управления (6) следует заменять соответствующими функциями, описывающими траекторию задаваемых поступательного и вращательного движений КП. В качестве примера рассматривалось движение КП по окружности. При этом траектория требуемого движения на плоской поверхности задавалась параметрическими уравнениями

$$\left.\begin{aligned} \bar{x}(t) &= r \sin \omega t + x_c \\ \bar{y}(t) &= r \cos \omega t + y_c \end{aligned}\right\}$$

с коэффициентами $r = 1.5\,м$, $\omega = 0.3\,рад/с$, $y_c = 2\,м$, $x_c = 5\,м$. Движение КП по заданной окружности можно осуществить двумя способами.

Способ первый. Управление движением осуществляется только по двум независимым каналам путем формирования управляющих сигналов f_x и f_y . При этом принималось $h = 0$. На рисунке 12 приведен график траектории выведения КП из начального состояния (3) в режим установившегося движения с заданной угловой скоростью ω по окружности с радиусом r при $k_x = k_y = 1$ и отсутствии возмущений. На рисунке 13 приведен график управляемого движения КП при аналогичных условиях, но с увеличенными коэффициентами усиления $k_x = k_y = 10$.

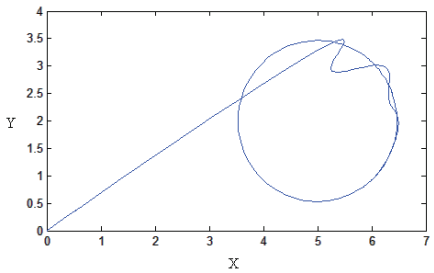


Рисунок 12 – Движение КП по траектории в виде окружности при $k_x = k_y = 1$.

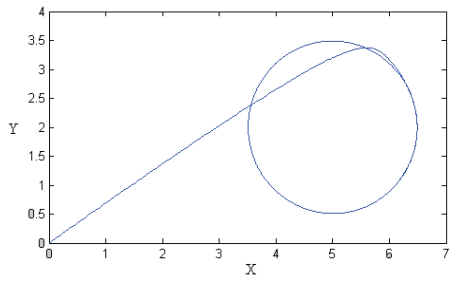


Рисунок 13 – Движение КП по траектории в виде окружности при $k_x = k_y = 10$.

Во втором случае управление КП осуществлялось по трем степеням свободы. При этом требуемое значение текущего углового положения КП определялось соотношением

$$\overline{\psi}(t) = arctg\frac{f_y(t)}{f_x(t)}.$$

На рисунках 14 и 15 приведены графики изменения координаты $\psi(t)$ при выведении КП на установившийся режим движения по траектории в виде окружности при $k_\psi = 1$ и $k_\psi = 10$ соответственно

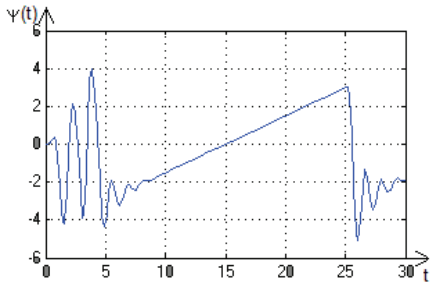


Рисунок 14 – Переходный процесс в канале управления координатой $\psi(t)$ при $k_\psi = 1$.

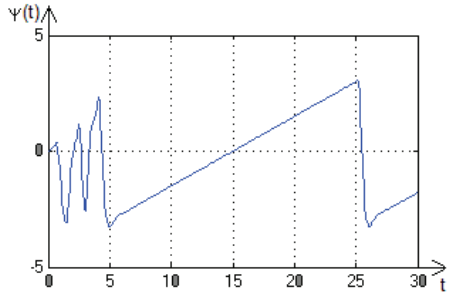


Рисунок 15 – Переходный процесс в канале управления координатой $\psi(t)$ при $k_\psi = 10$.

Как отмечалось, на вход приводов КП Robotino поступают управляющие сигналы в виде скоростей v_x, v_y движения платформы в горизонтальной плоскости, а также управляющий сигнал в виде угловой скорости ω вращательного движения мобильного робота. Поэтому для технической реализации сконструированного закона управления в рассматриваемом случае необходимо проинтегрировать вектор - функцию управления (9), компоненты которой имеют физический смысл силы и момента.

Выводы

При решении задачи аналитического конструирования адаптивного регулятора не использовались данные о параметрах объекта управления. Для синтеза закона управления движением КП достаточно лишь информации о порядке ее динамической модели.

Исследование проектируемых систем при наличии существенных параметрических и координатных возмущений показало, что они обладают выраженными робастными свойствами.

Список литературы

1. A. P. Aguiar and J. P. Hespanha. Position tracking of underactuated vehicles. In Proc. of the 2003 Amer. Contr. Conf., Denver, CO, USA, June 2003.
2. F. Alonge, F. D’Ippolito, and F.M. Raimondi. Trajectory tracking of underactuated underwater vehicles. In Proc. 40th IEEE Conf. on Decision and Contr., Orlando, Florida, USA, December 2001.
3. A. Behal, D.M. Dawson, W.E. Dixon, and Y Fang. Tracking and regulation control of an underactuated surface vessel with nonintegrable dynamics. IEEE Trans. on Automat. Contr., 47(3):495–500, March 2002.
4. Крутько П.Д. Новые технологии аналитического проектирования алгоритмического обеспечения систем управления движением // Управление, автоматизация и окружающая среда: Материалы международ. науч.-техн. конф., г. Севастополь, 8-13 сентября 2008 г. – Севастополь: Изд-во СевНТУ, 2008.
5. Антонов В.Н. Адаптивное управление в технических системах / В.Н. Антонов, В.А. Терехов, И.Ю. Тюкин.- С-Пб: Издательство С.-Петербургского университета, 2001.-244с.

Сведения об авторах

Краснодубец Леонид Андреевич - д.т.н., профессор, заведующий кафедрой технической кибернетики, Севастопольский национальный технический университет, ул. Университетская, 33, г.Севастополь, 99033, тел.(067)600-22-67, lakrasno@gmail.com.

Морозова Екатерина Федоровна – магистрант кафедры технической кибернетики Севастопольского национального технического университета, ул. Университетская, 33, г.Севастополь, 99033, тел.(050)924-39-25, kat_uk1@inbox.ru.

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ТА ОБЧИСЛЮВАЛЬНІ МЕТОДИ

УДК 004.9

МОДЕЛЮВАННЯ ПАРАМЕТРІВ ТРАНСПОРТНОЇ МЕРЕЖІ В СЕРЕДОВИЩІ АВТОМАТИЗОВАНОЇ СИСТЕМИ ПОШУКУ ОПТИМАЛЬНИХ РІШЕНЬ

Мокін В.Б., Романюк О.Н., Войтко В.В., Сторчак В.Г., Гавенко О.В.

Анотація: Розроблено новий метод оптимізації параметрів руху автомобільного транспорту з урахуванням межових змін параметрів. Проведено моделювання параметрів транспортних потоків у середовищі автоматизованої системи пошуку оптимальних рішень.
Ключові слова: транспортна мережа, моделювання, автоматизована система, граф.
Аннотация: Разработан новый метод оптимизации параметров движения автомобильного транспорта с учетом предельных изменений параметров. Проведено моделирование параметров транспортных потоков в среде автоматизированной системы поиска оптимальных решений.
Ключевые слова: транспортная сеть, моделирование, автоматизированная система, граф.
Abstract: A new method for parameter optimization of road traffic, taking into account boundary changes in the parameters. The modeling of traffic flow parameters in the environment of the search for optimal solutions.
Keywords: transportation network, modeling, automated system, graph.

Вступ

Сучасний етап розвитку інформаційних технологій дозволяє оптимізувати транспортні потоки у процесі моніторингу та управління розподілом завантаженості мережі автодоріг міста. Важливу роль тут відіграє впровадження новітніх систем комп’ютерного керування транспортними потоками. Тому актуальною є задача розробки та дослідження нових автоматизованих систем пошуку оптимальних рішень у процесі розв’язання транспортних задач [1-4].

Мета дослідження

Метою дослідження є оптимізація транспортних потоків сучасної мережі автодоріг міста. Під об'єктом дослідження розуміємо процеси формування та розподілу транспортних потоків. Предметом дослідження постають методи оптимізації руху транспортних засобів.

Постановка задачі

Основними задачами роботи є розробка та дослідження режимів роботи автоматизованої системи пошуку оптимальних шляхів у процесі керування розподілом завантаження мережі автодоріг шляхом моделювання параметрів транспортних потоків. Транспортна мережа формалізується у вигляді орієнтованого графа, де вершини графа презентують наявні перехрестя, а дуги графа – дороги між ними [1-4].

Оптимізація транспортних потоків передбачає реалізацію методу ефективного розподілу транспортних засобів, який орієнтований на виконання двох базових етапів розв’язання задачі. На першому етапі шляхом моніторингу мережі створюється її графічна модель, задаються вузлові параметри з урахуванням динамічної зміни їх значень у режимі реального часу та визначаються критерії оптимальності пошукових процесів. Другий етап акумулює розрахункові операції у процесі прийняття оптимальних рішень у середовищі автоматизованої системи та формує низку рекомендацій, спрямованих на оптимізацію транспортних потоків.

Необхідно розробити новий метод оптимізації параметрів руху автомобільного транспорту, який дозволить розрахувати за яких змін параметрів транспортної мережі можливе досягнення оптимальних або максимально допустимих значень тривалості руху транспорту по заданому маршруту. Також, слід розробити програмне забезпечення для проведення моделювання за допомогою цього методу на практиці.

Розробка засобів розв’язання транспортних задач

Моделювання параметрів транспортної мережі здійснюється в середовищі автоматизованої системи пошуку оптимальних рішень, яка базується на використанні засобів теорії графів у процесі формування математичної моделі та дослідження параметрів транспортних потоків. Побудова адекватної транспортної моделі передбачає моніторинг стану мережі автодоріг з урахуванням динамічної зміни значень її параметрів. Розроблені засоби автоматизації процесу розв’язання транспортних задач базуються на побудові та ідентифікації гібридної моделі, в основу якої покладено ймовірнісний підхід з елементами моделей аналогів та принцип слідування за лідером. Для кожного типу елементів

транспортної мережі передбачається побудова та ідентифікація окремої моделі, яка дозволяє адекватне відтворення процесів формування транспортних потоків з урахуванням залежностей між густиною наявних транспортних засобів і допустимою швидкістю руху автомобілів.

Серед базових характеристик транспортної мережі виділимо:

P – *пропускну здатність* (авт./год) автодороги;

H – *щільність* – кількість автомобілів на одиницю довжини автодороги (авт./км);

I_R – *інтенсивність руху* – кількість автомобілів, що проїжджає через переріз дороги за одиницю часу (авт./год);

I_z – *інтенсивність утворення заторів* (авт./год).

Очевидно, що $I_z = I_R - P$. При $I_z>0$ довжина затору буде збільшуватися, а при $I_z<0$ – зменшуватися на величину I_P*t. У випадку, коли $I_z=0$, усі машини встигатимуть вчасно подолати досліджувану ділянку дороги, що дозволить уникнути явища формування заторів.

Світовий досвід показує, що повністю вирішити проблему формування заторів у великих містах складно, проте розробка і використання новітніх засобів оптимізації транспортного руху дозволяє обґрунтовано прогнозувати значне підвищення експлуатаційних характеристик наявних транспортних мереж.

Одним із найбільш важливих структурних елементів транспортних мереж є перехрестя, оптимізація розподілу транспортних засобів на якому відіграє ключову роль у формуванні інтегрального показника сукупної ефективності роботи загальної транспортної мережі.

До головних засобів оптимізації руху транспорту на перехрестях відносимо:

- методи оптимізації транспортних потоків за рахунок їх оперативного перерозподілу на обрані гілки мережі;
- регулювання режимів роботи світлофорів;
- систему позиціювання дорожніх знаків;
- можливі засоби розширення робочої зони доріг.

Зауважимо, що планова реалізація останнього фактору з урахуванням показників фінансових витрат у реальній транспортній системі часто є неможливою. Тому, особливої уваги заслуговує дослідження перших трьох груп засобів оптимізації транспортного руху на перехрестях.

Серед базових критеріїв оптимальності при дослідженні параметрів транспортних потоків на перехрестях виділимо максимальну інтегральну пропускну спроможність перехрестя, максимальну пропускну спроможність автодорожніх ліній на заданих напрямках руху, мінімально можливе значення показника простою автомобіля на перехресті; мінімальне значення інтенсивності утворення заторів на заданих (чи усіх) напрямках транспортного руху на перехресті.

Структурна схема перехрестя (рис. 1) дозволяє записати його узагальнену математичну модель, де X_O, X_S, X_Z, X_R є множинами параметрів, за допомогою яких можна описати методи оперативного перерозподілу транспортних потоків за результатами пошуку оптимального шляху, роботу світлофорів, систему позиціювання знаків та обґрунтування рекомендованої ширини робочі зони доріг. Введемо інтегральний критерій оптимальності y: $y=F(X_O,X_S,X_Z,X_R)$.

Процес оптимізації руху транспорту на перехресті зводиться до знаходження таких аргументів критерію y, при яких виконується умова:

$$y \geq K_{opt},\ y=F(X_O,X_S,X_Z,X_R), \tag{1}$$

де K_{opt} – це гранично допустиме значення інтегрального критерію оптимальності y. Якщо $y < K_{opt,}$ то допустимого значення критерію оптимальності досягти неможливо.

Для розв'язання поставленої задачі, спрямованої на досягнення умови (1), пропонується використати відомий підхід до розв'язання оберненої задачі чутливості, орієнтованої на визначення рівня можливої зміни значень параметрів досліджуваної системи для досягнення заданих межових значень критерію оптимальності [5].

У даному випадку обернену задачу чутливості пропонуємо розв'язувати з використанням принципу рівних впливів [6], за яким часткові диференціали $\frac{\partial y}{\partial x_j}\cdot\Delta x_j$, де ($_{j=\overline{1,n}}$) однаково впливають на утворення загального відхилення цільової функції Δy:

$$\frac{\partial y}{\partial x_1}\Delta x_1 = \frac{\partial y}{\partial x_2}\Delta x_2 = \ldots = \frac{\partial y}{\partial x_n}\Delta x_n = \frac{\Delta y}{n} \tag{2}$$

або $\Delta y_1 = \Delta y_2 = \ldots = \Delta y_n = \frac{\Delta y}{n}$, де Δy_j — частка відхилення Δy, що спричиняється Δx_j, $j = \overline{1, n}$.

Подамо зв'язок між оцінюваною функцією і параметрами системи у вигляді:

$$y = \sum_{i=1}^{m} a_i \prod_{j=1}^{n} x_j^{\alpha_{ji}} \tag{3}$$

де m — кількість доданків полінома (3); n — кількість змінних (параметрів) x_j; a_i — додатні коефіцієнти функціоналу; α_{ji} — постійні коефіцієнти.

Для визначення відхилення Δx_j скористаємось безрозмірною формою запису, що отримується шляхом ділення цільової функції (3) на її оптимальне значення (y_o):

$$y_* = \sum_{i=1}^{m} \varpi_{io} \prod_{j=1}^{n} x_{j^*}^{\alpha_{ji}}, \tag{4}$$

де $y_* = \frac{y}{y_o}$, $x_{j^*} = \frac{x_j}{x_{jo}}$ — відносні значення відповідно функції та параметрів; ϖ_{io} — оптимальна пронормована вага і-го доданку полиному, знайдена за методом інтегральних аналогів [7]: $\varpi_{io} = \frac{a_i \prod_{j=1}^{n} x_j^{\alpha_{ji}}}{y}$, $i = \overline{1, m}$. Такий перехід від абсолютної системи одиниць до відносної уможливлюється використанням правила Фур’є [7], за яким усі члени рівняння, що описує будь-яке фізичне явище і подане у вигляді сум однорідних функцій в абсолютній системі одиниць, повинні мати однакову розмірність. За принципом рівних впливів вважаємо складові y_{j^*}, які спричиняються впливом x_{j^*}, $j = \overline{1, n}$, рівними і визначаємо їх за виразом:

$$y_{j^*} = \frac{y_o + \frac{\Delta y}{n}}{y_o} = 1 + \frac{1}{n}(y_* - 1). \tag{5}$$

Для знаходження x_{j^*} складаємо систему рівнянь:

$$1 + \frac{j}{n}(y_* - 1) = \sum_{i=1}^{m} \varpi_{io} \prod_{k=1}^{j-1} x_{k^* з}^{a_{ki}} \prod_{k=j+1}^{n} x_{k^* o}^{a_{ki}} \cdot x_{j^*}^{\alpha_{ji}}, \quad j = \overline{1, n} \tag{6}$$

де $x_{k^* з}$ — відносні значення змінних, знайдені з попередніх рівнянь системи (6); $x_{k^* o} = 1$ — відносні оптимальні значення змінних; x_{j^*} — шукане відносне значення змінної x у j-му рівнянні; y_* — задане відносне значення відхилення цільової функції від оптимуму.

Система рівнянь (6) виявляє залежності впливів змінних параметрів на значення цільової функції. Перше рівняння такої системи визначає відносне відхилення першої шуканої змінної x_{1^*} даної цільової функції для формування відхилення y_{1^*}. Кожне наступне рівняння простежує вкраплення чергових шуканих змінних x_{j^*}, $j = \overline{2, n}$ і відповідний приріст узагальненого критерію оптимальності y_{j^*}, $j = \overline{1, n}$.

Розглянемо приклад моделі оптимізаційного розподілу транспортних засобів на перехресті, керованому світлофорами, у випадку, коли поворот наліво заборонений для кожного напрямку руху. Серед наявної системи обмежень параметрів моделі виділимо умову, яка для кожного напрямку визначає інтенсивність руху транспорту не меншою від пропускної здатності дороги (рис. 2).

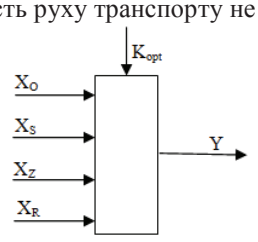


Рисунок 1 – Структурна схема перехрестя

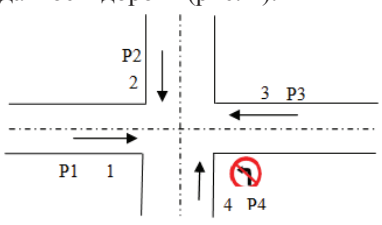


Рисунок 2 – Приклад моделі оптимізації розподілу транспортних засобів на перехресті

За визначенням пропускної здатності маємо:

$$P_{p.}=K_p/t; \qquad P_{pmax.}=Max(P_p),$$

де K_p – сумарна кількість машин, які проїжджають через перехрестя за фіксований проміжок часу t, який дорівнює циклу роботи світлофора. Через $t_з$ позначимо робочий час відображення зеленого світла для напрямків 1 та 3, тоді час активації зеленого світла для напрямків 2 та 4 буде дорівнювати $(t-t_з)$.

Кількість машин, які проїжджають за період циклу роботи світлофора, визначається виразом:

$$K_p=(P_1+P_3)t_з+(P_2+P_4)(t-t_з).$$

Тоді для розрахунку пропускної здатності дороги отримаємо вирази:

$$P_{p.}= ((P_1+P_3)t_з+(P_2+P_4)(t-t_з))/t=(P_1+P_3)t_з/t+(P_2+P_4)(1-t_з/t).$$

$$P_p= P_2+P_4+(P_1+P_3- P_2-P_4)t_з/t .$$

$$P_{pmax.}=Max(P_2+P_4+(P_1+P_3- P_2-P_4)t_з/t)=Max(P_2+P_4)+ Max((P_1+P_3- P_2-P_4)t_з/t).$$

Оскільки $Max(P_2+P_4)= P_2+P_4$, то формула визначення максимальної пропускної здатності матиме вигляд:

$$P_{pmax.}= P_2+P_4+ Max((P_1+P_3- P_2-P_4)t_з/t).$$

Звідки $Max((P_1+P_3- P_2-P_4)t_з/t)= \begin{Bmatrix} P_1+P_3-P_2-P_4, при\ P_1+P_3-P_2-P_4>0\ і\ t_з=t \\ 0, при\ P_1+P_3-P_2-P_4\le 0\ і\ t_з=0 \end{Bmatrix}$.

Тоді $P_{pmax.}= \begin{Bmatrix} P_1+P_3, при\ P_1+P_3-P_2-P_4>0\ і\ t_з=t \\ P_2+P_4, при\ P_1+P_3-P_2-P_4\le 0\ і\ t_з=0 \end{Bmatrix}$, тобто $P_{pmax}=Max(P_1+P_3,P_2+P_4)$.

Отже, максимально можлива пропускна здатність роботи перехрестя дорівнює максимальному значенню суми пропускних здатностей доріг протилежних напрямків руху. Якщо інтенсивність руху машин на кожному напрямку не перевищує пропускної здатності траси, то для досягнення максимальної пропускної здатності перехрестя зелене світло світлофора повинно горіти в межах повного робочого циклу. Інтенсивність прибування автомобілів до перехрестя визначається виразом: $I_{pp}= I_{R1+} I_{R2}+ I_{R3}+ I_{R4.}$ Тоді сумарна інтенсивність утворення затору на перехресті описується виразом: $I_{zp}= I_{pp} -P_{pmax.}$ У випадку, коли визначений показник $I_{zp}>0$, сумарна довжина затору на перехресті буде збільшуватися, що є обґрунтованим сигналом для використання різних факторів впливу на оптимізацію руху транспортних засобів у досліджуваній мережі автодоріг.

Транспортна мережа подається у вигляді орієнтованого графа (рис. 3), де вершини графа презентують наявні перехрестя, а дуги – дороги між ними. Інформація про дозвіл руху автотранспорту обраною віткою мережі зберігається в базі додаткових параметрів графа. Такий підхід спрощує процес маніпулювання графічною моделлю транспортної мережі в середовищі автоматизованої системи пошуку оптимальних рішень.

Розподіл автомобілів на перехресті з номером 5 можна подати у вигляді двомірного масиву ймовірностей R розміром 4x3 (рис. 4).

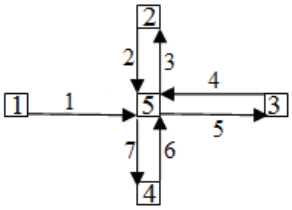


Рисунок 3 – Подання фрагменту транспортної

Номер дуги	3	5	7
1	R11	R12	R13
2	R21	R22	R23
4	R31	R32	R33
6	R41	R42	R43

Рисунок 4 – Розподіл автомобілів на мережі за допомогою орієнтованого графа перехресті №5

Очевидно, що сума ймовірностей кожної стрічки таблиці (див. рис. 4) дорівнює одиниці, оскільки кількість машин, які в'їжджають в область перехрестя, однакова кількості машин, які її покидають. Якщо у транспортній мережі по дорозі рух заборонений, то у відповідну комірку масиву R записується нуль.

Розроблено програмне забезпечення для проведення моделювання за допомогою запропонованого методу на практиці (рис. 5).

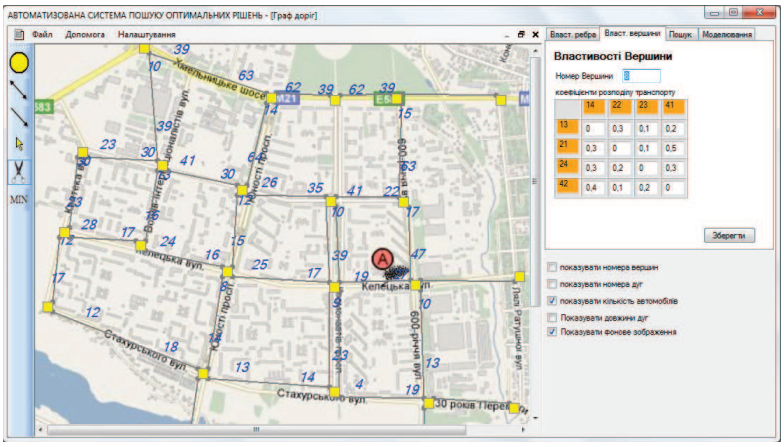


Рисунок 5 – Візуалізація результатів моделювання процесу розподілу автотранспорту в мережі

Висновок

Розроблено новий метод оптимізації параметрів руху автомобільного транспорту, який, на відміну від існуючих, використовує математичний апарат методу межових змін параметрів для розв’язання оберненої задачі оптимізації графа, що дозволяє розрахувати за яких змін параметрів (законів руху автомобілів на під’їзних дорогах, геометричних параметрів мережі тощо) транспортної мережі можливе досягнення оптимальних або максимально допустимих значень тривалості руху транспорту по заданому маршруту. Метод дозволить, також, моделювати та оптимізувати параметри самої транспортної мережі для максимізації її пропускної здатності у найбільш важливих напрямках. Для формалізації впливу параметрів світлофора на маршруті слідування транспорту на прикладі перехрестя з 4-ма входами/виходами виведено співвідношення для розрахунку максимальної пропускної здатності перехрестя для заданих параметрів його входів, що дозволяє розширити можливості застосування запропонованого методу оптимізації параметрів руху автомобільного транспорту. Розроблено програмне забезпечення для проведення моделювання за допомогою запропонованого методу на практиці.

Список літератури

1. Семенов В. В. Математическое моделирование динамики транспортных потоков мегаполиса. / В. В. Семенов. – М. : Наука, 2004. – 45с.
2. Прокудін Г.С. Моделі і методи оптимізації перевезень у транспортних системах. / Г. С. Прокудін. – К.: Нац. Тр-т Ун-т, 2006. – Монографія. – 224 с.
6. 3. Хейт Ф. Математическая теория транспортных потоков. / Ф. Хейт. – М. : Мир, 1996. – 286с.
4. Смирнов Н.Н. Математическое моделирование автотранспортных потоков / Н.Н. Смирнов. – М. : мех-мат, МГУ 1996. – 212с.
5. Демидович Б. П. Основы вычислительной математики. / Б. П. Демидович, И. А. Марон — М.: Наука, 1970. — 664с.
6. Войтко В. В. Один з методів розв’язання оберненої задачі чутливості / В. В. Войтко, І.С. Сербін // Вісник Вінницького політехнічного інституту — №6 — Вінниця, 2001. – С. 129-134 .
7. Веников В. А. Теория подобия и моделирования. / В. А. Веников. — М.: Высшая школа, 1976. — 480с.

Відомості про авторів

Мокін Віталій Борисович – д.т.н., професор, завідувач кафедри моделювання та моніторингу складних систем, vbmokin@gmail.com.

Романюк Олександр Никифорович – д.т.н., професор, перший проректор з наукової-педагогічної роботи по організації навчального процесу та його науково-методичного забезпечення, romanyuk@vstu.vinnica.ua.

Сторчак Володимир Григорович – аспірант кафедри моделювання та моніторингу складних систем, storchak@rambler.ru.

Войтко Вікторія Володимирівна – к.т.н., доцент, vojtko@vstu.vinnica.ua.

Гавенко Олег Віталійович – кафедра програмного забезпечення, магістрант з комп’ютерних наук, Gavenko2007@yandex.ru.

Вінницький національний технічний університет

УДК 621.396.6

ІМІТАНСНА ЛОГІКА

Л.Б. Ліщинська, М.А. Філинюк

Анотація: Запропонована імітансна логіка, в основі якої логічний стан схеми визначається характером імітансних RLC-параметров. Для побудови імітансних логічних схем пропонується використовувати властивості однотранзисторних узагальнених перетворювачів імітансу, що працюють у лінійному режимі. Оцінка потенційної швидкодії імітансних логічних схем складає порядку 10^{-9}–10^{-12}с і обмежується пролітними явищами у транзисторних структурах.

Анотация: Предложена иммитансная логика, в основе которой логическое состояние схемы определяется характером иммитансных RLC-параметров. Для построения иммитансных логических схем предлагается использовать свойства однотранзисторных обобщённых преобразователей иммитанса, работающих в линейном режиме. Оценка потенциального быстродействия иммитансных логических схем составляет порядка 10^{-9}–10^{-12} с и ограничивается пролётными явлениями в транзисторных структурах.

Abstract: Immittance logic is offered, in basis of which the logical state of chart is determined character of immittance RLC-parameters. For the construction of immittance logical charts it is suggested to use properties of the onetransistor generalized transformers of immittance, workings in the linear mode. Estimation of potential fast-acting of immittance logical charts makes order of 10^{-9}–10^{-12}s and limited to the flight phenomena in transistor structures.

Ключові слова: імітанс, імітансна логіка, узагальнений перетворювач імітансу, позитивна логіка, негативна логіка.

Вступ

Прогрес сучасного суспільства багато в чому визначається успіхами розвитку інформатики і обчислювальної техніки, в основі якої лежить булева алгебра, що заснована на двійковій формі представлення інформації. Технічна реалізація таких пристроїв базується на кодуванні інформації у вигляді стрибків постійного струму або напруги. Враховуючи, що такі сигнали отримали найменування – відеосигнали, логічні схеми, які їх використовують, можна розглядати як відеоімпульсні логічні схеми. Одночасно з логічними відеоімпульсними схемами, створювалися радіочастотні логічні схеми, що використовують як інформаційний параметр частоту, фазу і амплітуду гармонійних сигналів [1–5]. Їх ефективність теоретично, за швидкодією і апаратними витратам, в випадку реалізації багатозначної логіки, перевищила логічні відеоімпульсні схеми двійкової логіки. Але практично, через відсутність відповідної мікроелектронної бази НВЧ діапазону, цей напрям не набув широкого поширення. Хоча використання таких схем в сигнальних процесорах залишається перспективним.

Ще одним перспективним напрямом побудови логічних схем є використання оптичних логічних схем [6]. Вони також теоретично мають ряд істотних переваг і потенційних можливостей. Проте їх технічна реалізація далека від досконалості. Наприклад, за такими параметрами, як вартість, ступінь інтеграції та ін.

Здійснюються дослідження щодо створення логічних схем на інших (не електричних) фізичних ефектах (акустичних, пневматичних, хімічних тощо), що вважається актуальним, оскільки досягнуті на сьогоднішній день високі технічні параметри логічних відеоімпульсних схем при їх подальшому поліпшенні підійшли до своєї технологічної межі (за швидкодією; за потужністю, яка розсіюється). У зв'язку з цим, створення і розвиток логічних схем на основі нового інформаційного базису, може відкрити перспективу подальшого вдосконалення інформаційної техніки, наприклад, спеціального призначення.

Мета і постановка задач дослідження

Метою дослідження є розгляд можливості побудови логічних схем на основі нового координатного базису – імітансу.

Для досягнення цієї мети необхідно вирішити такі завдання:

– здійснити теоретичне обґрунтування створення логічних елементів на основі нового координатного базису – імітансу;

– розробити структурну реалізацію функціонально–повної системи імітансних логічних елементів;

– розглянути основи технічної реалізації імітансних логічних схем та оцінити їх швидкодію.

Теоретичне обґрунтування

Під імітансом W розуміється повний опір УПІ (імпеданс) $Z = \mathrm{Re}\,Z + j\,\mathrm{Im}\,Z$, або повна провідність (адмітанс) $Y = \mathrm{Re}\,Y + j\,\mathrm{Im}\,Y$. У загальному випадку це комплексні величини, що мають дійсну $\mathrm{Re}\,W$ і уявну $\mathrm{Im}\,W$ частотно–залежні частини [7].

Дійсна частина є диференціальним параметром: опором $R = \mathrm{Re}\,Z = \partial u / \partial i$, або провідністю $G = \mathrm{Re}\,Y = \partial i / \partial u$, які можуть бути як позитивними $R^{(+)}\left(G^{(+)}\right) > 0$, так і негативними $R^{(-)}\left(G^{(-)}\right) < 0$.

Уявна частина також є диференціальним параметром і характеризується диференціальною ємкістю $C = 1/p\,\mathrm{Im}Z = \mathrm{Im}Y/p$ або диференціальною індуктивністю $L = 1/p\,\mathrm{Im}Y = \mathrm{Im}Z/p$, які можуть бути як позитивними $C^{(+)}\left(L^{(+)}\right) > 0$, так і негативними $C^{(-)}\left(L^{(-)}\right) < 0$ [8].

Таким чином, у загальному випадку стан електричного ланцюга на змінному струмі можна характеризувати такими диференціальними параметрами: $R^{(+)}, R^{(-)}, G^{(+)}, G^{(-)}, C^{(+)}, C^{(-)}, L^{(+)}, L^{(-)}$.

Виходячи з цього, логічний стан схеми пропонується характеризувати не величиною струму або напруги, а характером імітансного параметра, що дозволить запропонувати шість варіантів імітансного представлення логічного «нуля» і «одиниці» (табл. 1).

При цьому слід звернути увагу, що відповідний логічний рівень описується не кількісною величиною імітансного параметра, а лише його характером або знаком, що підвищує перешкодозахищеність запропонованої системи.

При виборі варіантів не розглядалися комбінації, коли при реалізації логічних «0» або «1» імітансний параметр набуває нульового значення, оскільки при реальному втіленні цей варіант не реалізовується і погіршує параметри системи.

Враховуючи, що не існує реальних електричних ланцюгів, що володіють чисто дійсним ($\mathrm{Im}W = 0$) або чисто уявним ($\mathrm{Re}W = 0$) імітансом, запропоновані у табл. 1 варіанти є такими, що ідеалізуються і можуть бути реалізовані лише при певних допущеннях, наприклад на низьких частотах, де можна ігнорувати уявною складовою імітансу ($\mathrm{Im}W \approx 0$) або при використанні високодобротних імітансів, коли $\mathrm{Re}W \approx 0$.

Таблиця 1.

Варіанти імітансного представлення логічних «0» та «1»

Варіант 1		Варіант 2		Варіант 3	
Логічний рівень	Імітансний рівень	Логічний рівень	Імітансний рівень	Логічний рівень	Імітансний рівень
Позитивна логіка					
0	$R^{(-)}$	0	$G^{(-)}$	0	$C^{(+)}$
1	$R^{(+)}$	1	$G^{(+)}$	1	$L^{(+)}$
Негативна логіка					
0	$R^{(+)}$	0	$G^{(+)}$	0	$L^{(+)}$
1	$R^{(-)}$	1	$G^{(-)}$	1	$C^{(+)}$
Варіант 4		Варіант 5		Варіант 6	
Логічний рівень	Імітансний рівень	Логічний рівень	Імітансний рівень	Логічний рівень	Імітансний рівень
Позитивна логіка					
0	$L^{(-)}$	0	$C^{(-)}$	0	$C^{(-)}$
1	$L^{(+)}$	1	$C^{(+)}$	1	$L^{(-)}$
Негативна логіка					
0	$L^{(+)}$	0	$C^{(+)}$	0	$L^{(-)}$
1	$L^{(-)}$	1	$C^{(-)}$	1	$C^{(-)}$

Структурна реалізація

Функціональною повнотою володіє система логічних функцій «НІ», «І», «АБО». Їх структурну реалізацію розглянемо на прикладі першого варіанту позитивної імітансної логіки, коли логічній одиниці відповідає негативний дійсний імпеданс $"1 \to R^{(-)}"$, а логічному нулю – позитивний дійсний імпеданс $"0 \to R^{(+)}"$. У цьому випадку, структурні схеми основних логічних елементів мають вигляд, представлений на рис. 1.

Вони являють собою багатополюсники, в яких вихідний дійсний імпеданс $R_{вих}$ залежить від вхідного імітансу R_{Γ}. Кількісно ця залежність визначається коефіцієнтами перетворення імітансу T_i,

які у загальному випадку є комплексними величинами. Саме вибір і реалізація цих коефіцієнтів забезпечують реалізацію необхідних логічних функцій, представлених в табл. 2–4.

Таблиця 2.

Таблиця істинності імітансного логічного елемента «НІ»

$R_Г$	X	$R_{вих}$	Y
$R_Г^{(+)} > 0$	1	$-TR_Г < 0$	0
$R_Г^{(-)} < 0$	0	$TR_Г > 0$	1

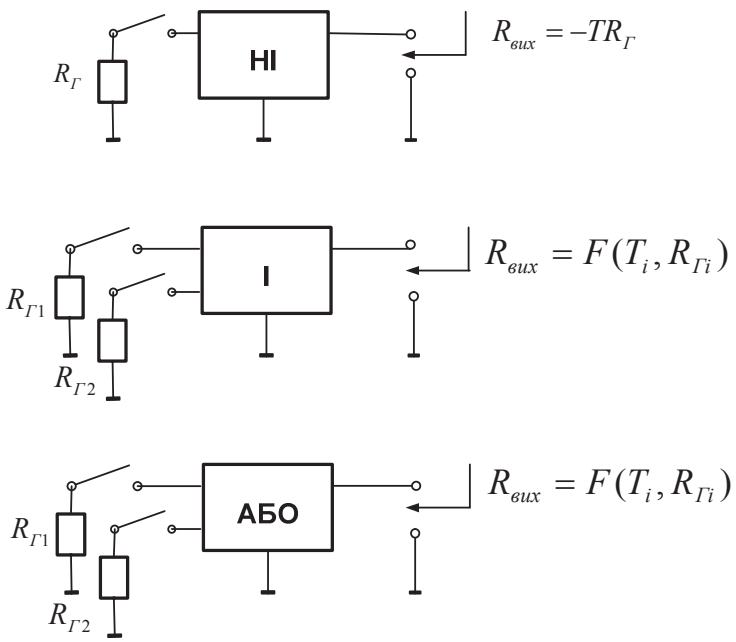


Рисунок 1 – Структурні схеми основних логічних елементів імітансної позитивної логіки, що використовує у якості інформаційного базису дійсний імітанс

Таблиця 3.

Таблиця істинності імітансного логічного елемента «І»

$R_{Г1}$	X1	$R_{Г2}$	X2	$R_{вих}$	Y
$R_{Г1}^{(-)} < 0$	0	$R_{Г2}^{(-)} < 0$	0	$R_{вих}^{(-)} < 0$	0
$R_{Г1}^{(-)} < 0$	0	$R_{Г2}^{(+)} > 0$	1	$R_{вих}^{(-)} < 0$	0
$R_{Г1}^{(+)} > 0$	1	$R_{Г2}^{(-)} < 0$	0	$R_{вих}^{(-)} < 0$	0
$R_{Г1}^{(+)} > 0$	1	$R_{Г2}^{(+)} > 0$	1	$R_{вих}^{(+)} > 0$	1

Таблиця 4.

Таблиця істинності імітансного логічного елемента «АБО»

$R_{Г1}$	X1	$R_{Г2}$	X2	$R_{вих}$	Y
$R_{Г1}^{(-)} < 0$	0	$R_{Г2}^{(-)} < 0$	0	$R_{вих}^{(-)} < 0$	0
$R_{Г1}^{(-)} < 0$	0	$R_{Г2}^{(+)} > 0$	1	$R_{вих}^{(+)} > 0$	1
$R_{Г1}^{(+)} > 0$	1	$R_{Г2}^{(-)} < 0$	0	$R_{вих}^{(+)} > 0$	1
$R_{Г1}^{(+)} > 0$	1	$R_{Г2}^{(+)} > 0$	1	$R_{вих}^{(+)} > 0$	1

Таким чином основою реалізації даних логічних елементів є узагальнені перетворювачі імітансу (УПІ) [9]. Існує велика кількість різновидів УПІ, але всіх їх можна розділити на дві групи: конвертори імітансу та інвертори імітансу.

У конвертора імітансу (КІ) вихідний перетворений імітанс $W_{вих}$ прямопропорційно залежить від імітансу $W_{Г}$, що підключається до його входу (перетворюваний імітанс)

$$W_{вих} = T_K W_{Г},$$

де $T_K = W_{вих}/W_{Г}$ – коефіцієнт конверсії імітансу.

У інвертора імітансу (ІІ) перетворений імітанс $W_{вих}$ зворотньо-пропорційно залежить від перетворюваного імітансу

$$W_{вих} = T_I / W_{Г},$$

де $T_I = W_{вих} W_{Г}$ – коефіцієнт інверсії імітансу.

Використовуючи в загальному випадку комплексний характер T_K і T_I можливо шляхом комбінації УПІ забезпечити реалізацію необхідної імітансної логічної функції. Наприклад, для реалізації імітансного логічного елементу «НІ» досить використовувати КІ з коефіцієнтом конверсії $T_K \prec 0$ (конвертор негативного імітансу). Реалізація інших логічних функцій вимагає більш складних структурних рішень і розробки методики синтезу таких структур.

Істотно розширити функціональні можливості і реалізацію таких структур дозволяє введення поняття керованих і некерованих УПІ. У простому випадку некерований УПІ представляє чотириполюсник, що має клеми для підключення перетворюваного імітансу $W_{Г}$ і клеми, між якими реалізується перетворений імітанс $W_{вих}$. Коефіцієнт перетворення імітансу, в ідеальному випадку, не залежить від зовнішнього впливу. Проте можлива реалізація УПІ, в якого коефіцієнт перетворення імітансу T залежить від управляючого впливу F_y, який може мати різну фізичну природу. Такі УПІ назвемо керованими. Значний інтерес представляють керовані УПІ, які в залежності від управляючого впливу F_y змінюють свої властивості з конвертора на інвертор імітансу і навпаки. Схематичне позначення такого УПІ представлено на рис. 2, таблиці перетворення його інформаційного параметра – у табл. 5–6, які використовуються при схемотехнічній реалізації імітансних логічний пристроїв.

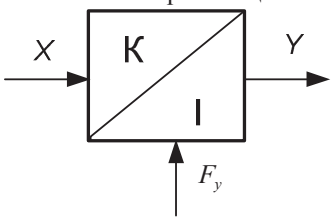


Рисунок 2 – Схематичне позначення керованого УПІ

Таблиця 5.

Таблиця перетворення інформаційного параметру керованим УПІ для позитивної логіки

K / I	F_y	X	Y
K	0	L / C	L / C
I	1	L / C	C / L

Таблиця 6.

Таблиця перетворення інформаційного параметру керованим УПІ для негативної логіки

К / I	F_y	X	Y
I	0	L / C	C / L
К	1	L / C	L / C

Технічна реалізація

Технічна реалізація УПІ можлива як на пасивних, так і на активних компонентах. Використання пасивних компонентів (відрізків лінії передачі, різних видів електричних і акустичних трансформаторів, феритів тощо) можливе в обмежених варіантах і, як правило, приводить до погіршення масогабаритних і технологічних характеристик. Крім того, такі УПІ не дозволяють реалізувати негативні дійсні імітанси, а отже і створити імітансні логічні пристрої з функціональною повнотою. У зв'язку з цим, перевагу слід віддати активному УПІ, що, зокрема, реалізовуються на базі напівпровідникових активних приладів.

При реалізації УПІ на базі активних напівпровідникових приладів спостерігається дві тенденції. Перша пов'язана з використанням, перш за все, підсилювальних властивостей транзисторів і зовнішніх, часто перехресних, зворотних зв'язків [10]. Такі схемотехнічні УПІ внаслідок прояву зовнішніх паразитних зворотних зв'язків мають обмежений частотний діапазон, що обмежує швидкодію імітансних логічних схем. Крім того, наявність декількох транзисторів, що працюють в активному режимі, веде до невиправдано великого використання енергії.

Наявність в кристалі транзистора внутрішнього позитивного зворотного зв'язку довгий час розглядалась як паразитне явище. Проте, саме завдяки цьому явищу, з'явилася можливість використання чотириполюсника на базі транзистора як однокристального УПІ, здатного працювати на частотах в десятки і навіть сотні ГГц [11]. Виходячи з цього, можна вважати одним з перспективних напрямів реалізацію імітансних логічних схем саме на базі однотранзисторних УПІ.

Проілюструємо можливість такої реалізації імітансного логічного елементу «НІ» на базі конвертора імітансу, що реалізовується на основі одноперехідного транзистора. Вибір відносно низькочастотного транзистора обумовлений необхідністю зменшення впливу паразитних реактивних елементів і ставить за мету не виявлення потенційних можливостей, а демонстрацію схемотехнічної реалізуємості. Розглянемо імітансний логічний елемент на базі одноперехідного транзистору, який включений за схемою із загальною першою базою (рис. 3.).

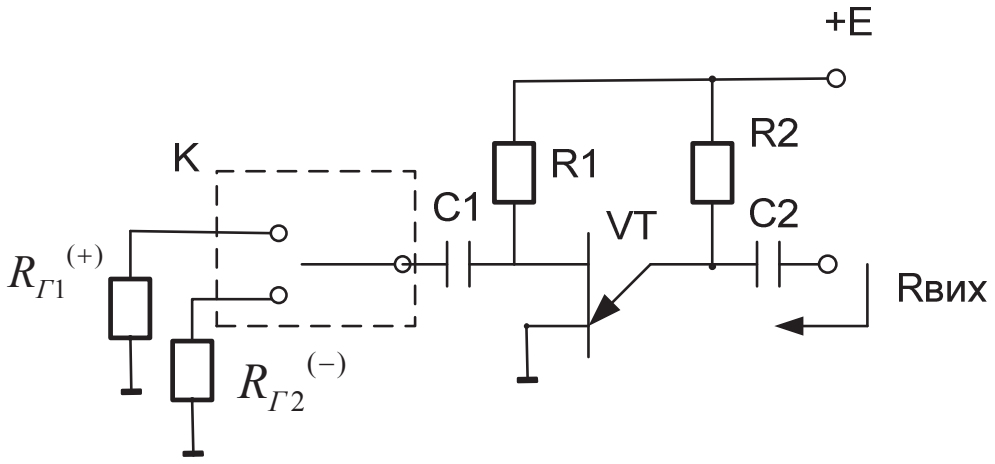


Рисунок 3 – Принципова електрична схема імітансного логічного елемента «НІ» на базі одноперехідного транзистора

При великих струмах емітера диференціальний опір емітера $r_E \to 0$ і опір між емітером і першою базою буде рівний [12]

$$R_{ЕБ1} = (1-\alpha)R_Г, \tag{1}$$

де α – коефіцієнт передачі транзистора по струму; $\alpha = i_{Б2}/i_E$; $i_{Б2}, i_E$ - струми, що протікають по виводу другої бази і емітера.

Коефіцієнт передачі по струму одноперехідного транзистора $\alpha > 1$. Звідки $R_{ЕБ1} = -(\alpha - 1)R_Г$.

Таким чином, при «підключенні» між першою і другою базою транзистора позитивного дійсного імітансу $R_{Г1}^{(+)}$ (логічна «1») вихідний імпеданс схеми $R_{вих} = -(\alpha - 1)R_{Г1} < 0$ (логічний «0»). У випадку, якщо між першою і другою базою транзистора «підключений» негативний дійсний імітанс $R_{Г2}^{(-)}$ (логічний «0») вихідний імітанс схеми $R_{вих} = (\alpha - 1)R_{Г2}^{(-)} > 0$ (логічна «1»). Розглянуті режими відповідають логічній функції «НІ» (табл. 2). Введений в схему комутатор «К» є умовною ланкою, що імітує появу на вході схеми відповідного імітансу.

Для оцінки потенційної швидкодії схеми врахуємо частотну залежність коефіцієнта передачі $\overline{\alpha}$ у виразі (1). Використовуючи однополюсну апроксимацію для $\overline{\alpha} = \alpha_0/(1 + j\Omega_\alpha)$, де α_0 – низькочастотне значенні коефіцієнта передачі транзистора по струму; Ω_α – приведена частота; $\Omega_\alpha = f/f_\alpha$; f_α - гранична частота одноперехідного транзистора, знаходимо коефіцієнт конверсії схеми

$$T_K = \frac{1 - \alpha_0 + j\Omega_\alpha}{1 + j\Omega_\alpha}.$$

Звідки, фазова затримка перетворення дійсного імітансу рівна

$$\phi_{T_K} = arctg\left(\operatorname{Im} T_K / \operatorname{Re} T_K\right) = arctg\left(\frac{\Omega_\alpha \alpha_0}{1 - \alpha_0 + \Omega^2}\right).$$

Враховуючи, що $\alpha_0 > 1$, а $\Omega^2 << 1$, знаходимо

$$\phi_{T_K} \approx arctg\left(\frac{\alpha_0}{\alpha_0 - 1}\Omega_\alpha\right) = 2\pi f t_З, \tag{2}$$

де $t_З$ – часова затримка процесу перетворення імітансу.

Вирішуючи (2) відносно $t_З$, знаходимо

$$t_З \approx \frac{1}{2\pi f_\alpha} arctg\frac{\alpha_0 \Omega_\alpha}{(\alpha_0 - 1)}. \tag{3}$$

З (3) витікає, що чим більше гранична частота f_α транзистора, тим вище швидкодія даного елементу. Наприклад, при використанні одноперехідного транзистора КТ 117А, що має $\alpha = 3$ і $f_\alpha = 200 МГц$ на частоті 100 *МГц*, знаходимо $t_З = 1,2$ *нс*. Вказана затримка співрозмірна за величиною з часовою затримкою перемикання сучасних відеоімпульсних логічних схем. Але при цьому слід врахувати, що у даному випадку розглядається відносно низькочастотний транзистор. А найбільш високочастотні сучасні транзистори мають граничну частоту в 10^3 рази більшу, що дозволяє спрогнозувати одну з переваг даних імітансних логічних схем, високу швидкодію, що обмежується, перш за все, часом прольоту носіїв струму, і не обмежується перехідними процесами у напівпровідниковій структурі.

Висновки

1. В основі імітансної логіки пропонується визначати логічний стан схеми характером імітансних параметрів: $R^{(+)}, R^{(-)}, G^{(+)}, G^{(-)}, C^{(+)}, C^{(-)}, L^{(+)}, L^{(-)}$, що дозволяє запропонувати шість варіантів представлення логічних «0» та «1».

2. На відміну від відеоімпульсних і радіочастотних логічних схем, які використовують нелінійні властивості напівпровідникових приборів, для побудови імітансних логічних схем пропонується використовувати властивості однотранзисторних узагальнених перетворювачів імітансу, що працюють в лінійному режимі.

3. Оцінка потенціальної швидкодії імітансних логічних схем складає порядку 10^{-9}–10^{-12}с і обмежується пролітними явищами у транзисторних структурах.

4. Подальший розвиток імітансної логіки вимагає розвитку теорії синтезу імітансних логічних схем на базі різноманітних RLC–інформаційних параметрів, а також розвиток схемотехнічної реалізації таких схем на базі однокристальних УПІ, що використовують різні види інжекційно–пролітних транзисторних структур.

Список літератури

1. Кнорре К.Г., Тузов В.М., Шур Г.И. Фазовые и частотные СВЧ элементы. – М.: Сов. радио, 1975. – 352 с.
2. Кичак В.М. Радіоімпульсні логічні НВЧ елементи. – Вінниця: УНІВЕРСУМ–Вінниця, 1999. – 240 с.
3. Нейман М.С. О сверхвысокочастотной вычислительной прерывной автоматике // Электронные радиоимпульсные вычислительные системы дискретного действия. Труды МАИ. – Вып. 149. – С.3-8.
4. Молчанов А.А., Волкогонов В.П., Лоза Ю.Х., Яковега Г.Н. Проектирование многофункциональных интегральных схем. – К.: Техніка, 1984. – 143 с.
5. А.с. 863132 СССР. Радиочастотный логический элемент / Н.А. Філинюк, Ю.Г. Калиниченко // Бюл. Изобр. – 1982. - № 36.
6. Кожемяко В.П., Натрошвили О.Г., Мартынюк Т.Б., Имнаишвили Л.Ш. Оптоэлектронная схемотехника. – К.: УМК, 1988. – 276 с.
7. Сигорский В.П., Петренко А.И. Основы теории электротехнических схем. – К.: Техника, 1967. – 609 с.
8. Філинюк М.А. Основи негатроніка. – Вінниця: УНІВЕРСУМ–Вінниця, 2006. – Т.1. Теоретичні і фізичні основи негатроніка. – 456 с.
9. Філинюк М.А. Аналіз і синтез інформаційних пристроїв на базі потенційно–нестійких узагальнених перетворювачів імітансу. – Вінниця: ВДТУ, 1998. – 85 с.
10. Ионкин П.А., Миронов В.Г. Синтез RC схем с активными невзаимными элементами. – М.: Энергия, 1976. – 240 с.
11. Филинюк Н. А. Основы анализа и синтеза информационных устройств на базе инжекционно-пролетных эффектов : автореф. дис. на соискание уч. степени докт. техн. наук: спец. 05.13.05 «Элементы и устройства вычислительной техники и систем управления» / Филинюк Николай Антонович ; Ин-т кибернетики им. В. М. Глушкова АН УССР. – К. : 1984. – 47 с.
12. Лищинская Л.Б., Булыга И.В., Шведюк А.Г., Филинюк Н.А. Обобщенные преобразователи иммитанса на основе инжекционно-пролётной транзисторной структуры с общим истоком // Наукові праці ВНТУ. – 2008. – №2. – С. 1-18.

Відомості про авторів

Ліщинська Людмила Броніславівна – к.т.н., доцент, Вінницький торговельно–економічний інститут КНТЕУ, вул.. Соборна, 87, м. Вінниця, 21000, L_Fil1@mail.ru

Філинюк Микола Антонович – д.т.н., професор, завідувач кафедри проектування комп'ютерної та телекомунікаційної апаратури, Вінницький національний технічний університет, Хмельницьке шосе, 95, м. Вінниця, 21021, N_Fil_45@mail.ru

УДК 681.325.5

РЕКУРСИВНІ МЕТОДИ ТА АЛГОРИТМИ УНІВЕРСАЛЬНИХ АДИТИВНИХ ПЕРЕТВОРЕНЬ В АМ-СИСТЕМАХ ЧИСЛЕННЯ

О.Д. Азаров, О.І. Черняк

Анотація. Запропоновано рекурсивні методи та алгоритми, що дозволяють в *АМ*-системах числення виконувати універсальні адитивні перетворення на основі елементарних адитивних перетворень.
Аннотация. Предложены рекурсивные методы и алгоритмы, позволяющие в *АМ*-системах счисления выполнять универсальные аддитивные преобразования на основе элементарных аддитивных преобразований.
Annotation. The recursive methods and algorithms allowing in *AM*-numerical systems to carry out universal additive transformations by means of elementary additive transformations are offered.
Ключові слова: рекурсивні методи, *АМ*-система.

Вступ

При організації розподілених обчислень виникає проблема, що полягає у необхідності реалізації великої кількості інформаційних зв’язків між пристроями. Одним з відомих підходів до вирішення даної проблеми є порозрядне конвеєрне обробляння послідовних кодів, що дозволяє значно зменшити кількість інформаційних зв’язків без суттєвого зменшення продуктивності. Для вирішення широкого кола задач таке обробляння вимагає використання надлишкових систем числення, оскільки лише у них можливе порозрядне виконання усіх арифметичних операцій у єдиному потоці за рахунок обмеження перенесення у старші розряди. Від довжини перенесення залежать апаратні витрати пристроїв, що реалізують таке обробляння. Серед відомих надлишкових позиційних систем числення для порозрядного обробляння найбільш широко використовується знакорозрядна [1,2,3,4] з цифрами -1, 0, 1. Для передавання одного розряду у цій системі числення потрібно два інформаційних зв’язки. Відомо також порозрядне обробляння у кодах золотої пропорції з цифрами 0 і 1 [5]. Це дозволяє передавати один розряд коду через один інформаційний зв’язок.

Актуальність

Можна створити безліч систем числення з будь-якою кількістю цифр, що дозволяють виконувати порозрядне обробляння. Кожна з них має різну довжину перенесення при додаванні і тому потребує різних апаратних витрат для реалізації пристроїв. Отже, актуальною є задача визначення системи числення, що забезпечує мінімальні апаратні витрати на реалізацію порозрядних пристроїв при мінімальній кількості інформаційних зв’язків. Для вирішення даної задачі авторами запропоновано клас систем числення, що узагальнює відомі та дозволяє створювати нові системи числення з можливістю порозрядного обробляння [6]. Вони названі *АМ*-системами числення. Будь-яка *АМ*-система числення може бути описана за допомогою множини цифр $\{0...c_{k-1}\}$, основи системи числення w і адитивного співвідношення ${}^{t}A^{\tau,p} : w^{tp} = R^{\tau,p}$, де $R^{\tau,p} = \sum_{i=0}^{p} r_i \cdot w^{\tau i}$ - граничне значення ($r \in \{0...c_{k-1}\}$). Наявність адитивних співвідношень в *АМ*-системах числення дозволяє ввести операції адитивного перетворення кодів (*А*-перетворення), що змінюють код при збереженні його числового еквіваленту. *А*-перетворення здійснюють перенесення і запозичення та можуть використовуватись при додаванні і відніманні кодів. За напрямком перенесення *А*-перетворення поділяються на перетворення з перенесенням у старші розряди (*AL*) і перетворення з перенесенням у молодші розряди (*AR*). Для здійснення *А*-перетворень необхідно виконання певних умов. Необхідна умова *i*-го *AL*-перетворення: значення розрядів від 0-го до *i*-го повинно бути не меншим ніж відповідне граничне значення, а значення (*i*+*t*)-го розряду повинно бути меншим ніж старша цифра. Необхідна умова *i*-го *AR*-перетворення: значення розрядів від 0-го до *i*-го повинно бути не більшими ніж різниця між максимальним кодом у цих розрядах і відповідним граничним значенням, а значення (*i*+*t*)-го розряду повинно бути більшим нуля. За достатніми умовами виконання *А*-перетворення поділяються на елементарні (*E*), універсальні (*U*) та повні (*F*). Для виконання *EA*-перетворень достатньо, щоб у кожному розряді перетворюваного коду виконувались умови, що не допускали б переповнення розрядів чи виникнення у них від’ємних значень. Однак можливі випадки, у яких не всі розряди перетворюваного коду задовольняють достатній умові *EA*-перетворення, хоча у цілому значення коду задовольняє необхідній умові. Оскільки у таких випадках *EA*-перетворення не виконуються, то вони не можуть використовуватись як перенесення і запозичення при виконанні порозрядного додавання і віднімання. Для цього слід застосовувати *UA*-перетворення.

Достатні умови *UA*-перетворень подібні до необхідних умов. Відмінність полягає у тому, що для *i*-го *UA*-перетворення перевіряються розряди не від 0-го до *i*-го, а від (*i*-*tb*)-го до *i*-го. Перевірка достатніх умов та виконання *UA*-перетворень у загальному випадку є досить складними для реалізації. Тому актуальною є розробка спрощених методів та алгоритмів *UA*-перетворень в *АМ*-системах числення.

Мета

Метою даної статті є підвищення ефективності проектування пристроїв порозрядної обробки за рахунок виконання універсальних адитивних перетворень в *AM*-системах числення як послідовності елементарних адитивних перетворень.

Задачі

Для досягнення даної мети необхідно вирішити такі задачі:

1) проаналізувати усі можливі ситуації, що виникають при виконанні достатніх умов *UA*-перетворень;

2) для кожної ситуації визначити послідовність *EA*-перетворень.

Розв'язання задач

Методи виконання *UA*-перетворень на основі *EA*-перетворень зручно представити у рекурсивному виді. Для визначення послідовності *EA*-перетворень, що приводять до виконання *i*-го *UAL*-перетворення над розрядами коду від *i*-го до (*i*-τb)-го, необхідно розглянути три можливих випадки виконання достатніх умов *UA*-перетворень. У кожному з цих випадків вважається, що $x_{i+t} < c_{k-1}$. Перший випадок – якщо у розрядах від від *i*-го до (*i*-τb)-го виконується умова *i*-го *EAL*-перетворення, тобто для кожного *j*-го з порівнюваних розрядів коду від *i*-го до (*i*-τp)-го виконується $x_{i-j} \ge r_{\tau p-j}$:

$$\forall_{0<j<\tau p} (x_{i-j} \ge r_{\tau p-j}).$$

У даному випадку *i*-е *UAL*-перетворення зводиться до *i*-го *EAL*-перетворення:

$${}^{t}UAL_i^{\tau,p}(X_{i-\tau b}^{\tau b+t}) = {}^{t}EAL_i^{\tau,p}(X_0^{n-1}).$$

Після цього виконання *i*-го *UAL*-перетворення завершується. Другий випадок – якщо умова *EAL*-перетворення виконується тільки для групи розрядів від *i*-го до (*i*-τg)-го, де $g<p$, і для деякого $0 \le v \le g(\tau+1)/t$ значення (*i*-τg+tv)-го розряду цієї групи більше (-τg-tv)-го розряду граничного значення:

$$\forall_{0\le j\le g} (x_{i-\tau j} \ge r_{\tau(p-j)}) \wedge x_{i-\tau(g+1)} < r_{\tau(p-g-1)} \wedge \exists_{0\le v\le g(\tau+1)/t} (x_{i-\tau g+tv} > r_{\tau g-tv}).$$

У даному виконується безумовне (*i*-τg+t(v-1))-е *EAR*-перетворення, а потім знову рекурсивно виконується *i*-те *UAL*-перетворення:

$${}^{t}UAL_i^{\tau,p}(X_{i-\tau b}^{\tau b+t}) = {}^{t}UAL_i^{\tau,p}({}_{-}{}^{t}EAR_{i-\tau g+t(v-1)}^{\tau,p}(X_0^{n-1})).$$

Тут ${}_{-}EAR$ – безумовне *EAR*-перетворення, що виконується без перевірки умови. При виконанні безумовного *EAR*-перетворення за рахунок додавання у молодших розрядах може виникнути переповнення. Для зменшення величини переповнення другий випадок потрібно аналізувати, починаючи з v=0 і збільшуючи кожного разу v на одиницю. Третій випадок – якщо не виконується умова першого чи другого випадку для *i*-го *UAL*-перетворення. У даному випадку виконується (*i*-*t*)-е *UAL*-перетворення, а після його завершення знову рекурсивно виконується *i*-те *UAL*-перетворення.

$${}^{t}UAL_i^{\tau,p}(X_{i-\tau b}^{\tau b+t}) = {}^{t}UAL_i^{\tau,p}({}^{t}UAL_{i-t-\tau g}^{\tau,p}(X_{i-t-\tau(b+g)}^{\tau p+t})).$$

UAL-перетворення у третьому випадку знову може призвести до рекурсивного *UAL*-перетворення у молодших розрядах перетворюваного коду, і так далі, поки номер перетворення не менший ніж *i*-τb. Якщо після останнього *UAL* перетворення умова відповідного *EAL*-перетворення не виконується, то *i*-е *UAL*-перетворення завершується. Отже, рекурсивний метод *i*-го *UAL*-перетворення на основі *EA*-перетворень такий. Перед початком виконання перетворення глобальній змінній m присвоюється значення *i*. Далі викнуються такі дії.

п.1 Якщо виконується умова $m<i-\tau b$, то закінчити перетворення.

п.2 Якщо не виконується умова $\forall_{0<j<\tau p} (x_{m-j} \ge r_{\tau p-j})$, то перейти на п.4.

п.3 Виконати ${}^{t}EAL_m^{\tau,p}(X_0^{n-1})$ і закінчити перетворення.

п.4 Якщо для не виконується умова

$$\exists_{0\le g<p} (\forall_{0\le j\le g} (x_{m-\tau j} \ge r_{\tau(p-j)}) \wedge x_{m-\tau(g+1)} < r_{\tau(p-g-1)} \wedge \exists_{0\le v<g(\tau+1)/t} (x_{m-\tau g+tv} > r_{rg-tv})),$$

то перейти на п.6.

п.5 Виконати ${}_{-}{}^{t}EAR_{m-\tau g+t(v-1)}^{\tau,p}(X_0^{n-1})$, перейти на п.7.

п.6 Виконати ${}^{t}UAL^{\tau,p}_{m-t-\tau g}(X^{\tau b+t}_{m-t-\tau(b+g)})$.

п.7 Виконати ${}^{t}UAL^{\tau,p}_{m}(X^{\tau b+t}_{m-\tau b})$.

Більш детально універсальне *AL*-перетворення на основі елементарних *A*-перетворень представлено граф-схемою алгоритму на рис. 1.

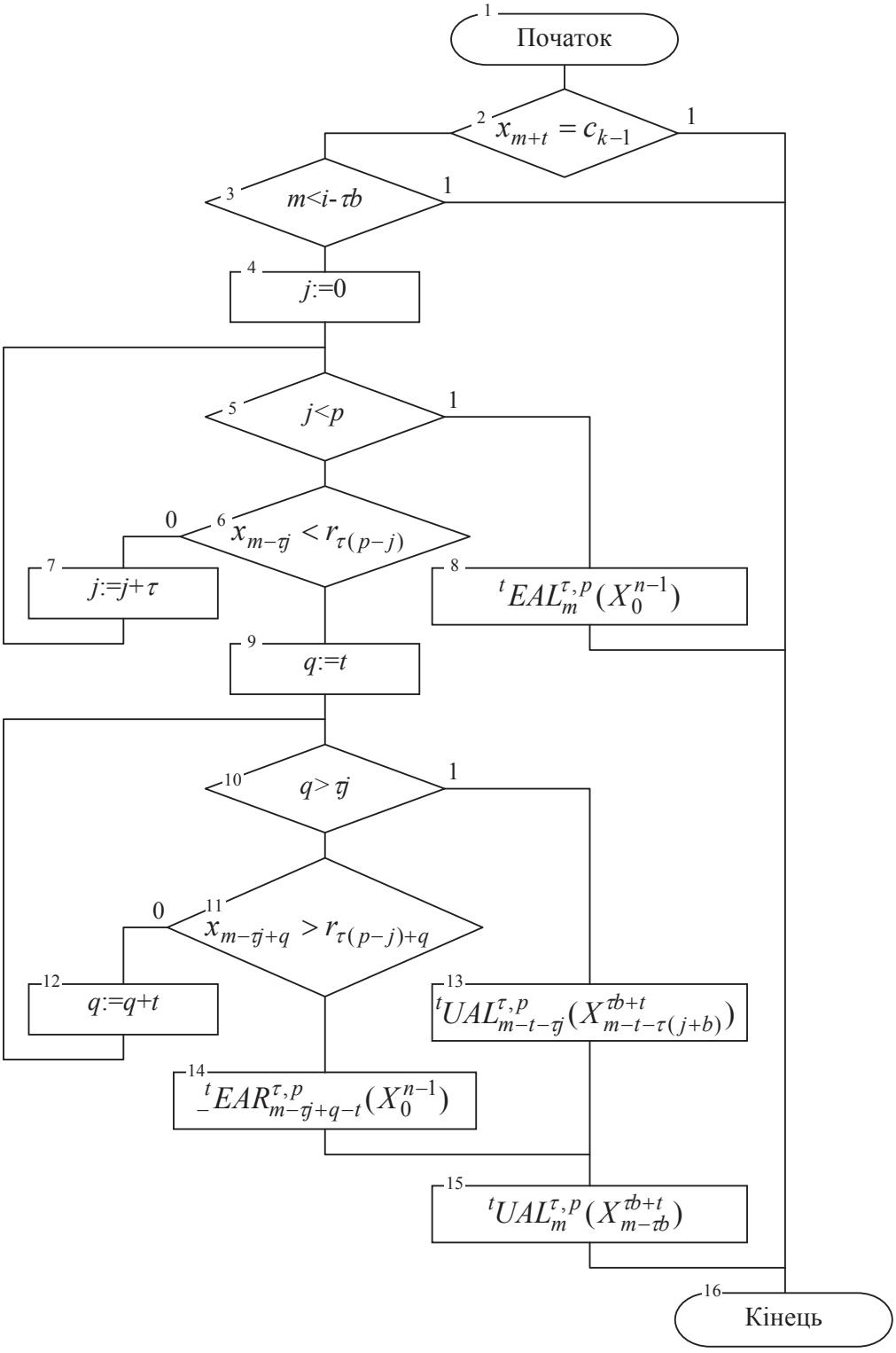


Рисунок 1 – Граф-схема алгоритму виконання *UAL*-перетворення на основі *EAL*- і *EAR*-перетворень

Для визначення послідовності *EA*-перетворень, що приводять до виконання *UAR*-перетворення над розрядами коду від *i*-го до (*i*-*τb*)-го, необхідно також розглянути три випадки. В кожному з цих випадків вважається, що $x_{i+t} > 0$. Перший випадок – коли в розрядах від *i*-го до (*i*-*τb*)-го виконується умова *EAR*-перетворення:

$$\underset{0<j<\tau p}{\forall}(x_{i-j} + r_{\tau p-j} \le c_{k-1}).$$

У даному випадку *i*-е *UAR*-перетворення зводиться до *i*-го *EAR*-перетворення:

$${}^{t}UAR_{i}^{\tau,p}(X_{i-\tau b}^{\tau b+t}) = {}^{t}EAR_{i}^{\tau,p}(X_{0}^{n-1}).$$

Після цього виконання *i*-го *UAR*-перетворення завершується. Другий випадок – якщо умова *EAR*-перетворення виконується тільки для групи розрядів від *i*-го до (*i*-*τg*)-го (де $g<p$) і для деякого $0 \le v \le g(\tau+1)/t$ значення (*i*-*τg*+*tv*)-го розряду цієї групи менше величини $c_{k-1} - r_{\tau g-tv}$:

$$\underset{0\le j\le g}{\forall}(x_{i-\tau j} + r_{\tau(p-j)} \le c_{k-1}) \wedge x_{i-\tau(g+1)} +$$

$$+ r_{\tau(p-g-1)} > c_{k-1} \wedge \underset{0\le v\le g(\tau+1)/t}{\exists}(x_{i-\tau g+tv} < c_{k-1} - r_{\tau g-tv}).$$

У даному випадку виконується безумовне (*i*-*τg*+*t*(*v*-1))-е *EAL*-перетворення, а потім знову рекурсивно виконується *i*-те *UAR*-перетворення:

$${}^{t}UAR_{i}^{\tau,p}(X_{i-\tau b}^{\tau b+t}) = {}^{t}UAR_{i}^{\tau,p}({}^{t}_EAL_{i-\tau g+t(v-1)}^{\tau,p}(X_{0}^{n-1})).$$

Тут $_EAL$ – безумовне *EAL*-перетворення, що виконується без перевірки умови. При виконанні безумовного *EAL*-перетворення за рахунок віднімання у молодших розрядах може виникнути від'ємне значення. Для зменшення абсолютної величини від'ємного значення другий випадок потрібно аналізувати, починаючи з $v=0$ і збільшуючи кожного разу *v* на одиницю. Третій випадок – коли не виконується умова першого чи другого випадку для *i*-го *UAR*-перетворення. У даному випадку рекурсивно виконується (*i*-*t*-*τg*)-е *UAR*-перетворення, а після його завершення знову рекурсивно виконується *i*-те *UAR*-перетворення.

$${}^{t}UAR_{i}^{\tau,p}(X_{i-\tau b}^{\tau b+t}) = {}^{t}UAR_{i}^{\tau,p}({}^{t}UAR_{i-t-\tau g}^{\tau,p}(X_{i-t-\tau(b+g)}^{\tau b+t})).$$

Отже, виконання *UAR*-перетворення у третьому випадку знову може призвести до рекурсивного виконання *UAR*-перетворення у молодших розрядах перетворюваного коду, і так далі, поки номер перетворення залишається не меншим ніж *i*-*τb*. Якщо після виконання останнього *UAR* перетворення умова відповідного *EAR*-перетворення не виконується, то *i*-е *UAR*-перетворення завершується.

Метод виконання *i*-го *UAR*-перетворення на основі *EA*-перетворень такий. На кожній ітерації виконується *UAR*-перетворення з номером *m*. Перед початком виконання перетворення глобальній змінній *m* присвоюється значення *i*. Далі виконуються такі дії.

п.1 Якщо виконується умова $m = i-\tau b$, то завершити перетворення.

п.2 Якщо не виконується умова

$$\underset{0<j<\tau p}{\forall}(x_{m-j} + r_{\tau p-j} \le c_{k-1}),$$

то перейти на п.4.

п.3 Виконати ${}^{t}EAR_{m}^{\tau,p}(X_{0}^{n-1})$ і завершити перетворення.

п.4 Якщо не виконується умова

$$\underset{0\le j\le g}{\forall}(x_{m-\tau j} + r_{\tau(p-j)} \le c_{k-1}) \wedge x_{m-\tau(g+1)} +$$

$$+ r_{\tau(p-g-1)} > c_{k-1} \wedge \underset{0\le v\le g(\tau+1)/t}{\exists}(x_{i-\tau g+tv} < c_{k-1} - r_{\tau g-tv}),$$

то перейти на п.6.

п.5 Виконати ${}^{t}_EAL_{m-\tau g+t(v-1)}^{\tau,p}(X_{0}^{n-1})$, перейти на п.7.

п.6 ${}^{t}UAR_{m-t-\tau g}^{\tau,p}(X_{m-t-\tau(b+g)}^{\tau b+t})$

п.7 $^{t}UAR_{m}^{\tau,p}(X_{m-\tau b}^{\tau b+t})$.

Більш детально універсальне *AR*-перетворення на основі елементарних *A*-перетворень представлено граф-схемою алгоритму на рис. 2.

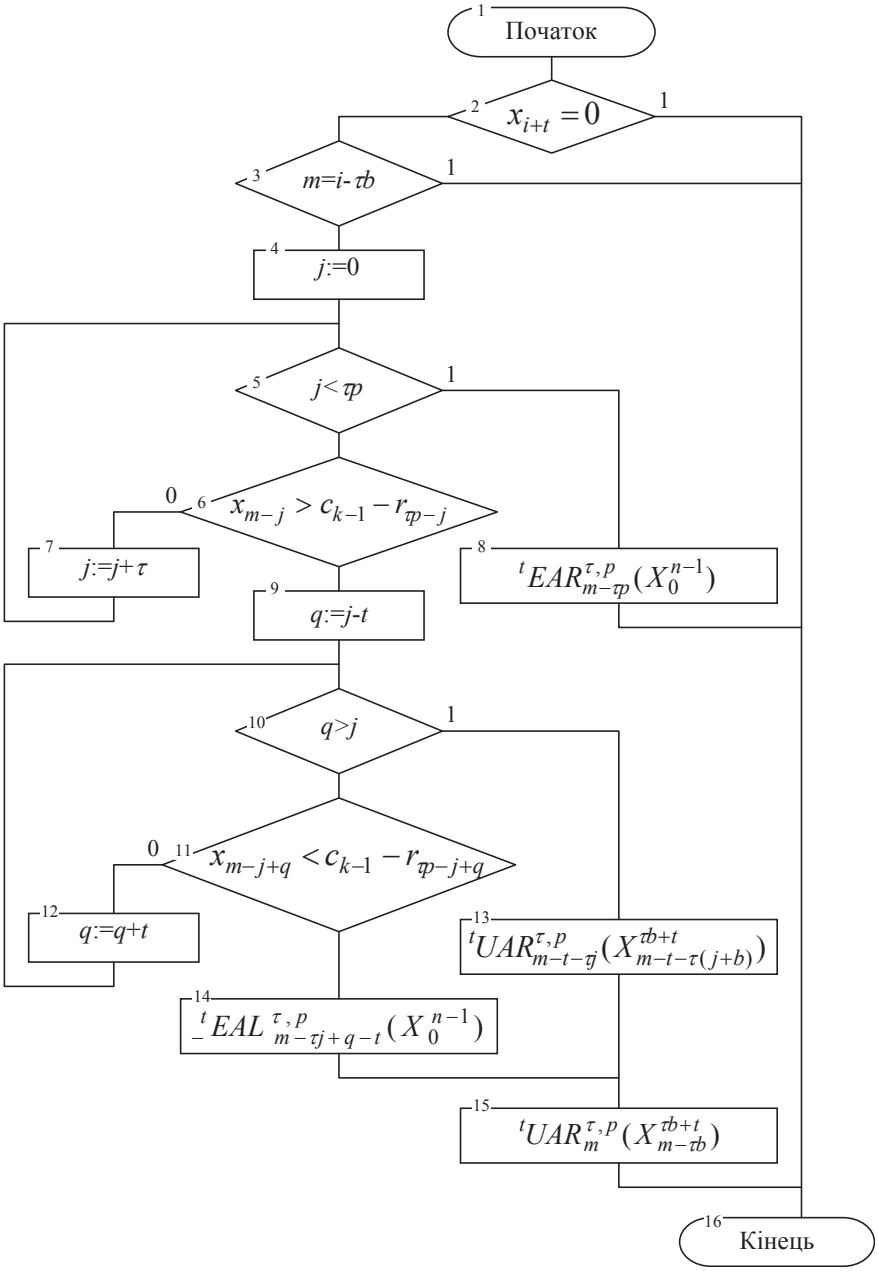


Рисунок 2 – Граф-схема алгоритму виконання *UAR*-перетворення на основі *EAR*- і *EAL*-перетворень

Нижче наведено приклади виконання *UAL*- та *UAR*-перетворень за допомогою *EA*-перетворень для деякої конкретної *AM*-системи числення.

Параметри *AM*-системи числення:

$$\left\{\begin{array}{l} C_4=\{0,1,2,3\}; \\ {}^{1}A^{1,3}:10000=3221 \end{array}\right\}.$$

Код числа: $X_0^9=111331101$.

$^{1}UAL_{5}^{1,3}$ -перетворення:

1	1	1	3	3	1	1	0	1	
				-1	+3	+2	+2	+1	$_^{1}EAR_{3}^{1,3}$
1	1	2	3	2	4	3	2	2	
		+1	-3	-2	-2	-1			$^{1}EAL_{5}^{1,3}$
1	1	3	0	0	2	2	2	2	

Код числа: $X_0^9 = 2020003322$.

$^{1}UAR_{6}^{1,3}$ -перетворення:

2	0	2	0	0	0	3	3	2	2	
					+1	-3	-2	-2	-1	$_^{1}EAL_{3}^{1,3}$
2	0	2	0	0	1	0	1	0	1	
			-1	+3	+2	+2	+1			$^{1}EAR_{6}^{1,3}$
2	0	1	3	2	3	1	1	0	1	

Як видно з проведеного аналізу, рекурсивні методи і алгоритми *UAL*- та *UAR*-перетвореннь дозволяють виконувати їх на основі *EA*-перетворень обох типів подібним чином. При виконанні кожного з *UA*-перетворень аналізується три випадки. У першому випадку *i*-те *UA*-перетворення зводиться до *i*-го *EA*-перетворення того ж типу. У другому випадку *i*-те *UA*-перетворення потребує попереднього виконання безумовного *v*-го *EA*-перетворення протилежного типу і подальшого рекурсивного повторення виконання *i*-го *UA*-перетворення того ж типу. У третьому випадку *i*-те *UA*-перетворення потребує рекурсивного виконання *UA*-перетворення того ж типу у молодших розрядах. Завершується рекурсивний алгоритм *i*-го *UA*-перетворення після виконання *i*-го *EA*-перетворення того ж самого типу або, якщо після аналізу не виконується жодна з описаних умов.

Висновок

Запропоновані рекурсивні методи і алгоритми дозволяють виконувати універсальні адитивні перетворення за допомогою скінченної послідовності елементарних адитивних перетворень у будь-якій *AM*-системі числення, що спрощує розробку таблиць переходів при проектуванні пристроїв порозрядного обробляння у цих системах числення.

Список літератури

1. Avizenis A. Binary-compatible signet-digit arithmetic. IN: AFIPS Conf Proc. – Vol. 26 – P1. – 1964 – P.663.
2. Ch. Frougny, On-line finite automata for addition in some numeration systems. Theoretical Informatics and Applications 33 (1999), 79–101.
3. Самофалов К.Г., Луцкий Г.М. Основы построения конвейерных ЭВМ.- Киев: Вища школа, 1981. – 234 с.
4. Каляев А.В. Многопроцессорные системы с программируемой архитектурой. – М.: Радио и связь, 1984. – 240 с.
5. Методи конвеєрної порозрядної обробки послідовних кодів золотої пропорції / О.І. Черняк, О.Д. Азаров // Вісник ВПІ. – 1996. - №1. – С. 14-17.
6. Системи числення з адитивними та мультиплікативними співвідношеннями між вагами розрядів / О.Д.Азаров, О.І.Черняк, П.О.Черняк // Вісник ВПІ. – 2001. - №1. – С. 58-64.

Відомості про авторів

Азаров Олексій Дмитрович, д.т.н., професор, директор інституту інформаційних технологій та комп’ютерної інженерії, завідувач кафедри обчислювальної техніки Вінницького національного технічного університету; 21021, м. Вінниця, Хмельницьке шосе, 95; т.(0432)43-90-02, azarov@lili.vstu.vinnica.ua.

Черняк Олександр Іванович, інженер кафедри обчислювальної техніки Вінницького національного технічного університету, 21021, м. Вінниця, Хмельницьке шосе, 95; т.(0432)43-90-02,, alexandr.chernyak@gmail.com.

УДК 004.75, 004.724.2

ВДОСКОНАЛЕНА МОДЕЛЬ ДЛЯ ВИЗНАЧЕННЯ ТОПОЛОГІЧНОЇ ЛОКАЛЬНОСТІ В МЕРЕЖІ ІНТЕРНЕТ

Порєв Г.В.

Анотація: Методи визначення топологічної відстані є ключовими для підвищення ефективності однорангових та розподілених мереж. На відміну від багатьох існуючих або запропонованих методів, які зазвичай використовують зовнішню інфраструктуру, ми розробили і вдосконалили суто обчислювальну клієнтську метрику CARMA, яка базується на попередньо обчислених топологічних структурах Інтернет. Як показала експериментальна перевірка, ця метрика досить добре відповідає дійсній виміряній топологічній відстані.

Abstract: The topological distance estimation methods are keys to efficiency of peer-to-peer and distributed networks. Contrary to many existing or proposed methods which usually employ external infrastructure, we have developed and improved purely computational client-side CARMA metric based on the precomputed topological structure of the Internet. As the experimental validation results show, this metric fairly well corresponds to the actual measured distance.

Анотация: Методы определения топологического расстояния являются ключевыми для повышения эффективности одноранговых и распределённых сетей. В отличие от многих существующих и предложенных методов, которые обычно используют внешнюю инфраструктуру, мы разработали и усовершенствовали чисто вычислительную клиентскую метрику CARMA, которая базируется на предвычисленных топологических структурах Интернет. Как показала экспериментальная проверка, эта метрика достаточно хорошо соответствует реальному измеренному топологическому расстоянию.

Ключові слова: Інтернет, топологія, метрика відстані.

Вступ

В порівнянні з початком 20-го сторіччя, коли однорангові мережі тільки з’явилися, їх розвиток і різноманітність вирішуваних задач значно розширено. На сьогоднішній день існує ціла множина похідних від однорангових мереж технологій, зокрема файловий обмін, миттєві текстові повідомлення, VoIP, мультимедіа на вимогу, доставка програмного забезпечення тощо.

Різноманітні дослідження структури трафіку Інтернет вказують, що від 30% до 50% згенерованого користувачами трафіку має відношення до однорангових технологій, включаючи також USENET та нефайлообмінні служби. Хоча ідеологія однорангових мереж полягає, зокрема, в послабленні навантаження на централізовані серверні ферми, притаманна асиметрія каналів зв’язку у кінцевих користувачів змусила провайдерів Інтернет збільшити витрати на обслуговування та оновлення обладнання «останньої милі» для збереження якості надання послуги на належному рівні. Деякі оператори зв’язку також ввели непопулярні рішення, які виявляють та примусово знижують якість або швидкість каналу для трафіку, який визнано належним до однорангових мереж, що, в першу чергу, впливає на файловий обмін.

З цих причин дослідники шукають шляхи для оптимізації функціонування однорангових мереж в контексті регуляції трафіку, вважаючи потенціалом для оптимізації притаманну Інтернет кластерну структуру. Загальна ідея полягає в збільшенні швидкості обміну даними в середині певного сегменту мережі при мінімізації трафіку між такими сегментами. В даному випадку, масштаб поняття «сегмент» не визначено, оскільки звичайно можна виділити більш ніж один рівень сегментації Інтернет.

Деякі дослідники запропонували низку рішень, які передбачають побудову зовнішньої (по відношенню до шару однорангової мережі) інфраструктури, що спеціалізується на відстеженні умов внутрішньо-мережевих та між-мережевих каналів зв’язку, запам’ятовує явно знайдені шляхи маршрутизації та на основі цих даних видає прогнози маршрутизації та іншу інформацію. Такими рішеннями є, наприклад, P4P [1] та iPlane[2]. Інші пропозиції, такі як [3] передбачають пряме втручання в трафік однорангової мережі для приведення його у відповідність до політики оператора зв’язку.

В даній роботі ми пропонуємо інший підхід, який не потребує зовнішньої негарантованої інфраструктури для обчислення топологічних даних. Такий підхід був окреслений в попередній роботі, зокрема [4], де топологічна відстань була розглянута в контексті метричної дистанції для оптимізації побудови мінімальних покриваючих дерев.

Таким чином, метою даної роботи є вдосконалення як структурної моделі мережі Інтернет так і алгоритму знаходження метричного показника локальності для підвищення ефективності організації та функціонування однорангових мереж.

Огляд попередніх доробок

В роботі [4] метрична модель CARMA (Combined Affinity Reconnaissance Metric Architecture) надавала можливість оцінювати локальність між парою довільних адрес IPv4. Основною відмінністю цієї моделі від аналогів було те, що оцінка виконувалась суто внутрішніми обчисленнями, без прямих вимірювань або звертань до зовнішніх служб. З огляду на це, CARMA можна вважати наближеною математичною моделлю топології мережі Інтернет в цьому конкретному контексті відносної локальності мережних вузлів.

CARMA розпочинає роботу з попереднього завантаження структурної інформації з публічно доступних служб — регіональних Інтернет-реєстрів (PIP) та перетворення її на внутрішню графо подібну

структуру. На відміну від рішень, що базуються на інфраструктурі PlanetLab або таких, що використовують функції RouteViews, служби PIP є обов'язковими, публічними та суттєвими для функціонування Інтернет, і, таким чином, значно надійнішими. Файли, важливі для функціонування CARMA включають множину *delegated-latest-** з переліком зареєстрованих адрес та автономних систем та інші файли з бази даних WHOIS-сервера, які містять інформацію про під-діапазони та множини автономних систем (MAC).

Оскільки швидкість приросту бази даних PIP порівняно низька, системі CARMA буде достатньо оновлювати актуальні дані лише один раз на кілька днів. Як тільки дані завантажені, CARMA будує модель наближення топології Інтернет з деякими спрощеннями, яка містить наступні 4 структурні рівні: а) діапазони IPv4 розділені на б) під-діапазони і одночасно оголошені як належні до в) автономних систем (AC), які, в свою чергу, об'єднані в г) MAC. В рамках побудованої моделі вважається, що сутності нижчого рівня є явно з'єднаними між собою через сутність вищого рівня, в якій вони оголошені.

На основі такої структури та отримавши на вході пару довільних IPv4 адрес CARMA може обчислити їх взаємну локальність, яка може бути виражена в чотирьох класах:

- *Під-діапазон (subrange)* — визначає присутність другого вузла в тому самому під-діапазоні IPv4, що і перший, у відповідності до його визначення в базі даних під-діапазонів WHOIS сервера. Здебільшого це означає присутність обох вузлів в сегменті мережі, за яку відповідає один і той самий маршрутизатор. Наприклад, це можуть бути кінцеві користувачі, підключені до точки присутності оператора зв'язку або вузли всередині мережі університету, які зазвичай мають підключення лише через одного оператора зв'язку.
- *Діапазон (range)* — визначає присутність другого вузла в тому самому діапазоні IPv4, що і перший, у відповідності до його визначення в базі даних делегацій (*delegated-latest-**). Здебільшого це означає, що обидва вузли знаходяться у внутрішній мережі відділу або невеликої організації.
- *AC (as)* — визначає присутність другого вузла в адресному просторі, який анонсовано тією ж самою AC, що і для першого вузла. Хоча це не гарантує такої ж однозначної та безпосередньої зв'язності, як два попередніх класи, пакети між вузлами навряд чи вийдуть за межі корпоративної мережі або AC, оскільки AC є основною одиницею маршрутизації в Інтернет [5], і тому вони будуть оброблені та передані всередині мережі оператора зв'язку (пакети, адресовані на ті самі AC, керуються тими самими правилами маршрутизації впродовж всього шляху).
- *MAC (asset)* — визначає що вхідні вузли належать до різних автономних систем, об'єднаних в одну множину. Така множина може виявитися точкою обміну трафіком (IX), якщо кількість учасників точки обміну досить велика. Перевага такого визначення не одразу очевидна, але потрібно згадати, що в країнах де інфраструктура Інтернет лише розвивається, різниця в якості каналу та видимій швидкості передачі може прямо залежати від такого класу локальності та сягати кількох порядків.
- *Горизонт (distant)* — показує, що відносна локальність двох мережних вузлів не може бути однозначно обчислена і тому вони вважаються такими, що розташовані топологічно далеко.

Перевага, яку може надавати CARMA за допомогою цих класів, полягає, наприклад, в перевпорядкуванні списку вузлів, за допомогою яких даний вузол вносить себе в множину обміну або з яких починає запитувати потрібну інформацію. В принципі, вибираючи вузли з найближчим класом локальності, програмне забезпечення вузла може зменшити час відклику та збільшити швидкість обміну.

Аналіз швидкодії CARMA

Для оцінки швидкодії моделі CARMA було виконано ряд тестових запусків з вузлів, що мають IP адреси користувацького простору ВАТ «УкрТелеком», при цьому другі вузли пари бралися з 6 випадкових зграйок BitTorrent з одного з найбільш популярних публічних трекерів — RuTracker (www.rutracker.org). Розподіл результатів наведено на рис.1.

Очевидно висока кількість вузлів що попали під клас MAC пояснюється тим, що історично RuTracker має досить значну користувацьку базу з України. Модель CARMA змогла виявити до 19% таких вузлів, оскільки більшість їх з топологічної точки зору «належить» двом основним точкам обміну даних — UA-IX та DataGroup-IX [6]. Це дає на виході моделі CARMA клас MAC для всіх таких адрес, якщо перший вузол також «належить» до такої точки обміну.

Завдяки усталеній технологічній та бізнесовій практиці провайдерів-учасників українських точок обміну, полоса пропускання та ціна трафіку зовні та всередині UA-IX можуть значно відрізнятися, до кількох порядків. З огляду на це, ми вважаємо, що модель CARMA має значний потенціал оптимізації. Якщо правила відбору віддалених вузлів будуть формуватися у відповідності до класів локальності моделі CARMA, наприклад, в мережах типу BitTorrent, для вузлів що працюють в адресному просторі UA-IX або аналогічній національній конфігурації, значно менше вузлів повинні будуть з'єднуватися зовні своїх точок обміну і значно більше вузлів зможуть вибрати партнерів для обміну даними з тих, у кого більше шансів мати швидший канал зв'язку.

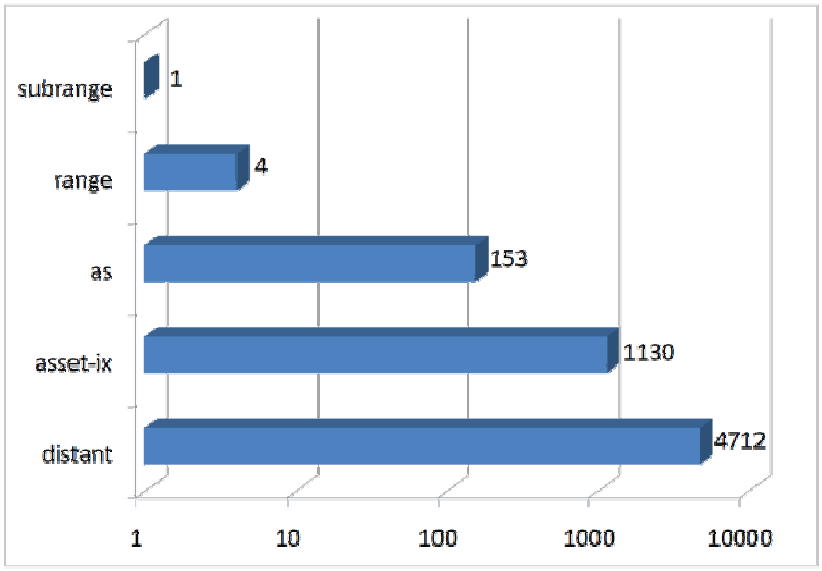


Рисунок 1 — Розподіл класів локальності по RuTracker

Вдосконалення моделі CARMA

Результати, отримані в серії тестів лише щодо трекера RuTracker не можуть бути визначальними. В подальших тестах ми вирішили взяти множини адрес зі зграйок BitTorrent щонайменше з двох трекерів — RuTracker та TorrentsNetUa (www.torrents.net.ua). Ці трекери відрізняються в одному ключовому аспекті, важливому для перевірки та підкреслення переваг моделі CARMA — відсоток вузлів в адресному просторі тієї самої точки обміну трафіком, що і вузол спостереження. За нашими приблизними оцінками, RuTracker має біля 20% а TorrentsNetUA має біля 95% українських користувачів он-лайн в будь-який момент. Якщо CARMA зможе експериментально підтвердити таку переважну кількість, це буде першим підтвердженням коректності моделі.

Однак подальші тести виконані з адресного простору UA-IX на зграйках TorrentsNetUa показали, що кількість вузлів, що були віднесені до класу «горизонт» значно менша, ніж очікувалося (рис.2).

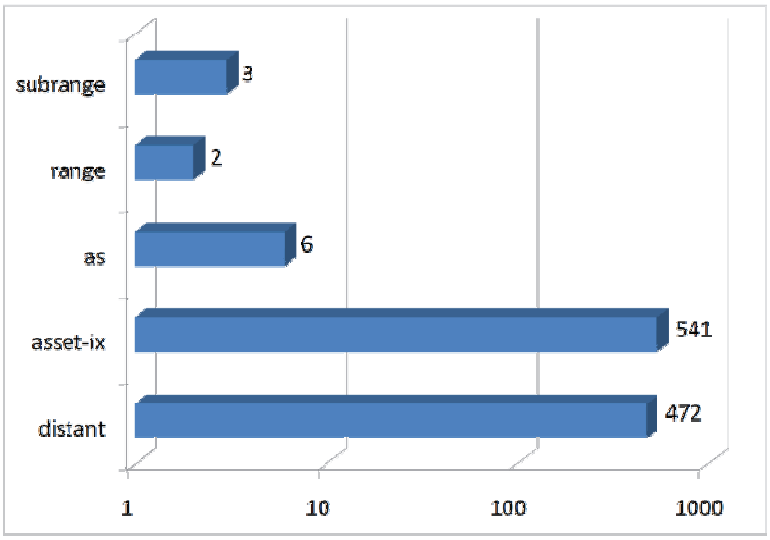


Рисунок 2 — Розподіл класів локальності по TorrentsNetUa

Вихідними даними цього графіку є випадково вибрані 1024 вузли зі зграйки на трекері TorrentsNetUa. Припускається, що переважна більшість з них працює в адресному просторі UA-IX, але тільки 54% були визначені як такі моделлю CARMA.

Така очевидна невідповідність зумовила пошук шляхів підвищення точності визначення класів. Оскільки модель CARMA демонструє досить високу швидкість (приблизно 10 пар адрес за секунду) обчислення класів навіть на комп’ютерах середнього рівня, нами було запроваджено додаткові перевірки

локальності в рамках згенерованої топологічної моделі, зокрема для двох додаткових різновидів взаємодії MAC.

Додаткові два класи локальності, що знаходяться вище «AC» та нижче, ніж «горизонт», такі:

- *MAC-Зв'язок (asset-link)* — показує, що вхідні адреси належать різним MAC, але принаймні один з цих MAC включає в себе інший.
- *Магістраль (backbone)* — визначає, що автономні системи, до яких належать задані діапазони IP-адрес, складають різні MAC, і існує третя MAC, яка включає в себе ці дві.

Ці два додаткові класи були введені також і на основі детального аналізу структури та властивостей точки обміну UA-IX, який наведено в роботі [6].

Експериментальна перевірка

Тепер, коли модель CARMA може віднести пару адрес до одного з семи класів локальності, її ефективність не може бути оцінена просто з видимої відповідності між прогнозом та розподілом класів в експерименті.

З цієї причини ми вирішили використати модифікований метод ICMP *traceroute* у вигляді модуля в програмі пакетної обробки CARMA. Модуль вважає що перша адреса на вході є власною і виконує серію запитів ICMP *traceroute* щодо другої адреси. Модифікований метод ICMP *traceroute* відрізняється від стандартного таким чином:

- Протокол ICMP — аналогічно Windows-версії засобу, метод використовує ICMP замість UDP в аналогічному засобі середовища GNU. Здебільшого це зумовлено тим, що середовищами для запуску CARMA є Windows x64 та x86, де частиною IP Helper API є легкі в використанні та з програмової точки зору інтерфейси ICMP.
- Не проводиться перетворення DNS імен — адреси проміжних вузлів шляху не перетворюються в імена вузлів, і самі адреси не виводяться, оскільки в даному випадку інтерес представляє лише їх кількість.
- Адаптивна перевірка — на відміну від засобу командного рядку, час чекання відклику задано одною секундою. Якщо останній вузол, що відповів на запит не є другим вузлом вихідної пари і перевищення часу очікування сталося будь-де на протязі маршруту, весь процес вимірювання кількості вузлів запускається спочатку з першого проміжного вузла. Однак це не може трапитися більш ніж тричі. Якщо цільового вузла не досягнуто на другу і третю спробу, за кількість вузлів переходу приймається найбільше число переходів з трьох спроб. Це ефективно прибирає наслідки випадкових мережних затримок, які можуть викликати тимчасове перевищення часу очікування.

Модифікований метод ICMP *traceroute* було перевірено з користувацького підключення типу ADSL всередині UA-IX, яке характеризується середнім часом відклику біля 50 мілісекунд до найближчого оточення з точки обміну та біля 300 мілісекунд з інших вузлів. Експерименти показали, що середня тривалість циклу вимірювання кількості переходів для вузла становить не більше 0.5 секунди, якщо вузол відповідає на запит; не більше 5 секунд, якщо вузол не відповідає на запити до часу очікування; не більше 3 секунд, якщо маршрутизатори повертають помилку досяжності. Однак, в середньому, вимірювання кожної тисячі пар вузлів займає близько 1 години.

Модуль було інтегровано в програму пакетної обробки CARMA таким чином, щоб для кожної пари після обчислення класу локальності було проведено вимірювання реальної кількості переходів з подальшим записом у спільний протокол експерименту. Результат вимірювання кількості переходів не заноситься в протокол тільки у випадках, коли виявлено петлю маршрутизації (приблизно 2 випадки на 1000 пар адрес).

Щоб отримати різноманітний розподіл класів локальності CARMA, як і розподіл кількості переходів, нами було використано обидва публічних трекери, як показано вище. Однак, обсяг та зграйок серйозно відрізнявся — RuTracker зміг надати зграйку в 3610 адрес, тоді як TorrentsNetUa мав певні труднощі зі звітуваннями зграйок обсягом вище 1000 адрес. Причиною цього є те, що на RuTracker дозволяється (а з недавнього часу є обов'язковим) включення механізмів розповсюдження DHT та PEX. Ці механізми протоколу BitTorrent [7] дають можливість отримувати повний список адрес безпосередньо від вузлів, які вже приймають участь в зграйці замість часткових (обмежених в кількості вузлів у відповіді) запитів до трекеру, які можуть суттєво навантажувати його і спричиняти збої в обслуговуванні.

Тестова зграйка з RuTracker складається з 3610 вузлів, з яких, як вважається, біля 20% має належати до простору UA-IX. TorrentsNetUa звітував зграйку обсягом 891 вузол, з яких практично всі належать до адресного простору UA-IX. Кожний вузол був оброблений вдосконаленою моделлю CARMA з адресою локального вузла (також всередині простору UA-IX). Потім було вимряно умовну відстань між цими вузлами в термінах кількості перехідних вузлів.

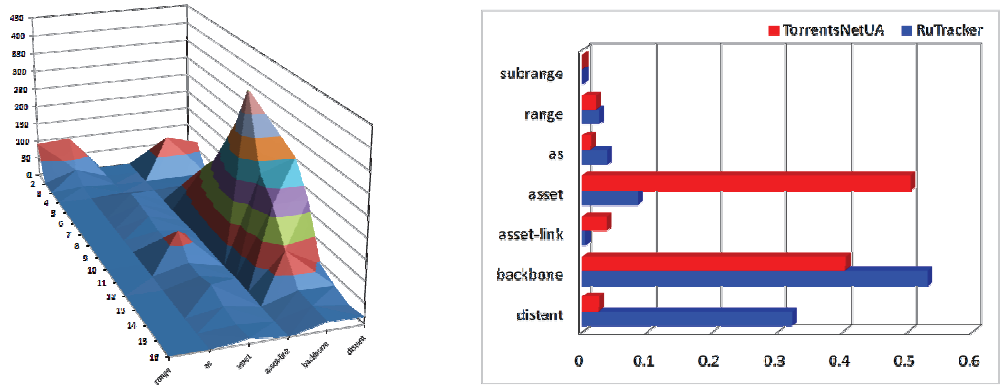
Експеримент тривав 5 годин. Щоб надати уявлення про характер результатів, їх значна частина наведена та нормалізований розподіл класів локальності наведені на рис.3.

Аналіз результатів та висновки

Аналіз отриманих результатів показав два суттєві покращення в функціонуванні моделі CARMA.

По-перше, розподіл класів по TorrentsNetUa показав, що кількість вузлів в класі «горизонт» зменшилася до 2.5% в порівнянні з попередніми експериментами, що підтверджує зроблений нами раніше прогноз.

Однак, потрібно відзначити, що в деяких випадках цей клас може означати вузол, який не оперує в адресному просторі UA-IX, але його адресний блок зареєстровано в Україні — якщо, наприклад, інформація в базах даних PIP є некоректною. В той же час можлива і обернена ситуація — в даному випадку не всі вузли, що отримали клас «магістраль» належать адресному простору українських точок обміну завдяки наявності кількох швидкісних каналів зв’язку між ними та російськими точками обміну через інфраструктуру TeliaSonera AB.



а) б)

Рисунок 3 — а) Характер відповідності обчислених класів локальності виміряній кількості перехідних вузлів; б) — нормалізований розподіл класів локальності в експерименті.

По-друге, помітний малий відсоток вузлів, що отримали клас «MAC-зв’язок» для обох трекерів може свідчити про порівняно низьку кількість кросс-AC зв’язків, які не пов’язані спільною MAC більш високого рівня, або загальну тенденцію до побудови ієрархічних політик маршрутизації в національних конфігураціях мережі Інтернет, що відповідає рекомендаціям, даним в роботі [6].

Досить несподіваним став факт відсутності вузлів з класом «під-діапазон». Це могло статися або через невеликий обсяг виборки, або завдяки тому факту, що в базі даних PIP існує перекриваючий діапазон з тими ж параметрами — в цьому випадку модель CARMA повертає більш високий клас локальності «діапазон».

Із застереженнями, що числове представлення обчисленого класу локальності, на відміну від кількості вузлів переходу, не має ніякого фізичного значення та з розумінням, що природа цих двох параметрів є повністю відмінною, з рис.3а можна зробити висновок про високий рівень відповідності обчисленого класу локальності в моделі CARMA та топологічної відстані в мережі.

Дана робота є логічним розвитком попередніх доробок авторів [4] з частковим залученням механізмів 2-го рівня загальної моделі CARMA. В даній роботі, зокрема, було виконане вдосконалення як структурної моделі мережі Інтернет так і алгоритму знаходження метричного показника локальності шляхом введення двох додаткових класів локальності та операцій для їх знаходження, що дає підгрунтя для подальшої розробки та ефективного використання комбінованої метрики локальності в рамках загальної моделі CARMA.

Впровадження повної реалізації загальної моделі CARMA надає механізми для оптимізації як існуючих розподілених і однорангових мереж так і розробки оптимальної архітектури нових мереж.

Робота виконана в рамках гранту президента України для підтримки наукових досліджень молодих вчених, НДР GP/F27-0040 «Дослідження регуляційних та рейтингових методів в однорангових мережах типу BitTorrent» договір № Ф27/8-2010 від 12 лютого 2010.

Список літератури

1. H. Xie, A. Krishnamurthy, A. Silberschatz, and R. Y. Yang, “P4P: Explicit communications for cooperative control between P2P and network providers”, 2007: http://www.dcia.info/documents/P4P_Overview.pdf
2. H. Madhyastha, T. Isdal, M. Piatek, and C. Dixon, “iPlane: An information plane for distributed services” In Proceedings of the 7th USENIX Symposium on Operating Systems Design and Implementation. USENIX, 2006, pp. 367–380.

3. T. Karagiannis, P. Rodriguez, and K. Papagiannaki, “Should internet service providers fear peer-assisted content distribution?” in IMC’05: Proceedings of the 5th ACM SIGCOMM conference on Internet measurement. New York, NY, USA: ACM, 2005, pp. 63–76.
4. G. Poryev, H. Schloss, and R. Oechsle, “CARMA based MST approximation for multicast provision in P2P networks” In Proceedings of the Sixth International Conference on Networking and Services (ICNS 2010). Cancun, Mexico: IEEE, 2010, pp. 123–128.
5. J. Hawkinson and T. Bates, “Guidelines for creation, selection, and registration of an Autonomous System (AS)”. RFC 1930 (Best Current Practice), Internet Engineering Task Force, mar 1996. http://www.ietf.org/rfc/rfc1930.txt
6. Фурашев В., Зубок В., Ланде Д. «Параметры украинского сегмента интернет как сложной сети» // Открытые информационные и компьютерные технологии: Сб. научн. трудов. Вып. 40.—Харьков: Нац. аэрокосм. ун-т. "ХАИ", 2008.—С.235-242.
7. B. Cohen, “Incentives build robustness in BitTorrent” in the First Workshop on the Economics of Peer-to-Peer Systems, 2003.—pp.140-144.

Відомості про авторів

Порєв Геннадій Володимирович — к.т.н., старший науковий співробітник, докторант Національного технічного університету України «КПІ», пр.Перемоги, 37, м.Київ, 03056, тел.. 258-52-09, +380-68-321-345-1, core@barvinok.net

УДК 681.324

СРАВНИТЕЛЬНЫЙ АНАЛИЗ РЕЗУЛЬТАТОВ РАЗЛИЧНЫХ ПОДХОДОВ К МОДЕЛИРОВАНИЮ СИСТЕМЫ ЗАЩИТЫ И НФОРМАЦИИ В РАСПРЕДЕЛЕННЫХ СЕТЯХ ОБСЛУЖИВАНИЯ

Б.Г. Исмайлов

Аннотация: Проведены сравнительный анализ результатов математических и имитационных методов решения задачи определения оптимальной программно-технической структуры систем защиты информации. Анализ результатов показывает, что они отличаются в пределах (2-10)%, а степень адекватности аналитической модели к исследуемому объекту увеличивается с уменьшением нагрузки сети.

Abstract: Comparison of mathematical and simulation methods to finding the optimal structure of systems of information protecting are carried out. Analysis of the results of the both methods shows that they differ in the interval of 2-10% and adequacy of analytical model increase with decreasing network’s load.

Актуальность проблемы

В данной работе проводится сравнительный анализ результатов различных подходов к решению проблемы определения оптимальной программно-технической структуры систем защиты информации (СЗИ) [1]. Такие системы создаются между различного рода распределенными сетями - с одной стороны и глобальной сетью с другой. Они инспектируют и фильтруют проходящие через них информацию. Это своего рода межсетевой шлюз (gateway), ориентированные на функции информационной защиты сетей. Такая структура межсетевых соединений позволяет резко снизить угрозу несанкционированного доступа в локальную и распределенную сеть за счет использования способа маскарада (masquerading), когда весь исходящий из РС трафик посылается от имени СЗИ, делая РС практически «невидимой». Методы передачи информации в таких сетях рассматриваются в [2-4].

Анализ показывает, что перечисление возможных угроз сети практически невыполнимо из-за огромного их количества. Поэтому на основе некоторых характерных особенностей, таких как запрограммированные или незапрограммированные действия, засорение почтового ящика, нахождение черных ходов, загромождение канала, выведение из строя компьютера и т.д. в [1] предлагается классификация возможных угроз сети, как троянский конь (ТК), вирусы (В), сообщение засоряющие почтовый ящик (СЗПЯ), черные ходы (ЧХ), атаки нацеленные на отказ сервиса (АНОС) и некорректно работающие программы (НРП).

Задача определения характеристик систем защиты информации (СЗИ) в РС решается на основе приведенной классификации возможных угроз сети. Исходя из практических соображений можно принять, что число злоумышленных программ (сообщений) составляет треть от общего числа сообщений.

Распределенные сети, функционирующие в условиях большой интенсивности при пуассоновском потоке с групповым поступлением, рассмотрены в [5] . Другие модели, связанные с развитием методов маршрутизации и управления потоком, анализируются в [6]. В целях обеспечения информационной безопасности сетей необходимо организовать некоторую систему защиты, включающие современные технические средства[7] для контроля передаваемой информации и их комплексной защиты от различных воздействий. Эти системы содержать некоторый комплекс программ, требующий определенный объем памяти:

- программы, осуществляющие криптографические шифрование почтовых сообщений, таких как Pretty Good Privacy (PGP);
- утилиты, позволяющие обнаруживать и уничтожать «шпионские» программы , например Ad – aware X cleaner;
- брандмауэры, обнаруживающие и блокирующие несанкционированный доступ к компьютеру, не допускающие попадания на жесткий диск «мусора», «шпионского» программного обеспечения и «троянов»;
- антивирусные программы (Antivral Toolkit, Kaspresky antivirius, Dr.Web и др.) и различные утилиты, направленные на борьбу с конкретными вирусами.

Проведение периодического анализа эффективности СЗИ является одной из основных задач в РС [7]. Путем совершенствования и оптимизации характеристик систем защиты можно обеспечить достаточно высокую их эффективность, труднопреодолимую даже опытному злоумышленнику. Решение этих проблем требуют разработки соответствующих математических моделей и метод их анализа.

В доступной литературе известны некоторые попытки решения указанных проблем [2-4]. Они в основном используют методы защиты информации на уровне имени и пароля пользователя, которые недостаточна для предотвращения входа в сеть посторонних лиц. Кроме того большое число потенциальных каналов проникновения в сеть еще больше усложняет защиту информации в сети. В отличие от указанных работ здесь предложен альтернативный подход к решению задачи по определению оптимальной программно-технической структуры СЗИ. Этот подход основан на принципах теории

систем массового обслуживания (СМО), учитывающих характеристики воздействя возможных угроз на функционирование сети.

Целью данного исследования является сравнительный анализ результатов математических и имитационных моделей систем защиты информации в распределенных сетях обслуживания.

Постановка задачи

Проблема состоит в разработке эффективного подхода к решению задачи сравнительного анализа результатов расчета и оптимизации функциональной структуры системы защиты информации в РС. В качестве математической модели СЗИ может служить следующая многоканальная СМО, состоящая из N компьютеров. Компьютеры характеризуются, в основном, интенсивностями обслуживания, распределенные по экспоненциальному закону, при этом на вход системы поступает пуассоновский поток сообщений с интенсивностью λ_P, время обслуживания подчиняется экспоненциального, постоянного и эрлангового закона распределения. Поступающая информация фильтруется и распределяется по сети.

Функционирование сети может быть нарушено со стороны злоумышленников и восстанавливаться с помощью комплекса программ, как во время передачи сообщений, так и в промежутке времени, когда передача сообщений не производится. Предполагается, что время передачи информации (T_{ci}), время исправной работы сети ($T_i = 1/d_i$), время ее восстановления (T_{ni}) и время старения информации ($T_D = 1/\nu$) в условиях возможных угроз со стороны злоумышленников распределены по экспоненциальному закону с параметрами μ_i, c_i, d_i, ν соответственно.

Показателем эффективности является математическое ожидание вероятности потери от несвоевременного распределения сообщений после фильтрации по компьютерам сети, т.е. требуется решить следующую задачу [1]:

$$M\left[\bar{P}\right] = \min \sum_{i=1}^{N} p_i q_i \tag{1}$$

при ограничениях

$$1 - \sum_{i=1}^{N} p_i = 0, \quad 0 \le P_i \le \mu_i k_i / \lambda, \quad i = \overline{1, N}, \tag{2}$$

$$\text{где } \mu_i = 1/T_{ci},\ h(p) = 1 - \sum_{i=1}^{N} p_i,\ q_i(p) = p_i,\ q_2(p) = p_i \le \frac{\mu_i k_i}{\lambda}, \tag{3}$$

$$q_i = \frac{(1 - p_i \lambda h_{1i})}{(1 - p_i \lambda h_{1i} + \nu_i h_{1i}}, \ h_{1i} = \frac{1}{\mu_i k_i}, \ \nu_i = \nu\left[k_i + \frac{(1-k_i)(\nu - p_i\lambda)}{\nu(1+\nu T_{ni}}\right], \ i = \overline{1, N}. \tag{4}$$

Здесь приняты следующие обозначения: $\bar{P}$ - вероятность потери от несвоевременного распределения сообщений по компьютерам после фильтрации в РС, p_i - вероятность потери от непопадания сообщений для передачи на i-й компьютер, q_i - вероятность потери в i-м компьютере от несвоевременной доставки сообщений, k_i - коэффициент готовности, T_{ni} - среднее время простоя.

Для решения задачи (1)-(4) в [1] предлагается использовать метод обобщенного приведенного градиента [8], на основе который разработаны алгоритмы расчета и оптимизации характеристик СЗИ относительно выбранного критерия качества. Этот метод позволяет исследовать поведение СЗИ при любых диапазонах изменения структурных и нагрузочных параметров модели.

Анализ результатов математической модели СЗИ в РС

С целью расчета характеристик СЗИ на основе разработанного алгоритма проведены объемные вычислительные эксперименты в широких диапазонах изменения как структурных, так и нагрузочных параметров модели. Так, для исходных данных $1/\mu_i = (0.042; 0.042; 0.063)$, $k_i = (0.97; 0.79; 0.81)$, $T_{n_i} = (0.6; 0.9; 1.0)$, $\nu = 0.5 (T_D = 2)$ исследованы зависимости $p_i = f(\lambda), i = \overline{1,3}$. Соответствующие результаты показанны в табл.1.

Таблица 1. Зависимости $p_i = f(\lambda), i = \overline{1,3}$.

λ	p_1	p_2	p_3
1/1000	1.00	0	0
1/2000	0.80	0.20	0
1/3000	0.78	0.22	0
1/4000	0.76	0.24	0
1/5000	0.71	0.29	0
1/6000	0.73	0.27	0
1/7000	0.69	0.21	0.10
1/8000	0.65	0.23	0.12
1/9000	0.62	0.25	0.13
1/10000	0.59	0.27	0.14
1/11000	0.56	0.29	0.15

В силу допустимых потерь информации полученные в табл.1, определены основные характеристики многоканальной СМО для экспоненциального, постоянного и Эрлангового времени обслуживания. Здесь основные характеристики являются L_q - длина очереди, L_s - количество сообщений в системе, τ_q - время ожидания сообщений в очереди, τ_s - время пребывания сообщений в системе (см. табл.2-4)..

Таблица 2. Результаты для экспоненциального времени обслуживания

$a = \lambda / \mu N$	L_q	L_s	τ_q	τ_s
0.95	7.75	8.23	10856	13210
0.67	0.53	2.34	795	3497
0.46	0.14	1.92	172	2708

Таблица 3. Результаты для постоянного времени обслуживания

$a = \lambda / \mu N$	L_q	L_s	τ_q	τ_s
0.95	1.138	2.69	1998	3805
0.67	0.46	2.30	547	3057
0.46	0.087	1.82	117	2642

Таблица 4. Результаты для Эрлангового времени обслуживания

$a = \lambda / \mu N$	L_q	L_s	τ_q	τ_s
0.95	3.72	4.21	3850	7250
0.67	0.63	2.74	352	1497
0.46	0.17	1.35	94	1308

Эти результаты подтвердили теоретические ожидания относительно поведения функции потери от несвоевременного распределения сообщений по компьютерам после фильтрации в РС, с определенной схемой маршрута распределения сообщений в сети в условиях возможных угроз со стороны злоумышленников.

При построении СЗИ в РС, наряду с аналитическими методами моделирования зачастую используют и метод имитационного моделирования [9]. Последний метод представляет разработчикам возможность исследования объектов практически любой сложности. При этом имитационное моделирование используется как составная часть системы автоматизации проектирования на этапах эскизного и технического проектирования. К числу наиболее широко распространенных инструментальных средств, обеспечивающих поддержку принятия решения, относится General Purpose Simulation System (GPSS) [9].

Анализ результатов имитационной модели СЗИ в РС

Рассматриваются варианты имитационных моделей СЗИ в РС, имеющие следующие схемы маршрутизации сообщений двух типов. Маршрут обработки сообщений для первого типа выполняется по функциям (операции по защиты информации) 1-3, а для сообщений второго типа выполняется по функциям (операции по защиты информации) 4-6. Распределение выполняемые функции защиты информации по компьютерам $k_i, i = \overline{1,3}$ приведены в табл.5. Интервалы времени между поступающими сообщениями и времена их выполнения приведены в табл.6 и 7.

Таблица 5. Распределение функции защиты информации по компьютерам $k_i, i = \overline{1,3}$

Функции	Функция 1	Функция 2	Функция 3	Функция 4	Функция 5	Функция 6
варианты	k_1	k_2	k_3	k_1	k_2	k_3
1	k_1	k_2	k_3	k_3	k_2	k_1
2	k_1	k_2	k_3	k_3	k_1	k_2
3	k_1	k_2	k_3	k_1	k_2	k_3
4	k_1	k_2	k_3	k_2	k_1	k_3
5	k_1	k_2	k_3	k_2	k_3	k_1
6	k_2	k_1	k_3	k_1	k_2	k_3
7	k_2	k_1	k_3	k_1	k_3	k_2
8	k_2	k_1	k_3	k_2	k_1	k_3
9	k_2	k_1	k_3	k_2	k_3	k_1
10	k_2	k_1	k_3	k_3	k_1	k_2
11	k_2	k_1	k_3	k_3	k_2	k_1
12	k_3	k_1	k_3	k_1	k_2	k_3

Требуется определить среднюю загрузку каждого компьютера, среднее время обработки сообщений каждого типа, длина очередей на обработку для компьютеров, объем памяти, необходимый для данного потока сообщений. В модели таймер настроен на выполнение моделирования в течение установленного модельного времени. При необходимости таймер должен быть откорректирован. После прогона модели имитации СЗИ в РС получены результаты, которые показаны в табл.8-10.

Таблица 6. Интервалы времени поступления сообщений

Вариант	Интервал времени поступления сообщений первого типа	Интервал времени поступления сообщений второго типа
	30 ± 5	20 ± 5
1	25 ± 4	25 ± 6
2	20 ± 3	30 ± 7
3	15 ± 5	35 ± 8
4	10 ± 4	20 ± 5
5	30 ± 5	10 ± 3
6	15 ± 4	15 ± 6
7	30 ± 10	15 ± 3
8	20 ± 5	20 ± 5
9	25 ± 4	10 ± 3
10	45 ± 5	15 ± 5
11	20 ± 4	15 ± 3
12	10 ± 3	15 ± 5

Таблица 7. Интервалы времени выполнения функции защиты информации (мин)

Вариант	1	2	3	4	5	6
	5 ± 2	2 ± 4	10 ± 3	7 ± 3	15 ± 5	15 ± 5
1	20 ± 4	5 ± 2	15 ± 5	15 ± 5	7 ± 3	10 ± 3
2	10 ± 3	15 ± 3	5 ± 2	20 ± 4	10 ± 3	7 ± 3
3	18 ± 3	10 ± 3	12 ± 5	20 ± 4	25 ± 8	12 ± 4
4	12 ± 5	15 ± 5	18 ± 3	10 ± 3	5 ± 2	20 ± 4
5	15 ± 5	20 ± 4	10 ± 3	18 ± 3	12 ± 5	2 ± 4
6	10 ± 3	25 ± 8	5 ± 2	14 ± 5	18 ± 3	15 ± 5
7	15 ± 5	12 ± 5	20 ± 4	5 ± 2	10 ± 3	18 ± 3
8	20 ± 4	18 ± 3	10 ± 3	7 ± 3	15 ± 5	25 ± 8
9	10 ± 3	15 ± 5	10 ± 3	12 ± 5	5 ± 2	20 ± 4
10	25 ± 8	5 ± 2	12 ± 5	7 ± 3	10 ± 3	15 ± 5
11	20 ± 4	10 ± 3	15 ± 5	5 ± 2	12 ± 5	25 ± 8
12	12 ± 5	20 ± 4	25 ± 8	15 ± 5	5 ± 2	10 ± 3

Таблица 8. Средняя загрузка компьютеров (в %)

компьютер	в течение 8ч.	в течение недели
k_1	48	52
k_2	92	98
k_3	88	96

Таблица 9. Максимальная длина очередей к компьютерам

компьютер	в течение 8ч.	в течение недели
k_1	1	1
k_2	12	59
k_3	2	3

Таблица 10. Среднее время обработки сообщений на компьютерах (в мин.)

компьютер	в течение 8ч.	в течение недели
k_1	5.9	6.31
k_2	16.33	16.99
k_3	12.38	13.18

По результатам моделирования можно сделать вывод о том, что общее число обработанных сообщений в течение 8 часов составляет 40, в в течение рабочей недели составляет 142. Эти данные позволяют рассчитывать необходимого объема памяти для СЗИ в РС. При этом первый компьютер k_1 загружен на 50% и перегружен компьютер k_2 (об этом говорит средний процент использования 98% и длина очереди 59). При этом компьютер k_3 загружен оптимально. Отметим, что для повышения эффективности функционирования СЗИ в РС при данном потоке сообщений можно использовать два компьютера.

С целью определения оптимальной структуры СЗИ в РС при заданном потоке сообщений можно продолжить прогона модели. Кроме того, если структуру сети менять нельзя, то используя возможности языка моделирования GPSS можно подобрать такой поток сообщений, который позволил бы возможность загружать сети оптимально.

Разработка алгоритма анализа и сравнение результатов математических и имитационных методов

С целью сравнительного анализа результатов математических и имитационных методов разработан алгоритм, который имеет следующие шаги.

Шаг 1. Построение модели имитации для различных случаев.

Шаг 2. Выполнение процесса имитации при нормальных условиях, получение различных вариантов и обоснование модели.

Шаг 3. Сравнение результатов математических и имитационных моделей.

Шаг 4. Если результаты математических и имитационных моделей сходятся, выполняются имитации для пиковой нагрузки. В противном случае система расширяет свои возможности (т.е. увеличивается значения структурных параметров).

Шаг 5. Осуществляется процесс тестирования (построение и обработка РС и проверка всех функций).

Шаг 6. Выполнение имитации при нормальных условиях и построение для получения различных вариантов.

Шаг 7. Проверка сходимости результатов. Если они сходятся, сеть проверятся дополнительно в условиях пиковых нагрузок. В противном случае сеть расширяет свою возможность и осуществляется переход к четвертому шагу.

Сравнения результатов математических и имитационных моделей осуществляется так:

$$\Delta P = \left[\frac{(P^{\bullet} - P)}{P} \right].100\,\%$$

где $P^{\bullet}$, P – значения характеристик математических и имитационных моделей, соответственно.

На основе разработанного алгоритма анализ результатов полученных после прогона обоих модели показывает, что они отличаются в пределах 2-10%, а степень адекватности математической модели к исследуемому объекту увеличивается с уменьшением значения нагрузки $a = \lambda / \mu N$.

Выводы и рекомендации

Проведен сравнительный анализ результатов математических и имитационных моделей системы защиты информации в распределенных компьютерных сетях. Проведены вычислительные эксперименты на основе разработанных алгоритмов и получены численные результаты, которые могут быть использованы при построении СЗИ в РС производственного назначения.

Анализ результатов обоих модели показывает, что они отличаются в пределах 2-10%, а степень адекватности математической модели к исследуемому объекту увеличивается с уменьшением нагрузки $a = \lambda / \mu N$.

Список литературы

1. Исмайлов Б.Г. Исследование характеристик систем защиты информации распределенной сети // Автоматика и вычислительная техника. Рига.: 2006, №3. с.51-59, {ISSN 0132-4160}.
2. Герасименко В.А., Малюк А.А. Основы защиты информации. М.: ППО «Известия» УДПРФ. 1997. 372 с., {ISSN 002-2197}.
3. Герасименко В.А. Защита информации в автоматизированных системах обработки данных. Развитие, итоги, перспективы // Зарубежная радиоэлектроника, 1993, №3, с.3-21.{ISSN3.3-14/111-93}.
4. Грушо А.А., Тимонина Е.Е. Теоретические основы защита информации. М.: Издательство Агентства «Яхтсмен», 1996. 130с. {ISBN 11-96}
5. Алиев А.А., Исмайлов Б.Г. Анализ характеристик многопотоковых сетей обслуживания // Радиоэлектроника, информатика и управление, Запорожье.: 2001, №2. с.66-69 {ISSN 01886}
6. N.F. Maxchemchuk, ,M.Ei.Zarki. Routing and flow control in high–speed wide area networks // Proc. of the IEEE. 1990, Vol. 78, N1. p. .204-221. {ISBN 01890}.
7. Алябев С.В. Проблемы защиты информации в сети промышленного предприятия // Сб.трудов ПУКИ, Выпуск 8 (Воронеж).Центральное Черноземное книжное издательство, 2003. с.69-70. {ISBN 5.7458-0575-7}.
8. Химмелблау Д. Прикладное нелинейное программирование.- М.: Мир,1975.,540с. {ISBN 01/75}.
9. Шрайбер Т.Дж. Моделирование на GPSS.- М.: Машиностроение,1980,-592с. {ISSN 02880}.

Сведения об авторах

Исмайлов Балами Гасым оглы, доцент кафедры Информатики Сумгаитского Госдарственного Университета, гор. Сумгаит 43-й квартал. Email: Balemi@rambler.ru

УДК 004.75

ШВИДКИЙ АЛГОРИТМ ВИЗНАЧЕННЯ ВПЛИВУ НЕПОЛАДОК ЕЛЕМЕНТІВ МЕРЕЖІ ДОСТУПУ НА ЯКІСТЬ ОБСЛУГОВУВАННЯ КЛІЄНТІВ

Теленик С.Ф., Ролік О.І., Ясочка М.В., Малюгін Д.В.

Анотація: Розглядається важлива для великих операторів телекомунікаційних послуг проблема визначення впливу відмов елементів мережі доступу на якість обслуговування клієнтів. Пропонується підхід до швидкого розв’язання проблеми в умовах жорстких реальних вимог операторів до часу реагування на неполадки у мережах, в основу якого покладена ідея комбінування продукційного і алгоритмічного підходів. Наводиться відповідний алгоритм, побудований на основі роботи з продукційними системами, які забезпечують гнучкість моделі розповсюдження і оцінювання впливу неполадок елементів мережі на якість обслуговування клієнтів, і прийомах роботи з графами, які забезпечують дотримання часових вимог. Реалізація алгоритму базується на збереженні даних про мережу та їх обробленні у оперативній пам’яті. Наводяться експериментальні дані дослідження запропонованого алгоритму.

Аннотация: Рассматривается важная для больших операторов информационно-телекоммуникационных услуг проблема определения влияния отказов элементов сети доступа на качество обслуживания клиентов. Предлагается подход к быстрому решению проблемы в условиях жестких реальных требований операторов ко времени реакции на отказы в сетях, в основу которого положена идея комбинирования продукционного и алгоритмического подходов. Приводится соответствующий алгоритм, построенный на основе работы с продукционными системами, которые обеспечивают гибкость модели распространения и оценки влияния отказов элементов сети на качество обслуживания клиентов, и приемах работы с графами, которые обеспечивают выполнение временных требований. Реализация алгоритма базируется на сохранении данных о сети и их обработке в оперативной памяти. Приводятся экспериментальные данные исследования предложенного алгоритма.

Abstract: The problem of estimation of net elements flow influence on client service quality at large communication operators is considered. The approach based on available solutions review and analysis of influence factor structure-technology and service-function intended for solving this problem at breakneck speed is proposed. The models of distribution and estimation of access net elements flaw influence in form of tasks on special graphs is elaborated. The algorithm, based on production system and methods of working with graphs, is proposed. Implementation of the algorithm is based on storing and processing access net data in operational memory. The results of proposed algorithm experimental investigation are described.

Вступ

Стрімкий розвиток інформаційних (ІТ) та комунікаційних (КТ) технологій призвів до глобальної інформатизації суспільства [1]. Одним із її проявів є створення густої сітки телекомунікаційних мереж як великомасштабного середовища надання операторами зв’язку різноманітних комунікаційних послуг своїм клієнтам. У таких мережах виділяються транспортні мережі та мережі доступу.

У мережах доступу великих операторів зв’язку може налічуватися декілька мільйонів клієнтів. Кожний з клієнтів має певний набір параметрів, належить до певного класу і має певний набір вимог до рівня обслуговування, зафіксований в угоді про рівень обслуговування SLA (Service Level Agreement).

Коли у мережі доступу з’являються неполадки в комунікаційних пристроях чи сервісах, у клієнтів знижуються показники якості обслуговування. Для того, щоб визначити політику оператора щодо взаємодії з клієнтами на період впливу неполадок, службі обслуговування клієнтів потрібні показники зменшення рівня обслуговування для кожного клієнта. Розрахунки зазначених показників залежать від багатьох параметрів, насамперед сервісів, технологій, які застосовуються оператором, типу обладнання, структури мережі. При цьому логіці здійснення розрахунків притаманна змінність. А час, за який необхідно виконати зазначені розрахунки, загалом досить обмежений.

Постає проблема представлення мереж доступу і організації розрахунків показників зменшення якості обслуговування клієнтів внаслідок неполадок в мережі у реальних умовах функціонування.

У статті пропонується підхід до швидкого розв’язання зазначеної проблеми на основі комбінування продукційного і алгоритмічного підходів, наводиться алгоритм одержання потрібних результатів у межах часових вимог з відповідним рівнем точності.

Сутність проблеми

Мережі операторів поділяються на два основних класи — магістральні мережі передавання даних і мережі доступу. Перші з них становлять ядро великих мережних інфраструктур і призначені для забезпечення зв’язку між великими телекомунікаційними вузлами або регіональними підмережами. Магістральні мережі використовують канали зв’язку з великими пропускними здатностями, об’єднують інформаційні потоки значної кількості підмереж і відрізняються підвищеною надійністю.

Мережі доступу становлять один з ключових компонентів інформаційної інфраструктури, за допомогою яких абонентам надаються різноманітні інформаційні та телекомунікаційні послуги. У мережах доступу деяких операторів може налічуватися до $n \cdot 1000000$ клієнтів, а витрати на створення, модернізацію і підтримку працездатності таких мереж складають вагому частку бюджету операторів. Ці мережі мають переважно ієрархічну будову, кількість рівнів якої залежить від кількості клієнтів і їх розташування, причому кількість рівнів мереж доступу деяких операторів може досягати 10. Для забезпечення вимог до надійності обслуговування клієнтів їм забезпечується доступ до ресурсів мережі

декількома шляхами. Загальна ієрархічна будова мережі оператора з основними і резервними лініями зв’язку детально описана у попередній праці авторів [2]. Її важливий фрагмент буде наведений нижче.

Таким чином, загальна структура мережі доступу в термінах теорії графів не є кореневим деревом, а становить ярусно-паралельну форму спеціального вигляду: на найвищому ярусі зосереджені ресурси, а на найнижчому, тобто у листках — клієнти; кожна вершина може мати n батьківських вершин з різних ярусів і m дочірніх вершин з різних ярусів. Але принципово важливим є той факт, що цей граф можна зобразити у вигляді об’єднання двох і більше кореневих дерев, які можуть перетинатися. Перетин цих графів визначає ті пристрої і лінії зв’язку, які використовуються для надання основних і додаткових шляхів доступу клієнтам.

У кожному вузловому комунікаційному пристрої кожна лінія зв’язку задіює певний порт. Таким чином, ці пристрої структуровані і неполадки можуть стосуватися усього пристрою в цілому або певної його підструктури. При цьому мова може йти про повну втрату функції пристрою (підструктури), чи втрату частини функцій пристрою (підструктури).

Тому оператор для кожного пристрою вводить спеціальний параметр оцінки рівня його відповідності нормам функціонування — severity level, який набуває значень в діапазоні від 0 до 100. Значення кожної позиції цього параметра для оператора чітко визначене, тобто оператор знає наслідки для клієнта кожного значення у зазначеному діапазоні.

Існує набір правил, за допомогою яких, знаючи значення цього показника для кожного комунікаційного пристрою мережі, структуру мережі, параметри технологій, сервісів і обладнання, можна розрахувати значення показника для клієнта. Зазначений набір правил досить громіздкий. Мережа доступу оператора може налічувати мільйони клієнтів, які можуть належати до різних класів за важливістю для оператора, наприклад, можна говорити щонайменше про клієнтів класів platinum, gold, silver, bronze. При цьому на розрахунок показників для усіх клієнтів відводиться обмежений час, який у деяких операторів становить не більше 15 хвилин на 1,5 млн. клієнтів.

Необхідне відповідне програмне забезпечення, яке, по-перше, буде відповідати за ведення необхідних баз даних і виконання розрахунків в обмежений час, а, по-друге, буде інтегрованим в системи рівня Operations Support System (OSS) для управління мережами доступу операторів зв’язку.

Інформація щодо мережі подається в неструктурованому в термінах графів вигляді. Задаються лише вузлові пристрої і пристрої наступного рівня, приєднані до їх відповідних портів, та інформація щодо пристроїв, портів і зв’язків.

Коли в системі трапляються неполадки, інформація про них надходить у вигляді нових значень параметрів оцінки рівня їх відповідності нормам функціонування. Тоді й необхідно врахувати нові значення показників обслуговування клієнтів з урахуванням наведених вище часових обмежень.

З урахуванням суттєвих особливостей логіки визначення впливу неполадок у мережах доступу, наявності значної кількості чинників впливу, великої вимірності мереж, значної кількості клієнтів і жорстких часових обмежень виникає складна практична проблема представлення інформації про структуру мережі і правила перерахунку показників і створення програмного інструментарію оцінювання впливу неполадок у мережах доступу операторів зв’язку на якість обслуговування клієнтів.

На цю проблему накладаються додаткові обмеження, пов’язані з тим, що періодично q разів на добу поступає інформація від інших компонентів систем рівня OSS управління мережами доступу операторів зв’язку про зміни в структурі мережі.

Тому виникає наукова проблема розроблення моделей і відповідних методів і алгоритмів перерахунку, здатних скласти основу програмного інструментарію, який дозволить вирішити наведену практичну проблему в рамках зазначених часових, технологічних та інших обмежень.

Огляд існуючих рішень

У цілому, аналізу впливу неполадок у мережах великих операторів інформаційно-комунікаційних послуг присвячено доволі багато праць, назвемо хоча б праці [2—6]. Вони присвячені різним аспектам розв’язання цієї актуальної проблеми. Більшість праць присвячені аналізу впливу відмов у наданні сервісів на бізнес-операції компаній, які надають сервіси або використовують IT для підтримки своїх бізнес-процесів [3—4]. Автори цих праць переважно концентруються на моделях і методах ефективного управління інцидентами і процесами відновлення сервісів. На їх основі розроблюються системи, метою яких є мінімізація впливу на бізнес зривів найважливіших сервісів. Базовою концепцією побудови цих систем є інтеграція управління виконанням бізнес-процесів з управлінням системою у цілому і забезпечення їх відповідною IT-інфраструктурою.

Деякі праці присвячені розробленню моделей, які визначають залежності сервісів від інших сервісів і ресурсів, а також параметрів якості сервісів у SLA клієнтів і поточного рівня використання сервісів. На основі цих моделей розроблюються процедури прийняття рішень для ефективного відновлення сервісів [5]. Значна частка праць присвячена моделям аналізу впливу неполадок у певних типах мереж, наприклад IP-телефонії [6].

Ще одним напрямком досліджень є розроблення моделей впливу на бізнес операційних ризиків, які є результатом змін, пов’язаних зі зривами у наданні сервісів на невизначений час [7].

Якщо для управління виконанням запропоновано багато моделей загального вигляду, наприклад [8—9], які знайшли втілення у реальних системах управління IT-інфраструктурою [10], то інформація щодо загальних моделей оцінювання впливу неполадок у мережах великих операторів інформаційно-комунікаційних мереж на якість обслуговування клієнтів у літературних джерелах відсутня. У статті [2] запропоноване рішення для задачі визначення впливу неполадок у мережі доступу на якість сервісів, які надаються операторами зв’язку клієнтам. Це рішення базується на визначенні системи правил поширення впливу неполадок у мережі на якість сервісів і загальної схеми застосування цієї системи, яка враховує структуру мережі і особливості сервісів.

Але виникає потреба у розвитку зазначених положень, доведенні їх до рівня завершеного підходу до оцінювання якості сервісів на основі використання правил. Крім того, необхідно розробити конкретний алгоритм розрахунків параметрів впливу і експериментально перевірити його працездатність та можливість одержання потрібних результатів у прийнятний для реальних мереж доступу час.

Постановка проблеми дослідження

Нехай задана інформація про складові мережі доступу і їхні взаємозв’язки, час між змінами в інформації про складові мережі доступу і їхні взаємозв’язки і час між надходженням повідомлень про неполадки в мережі і/або запитами від користувачів щодо визначення рівня їхнього обслуговування.

Необхідно розробити спосіб представлення інформації про мережу доступу, розробити і експериментально дослідити алгоритм розрахунку показників рівня обслуговування клієнтів відповідно до заданих змін у показниках функціонування вузлів.

Підхід до розв’язання проблеми

Очевидно, ефективно розв’язати поставлену проблему можна лише за наявності продуманого представлення мережі доступу і відповідного способу визначення ступеня впливу неполадок на якість обслуговування клієнтів оператора.

Враховуючи змінність мереж, їх обладнання і технологій, мінливість політики операторів щодо базових принципів оцінювання впливу, спосіб оцінювання не доцільно базувати на основі алгоритмічного підходу. Варто використовувати системи, побудовані на правилах — продукційні системи. Правила таких систем (продукції) формують базу знань, а механізм виведення визначає порядок застосування правил, використовуючи робочу область, у якій накопичуються дані щодо поточного стану системи. Таким чином, дані (правила і робоча область) відокремлені від процедур (механізм виведення) і необхідність перепрограмування системи при зміні окремих правил не виникає.

З іншого боку, оскільки продукційні системи загалом побудовані на символьних перетвореннях, їх застосування не може забезпечити виконання розрахунків в умовах жорстких часових обмежень.Тому найперспективнішим видається використання спрощеного представлення мережі, яке містить результати опрацювання правил продукційної системи у вигляді, готовому для прямих розрахунків. Найкращим варіантом видається таке опрацювання правил, коли умовні частини продукцій перевіряються заздалегідь, вирішуються можливі конфлікти і на відповідному шляху у графі зберігаються потрібні для розрахунку впливу значення коефіцієнтів поширення впливу неполадок. У цьому випадку визначення впливу зводиться до перегляду графу від вершин верхніх ярусів до листків, розпочинаючи з вершин з неполадками, і виконанні послідовності арифметичних операцій над коефіцієнтами поширення.

Дійсно, виходячи з різноплановості вимог, доцільно представлення мережі розбити на повне і спрощене представлення. Перше відповідає за склад і структуру мережі, користувачів, сервіси, обладнання і технології, характеристики елементів мережі. Звичайно, воно становить ядро підсистеми інвентаризації — центрального компонента системи управління оператора. У великих базах даних цієї підсистеми дані зберігаються і обробляються за допомогою промислових СКБД. Будемо вважати, що інформація про склад, зв’язки і характеристики елементів мережі знаходиться у таких таблицях:

Objects (поля, які використовуються алгоритмом визначення впливу неполадок: ідентифікатор об’єкта; ім’я об’єкта; тип об’єкта);

Reference (поля, які використовуються алгоритмом визначення впливу неполадок: ідентифікатор об’єкта; посилання на пов’язаний з ним об’єкт; ідентифікатор атрибута);

Parameters (поля, які використовуються алгоритмами визначення впливу неполадок: ідентифікатор об’єкта; ідентифікатор атрибута; значення атрибута).

Спрощене представлення мережі потрібне для швидкого аналізу і оцінювання впливу стану мережі на рівень якості обслуговування клієнтів. Обсяг спрощеного представлення повинен дозволяти цілком завантажувати його в оперативну пам’ять для виконання у ній усіх необхідних розрахунків. Спрощене представлення у вигляді дерева розповсюдження складається з двох частин:

1) опис структури мережі доступу і структури сервісів та служб, які їх підтримують: мережа описується системою елементарних ієрархічних підграфів, що дозволяє застосувати ефективні в умовах зміни структур формальні засоби і структури даних;

2) опис особливостей врахування впливу неполадок у мережі доступу на рівень якості обслуговування клієнтів: задаються елементи мережі, які впливають на якість функціонування наступних елементів, сегментів або сервісів мережі і визначають напрям і об’єм необхідних обчислень (попереднім заданням коефіцієнтів поширення впливу неполадок для виконання безпосередніх обчислень у разі виникнення неполадок).

Отже, виконання розрахунків здійснюється у дві фази. Підготовча фаза призначена для формування необхідних для виконання розрахунків даних. Вона виконується модулем формування спрощеного представлення, який застосовується одноразово для формування зазначеного представлення із повного. Спеціальний модуль зміни спрощеного представлення застосовується кожного разу, коли у мережі відбуваються зміни. Оскільки організація розрахунків здійснюється на основі системи продукцій, на цій фазі правила заздалегідь прив’язуються до структури мережі у спрощеному представленні.

Правила у нотації Бекуса-Наура мають такий вигляд:

<система_правил>::=[<продукція>]

<продукція>::=<клас>P<назва><умова>→<наслідок>

<клас>::=propagation rule|severitylevel rule

<умова>::=[{-} <проста_умова>]

<проста_ умова>::=<клас_об’єкту><ім’я><атрибут><предикат><значення>

<клас об'єкту>::= client | service | technology | device | port | resource

<наслідок>::= [<команда>|<формула>|<програма>]

<предикат>::=≠|=|<|>|≤|≥

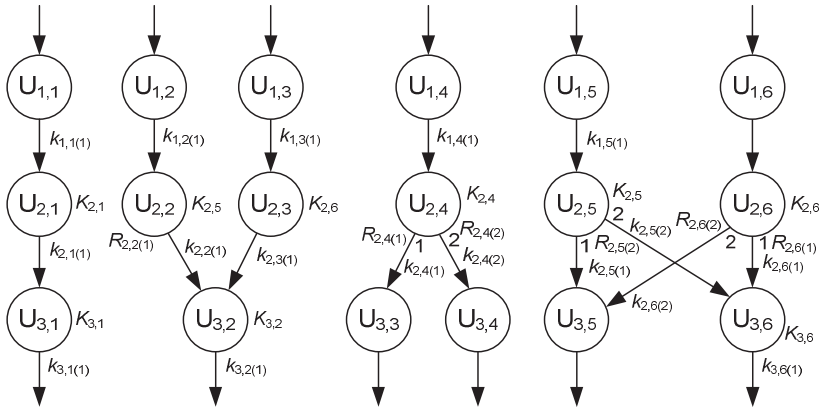
<атрибут>::=<ім’я_сервісу>|<ім’я_технології>|<тип_пристрою>|<параметр>|<тип_клієнта>|<ім’я_клієнта>

Параметрами можуть бути будь-які характеристики сервісів, технологій, пристроїв клієнтів. Значеннями можуть бути константи й імена. Команди виконують дії над елементами робочої області. Їх синтаксис залежить від типу (modify, remove, add).

Елементи робочої області (вони задаються кортежами повного представлення мережі доступу і елементами спрощеного представлення):

<елемент>::=(<клас об’єкта>Name<ім’я> [<атрибут><значення>])

Семантика продукційних систем традиційна, тобто ліві частини продукцій розглядаються як умови, а праві частини — як дії.



а) *б*) *в*) *г*)

Рисунок 1 — Можливі способи з’єднання вузлів мережі доступу і відповідні види правил

На другій фазі, у процесі безпосередніх обчислень підготовлене заздалегідь спрощене представлення опрацьовується алгоритмічно після надходження повідомлення про неполадку. Для цього можна використовувати швидкі алгоритми обходу графів.

Особливості застосування правил для аналізу впливу неполадок продемонструємо на правилах поширення впливу неполадок, коли в якості показників якості використовуються параметри продуктивності [2]. Ці правила наведені на рис. 1. Вони адекватно враховують структуру мережі доступу, основні структурні способи з’єднання її вузлів, особливості комунікаційних технологій, налаштування обладнання, залежність від величини коефіцієнтів зниження швидкості передавання у вхідних лініях, наявність резервування та інші параметри. Зроблені максимально простими для

підвищення швидкодії, ці правила формують основу відповідних продукцій. У наведених правилах для формального оцінювання впливу неполадок на вузли і абонентів мережі доступу введені:

1) коефіцієнт зниження швидкості передавання $k_{i,j(n)}$, який визначає, у скільки разів зменшується швидкість передавання в n-й лінії зв'язку, яка виходить із вузла $U_{i,j}$ (j-й вузол мережі на i-му рівні її ієрархії), відносно нормативного показника, і набуває значення у діапазоні [0, 1];

2) коефіцієнт $K_{i,j}$, який залежить від обумовленого неполадками стану вузла $U_{i,j}$ і визначає ступінь зниження швидкості передавання на виходах вузла $U_{i,j}$ відносно нормативного. Цей коефіцієнт також набуває значення у діапазоні [0, 1], причому $K_{i,j} = 0$ означає, що неполадки унеможливлюють трафік через вузол $U_{i,j}$, а $K_{i,j} = 1$ говорить про те, що неполадки у вузлі $U_{i,j}$ не впливають на швидкість передавання.

При виявленні неполадок у мережі виконується їх локалізація і визначається вплив на абонентів шляхом підрахунку за допомогою відповідних правил значень коефіцієнтів $K_{i,j}$ для потрібних вузлів.

Спосіб а). При справному вузлі $U_{2,1}$ швидкість доступу у вихідній лінії визначається вхідною швидкістю і обмеженнями на швидкість передавання і приймання, встановленими для портів вузла $U_{2,1}$ при його конфігуруванні. Якщо ж вузол $U_{2,1}$ несправний, то можливі два випадки. У першому випадку, коли коефіцієнт $K_{2,1} = 0$, швидкість у вихідній гілці дорівнює 0 незалежно від швидкості на вході. У другому — $K_{2,1} \neq 0$ і неполадка у вузлі $U_{2,1}$ призводить до падіння швидкості у вихідній лінії на коефіцієнт $K_{2,1}$. Тоді коефіцієнт падіння швидкості відносно номінальної у вихідній лінії зв'язку вузла $U_{2,1}$ буде визначатися таким чином

$$k_{2,1(1)} = K_{2,1} \times k_{1,1(1)}. \quad (1)$$

У свою чергу, коефіцієнт $k_{3,1(1)}$ падіння швидкості на виході вузла $U_{3,1}$ буде визначатися за допомогою такого виразу

$$k_{3,1(1)} = K_{3,1} \times k_{2,1(1)} = K_{3,1} \times K_{2,1} \times k_{1,1(1)}. \quad (2)$$

Відповідні цьому способу правила поширення розіб'ємо на таких дві групи:

а) правила прямого поширення впливу неполадок:

IF tt.nod = 1 ∧ type.nod = '…' ∧ class.nod = '…' ∧ type.child.nod = '…' *THEN* tt.child.nod = 1

IF tt.nod = 1 ∧ type.nod= '…' ∧ class.nod = '…' ∧ type.reference.nod = '…' ∧ type.nod.reference = '…' ∧ attribute.nod.reference = '…' *THEN* tt.reference.nod = 1

б) правила зворотного поширення впливу неполадок:

IF tt.nod = 1 ∧ type.nod = '…' ∧ class.nod = '…' ∧ type.parent.nod. = '…' *THEN* tt.parent.nod = 1

IF tt.nod = 1 ∧ type.nod = '…' ∧ class.nod = '…' ∧ type.nod = '…' ∧ type.nod.reference = '…' ∧ attribute.reference.nod = '…' *THEN* tt.reference.nod = 1

Для спрощення у цих правилах не визначаються імена класів об'єктів і значення атрибутів. Тут прості умови tt.nod = 1, tt.child.nod = 1, tt.parent.nod = 1 і tt.reference.nod = 1 означають, що ознака наявності неполадки приписана відповідно вузлу, дочірній вершині вузла, батьківській вершині вузла, вершині, на яку у вузлі є посилання. Інші прості умови зрозумілі із контексту.

Правила формування severity level мають ту ж структуру, але додатково встановлюють коефіцієнт впливу неполадки у відповідному вузлі (дочірньому чи батьківському, тому, що посилається, чи тому, на який посилаються).

Спосіб б). Якщо у вузол, наприклад, $U_{3,2}$, ведуть декілька ліній зв'язку, які використовуються для розподілення навантаження, а не для резервування, то при $K_{3,2} = 1$ для оцінювання впливу необхідно застосовувати вираз

$$k_{3,2(1)} = \begin{cases} k_{2,2(1)} + k_{2,3(1)}, \text{ якщо } k_{2,2(1)} + k_{2,3(1)} \le 1, \\ 1, \text{ якщо } k_{2,2(1)} + k_{2,3(1)} > 1. \end{cases}$$

Якщо коефіцієнт $K_{3,2}$ не дорівнює 1, то виконується відповідна корекція коефіцієнта $k_{3,2(1)}$.

Відповідні цьому способу правила поширення мають такий же вигляд, як і для попереднього випадку. Правила формування severity level повинні мати такий вигляд:

IF tt.nod_1 = 1 ∧ severity.nod_1 = k ∧ type.nod_1 = '…' ∧ class.nod_1 = '…' ∧ type.child.nod = '…' ∧ severity.nod_2 = m ∧ type.nod_2 = '…' ∧ class.nod_2 = '…' *THEN* tt.child.nod = 1 ∨ severity.child.nod = $f(k+m)$.

Тут використовуються ті ж прості умови, що і для способу а, але nod_1 і nod_2 позначають двох різних батьків. Змінні severity.nod_1 і severity.nod_2 набувають значення severity level для двох різних батьків. Тут і надалі за допомогою функції $f(x)$ підкреслюється необхідність розрахунку коефіцієнтів за наведеними вище формулами.

Спосіб в). Нехай $R_{i,j(n)}$ — це коефіцієнт, який визначає розподіл полоси пропускання між n вихідними лініями зв'язку j-го вузла i-го ієрархічного рівня. Цей коефіцієнт набуває значення із діапазону [0, 1], причому, якщо швидкість доступу у вхідній лінії повністю розподіляється між портами

вихідних ліній без додаткових обмежень, які встановлюються при конфігуруванні вузла $U_{2,4}$, то приймається

$$\sum_{n=1}^{N_{i,j}} R_{i,j(n)} = 1.$$

Тоді коефіцієнт падіння швидкості у 1-й та 2-й лініях зв’язку на виході вузла $U_{2,4}$ буде визначатися таким чином:

$$k_{2,4(1)} = R_{2,4(1)} \times K_{2,4} \times k_{1,4(1)},$$
$$k_{2,4(2)} = R_{2,4(2)} \times K_{2,4} \times k_{1,4(1)}.$$

Відповідні цьому способу правила поширення мають такий же вигляд, як і для попередніх випадків. Правила формування severity level повинні мати такий вигляд:

IF tt.nod = 1 ∧ severity.nod = k ∧ type.nod_1 = ‘…’ ∧ class.nod_1 = ‘…’ ∧ type.child.nod = ‘…’ ∧ type.child.nod = ‘…’ ∧ class.child.nod = ‘…’ *THEN* tt.child.nod = 1 ∨ severity.child.nod = $f(k)$.

Спосіб г). Для підвищення надійності у мережах доступу використовуються резервні лінії зв’язку. На рис. 1*г* такими лініями є лінії 2, які виходять із вузлів $U_{2,5}$ і $U_{2,6}$, тоді як лінії 1 є основними. Перемикання на резерв в IP-мережах виконується автоматично при виході з ладу основної лінії зв’язку чи основного вузла, а розподіл швидкостей в основній і резервній лініях прописується у вузлових елементах при їх конфігуруванні. Тоді для оцінювання коефіцієнта $k_{3,6(1)}$ необхідно застосовувати вираз

$$k_{3,6(1)} = \begin{cases} K_{3,6} \times k_{2,6(1)}, \text{ якщо } K_{2,6} \text{ дорівнює } 1, \\ K_{3,6} \times k_{2,6(2)}, \text{ якщо } K_{2,6} \text{ дорівнює } 0, \text{ а } K_{2,5} \text{ дорівнює } 1, \\ 0, \text{ якщо } K_{2,6} = K_{2,5} = 0. \end{cases}$$

У свою чергу, для оцінювання коефіцієнта $k_{2,5(2)}$ необхідно застосовувати вираз

$$k_{2,5(2)} = \begin{cases} R_{2,5(2)} \times K_{2,5} \times k_{1,5(1)}, \text{ якщо } K_{2,6} = 0, \\ 0, \text{ якщо } K_{2,6} \neq 0. \end{cases}$$

Природно, що використання частини канальних ресурсів на підтримку резервної лінії може призвести до падіння швидкості в основній лінії. Тоді для оцінювання коефіцієнта $k_{2,5(1)}$ необхідно застосовувати вираз

$$k_{2,5(1)} = \begin{cases} R_{2,5(1)} \times K_{2,5} \times k_{1,5(1)}, \text{ якщо } K_{2,6} = 0, \\ K_{2,5} \times k_{1,2(1)}, \text{ якщо } K_{2,5} \neq 0. \end{cases}$$

Таким чином, стан одного з вузлів, наприклад, $U_{2,6}$ може здійснювати вплив на коефіцієнти падіння швидкості у вихідних лініях інших вузлів, у нашому випадку $U_{2,5}$, які знаходяться на тому ж чи на вищих ієрархічних рівнях відносно вузла, про який іде мова, і не зв’язаних з ним фізично.

Відповідні цьому способу правила повинні належати такій множині:

IF tt.nod = 1 ∧ severity.nod = k ∧ type.nod = ‘…’ ∧ class.nod = ‘…’ ∧ type.child.nod = ‘…’ ∧ sign.reserve_parent = 1 ∧ output.nod.reserve_parent = m ∧ type.reserve_parent = ‘…’ ∧ class.reserve_parent = ‘…’ ∧ type.child.reserve_parent = ‘…’∧ output.child.reserve_parent = n *THEN* tt.child.nod = 1 ∨ severity.child.nod = f(k,m) — якщо є резервна лінія зв’язку і для неї виділений запас продуктивності;

IF tt.nod = 1 ∧ severity.nod = k ∧ type.nod = ‘…’ ∧ class.nod = ‘…’ ∧ type.child.nod ‘…’ ∧ sign.reserve_parent = 1 ∧ output.nod.reserve_parent = m ∧ type.reserve_parent = ‘…’ ∧ class. reserve_parent = ‘…’ ∧ type.child.reserve_parent = ‘…’ ∧ output.child.reserve_parent = n *THEN* tt.child.nod = 1 ∨ severity.child.nod = $f(k+m)$ ∨ tt.child.reserve_parent = 1 ∨ severity.child.reserve_parent = $f(n-m)$ — якщо є резервна лінія зв’язку, але для неї не виділений запас продуктивності і резерв підтримується за рахунок дочірніх вершин резервного батька.

Тут додаткові прості умови sign.reserve_parent = 1 і tt.child.reserve_parent = 1 відповідно означають, що існує зв’язок з резервним батьком і ознака наявності неполадки приписана дочірній вершині резервного батька. Змінні severity.nod, severity.child.nod, severity.child. reserve_parent набувають значення severity level для вузла, дочірньої вершини і дочірньої вершини резервного батька. Змінні output.nod.reserve_parent і output.child.reserve_parent набувають значення коефіцієнтів зниження швидкості передавання, які визначені вище.

Прив’язавши правила до структури мережі доступу, тобто визначивши задіяні правила для вершин графу і включивши відповідні коефіцієнти поширення впливу неполадок у спрощене представлення, другу фазу можна звести до перерахунку впливу неполадок на основі обходу графів спеціального вигляду. Дерево розповсюдження впливу становить, по суті, структуру даних для організації роботи алгоритму обходу графів спеціального вигляду. У процесі обходу зазначений алгоритм повинен також виконати розрахунок оцінок впливу неполадок на рівень обслуговування

клієнтів оператора на основі збережених у вершинах спрощеного представлення коефіцієнтів поширення.

Таким чином, постає потреба у алгоритмі, який повинен інтегрувати такі узагальнені процедури: структуризація повного представлення мережі і побудова каркасу дерева розповсюдження впливу; вибір продукцій для вузлів дерева розповсюдження впливу; визначення коефіцієнтів перерахунку для вузлів дерева розповсюдження впливу; безпосередній перерахунок впливу. До того ж, цей алгоритм повинен здійснювати розрахунки в умовах жорстких часових обмежень.

Швидкий алгоритм оцінювання впливу

Розглянемо ідеї, які можна покласти у реалізацію кожної із цих процедур, з метою мінімізації часу розрахунків, не випускаючи можливості здійснити це з використанням якомога менших обсягів пам’яті.

Призначення процедури структуризації повного представлення мережі і побудова каркасу дерева розповсюдження впливу полягає у формуванні на основі повного представлення мережі доступу спеціальної ярусно-паралельної форми. Таким чином, буде здійснено перехід до структурованого спрощеного представлення, у якому мережа доступу задається у термінах піддерев і шляхів. Воно є основою візуалізації мереж доступу з точки зору аналізу впливу неполадок.

Оскільки у алгоритмі передбачається використовувати порівневий обхід графа, у цій процедурі побудови графа кожному вузлу необхідно приписати певний рівень, виходячи із кількості проміжних елементів між ним і коренем. Такий граф зображений на рис. 2. Одне із обмежень такого підходу полягає в тому, що на кожному шляху від рівня 1 до вершин рівня користувачів повинні бути визначені вузли на кожному проміжному рівні. Таким чином, для зв’язків, які поєднують вершини несуміжних рівнів, наприклад, для відображення безпосереднього зв’язку між вузлами рівня 1 і рівня *N*, необхідно вводити фіктивні вузли на кожному проміжному рівні (на рис. 2 фіктивні вузли $U_{2,2}$ і $U_{2,5}$ зображені пунктирними лініями). Розрахунки для цих вузлів не виконуються і вони не впливають на показники мережі доступу.

Граф доцільно зберігати в оперативній пам’яті у вигляді списку ребер, кожний елемент якого містить інформацію про початкову і заключну вершини відповідного ребра, а також додаткову інформацію (тип, коефіцієнти і т. п.), необхідну для розрахунків. Для вузлів будемо використовувати окремий масив, кожний елемент якого містить інформацію про рівень, якому належить вузол, коефіцієнт неполадки і місце для тимчасової інформації, яка з’являється при виконанні розрахунків для вузла.

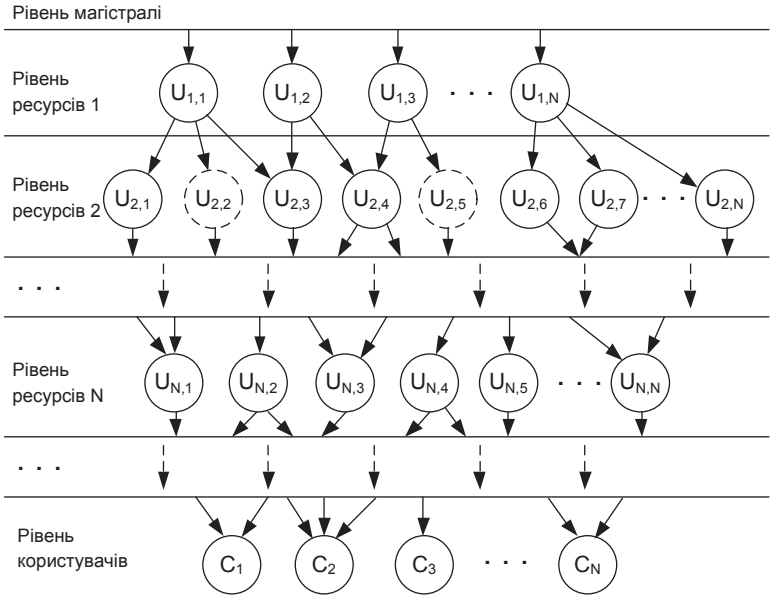


Рисунок 2 — Загальний вигляд графа розповсюдження впливу неполадок

В основу здійснення вибору правил покладемо механізм виведення у продукційних системах. Функціонування механізму виведення зводиться до послідовного виконання операцій перевірки умов правил, доки не буде визначено задіяні правила для кожної вершини графу. Правила розбиті на прості умови, кожній з яких приписаний унікальний ідентифікатор. Самі ж правила становлять собою колекції простих умов, поєднаних логічним оператором кон’юнкції. Таким чином, при зчитуванні чергового атрибута вузла значення атрибута порівнюється з простою умовою відповідного типу. При виконанні умови відповідний ідентифікатор додається у колекцію вузла. При цьому перевіряється, чи не

спрацювало якесь правило. У випадку спрацювання правила колекція об’єкта очищується, а на вихідних ребрах вузла проставляються коефіцієнти відповідно до зазначеного правила. Такий підхід дозволяє виконати необхідні розрахунки за один перегляд таблиці атрибутів, а також мінімізувати використання оперативної пам’яті. Для вирішення конфліктних ситуацій застосовується механізм пріоритету правил. Правила переглядаються у певному порядку, перше з них, умова якого виконується, застосовується.

Визначення коефіцієнтів перерахунку для вузлів дерева розповсюдження впливу здійснюється шляхом формування необхідних для визначення впливу коефіцієнтів поширення згідно з правими частинами задіяних для вершини правил.

Розрахунок впливу неполадок здійснюється порівнево. Таким чином, при виконанні розрахунків для поточного вузла ми впевнені, що виконані розрахунки для усіх вузлів попереднього рівня, які впливають на поточний вузол.

Отже, загальна ідея забезпечення ефективності перерахунку показників впливу неполадок елементів мережі доступу на інші елементи мережі і, насамкінець, на якість обслуговування клієнтів оператора зв’язку полягає у тому, щоб перейти від повторюваного послідовного перегляду робочої області і області правил до комбінованого застосування таких трьох прийомів:

1. індексація елементів робочої області: це дозволить підключити до усіх простих умов ті елементи робочої області, що їм задовольняють, і після цього переглядати тільки ці списки. Для кожного правила створюється лічильник кількості простих умов, початковим значенням якого є 0. При зміні робочої області до його значення додається або віднімається 1 в залежності від того, додається чи видаляється елемент робочої області, що задовольняє якійсь простій умові відповідного правила. Таким чином, порівняння лічильника з кількістю простих умов правил дозволяє швидко відібрати правила, умови яких виконуються. З кожною простою умовою зберігається список елементів робочої області, яким вона задовольняє. При зміні елементів робочої області списки коригуються. При інтерпретації переглядаються тільки списки і готується перелік конфліктних правил;

2. структурування системи правил: в пам’яті будується дерево описання простих умов, вершини якого відповідають ознакам простих умов, а листки — тим продукціям, прості умови яких виконуються на шляху від кореня. Система підтримує пам'ять про ті вузли, що відповідають даному стану робочої області, і модифікується при нових елементах робочої області;

3. прив’язування правил до вузлів спрощеного представлення мережі доступу: посилання на правила і лічильники зберігаються у кожному вузлі спрощеного представлення мережі доступу.

Коли дерево розповсюдження впливу побудоване, тим самим визначений загальний характер виконання розрахунків. Але існують альтернативи виконання розрахунків і після побудови дерева. По-перше, спільне використання списків, дерев і спрощеного представлення мережі доступу потенційно може забезпечити високу ефективність перерахунку показників впливу неполадок. По-друге, ще більше підкреслює сподівання на високу ефективність та обставина, що спрощене представлення можна розташувати у оперативній пам’яті. Отже, зникає необхідність у постійному звертанні до значно повільнішої зовнішньої пам’яті і, як наслідок, міцнішають сподівання на високу ефективність перерахунку. Тому доцільно навести загальний опис методу розрахунку впливу неполадок, який буде втілений у програмній реалізації.

Для швидкого аналізу впливу стану мережі на якість обслуговування користувачів був розроблений алгоритм, побудований на збереженні і опрацюванні спрощеного представлення у оперативній пам’яті. Він складається із двох підалгоритмів — формування і виконання спрощеного представлення.

Підалгоритм BS1 формування спрощеного представлення.

Вхід: Повне представлення (таблиці інвентарної бази) і система продукцій.

Вихід: Спрощене представлення.

Крок 1. Послідовний перегляд кортежів таблиці Objects і формування відповідних вершин і ребер спрощеного представлення.

Крок 2. Послідовний перегляд кортежів таблиці Reference і формування відповідних вершин і ребер спрощеного представлення.

Крок 3. Вибираємо черговий кортеж таблиці Parameters.

Крок 4. Знаходимо і позначаємо у спрощеному представленні у відповідній вершині просту умову, яка міститься у кортежі таблиці Parameters.

Крок 5. Знаходимо і позначаємо у таблиці продукцій у відповідних правилах просту умову, яка міститься у вибраному кортежі таблиці Parameters.

Крок 6. Перевіряємо умову: чи є продукція, усі прості умови якої виконані.

Якщо так, то реалізуємо правило у спрощеному представленні, формуючи в ньому відповідні поля у відповідних ребрах і вершинах, значення коефіцієнтів для розрахунку впливу неполадок, коли вони трапляться. Вилучаємо позначки простих умов у відповідній вершині спрощеного представлення. Вилучаємо позначки простих умов у відповідному правилі.

Якщо ні, то переходимо на крок 5.

Крок 7. Якщо усі кортежі таблиці Parameters розглянуті, то видаємо показники впливу неполадок на якість обслуговування клієнтів і кінець роботи, інакше — перехід на крок 3.

При зміні конфігурації мережі із інвентарної бази надходить відповідний сигнал (наприклад, від тригера в СУБД). У цьому випадку будуть виконані відповідні записи у списку ребер і масиві вершин, здійснений перерахунок підпорядкованих вузлів. Час перерахунку не перевищить часу розрахунку впливу неполадки. При додаванні нового користувача розрахунок здійснюється тільки для нього.

Підалгоритм BE1 виконання спрощеного представлення. Розглянемо алгоритм на прикладі, наведеному на рис.2. Нехай вузол $U_{1,3}$ вийшов з ладу на 50%. Робота алгоритму:

Крок 1. Розраховуємо вхідний коефіцієнт для вузла $U_{1,3}$. Оскільки вузол знаходиться на рівні 1, то вхідний коефіцієнт буде завжди дорівнювати 1.

Крок 2. Цьому вузлу встановлюється новий коефіцієнт неполадки, який дорівнює 0,5.

Крок 3. Перебираються вихідні ребра поточного вузла і розраховуються коефіцієнти пропускної здатності за формулою: $k = P * K$, де P — пропускна здатність на вході вузла, K — коефіцієнт неполадки вузла. Також, при обході ребра, на наступному (другому) рівні позначаються вузли, на які впливає ця неполадка. У нашому випадку це буде вузол $U_{2,4}$ і фіктивний вузол $U_{2,5}$.

Крок 4. Розрахунок першого рівня закінчений. Другий рівень робимо поточним.

Крок 5. У циклі перебираються вузли, позначені у п. 4. Оскільки вузол $U_{2,5}$ — фіктивний, то розрахунки в ньому не виконуються, а тільки позначаються вузли наступного рівня, на які може вплинути ця неполадка.

Крок 6. Розраховуємо неполадку у вузлі $U_{2,4}$. Перебираємо масив вхідних ребер і підсумовуємо коефіцієнти пропускної здатності (S). Вхідний коефіцієнт у вузлі розраховується за формулою: u = min(S; 1). Якщо для всіх постраждалих вузлів поточного рівня u = 1, то розрахунки припиняються, оскільки нижні рівні будутьзнаходитися у повністю функціональному стані і неполадка «згасне».

Крок 7. Переходимо до кроку 1, але на кроці 2 встановлюємо коефіцієнт неполадки уже поточного вузла. У нашому випадку для $U_{2,5}$ він дорівнює 1 (вузол повністю функціональний).

Крок 8. Дії повторюємо до тих пір, доки поточним не стане рівень користувачів.

Крок 9. Для кожного користувача перебираємо масив вхідних ребер і виконуємо для нього дії кроку 6.

Крок 10. Формуємо масив постраждалих користувачів з коефіцієнтом неполадки.

Резервні вузли враховуємо за допомогою правил, наведених вище для способу г) з’єднання вузлів мережі доступу. Якщо у постраждалого вузла є вхідні резервні ребра, то вплив неполадки може бути зменшений за рахунок потужності резервного ребра. При цьому, у деяких випадках, вплив неполадки може поширитися на дочірні вершини резервного батька.

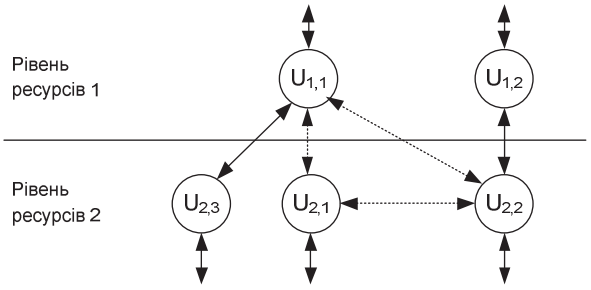


Рисунок 3 — Приклад петлі

Оскільки напрям ребра вказує на зростання рівня, а трафік може передаватися і в зворотному ієрархії напрямі, то може виникнути проблема розрахунку петель. Приклад петлі наведений на рис. 3.

Тут петля позначена пунктиром. Зв’язок на одному рівні (між $U_{2,1}$ і $U_{2,2}$) буде реалізований за допомогою двох ребер. Зв’язок вигляду $U_{1,1} <> U_{2,2}$ реалізувати практично неможливо. Тому необхідно замінити його на еквівалентний. Еквівалент фрагменту мережі з петлею наведений на рис. 4.

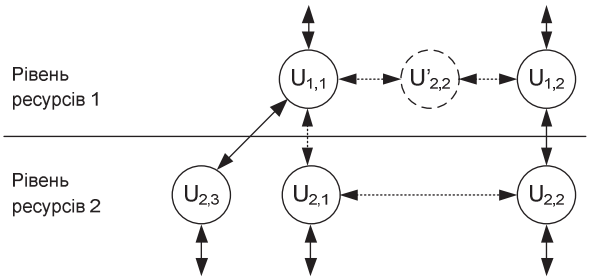


Рисунок 4 — Еквівалентне перетворення петлі

Таким чином, у нашій моделі будуть існувати зв’язки тільки вниз або на одному рівні, що дозволить успішно застосовувати описаний вище алгоритм роботи. Для еквівалентної заміни для перерахунків знадобиться коефіцієнт R для ребра між вузлами $U_{1,1}$ і $U'_{2,2}$, який буде вказувати на частку загальної вихідної потужності вузла $U_{2,2}$, що буде переключена на це ребро. Також нам знадобиться спеціальний еквівалентний вузол $U'_{2,2}$, який буде копією вузла $U_{2,2}$, але лише з одним виходом. Варто зазначити, що однорівневі зв’язки у графі можуть бути тільки резервними.

Результати експериментальних досліджень

Експерименти проводилися над різними варіантами структур мереж доступу, сервісів, кількостей клієнтів. При цьому характеристики обладнання і технології змінювалися. Алгоритм був оцінений стосовно часу і пам’яті, необхідних для одержання розв’язків. Експеримент проводився для двох підалгоритмів, реалізованих у вигляді окремих модулів однієї програми, окремо.

При тестуванні першого модуля будувалися дерева з різною кількістю елементів. Кожний елемент описувався 2 атрибутами — ім’ям і типом. Використовувалась обмежена кількість правил.

При тестуванні другого модуля експеримент проводився на випадково згенерованих деревах різних розмірів. Під час експерименту випадковим чином емулювалися виходи з ладу верхнього вузла і, таким чином, перераховувалося усе дерево.

Результати експерименту з під алгоритмом *BS1* (першим модулем програми) наведені у вигляді графіку, зображеного на рис. 5, і у таблиці 1.

Таблиця 1 — Результати дослідження першого модуля

Кількість вузлів	Час генерації дерева ($t_г$), мс	Час читання атрибутів ($t_а$), мс	Час обробки правил ($t_п$), мс	Загальний час ($t_{заг}$), мс
13620	876	613	12	1501
65025	3873	2737	58	6668
270645	10853	11616	106	22575
476268	32905	24432	3236	60573

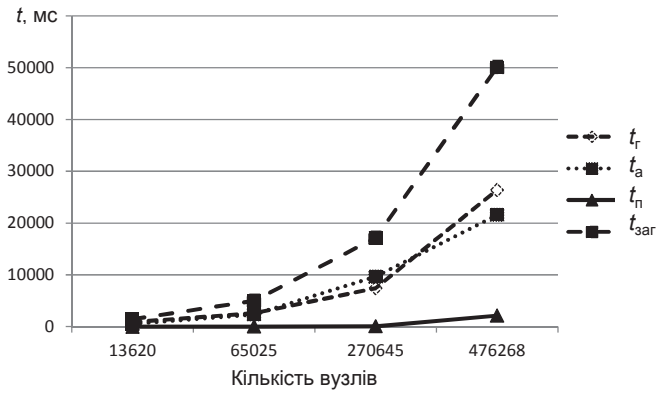


Рисунок 5 — Графік залежності часу побудови спрощеного представлення від кількості елементів

Результати експерименту з підалгоритмом *BE1* (другим модулем програми) наведені у вигляді графіків, зображених на рисунках 6, 7, і у таблиці 2. При цьому максимальна кількість ребер не перевищувала 1,5 млн.

Таблиця 2 — Результати дослідження другого модуля

Кількість вузлів	1007	9977	270645	476268
Об’єм пам’яті, що потребується, K	5332	10240	99288	200384
Час обчислення ($t_{об}$), ms	5	41	155	633

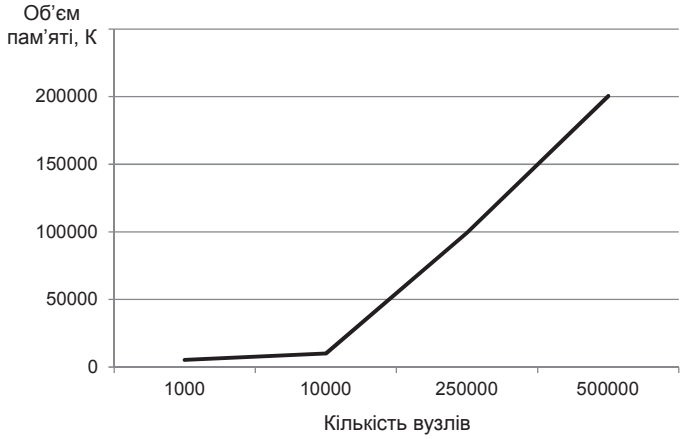


Рисунок 6 — Графік залежності об’єму оперативної пам’яті, що потребується, від кількості елементів мережі

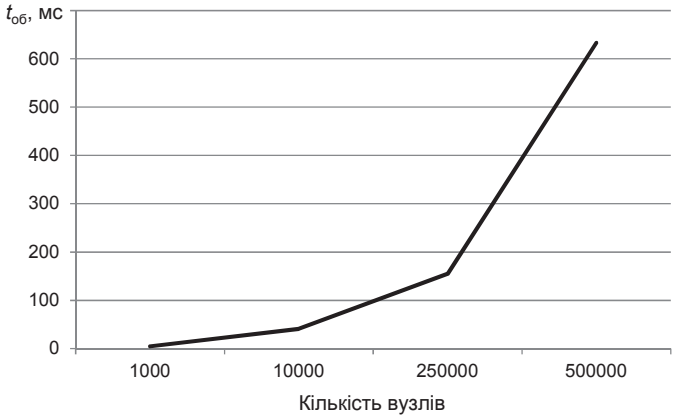


Рисунок 7 — Графік залежності часу розрахунку показників обслуговування від кількості елементів мережі

Експериментальні дані наведені без перерахунку петель.

Результати експериментів продемонстрували працездатність запропонованих алгоритмів. При цьому необхідні для розрахунків пам’ять і час не перевищують достатньо жорсткі вимоги операторів зв’язку.

Висновки

У статті запропоновано підхід до швидкого розв’язання проблеми визначення впливу відмов елементів мережі доступу на якість обслуговування клієнтів великих операторів інформаційно-телекомунікаційних послуг на основі комбінування продукційного і алгоритмічного підходів. Реалізація підходу базується на графо-орієнтованій моделі представлення мережі і сервісів, системі продукцій і відповідному механізмі виведення, а також алгоритмі організації процесу визначення впливу відмов елементів мережі доступу на якість обслуговування клієнтів на основі схем обходу графів. Реалізація алгоритму базується на ідеї збереження і опрацювання спрощеного представлення мережі у оперативній пам’яті.

Проведені експериментальні дослідження розробленого варіанту реалізації запропонованих моделі і алгоритму. Вони продемонстрували працездатність запропонованого варіанту реалізації алгоритму. Витрати часу і пам’яті, необхідні для одержання потрібних операторам результатів, не

перевищують реальних обмежень (нагадаємо, що час розрахунків у деяких операторів мереж доступу становить не більше 15 хвилин на 1,5 млн. клієнтів).

Подальший розвиток досліджень пов’язаний з розробленням моделей і алгоритмів, які враховують інші особливості мереж доступу, технології, що дозволить розширити множину ситуацій, в яких може застосовуватися розроблений підхід.

Список літератури

1. Павлов А.А. Информационные технологии и алгоритмизация в управлении / А.А. Павлов, С.Ф. Теленик. — К.: Техника, 2002. — 344 с.
2. Теленик С.Ф. Определение распространения влияния неисправностей в сети доступа на качество предоставляемых сервисов / С.Ф. Теленик, А.И. Ролик, М.М. Букасов, М.В. Ясочка // Вісник НТУУ «КПІ». Інформатика, управління та обчислювальна техніка. — К.: «ВЕК+». — 2009. — №50. — С. 164—173.
3. Biazetti A., Goldszmidt G. Integrating business performance management with IT management through impact analysis and provisioning // NOMS 2008 — 11th IEEE/IFIP Network Operations and Management Symposium, vol. 11, no. 1, April 2008, pp. 504—518.
4. Bartolini C., Stefanelli C., Tortonesi M. Business-impact analysis and simulation of critical incidents in IT service management // IM 2009 — 11th IFIP/IEEE International Symposium on Integrated Network Management, vol. 11, no. 1, June 2009, pp. 9—16.
5. Hanemann A., Schmitz D., Sailer M. Towards a framework for failure impact analysis and recovery with respect to service level agreements // IM 2005 — IFIP/IEEE International Symposium on Integrated Network Management, no. 1, May 2005, pp. 1189—1192.
6. Shankar S., Satyararayaman O. An Automated System for Analyzing Impact of faults in IP Telephony Networks// Network Operations and Management Symposium, 2006. NOMS 2006. 10th IEEE/IFIP, April 2006, p. 1—4.
7. Setzer T., Bhattacharya K. Ludwig H. Decision support for service transition management enforce change scheduling by performing change risk and business impact analysis // NOMS 2008 — 11th IEEE/IFIP Network Operations and Management Symposium, vol. 11, no. 1, April 2008, pp. 200—207.
8. Теленик С.Ф. Управління ресурсами центрів оброблення даних / С.Ф. Теленик, О.І. Ролік, М.М. Букасов, К.О. Крижова // Вісник ЛНУ імені Івана Франка. — 2009. — №11. — С. 103—119.
9. Теленик С.Ф. Моделі управління розподілом обмежених ресурсів в інформаційно-телекомунікаційній мережі / С.Ф. Теленик, О.І. Ролік, М.М. Букасов // Вісник НТУУ «КПІ». Інформатика, управління та обчислювальна техніка. — К.: Екотех. — 2006. — №44. — С. 243—246.
10. Теленик С.Ф. Система управління інформаційно-телекомунікаційною системою корпоративної АСУ / С.Ф. Теленик, О.І. Ролік, М.М. Букасов, Р.Л. Соколовський // Вісник НТУУ «КПІ». Інформатика, управління та обчислювальна техніка. — К.: «ВЕК+». — 2006. — №45. — С. 112—126.

Відомості про авторів

Теленик Сергій Федорович — д.т.н., професор, завідувач кафедри автоматики та управління в технічних системах, Національний технічний університет України «Київський політехнічний інститут», пр. Перемоги 37, м. Київ, 03056, тел. (044) 406-86-10, telenik@auts.ntu-kpi.kiev.ua.

Ролік Олександр Іванович — к.т.н., доцент кафедри автоматики та управління в технічних системах, Національний технічний університет України «Київський політехнічний інститут», пр. Перемоги 37, м. Київ, 03056, тел. (044) 406-86-10, rolick@auts.ntu-kpi.kiev.ua

Ясочка Максим Володимирович — аспірант кафедри автоматики та управління в технічних системах, Національний технічний університет України «Київський політехнічний інститут».

Малюгін Дмитро Володимирович — студент, Національний технічний університет України «Київський політехнічний інститут».

УДК 004.9+656

ТЕХНОЛОГІЯ ПОБУДОВИ ІНФОРМАЦІЙНОЇ МОДЕЛІ ТРАНСПОРТНОЇ МЕРЕЖІ МІСТА НА ОСНОВІ ГЕОІНФОРМАЦІЙНИХ МОДЕЛЕЙ ЇЇ ЕЛЕМЕНТІВ

В.Б. Мокін, В.Г. Сторчак

Анотація: Розглянута проблема управління транспортними потоками міста з використанням геоінформаційних технологій. Розроблено нову інформаційну технологію автоматизованої побудови геоінформаційної моделі транспортної мережі міста на основі бази знань моделей елементів цієї мережі з урахуванням їх вагомості для задачі заданого типу. Запропоновано ряд прийомів щодо застосування цієї технології на практиці. Проведено її практичну апробацію у м. Вінниця.

Аннотация: Рассмотрена проблема управления транспортными потоками города с использованием геоинформационных технологий. Разработана новая информационная технология автоматизированной построения геоинформационной модели транспортной сети города на основе базы знаний моделей элементов этой сети с учетом их значимости для задачи заданного типа. Предложен ряд приемов по применению этой технологии на практике. Проведено ее практическую апробацию в г. Винница.

Abstract: There had been considered the issue of controlling over the city transport flows using GIS technologies. There had been developed the new information technology for automated building the GIS model of city transport network on basic a knowledge database with models elements of the above network with the consideration of their relevancy for the task of the set type. There had been suggested a number of methods for practical approbation of this technology. Its practical approbation had been made in the city of Vinnytsya.

Ключові слова: геоінформаційні системи, інформаційна модель, управління транспортними потоками, транспортна мережа.

Постановка задачі та вихідні передумови

Проблема завантаженості доріг великою кількістю транспорту робить актуальним різні методи та технології оптимального управління транспортними потоками. Існує багато моделей, які описують динаміку транспортних потоків, які поділяють на детерміновані та стохастичні, наприклад: моделі Гріншилдса и Грінберга, Лайтхілла-Уізема, моделі слідування за лідером, клітинні автомати, моделі стрибаючих частинок, та інші [1]. Всі вони тією чи іншою мірою відображають характер та закономірності зміни параметрів потоків у транспортній мережі (ТМ) міста. Але при моделюванні цих параметрів, як правило, або детально вивчаються та описуються усі процеси на основі теорії масового обслуговування, або обробляються дані натурних спостережень параметрів транспортних потоків, у т.ч. з камер відеоспостереження [1, 2]. Однак, ці підходи не дозволяють гнучко та швидко враховувати нові фактори та взаємовплив одних факторів на інші, наприклад, необхідність врахування нових об’єктів транспортної інфраструктури або погіршення умов пересування (стан доріг, «пробки» тощо) транспортних засобів вимагає значного часу на ідентифікацію та оптимізацію раніше побудованої моделі цієї мережі. Прискорити розв’язання цієї задачі можна, попередньо побудувавши інформаційні моделі елементів ТМ та процесів, які впливають на зміну параметрів транспортних потоків у місті. Ці моделі повинні враховувати як статичні та динамічні параметри елементів, так і співвідношення, що описують зміну цих параметрів під впливом інших елементів ТМ. Також, слід розробити алгоритми та засоби автоматизованої обробки даних про елементи ТМ з метою збільшення оперативності та комплексності щодо прийняття рішень з управління дорожнім рухом (ДР) міста. Світовий досвід довів, що оптимальним є розробка інформаційних моделей таких елементів у вигляді геоінформаційних моделей (ГІС-моделей) [3-5]. Отже, є актуальною задача створення інформаційної технології автоматизованої побудови геоінформаційної моделі транспортної мережі міста на основі геоінформаційних моделей її елементів. Технологія повинна містити моделі, методи, алгоритми, прийоми та програмне забезпечення для побудови та ідентифікації типових елементів ТМ за реальними даними.

Методологія розв’язання задачі

Кожна ГІС-модель містить інформаційну (числові, текстові та інші атрибути чи параметри) та просторову (координати, геометрія) складову. Для коректного відображення просторових даних сучасна ГІС-модель враховує поняття топології, тобто просторових відношень між координатами об’єктів [4-6]. Наприклад, координати дорожньої розмітки повинні збігатись з координатами доріг, координати табличок до дорожніх знаків повинні збігатись з координатами знаків, виїзди із зон паркування повинні мати спільні точки з вулицями та ін. Пропонуємо інформаційні моделі елементів транспортної мережі зберігати у спеціальній базі знань, яка містить параметри об’єктів та правила, операції і відношення між цими даними та об’єктами. Операції, як правило, виконуються над числовими параметрами об’єктів, а відношенням («перетинаються», «не перетинаються», «є підоб’єктом» та ін.) відповідають певні алгоритми виконання операцій над просторовими даними цих об’єктів. Також у базі знань пропонуємо зберігати інформацію про атрибутивні відношення, які відображають вплив параметрів одних елементів ГІС-моделей на параметри інших. Структура такої узагальненої ГІС-моделі наведена на рис. 1.

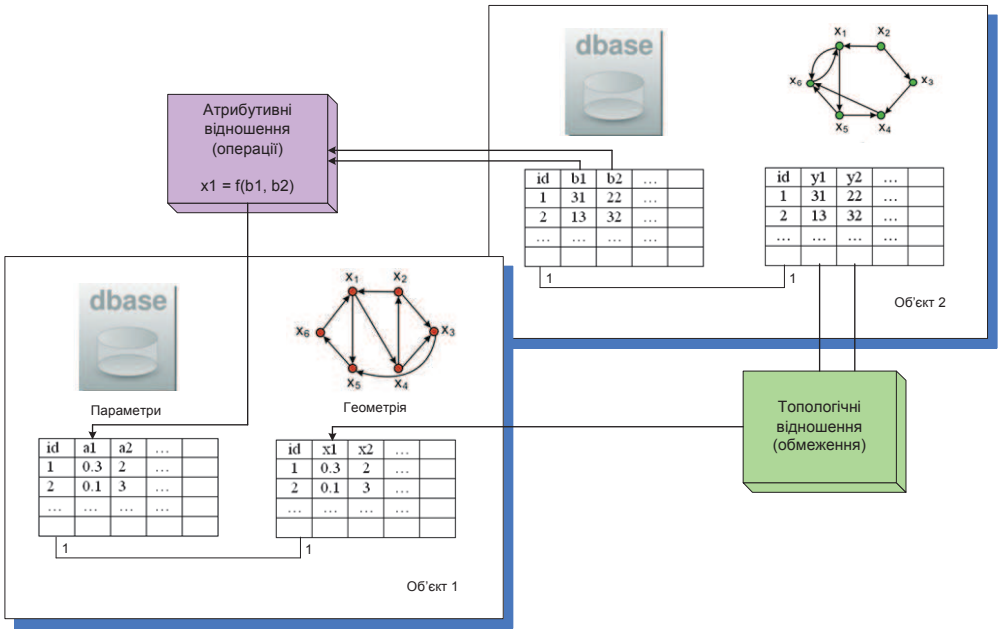


Рисунок 1 – Структура узагальненої геоінформаційної моделі елемента транспортної мережі

Наприклад, для ділянки вулиці з середньою швидкістю v установлення знаку «Обмеження швидкості» з параметром $k = 40$, який забороняє рух зі швидкістю вище 40 км/год, означає, що нове значення середньої швидкості v_2 для цієї ділянки вулиці визначається за співвідношенням:

$$v_2 = \begin{vmatrix} k, v > k, \\ v, v \le k. \end{vmatrix}$$

Більш складні співвідношення характеризують зміну пропускної здатності вулиці внаслідок дії засобів регулювання руху (світлофори з багатьма фазами переключення, нерегульовані перехрестя та ін.), появи ділянок дороги з підвищеною «ямковістю», виїздів та заїздів транспорту на зони паркування тощо.

Розроблена інформаційна технологія для реальної ТМ реалізується у такі етапи:

Етап 1. Ідентифікація ГІС-моделей основних елементів ТМ.

1. Визначення усіх факторів та основних елементів ТМ, які суттєво впливають на динаміку транспортних потоків міста і можуть бути важливими під час оптимізації ДР міста (пропускної здатності, організації «зелених хвиль», підвищення зручності проїзду маршрутного транспорту, паркування тощо).

2. Ідентифікація просторових моделей основних елементів ТМ: векторизація (нанесення на карту) усіх основних елементів ТМ як окремих шарів ГІС та редагування і збереження їх топологічних відношень, формування геометричних мереж тощо.

3. Формалізація та ідентифікація інформаційних моделей основних елементів ТМ: для кожного елемента здійснюється ідентифікація закономірностей щодо значень та динаміки зміни його параметрів шляхом виконання натурних спостережень, обробки даних відеоспостереження, опитування учасників ДР, вивчення технічної документації на засоби регулювання ДР тощо.

4. Визначення та збереження у базі знань атрибутивних відношень між параметрами інформаційних моделей основних елементів ТМ: на основі аналізу Правил дорожнього руху, експлуатаційно-будівельних норм об’єктів дорожньої мережі, закономірностей, геометричних параметрів та обмежень, які має вулична мережа міста, технічної документації на засоби регулювання ДР та інші елементи ТМ тощо визначаються співвідношення та алгоритми зміни параметрів одних елементів ТМ через параметри інших та обмеження параметрів одних елементів ТМ параметрами інших, які мають місце. Для реальних елементів, параметри яких не встановлюються нормативно-технічними документами та стандартами (перешкоди на дорогах («ямковість», ожеледиця та ін.), інтенсивність транспортного потоку на вулицях міста у різний час тощо відношення ідентифікуються шляхом виконання натурних спостережень, обробки даних відеоспостереження, опитування учасників ДР, аналізу протоколів ДТП тощо.

Етап 2. Ідентифікація ГІС-моделі ТМ та розв’язання прикладних задач, пов’язаних з оптимізацією ДР.

1. Визначення основних елементів, які впливають на ті чи інші параметри ТМ та які мають високу вагу під час розв’язання різних прикладних задач, пов’язаних з оптимізацією ДР. Наприклад, різні елементи слід враховувати під час розрахунку «зеленої хвилі» по різних, у т.ч. паралельних,

вулицях міста. А це може бути важливим під час ремонту вулиць чи різноманітних аварійних ситуацій. Складання довідника можливих прикладних задач та присвоєння значень ваги елементам ТМ, які відображають значущість впливу їх параметрів на успішний розв’язок цих задач. Відповідно, найменшу вагу матимуть ті елементи, параметри яких можна не враховувати під час розв’язання відповідної задачі.

2. Визначення прикладної задачі, яку слід розв’язувати в першу чергу, та встановлення її номеру P у довіднику можливих прикладних задач.

3. Визначення граничного значення ваги W, за якою відбирати найбільш впливові елементи ТМ (фактори).

4. Ідентифікація ГІС-моделі ТМ: підключення шарів ГІС з елементами, які мають вагу для задачі P вище заданого граничного значення W, та відключення інших шарів ГІС. Перевірка у базі знань виконання атрибутивних та просторових відношень та обмежень (операцій з параметрами та топології) між підключеними шарами ГІС та відповідний перерахунок значень параметрів, за необхідності.

5. Розв’язання поставленої прикладної задачі, пов’язаної з оптимізацією ДР. Візуалізація результатів розрахунку.

6. Якщо результат розрахунку збігається з експериментальними даними, тоді перейти до п. 7, інакше – враховувати більше факторів (елементів ТМ) шляхом зміни (збільшення або зниження) граничного значення ваги W з переходом до п. 4.

7. Якщо необхідно розв’язати нову задачу, тоді знайти її номер P у довіднику можливих прикладних задач та перейти на п. 3, інакше – на наступний пункт.

8. Завершити роботу.

Практична реалізація технології

Під час практичної реалізації запропонованої інформаційної технології пропонується враховувати такі особливості ТМ та прийоми роботи з ГІС:

1. Відомо, що типовими елементами ТМ та об’єктами транспортної інфраструктури міста, з точки зору регулювання дорожнього руху, є такі [1, 2]: вулиці, перехрестя, дорожні знаки, світлофори, дорожня розмітка, зони паркування, точки тяжіння (об’єкт або територія масового збирання людей: навчальні заклади, магазини, території масового відпочинку, великі громадські установи тощо), тимчасові об’єкти транспортної мережі (ділянки дороги, що ремонтуються, об’їзди тощо). Необхідно під час побудови їх ГІС-моделей та формалізації у базі знань використовувати уніфіковані умовні позначення та назви параметрів (атрибутів), щоб легше було потім їх обробляти.

а) б)

Рисунок 2 – Типове авторське програмне забезпечення на GISToolKit Free ГІС «Панорама» для зручної роботи з шарами ГІС ТМ та обробки їх даних і розв’язання прикладних задач: а) управління шарами ГІС;
б) аналіз даних в ГІС та побудова тематичних карт

2. Варто побудувати та зберігати у базі знань множину типових ГІС-моделей, які адаптуються під конкретні елементи ТМ, наприклад: «Вулиця із двобічним рухом», «Т-подібне перехрестя», «Забороняючий знак», «Трифазний світлофор», «Розмітка – суцільна лінія», «Точка тяжіння - Школа», «Ямковість дороги» тощо.

3. Варто розробити типове програмне забезпечення для зручної роботи з шарами ГІС ТМ та обробки їх даних і розв’язання прикладних задач, наприклад з використанням інструментарію GISToolKit Free ГІС «Панорама» [7] (рис. 2).

Дана технологія була успішно випробувана на вулицях міста Вінниці в рамках виконання робіт щодо створення пілотної версії Автоматизованої системи керування дорожнім рухом у межах Договору про науково-технічне співробітництво між Вінницькою міськрадою та Вінницьким національним технічним університетом № 28/2 від 1 березня 2010 року. Роботи виконувались на замовлення та за участі Державного підприємства «Спеціалізована монтажно-експлуатаційна дільниця ВДАІ УМВС України у Вінницькій області».

Висновки

Таким чином, розроблено нову інформаційну технологію побудови геоінформаційної моделі параметрів транспортної мережі, яка дозволяє більш гнучко та швидко враховувати моделі різних елементів мережі та взаємовплив параметрів і топології одних елементів на інші шляхом автоматизованої обробки моделей просторової та атрибутивної інформації геоінформаційних моделей елементів мережі з різних шарів ГІС, формалізованих та збережених у вигляді бази знань. Дістала подальший розвиток формалізація геоінформаційних моделей на прикладі елементів транспортної мережі шляхом врахування, поряд із просторовими відношеннями (топологією), атрибутивних відношень (операцій) над параметрами різних об’єктів ГІС. Запропоновано новий метод ідентифікації геоінформаційної моделі транспортної мережі (ТМ), який відрізняється від існуючих підходом щодо відбору елементів мережі, котрі найбільше впливають на параметри та топологію ГІС-моделі мережі, шляхом порівняння значень ваги цих елементів із певним граничним значенням. Успішна апробація розробленої інформаційної технології на вулицях міста Вінниці підтвердила доцільність її застосування.

Список літератури

1. Математическое моделирование динамики транспортных потоков мегаполиса – [Електронний ресурс] / Семенов В.В. — Режим доступу :
2. http://www.uran.donetsk.ua/~masters/2005/kita/shapovalova/library/semenov.pdf
3. Гаврилов А.А. Моделирование дорожного движения. / Гаврилов А.А. – Москва : Транспорт, 1980. – 189 с.
4. Комп’ютеризовані регіональні системи державного моніторингу поверхневих вод: моделі, алгоритми, програми. Монографія / Під ред. В. Б. Мокіна. — Вінниця: Вид-во ВНТУ “УНІВЕРСУМ-Вінниця”, 2005. — 315 с.
5. Mitchell A. – Environmental Systems Research Institute, Inc. The ESRI Guide to GIS Analysis. V.1: Geographic Patterns and Relationships: Redlands, USA, 1999. – 186 p.
6. Zeiler M. Modeling our World. – ESRI: Redlands, USA, 1999. – 202 p.
7. ArcGIS 9. Geodatabase Workbook. – ESRI: Redlands, USA, 2004.– 258 p.
8. Продукты КБ Панорама (РФ). Дорожное хозяйство. – [Електронний ресурс]. – Режим доступу : http://gisinfo.ru/projects/projects.htm#roads

Відомості про авторів

Мокін Віталій Борисович – доктор технічних наук, професор, завідувач кафедри моделювання та моніторингу складних систем Вінницького національного технічного університету.

Сторчак Володимир Григорович – аспірант кафедри моделювання та моніторингу складних систем Вінницького національного технічного університету, магістр електромеханіки.

УДК 004.8

ДОСЛІДЖЕННЯ МОЖЛИВОСТЕЙ ГЕНЕТИЧНОГО АЛГОРИТМУ В ЗАДАЧІ КЛАСТЕРИЗАЦІЇ КОРИСТУВАЧІВ МЕРЕЖІ INTERNET

С.М. Захарченко, Н.Р. Кондратенко, О.О. Манаєва

Вступ

На сьогоднішній день мережа Інтернет є універсальним глобальним інформаційним середовищем, а отже, невід’ємною частиною життя суспільства. Перед будь-яким Інтернет-провайдером постає задача тарифікації власних послуг відповідно до деяких критеріїв. Розв’язання цієї та інших задач покладається, як правило, на білінгову систему – програмний комплекс, що здійснює облік обсягу спожитих абонентом послуг, розрахунок та списання коштів відповідно до тарифів компанії [1].

Таким чином, однією з важливих складових будь-якої білінгової системи є розробка тарифних планів. Вона вимагає ретельного дослідження та формування компактних груп абонентів із подібними потребами.

Виділенням серед множини об’єктів деякої кількості однорідних підмножин, таких, щоб об’єкти всередині підмножин були в певному сенсі подібними, а об’єкти з різних підмножин – відмінними, займається задача кластеризації. Елементами підмножини можуть бути довільні об’єкти, які можна задати векторами характеристик. Самі ж групи прийнято називати кластерами.

Кластеризація має велику кількість практичних застосувань як в інформатиці, так і в інших галузях. Одним із найважливіших напрямків досліджень є її використання для аналізу даних та виділення прихованих закономірностей. До цього класу проблем належить і задача розподілу користувачів мережі Інтернет за їхніми вимогами до якості та характеру послуг, що надає провайдер.

Відомо, що для її розв’язання не існує універсальних методів, що дозволяють швидко знайти абсолютно точні рішення [2].

За способом розбиття на кластери розрізняють два типи алгоритмів: ієрархічні та неієрархічні [3]. Класичні ієрархічні алгоритми працюють тільки з категорійний атрибутами, коли будується повне дерево вкладених кластерів. Тут поширені агломеративні методи побудови ієрархій кластерів - в них проводиться послідовне об'єднання вихідних об'єктів і відповідне зменшення числа кластерів. Ієрархічні алгоритми забезпечують порівняно високу якість кластеризації і не вимагають попереднього задання кількості кластерів.

Неієрархічні алгоритми ґрунтуються на оптимізації деякої цільової функції, що визначає оптимальне в певному сенсі розбиття множини об'єктів на кластери. У цій групі популярні алгоритми родини k-середніх (k-means), які в якості цільової функції використовують суму квадратів зважених відхилень координат об'єктів від центрів шуканих кластерів. Кластери шукаються сферичної або еліпсоїдної форми. У канонічній реалізації мінімізація функції здійснюється на основі методу множників Лагранжа і дозволяє знайти тільки найближчий локальний мінімум.

Серед неієрархічних алгоритмів, не заснованих на відстані, слід виділити EM-алгоритм (Expectation-Maximization). У ньому замість центрів кластерів передбачається наявність функції щільності ймовірності для кожного кластеру з відповідним значенням математичного сподівання і дисперсією [3, 4].

Основним недоліком усіх цих алгоритмів є настільки великий обсяг обчислень, що, не зважаючи на високу швидкодію сучасних обчислювальних машин, розв’язання задачі залишається складним.

Тому для розв’язання цієї задачі доцільно використати інтелектуальні технології [5], зокрема генетичні алгоритми.

Постановка задачі

Поставимо задачу розбиття деякої множини користувачів провайдера інтернет-послуг на групи відповідно до певного набору характеристик: швидкості доступу, обсягу спожитого вхідного та вихідного трафіку. Таке розбиття дасть змогу оцінити потреби абонентів та відповідно до цього організувати максимально гнучку та пристосовану до ситуації на ринку систему тарифів провайдера.

Для розв’язання поставленої задачі запропонуємо генетичний алгоритм, що здійснює кластеризацію користувачів мережі Інтернет відповідно до виділених вище ознак та дослідимо його ефективність і швидкодію в порівнянні з алгоритмом граничного перебору на основі динамічного програмування.

Математична модель та методика дослідження

Нехай маємо набір абонентів $I=\{I_1, I_2, \ldots, I_n\}$ певного провайдера телекомунікаційних послуг. На основі множини показників, що описують цих абонентів, згрупуємо їх у підмножини (кластери)

таким чином, щоб абоненти, що входять до одного класу, були більш однорідними за характером споживання ними трафіку, схожими між собою більше, ніж із абонентами, що належать до інших класів.

Кожному з n абонентів поставимо у відповідність множину ознак (вимірів) $X_i=\{x_1, x_2, \ldots, x_p\}$, які будуть нести інформацію про характер споживання ним Інтернет-трафіку за деякий період. Таким чином, для множини I маємо множину векторів ознак $X=\{X_1, X_2, \ldots, X_n\}$, які повністю її характеризують. У такому записі окремих користувачів з множини I зручно представити у вигляді n точок у p-вимірному просторі ознак. Під подібністю/відмінністю будемо розуміти відносну геометричну близькість між такими точками у багатомірному просторі.

Нехай m – ціле число, менше за n. Задача кластеризації полягає в тому, щоб на основі даних, що містяться у множині I, розбити множину абонентів I на m кластерів.

Розв’язком задачі кластерного аналізу є розбиття, що задовольняє певному критерію оптимальності. Цей критерій може являти собою деякий функціонал, що відображає рівень бажаності різних розбиттів і групувань. Цей функціонал часто називають цільовою функцією [2]. За цільову функцію приймемо традиційну для задач кластеризації суму квадратів Евклідових відстаней між об’єктами в межах одного кластера:

$$d_{Eij} = \left(\sum_{l=1}^{m} \left(x_i^l - x_j^l \right)^2 \right)^{\frac{1}{2}}$$

В основі генетичних алгоритмів лежить теорія еволюції. Згідно з нею, кожен біологічний вид розвивається, поступово змінюється задля того, щоб якнайкращим чином пристосуватися до навколишнього середовища. Класичний генетичний алгоритм являє собою ітераційний процес, на кожній ітерації якого популяція послідовно піддається операціям відбору, схрещування та мутації (рисунок 1). Зупинивши ітераційний процес у певний момент та вибравши кращу особину з популяції, можна отримати достатньо прийнятний розв’язок задачі [5].

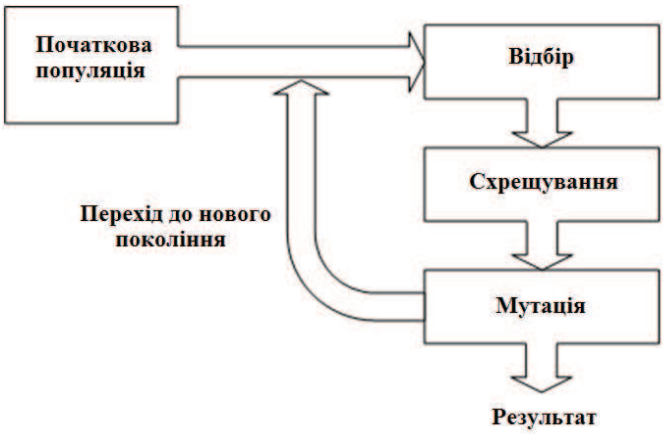


Рисунок 1 - Робота класичного генетичного алгоритму

Для формалізованого опису особини запропонуємо неоднорідну хромосому (рисунок 2):

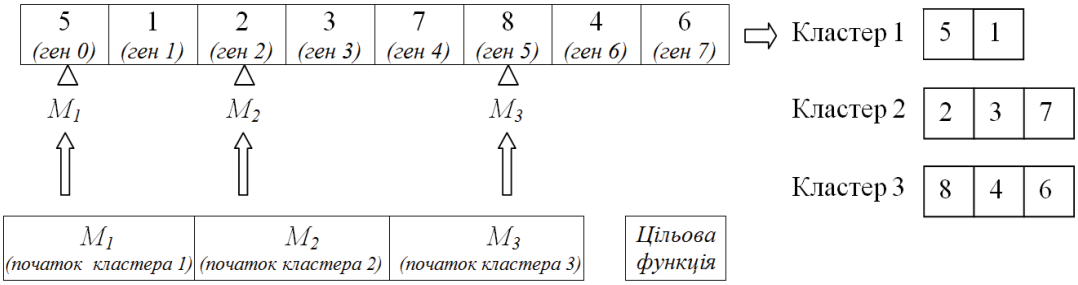


Рисунок 2 – Неоднорідна хромосома

Така хромосома містить дві складові, які попри зовнішню подібність мають принципово різний зміст. Перша складова являє собою лінійний масив розмірності n, що містить гени - номери об’єктів - у певній послідовності. Для виділення в межах хромосоми окремих кластерів введемо до складу нашої хромосоми другу складову – набір маркерів розбиття $M_1, \ldots M_m$. Значення маркеру $M_i = k$ сигналізує про

те, що з k-того гену основної складової починається новий кластер. Окрім того, кожній хромосомі з простору можливих розв’язків ставиться у відповідність значення цільової функції. В процесі отримання розв’язку задачі такі набори даних (проміжні розв’язки) будемо зберігати в робочій матриці.

Закономірно, що для хромосоми такого вигляду операції схрещування та мутації мають певні особливості. Їх буде розглянуто нижче при покроковому описі алгоритму.

Щоб розв’язати задачу кластеризації в описаній вище постановці, необхідно розробити послідовність операцій, які моделюють еволюційні процеси на основі математичних аналогів генетичної спадковості, мінливості та природного відбору.

1) Генерування початкової популяції. Початкова популяція містить n хромосом, згенерованих випадковим чином. При цьому як послідовність генів, так і характер (форма) розбиття задається випадково. Для кожного з n отриманих розбиттів розраховується цільова функція, і початкова популяція записується в робочу матрицю, займаючи її елементи з індексами від 0 до n-1.

2) Створення нащадків:

а) двоточкове схрещування (рисунок 3): хромосома випадковим чином розбивається на три частини, як показано на рисунку. Міняючи місцями крайні ділянки хромосоми, отримуємо новий розв’язок:

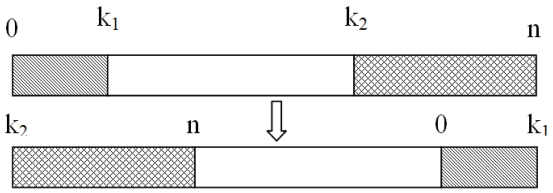


Рисунок 3 – Двоточкове схрещування

Хромосоми, отримані за механізмом двоточкового схрещування, записуються в робочу матрицю з індексами від n до 2n-1.

б) мутація 1: два випадково взяті гени однієї хромосоми міняються місцями (рисунок 4).

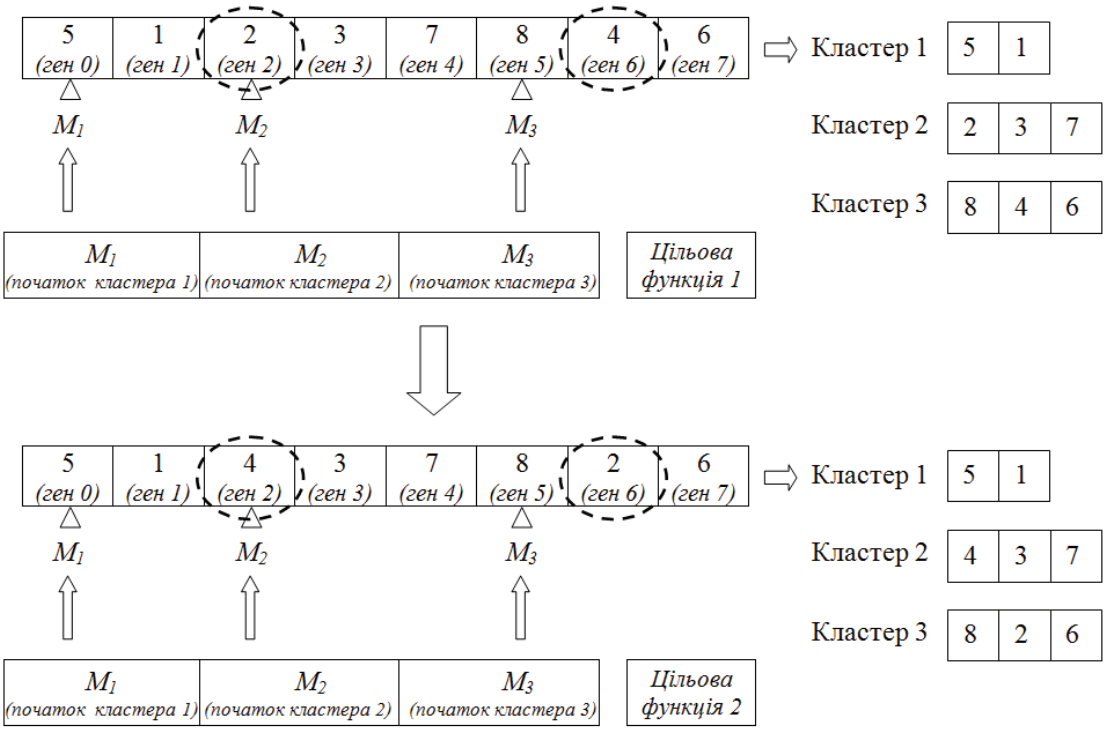


Рисунок 4 – Мутація в межах однієї форми розбиття

Таким чином, ми змінюємо положення лише одного об’єкту відносно кластерів. Цей вид мутації забезпечує мінливість у межах однієї форми розбиття. Мутанти, отримані за першим видом, розташовуються в робочій матриці з індексами від 2n до 3n-1.

в) мутація 2: для сформованої батьківської хромосоми випадковим чином генерується новий масив маркерів розбиття, при цьому порядок слідування об’єктів залишається незмінним (рисунок 5).

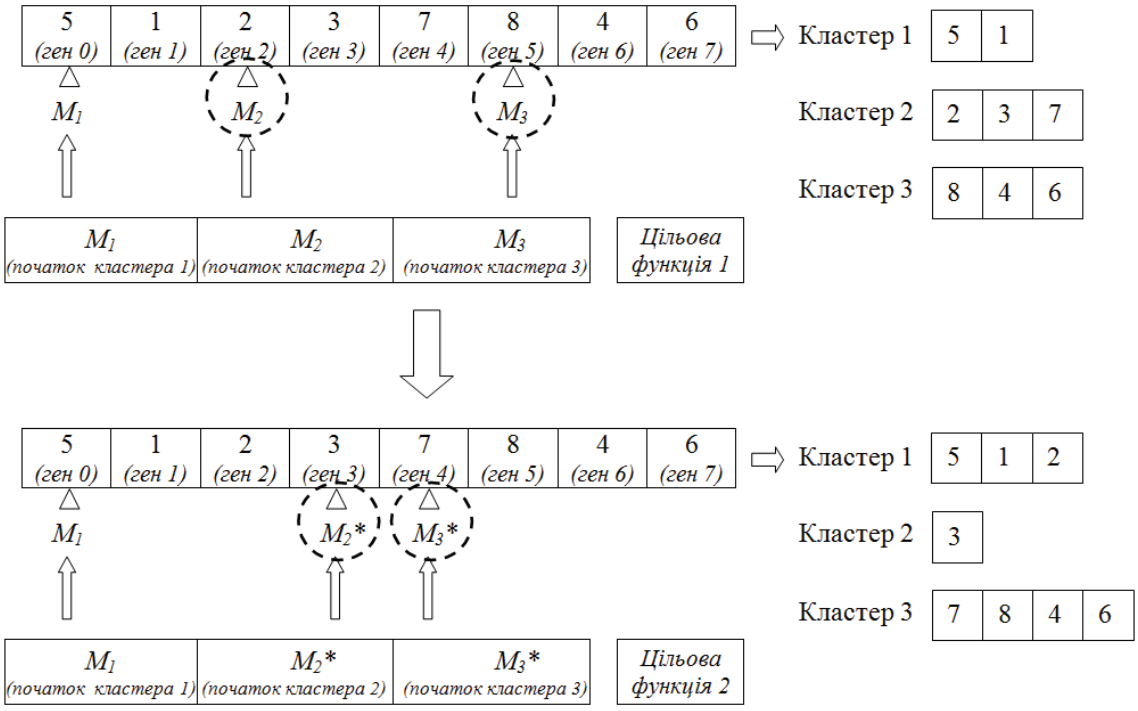


Рисунок 5 – Мутація форм розбиття

Цей вид мутації забезпечує змінність форм розбиття та дозволяє розширити область пошуку оптимального розв’язку. Мутанти за другим видом займають позиції з 3n до 4n-1 робочої матриці.

Для кожного з отриманих таким чином нащадків також підраховується цільова функція.

3) На цьому етапі реалізується механізм природного відбору в межах популяції. Застосуємо стратегію елітизму, яка полягає в тому, що особини з найбільшою пристосованістю гарантовано переходять у нову популяцію. Використання елітизму дозволяє прискорити збіжність генетичного алгоритму. Тому серед 4n особин популяції, хромосоми яких зберігаються в робочій матриці, відбираємо n за критерієм мінімальності цільової функції. Вони формують наступне покоління.

4) Якщо не виконується умова зупинки, очищуємо робочу матрицю та повторюємо попередню процедуру, починаючи з кроку 2. Ітераційний процес вважаємо завершеним, коли поточна сума значень цільової функції для кращих n особин перестає зменшуватись.

Таким чином, із покоління в покоління, сприятливі характеристики розповсюджуються по всій популяції. Схрещування найбільш пристосованих особин приводить до того, що досліджуються найбільш перспективні ділянки простору пошуку. В загальному підсумку популяція буде сходитись до оптимального розв’язку задачі або близького до оптимального. Перевага розробленого генетичного алгоритму полягає в тому, що він знаходить приблизні оптимальні розв’язки за відносно короткий час.

Комп’ютерний експеримент

Для дослідження було взято дані фірми-провайдера телекомунікаційних послуг «Хорс-Телеком». Розглядалася статистика доступу за добу 14.06.2010 р. для 50 користувачів. Для характеристики абонентів було обрано такі показники:

- швидкість передачі даних;
- повний обсяг вхідного трафіку за заданий період часу;
- повний обсяг вихідного трафіку за заданий період часу.

В результаті роботи програми отримуємо варіант розбиття набору вхідних векторів на 4 однорідні групи (таблиця 1):

Таблиця 1 – Результат розбиття множини абонентів на кластери

Номер кластеру	Кількість користувачів	Швидкість доступу	Співвідношення вхідного (Inbound) та вихідного (Outbound) трафіку
1	6	1 - 10 M	
2	23	128 k – 2 M	Inbound > Outbound
3	12	128 k – 1 M	Inbound >> Outbound
4	9	128 k – 1 M	Inbound ≤ Outbound

Таким чином, у вихідних даних можна виділити:

кластер 1 – невеликий сегмент користувачів із найвищою по вибірці швидкістю доступу;

кластер 2 – найбільший; у ньому представлені користувачі з середньою швидкістю доступу, в яких вхідний трафік перевищує вихідний;

кластер 3 – до нього потрапили переважно абоненти з невисокою швидкістю, що використовували лише вхідний напрямок передачі даних;

кластер 4 – складається з абонентів, для яких характерна перевага обсягу вихідного трафіку над вхідним; швидкість доступу переважно невисока.

Як уже зазначалося, дане розбиття є наближеним. Не зважаючи на це, в отриманих даних чітко прослідковується наявність компактних скупчень об’єктів, і на їх основі можна судити про характер логічних закономірностей, закладених у вихідних даних.

Проведемо дослідження часу роботи розробленого генетичного алгоритму порівняно з часом виконання алгоритму граничного перебору для різної кількості об’єктів у вихідній множині. Результати порівняння представимо на графіку, наведеному на рис. 6. По осі абсцис будемо відкладати кількість об’єктів у вихідній множині, по осі ординат – час роботи програми.

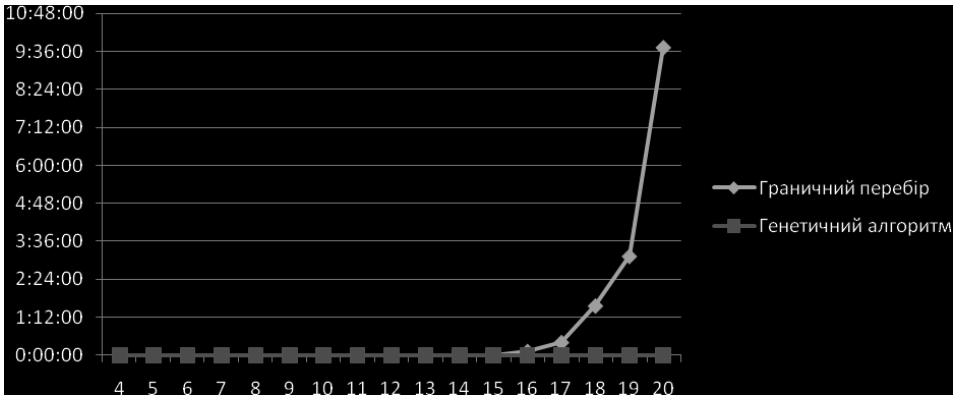


Рисунок 6 – Порівняння алгоритмів за часом виконання

Бачимо, що в задачах невеликої розмірності (до 20 користувачів) генетичний алгоритм за часом виконання показує результат, неспівровзмірно малий порівняно з алгоритмом граничного перебору. Час виконання останнього починає стрімко зростати вже при кількості n=15 користувачів.

При цьому в задачах більшої розмірності генетичний алгоритм поводить себе таким чином, як показано на рис. 7: помітне зростання часу виконання програми спостерігається лише тоді, коли кількість користувачів перевищує 50. Проте навіть для значно більшого обсягу вхідних даних (100 користувачів і більше) час виконання надзвичайно малий порівняно з аналогічним показником для алгоритму граничного перебору.

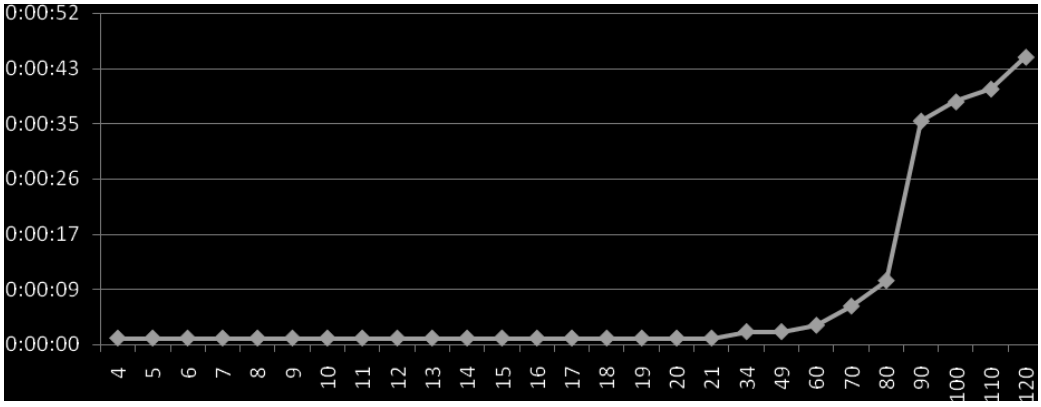


Рисунок 7 – Залежність часу виконання генетичного алгоритму від розмірності задачі

Слід зазначити, що експеримент поставлено на ПЕОМ із середніми на сьогоднішній день характеристиками: AMD Athlon 64 X2, 2,4 ГГц; 2 ГБ ОЗП.

Отримані дані дають змогу стверджувати, що за показником швидкодії генетичний алгоритм кластеризації має значну перевагу над алгоритмом граничного перебору на основі динамічного

програмування. Вже для 21 об’єкта вихідної множини час розв’язання задачі за допомогою методу граничного перебору не вкладається в жодні розумні межі. Водночас розроблений генетичний алгоритм дає розв’язок при значно більшій розмірності задачі за цілком прийнятний час.

Специфіка Інтернет-технологій порівняно з іншими вимагає передусім високої швидкодії програмних засобів, а також можливості нарощування розмірності задачі в широких межах. За таких умов доцільно надати перевагу запропонованому наближеному методу кластеризації.

Висновки

У статті запропоновано підхід для розв’язання задач кластеризації користувачів мережі Інтернет. Розроблено генетичний алгоритм для розв’язання поставленої задачі та відповідне програмне забезпечення для його практичного застосування. Задля врахування специфіки задачі та підвищення ефективності роботи генетичного алгоритму було застосовано неоднорідні хромосоми. Відповідно до цього було внесено суттєві модифікації до перебігу класичних процедур схрещування та мутації.

Розроблений алгоритм було досліджено на швидкодію в порівнянні з алгоритмом граничного перебору та показано значну його перевагу за цим показником.

Подальші дослідження будуть пов’язані з розвитком інтелектуального підходу до розв’язання даної задачі та використанням нечітких множин із метою подолання зашумленості вхідних даних.

Список літератури

1. Муссель К. Предоставление и биллинг услуг связи. Системная интеграция. – М.: Эко-Трендз, 2003. – 319 с.
2. Дюран Б., Оделл П. Кластерный анализ. - М.: Статистика, 1977. – 128 с.
3. Матвеев Ю.Н. Основы теории систем и системного анализа: учебное пособие. – Тверь: ТГТУ, 2007. – 100 с.
4. Мандель И.Д. Кластерный анализ. – М.: Статистика, 1988. – 176 с.
5. Ротштейн А.П. Интеллектуальные технологии идентификации: нечеткие множества, генетические алгоритмы, нейронные сети. – Винница: «УНИВЕРСУМ-Винница», 1999. – 320с.

Відомості про авторів

Захарченко Сергій Михайлович, к.т.н., доцент кафедри ОТ, Вінницький національний технічний університет, Хмельницьке шосе, 95, Вінниця, 21021, Україна, тел.: (0432) 43-70-41.

Кондратенко Наталія Романівна, к.т.н., доцент, професор кафедри ЗІ, Вінницький національний технічний університет, Хмельницьке шосе, 95, Вінниця, 21021, Україна, тел.: (0432) 59-83-79, e-mail: kondrn@yandex.ru.

Манаєва Ольга Олексіївна, студентка, Вінницький національний технічний університет, Хмельницьке шосе, 95, Вінниця, 21021, Україна, тел.: (0432) 56-33-42, e-mail: sleeepy@meta.ua.

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

УДК 004.932 + 004.931

ВИСОКОШВИКІСНИЙ МЕТОД ДЛЯ ВІДСТЕЖЕННЯ МІТОК У СКАНОВАНИХ ДОКУМЕНТАХ

Ремінний О.А.

Анотація: В даній роботі вперше запропоновано методику пошуку міток в сканованих документах за допомогою методу кругових бінарних обчислень. Підхід тестується на базі міток, що виділяються на вхідному зображенні за рахунок кіл, що розміщені в чотирьох кутах мітки. Наведено результати обробки вхідних зображень с точки зору швидкодії та ймовірності коректної обробки вхідного зображення.

Анотация: В данной работе впервые предложена методика поиска меток в отсканированных документах с помощью метода круговых бинарных вычислений. Подход тестируется на базе меток, выделяемых на входном изображении за счет кругов, которые размещены в четырех углах метки. Приведены результаты обработки входящих изображений с точки зрения быстродействия и вероятности корректной обработки входного изображения.

Abstract: In this paper the technique of the labels search in scanned documents using the circular binary calculations is proposed. The approach is tested on the basis of labels that are allocated on the input image with circles, which are placed in four corners of the label. The results of the input images processing are described from the standpoint of performance and the input image correct processing probability.

Ключові слова: сканування, відстеження міток, класифікація зображень.

Вступ

Сучасний стрімкий розвиток цифрових технологій призводить до того, що велика кількість паперових документів (книг, зображень, тощо), потребують приведення до цифрового формату. Найбільш розповсюджений варіант обробки таких документів є сканування з подальшим розпізнаванням, класифікацією і т.п.

На практиці зустрічаються задачі, пов‘язані з накладенням певих міток на вхідні зображення документів. Такі задачі пов‘язані з великими кількостями документів, які в подальшому потрібно буде автоматично класифікувати.

Обробка зображень завжди пов’язана зі значними апаратними затратами. В задачі класифікації зображень ці апаратні затрати можна поділити на два типи: затрати, повязані з виділенням об‘єкту на фоні та затрати на класифікацію виділеного об‘єкту.

В роботі розглядаються задачі класифікації об’єктів за їх формою. Прикладом може бути класифікація об’єктів на морській поверхні, знятих спектральною камерою з повітря, класифікація об’єктів, що рухаються по вулиці. До існуючих методів можна віднести метод медіанних осей[1], розпізнавання на основі порівняння виділених характеристик[2-5], наприклад кутових характеристик[2].

В [6],[7] запропоновано метод класифікації зображень на основі аналізу форми шуканих об’єктів. Основною перевагою над представленими методами є швидкість обробки вхідного зображення, коли для виділення інформації про все зображення достатньо зробити лише один прохід через всі пікселі. Крім того, мінімальною є і кількість даних, яка буде порівнюватись з еталонами, що також призводить до виграшу в швидкодії.

В статті розглянута задача відстеження міток на вхідному зображенні за допомогою методу кругових бінарних обчислень. Така задача може поставати при обробці великої кількості відсканованих документів, що є досить актуальним в наш час інформатизації документообігу. Представлений метод для розпізнавання об’єктів на основі геометричної відповідності розподілів точок, що належать об’єкту в геометричних фігурах, які окреслюють об’єкт. Метод працює через розбивання області знаходження об’єкта на частини, та обробку статистичної інформації, яку містять ці частини.

Метою роботи є використання методу кругових бінарних обчислень для підвищення швидкості знаходження міток на вхідному зображенні сканованого документу.

Для обробки вхідного зображення спочатку виділяєтються класифікатори для зафарбованих та незафарбованих елементів міток. Надалі відбувається обробка: сегментація, класифікація, перевірка класифікованих елементів на відповідність вхідній мітці.

Математична модель методу

В системах реального часу швидкість обчислень для ідентифікації об‘єкту є критчною величиною. До проблем які є вторинними розпізнавання можна віднести віддаленість та кут повороту об’єктів. Тому виникає потреба у простих, однак достатньо потужних методах класифікації.

Коли класифікація стосується об‘єктів, які належать до певного скінченного алфавіту, доцільно використовувати методи, що забезпечують максимальну швидкодію через аналіз тільки деяких характеристик об‘єкту. В [6] запропоновано метод класифікації зображень на основі аналізу форми шуканих об’єктів, який демонструє значну швидкодію при обробці вхідних зображень.

В[6] наведено варіант збору інформації про форму об’єкту за допомогою методу бінарних кругових обчислень. Наведемо коротку суть даного методу.

Спочатку знаходиться центр мас чорно-білого зображення предмету за формулою (1) – він і буде центром кола.

$$\begin{cases} X = \dfrac{x_1 \cdot m_1 + x_2 \cdot m_2 + \ldots x_n \cdot m_n}{m_1 + m_2 + \ldots m_n} \\ Y = \dfrac{y_1 \cdot m_1 + y_2 \cdot m_2 + \ldots y_n \cdot m_n}{m_1 + m_2 + \ldots m_n} \end{cases} \quad (1)$$

де x_i y_i – координати точки з масою m_i.

Оскільки вхідний файл ми вважаємо бінарним, кожна точка може приймати лише забарвлений ($m_i = 1$) або незабарвлений ($m_i = 0$) стан.

Для знаходження радіусу кола, в які буде вписано об’єкт, запропоновано наступний метод. Від центру мас в k напрямках будуються вектори. Потім обирається m підмножина векторів з найдовшою довжиною проходження через об’єкт (через точки, які є забарвленими). Значення довжини радіусів знаходиться як:

$$R = \frac{1}{m}\sum_{i=1}^{m} r_i, r_i \in (O(n))_m \quad (2)$$

Для знаходження радіусу кожного додаткового кола всередині зовнішнього використовується формулу()

$$R_i = R_{i-1} \cdot \cos(45°) \quad (3)$$

де R_i – радіус шуканого внутрішнього кола;
R_{i-1} - радіус попередньо обчисленого кола;

Об’єкт кола представимо у вигляді залежності від пікселів, що йому належать:

$$C(P) = \frac{\sum p_{i,j}}{\sum p_{i,j}\,|_{p=1}}, \quad (4)$$

де $p_{i,j}$ –піксель, що належить колу,
$p=1$- забарвлений піксель.

В залежності від кількості класів об’єктів та їх складності вводиться будь-яка кількість внутрішніх кіл як характеристик – $C(P_1)$,$C(P_2)$, $C(P_3)$…. Даний варіант значно швидший за інші методи виділення місцевих особливостей об’єкту [6-7].

Результати аналізу конкретного $C(P_i)$ об’єкту порівнюватимуться з еталонними $C(P_i)$’ значеннями для кожного з класів, в результаті чого знаходитимуться найбільш відповідні класи.

Практична реалізація методу для виділення міток на сканованих документах

Для конкретності можна розглянути зображення, наведене на рис. 1. Поверх тексту накладено мітку, основними характеристиками якої є зафарбовані та прозорі круги у кутах мітки. Зображення містить у собі велику кількість геометричних примітивів, які представляють собою літери англійського алфавіту. Також там можуть бути розміщені зображення і т.п.

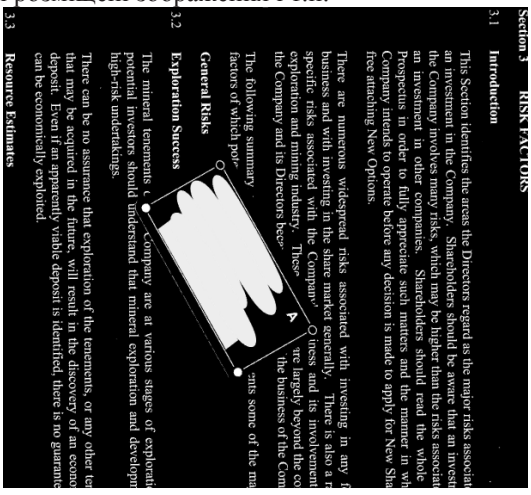


Рисунок 1 – Зображення відсканованої сторінки тексту з накладеною на текст міткою

Щодо проміжних кроків можна віднести виділення тих об‘єктів, які належать зображенню завдяки а) фільтруванню за розміром; б) завдяки використанню того факту, що мітка формується завдяки сполученню чотирьох символів визначеного алфавіту і в загальному випадку утворює прямокутник.

Початковий аналіз показує, що загалом мітка виділяється з загального алфавіту лише завдяки утворенню прямокутника з вхідних елементів.

За умовою, нам не можна привязуватись до елментів сканованої сторінки, т.я. на ній можуть бути тексти різного розміру та рисунки. Мітка також може бути різного розміру, єдине, що залишається в мітці сталим – це символи та їх розмір в кутах мітки.

В нашому випадку маємо два типи символів в кутках – зафарбоване та незафарбоване коло. Для того щоб виділити ці кола пропонується провести сегментацію з фільтрацією обєктів по розміру.

Потім з множини об‘єктів, які залишились, потрібно обрати ті, які відповідають по формі нашим цільовим об‘єктам. Їх буде описувати наступний алфавіт (його можна обмежити лише тими зображеннями, які можуть належати міткам (C_1 для зафарбованих зображень та C_2 для незафарбованих)):

$$\begin{cases} C_1(P) >= 0.9, \\ C_2(P) <= 0.1. \end{cases} \quad (5)$$

Останнім кроком є вибір тих об‘єктів, які утворюють прямокутник. Для цього достатньо скористатись простими математичними перетвореннями, які визначать паралельність прямих, утворених лініями.

З точки зору пошуку цільових об‘єктів за формою, нас найбільше цікавить 2 та 1 етапи. Алгоритм сегментації був обраний з бібліотеки AForge [8] – BlobCounter. Він дозволяє виділити на вхідному зображенні всі області в заданому інтервалі розмірів. Подальша обробка – збір статистичних данних у виділених елементах по формулі (4) і порівняння їх з величинами, приведеними в (5).

На персональному компьютері з CPU 2GHz операція сегментації вхідного зображення розміром 3904x5656 відбувається в середньому за 5 секунд (найдовша операція). Подальше коректне визначення кіл, які належать міткам, займає приблизно 100 мілісекунд.

Висновки

В даній статті вперше запропоновано метод вирішення задачі пошуку цільових об‘єктів на зображеннях за допомогою методу класифікації об’єктів за кореляційними ознаками відповідності розподілів зафарбованих точок, що належать об’єкту всередині геометричних фігур, які окреслюють об’єкт. Як геометрична фігура для вирішення конкретної задачі було обрано коло, т.я. воно максимально співпадає з формою цільових об‘єктів пошуку. Оскільки у роботі використовувалась прив‘язка до такої величини як розмір цільового об‘єкту, довелося знехтувати однією з важливих властивостей цільового методу – інваріантності до розміру об‘єкту.

Під час проведення експерименту було досягнуто високої швидкодії(визначення кіл, які належать міткам, займає приблизно 100 мілісекунд) та високої імовірності коректної ідентифікації кутових об‘єктів міток(мітки коректно розпізнавались на 17 зразках із 20 – ймовірність коректного розпізнавання 85%). В подальшому планується продовжити роботу з наведеним методом для вдосконалення його стійкості по відношенню до присутності шуму на зображенні та підвищенню точності.

Список літератури

1. S. C. Zhu. Stochastic Jump-Diffusion process for computing Medial Axes in Markov Random Fields/ S. C. Zhu // IEEE Trans. on PAMI. - Vol. 21, No.11. – 1999. – P. 1158-1169.
2. Mikolajczyk K. Shape recognition with edge-based features/ Mikolajczyk K., A. Zisserman, C. Schmid// Proceedings of the British Machine Vision Conference. – 2003.
3. Agarwal S. Learning a sparse representation for object detection/S. Agarwal, D. Roth// In Proc. ECCV. - Vol.4. - 2002. - P.113–130.
4. Amores J. Fast spatial pattern discovery integrating boosting with constellations of contextual descriptors/J. Amores, N. Sebe, and P. Radeva// In Proc. CVPR. - Vol.2.- 2005. - P.769–774.
5. Bouchard G. Hierarchical part-based visual object categroization/ G. Bouchard B. Triggs// In Proc. CVPR. - 2005. - P.710–715.
6. Kvetny R.N. Binary circular calculations method for the objects classification using their form/ Kvetny R.N., Reminnyi O.A.// Aplikovane vedecke novinki. – 2009. – P 54 - 60.
7. Квєтний Р.Н. Високошвидкісний метод класифікації зображень/ Р. Н. Квєтний Р.Н., Ремінний О. А. // "Оптико-електронні інформаційно-енергетичні технології". – 2009. – С 5 - 10.
8. Kirilow A. AForge.NET open source framework [Електронний ресурс] / Kirilow A. – Режим доступу: http://www.codeproject.com/KB/recipes/aforge.aspx.

Відомості про авторів

Ремінний Олександр Андрійович - аспірант кафедри АІВТ, Вінницький національний технічний університет, E-Mail: Alexander.Reminniy@gmail.com

004.89+656.2

МОДИФІКОВАНИЙ АЛГОРИТМ ПОБУДОВИ FP-ДЕРЕВ ПРИ АНАЛІЗІ НАДЗВИЧАЙНИХ СИТУАЦІЙ НА ЗАЛІЗНИЦІ

Савчук Т.О., *Юхимчук* С.В. *, Вознюк А.В.*

Анотація. В роботі запропоновано модифікований алгоритм побудови FP-дерев, який використовується при аналізі надзвичайних ситуацій, що виникають при перевезенні небезпечних вантажів залізничним транспортом. До модифікації належить: введення небінарних атрибутів, що збільшують покриття надзвичайних ситуацій кожним правилом, та використання більш складних перетинів множин атрибутів ситуації порівняно з правилами, що оперують бінарними атрибутами; введення характеристики інформативності правила.
Ключові слова: асоціативні правила, надзвичайні ситуації, пошук закономірностей, модифікований алгоритм, побудова FP-дерева, небінарний атрибут, інформативність правила.
Аннотация. В работе предложен модифицированный алгоритм построения FP-деревьев, который используется при анализе чрезвычайных ситуаций, возникающих при перевозке опасных грузов железнодорожным транспортом. К модификации принадлежат: введение небинарных атрибутов, которые увеличивают покрытие чрезвычайных ситуаций каждым правилом,а также использование более сложных пересечений множественных чисел атрибутов ситуации в сравнении с правилами, которые оперируют бинарными атрибутами; введение характеристики информативности правила.
Ключевые слова: ассоциативные правила, чрезвычайные ситуации, поиск закономерностей, модифицированный алгоритм, построение FP-дерева, небинарный атрибут, информативность правила.
Annotation. The modified algorithm of construction of FP-trees, which is used for the analysis of extraordinary situations which arise up at transportation of dangerous loads a railway transport is in-process offered. It is belonged to modification: introduction of unbinary attributes, which increase coverage of extraordinary situations every rule and the use of more difficult crossings of plurals of attributes of situation by comparising to rules which operate binary attributes; introduction of description of informing of rule.
Keywords: associative rules, extraordinary situations, search of conformities to law, modified algorithm, construction of FP-tree, unbinary attribute, informing of rule.

Вступ

Найбільш значущими і необхідними завданнями прогнозування надзвичайних ситуацій (НС) на залізничному транспорті є ймовірність виникнення кожного з шляхів розвитку таких ситуацій і, відповідно, їх масштабів, розмірів зон НС. Можливі тривалі наслідки при виникненні НС певних типів, масштабів, часових інтервалів або їх певних сукупностей; потреби сил і засобів для ліквідації прогнозованих надзвичайних ситуацій [1].

В цілому результати моніторингу та прогнозування є вихідною основою для розробки довгострокових, середньострокових і короткострокових прогнозів для прийняття відповідних рішень щодо попередження та ліквідації надзвичайних ситуацій [1, 2].

Значні обсяги сучасних баз даних викликали стійкий попит на нові масштабовані алгоритми аналізу даних. Одним з ефективних методів виявлення знань стали алгоритми пошуку асоціативних правил [3].

Доцільність застосування пошуку асоціативних правил до аналізу НС на залізничному транспорті

Безпека функціонування небезпечних об’єктів залежить від багатьох чинників: фізико-хімічних властивостей сировини, напівпродуктів і продуктів, від характеру технологічного процесу, від конструкції і надійності устаткування, умов зберігання і транспортування радіоактивних вантажів, стану контрольно-вимірювальних приладів і засобів автоматизації, ефективності засобів протиаварійного захисту і т.д.[3].

Потрібно аналізувати задачу та приймати необхідні рішення на основі неповних, нечітких даних, то традиційні методи не зможуть справитися з поставленою задачею. Виникає необхідність використання саме інтелектуальних систем та методів. Всі відхилення на протязі процесу розвитку надзвичайної ситуації, що можуть змінити його хід чи вплинути на якість вихідного результату знаходяться в відомих рамках, які також визначаються в процесі аналізу.

Таким чином, створюється статистична стабільність [3, 4], що обумовлює застосування технологій Data Mining при знаходженні закономірностей між зв’язаними подіями в базах даних та відповідає задачі пошуку асоціативних правил [5].

Постановка задачі

Перед побудовою асоціативних правил та подальшим їх аналізом необхідно визначити об’єкти, що будуть використовуватись.

Для побудови асоціативних правил достатньо бази даних, кожен кортеж якої описує конкретну надзвичайну ситуацію. Кортеж містить набір значень атрибутів надзвичайної ситуації, які описують наступні об’єкти: небезпечний вантаж ($HB\{pp_1,...,pp_m\}$), оточуюче середовище ($C\{c_1,...,c_p\}$), контейнер ($K\{k_1,...,k_n\}$), надзвичайна ситуація ($HC\{h_1,...,h_t\}$).

Усі атрибути можуть приймати як числові значення, так і представлятись у вигляді проміжку значень.

Отже, задачу аналізу НС, що базується на пошуку асоціативних правил. Можна сформувати таким чином: на основі множини значень атрибутів надзвичайних ситуацій $HB\{pp_1,...,pp_m\}$, $K\{k_1,...,k_n\}$, $C\{c_1,...,c_p\}$, $HC\{h_1,...,h_t\}$ необхідно визначити асоціативні правила, що описують надзвичайні ситуації та застосувати їх для аналізу вхідних даних, якими є значення атрибутів конкретної НС, що аналізується.

Модифікований алгоритм побудови FP-дерев

Завдання знаходження асоціативних правил розбивається на дві підзадачі:

1. Знаходження всіх наборів елементів, які задовольняють порогу minsupport. Такі набори елементів називаються часто зустрічаючими (частими наборами).

2. Генерація правил з наборів елементів, знайдених згідно п.1. з вірогідністю, що задовольняє порогу minconfidence.

Для опису небінарних величин використано апарат алгебри кінцевих предикатів, тобто якщо ознака X_j об’єкту може приймати значення з множини $A=\{a_1,a_2,...,a_n\}$, то така ознака може описуватись за допомогою кінцевого предикату наступним чином:

$$X_j^{a_i} = \begin{cases} 1, & якщо \quad X_j = a_i, \\ 0, & якщо \quad X_j \neq a_i. \end{cases}$$

Зазначимо, що мають місце наступні тотожності:

$$X^{a_1} \vee X^{a_2} \vee ... \vee X^{a_n} \equiv 1,$$
$$X^{a_i} X^{a_j} \equiv 0,$$

які дозволяють не використовувати окремі описи для асоціації з запереченням, тобто $\overline{X^{a_i}} = X^{a_1} \vee ... X^{a_{i-1}} \vee X^{a_{i+1}} ... \vee X^{a_n}$.

У випадку бінарних ознак для їх опису використовується унарний предикат, що розрізняє два значення – 0 и 1: X^0 и X^1.

Двох характеристик – підтримки й рівня довіри недостатньо для оцінки асоціації. У доповнення до підтримки й рівня довіри будемо використовувати таку характеристику, як рівень покращення, яка визначається таким чином:

$$Imp(X \to Y) = \frac{Conf(X \to Y)}{Sup(Y)}.$$

Розглянемо характеристики асоціативного правила з точки зору теорії ймовірностей:

$$Sup(X \to Y) = P(XY);$$

$$Conf(X \to Y) = \frac{Sup(X \to Y)}{Sup(X)} = \frac{P(XY)}{P(X)} = P_X(Y);$$
$$Imp(X \to Y) = \frac{Sup(X \to Y)}{Sup(X)Sup(Y)} = \frac{P(XY)}{P(X)P(Y)}.$$

Тоді обмеження, що накладаються на характеристики асоціації, визначаються як:

1) $0 \le Sup(X \to Y) \le 1;$

2) $Sup(X \to Y) \le Conf(X \to Y) \le 1;$

3) $$\frac{(Conf(X \to Y))^2}{Sup(X \to Y) \cdot Conf(X \to Y) + Conf(X \to Y) - Sup(X \to Y)} \le Imp(X \to Y) \le \frac{Conf(X \to Y)}{Sup(X \to Y)}.$$

Після знаходження асоціативних правил необхідно визначати доцільність та можливість їх використання. Для цього використовуватимемо оцінку інформативності асоціації, що розраховується на основі стандартних характеристик асоціативного правила. Повна взаємна інформація у загальному випадку може бути представленою таким чином:

$$I_{X \leftrightarrow Y} = \sum_{i=1}^{n} \sum_{j=1}^{m} P_{ij} \log_2 \frac{P_{ij}}{p_i r_j};$$

де $P_{ij} = P((X \sim x_i)(Y \sim y_i))$ – ймовірність того, що X знаходиться в стані x_i, а Y – в стані y_j;

$p_i = P(X \sim x_i)$ – ймовірність того, що X знаходиться в стані x_i;

$r_i = P(Y \sim y_i)$ – ймовірність того, що Y знаходиться в стані y_j.

Для асоціативних правил взаємна інформація буде визначатися як:

$$I_{X \leftrightarrow Y} = \sum_{i=1}^{n} \sum_{j=1}^{m} Sup(x_i \to y_j) \log_2 Imp(x_i \to y_j).$$

Правило процесу побудови FP-дерев: якщо для чергового об’єкта в дереві зустрічається вузол, ім'я якого збігається з іменем об’єкта, то об’єкт не створює нового вузла, а індекс відповідного вузла в дереві збільшується на 1. В іншому разі для цього об’єкта створюється новий вузол і йому присвоюється індекс 1.

Розглянемо генерацію агрегованих асоціативних правил на основі об’єднання гілок дерева покриттів, що полягає у наступному. Якщо потребується знайти узагальнені покриття, які містять значення ознаки P^{t_1} и P^{t_2}, то визначається список гілок дерева покриттів з даними значеннями ознаки P. Ці гілки групуються наступним чином: в одну групу входять ті, що містять ознаки на однакових ярусах дерева покриттів. В цьому списку існують гілки, які мають однакові префікси: $S_1\alpha_1, S_1\alpha_2, ...S_1\alpha_i$ (такі гілки обов’язково існують, тому що всі гілки містять хоча б один загальний префікс – нульову строку λ). У кожній групі можуть існувати гілки з різними загальними префіксами. Знаходимо $S = \max\{S_j\}$, $(j = \overline{1,k})$, де k – кількість різних однакових префіксів в даній групі гілок. Ці гілки об’єднуються по принципу: $S(\alpha_1 + \alpha_2 + ... + \alpha_i)$.

Рядок ($\alpha_1 + \alpha_2 + ... + \alpha_i$) формується таким чином: між значеннями ознаки з одного ярусу ставиться знак $\vee$. Отримана агрегована гілка додається в свою групу. Далі цей процес продовжується ітеративно до тих пір, поки існують гілки з однаковими префіксами.

Рівень підтримки такого об’єднання дорівнює сумі рівнів підтримки гілок, що об’єднуються.

Запропонований метод розрахунку дозволяє розкладати агреговане асоціативне правило [4, 5] на прості й знаходити його рівень підтримки через рівні підтримки простих асоціацій, які легко знаходяться з побудованого дерева покрить. Наприклад, правило $(Z^{c1} \vee Z^{c2}) \to X^{a3}(Y^{b1} \vee Y^{b2})$, що отримано з покриття $X^{a3}(Y^{b1} \vee Y^{b2})(Z^{c1} \vee Z^{c2})$ (*Sup=3*) може бути розкладено, як показано на рисунку 1.

Відмітимо, що розкладання агрегованого асоціативного правила може проводитися як зліва направо, так і навпаки. Даний підхід дозволяє не використовувати додаткове сканування БД для знаходження параметрів узагальнених асоціацій.

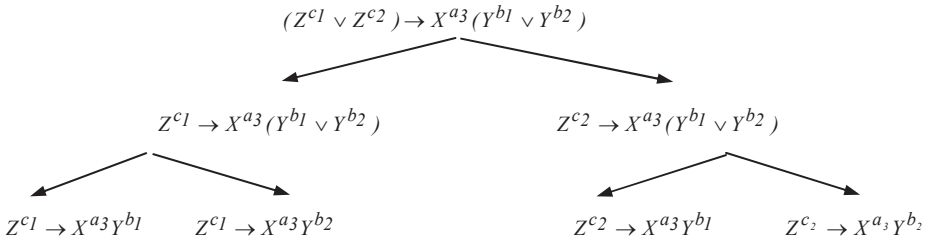


Рисунок 1 – Розкладання агрегованого асоціативного правила

Висновки

Таким чином, використання модифікованого алгоритму побудови FP-дерев, що базується на введенні інформативності небінарних асоціативних правил та подальша їх агрегація наведеним в роботі способом дозволяє, окрім розв’язання поставленої задачі, значно збільшити ефективність та швидкодію обробки надзвичайної ситуації. При цьому, одним з ефективних методів побудови асоціативних правил, що використовуватимуться для аналізу надзвичайних ситуацій є алгоритм побудови FP-дерев. При побудові небінарних асоціативних правил та розгляді їх з точки зору теорії ймовірностей вВводиться додаткова характеристика – “інформативність” правила.

Список літератури

1. С. В. Юхимчук, М. Д. Кацман «Моделі автоматизації вироблення рекомендацій керівнику гасіння пожежі на залізничному транспорті», Універсум-Вінниця, 2008 р.
2. М. Д. Кацман, Г. Б. Кононов, І. В. Діденко, Н. В. Огороднійчук. Ліквідація пожеж на залізничному транспорті. Київ, видавництво «Основа» 2006 р. 215 с.
3. Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И. Методы и модели анализа данных: OLAP и Data Mining. – СПб.: БХВ-Петербург, 2004 – 336с.
4. Nicolas Pasquier, Yves Bastide, Rafik Taouil, Lotfi Lakhal Discovering Frequent Closed Itemsets for Association Rules.
5. Игнатов Д, Методы бикластеризации для анализа интернет-данных.

Відомості про авторів

Савчук тамара Олександрівна – к.т.н., професор кафедри комп’ютерних наук, Україна, м. Вінниця, Вінницький національний технічний університет, savchtam@gmail.com.

Юхимчук Сергій Васильович – д.т.н., професор кафедри комп’ютерних наук, Україна, м.Вінниця, Вінницький національний технічний університет.

Вознюк Андрій Вікторович – магістрант кафедри комп’ютерних наук, Україна, м. Вінниця, Вінницький національний технічний університет, kenny@ukr.net.

УДК 681.5.017:616-71

СИСТЕМА ДІАГНОСТУВАННЯ ФУНКЦІОНАЛЬНОГО СТАНУ ЛЮДИНИ НА ОСНОВІ ІНВАРІАНТНОГО ПЕРЕТВОРЮВАЧА ПАРАМЕТРІВ БАТ

Роїк О.М., Яремко С.А.

Анотація: Запропонована структурна схема системи діагностування функціонального стану людини на основі первинного інваріантного перетворювача параметрів БАТ, що дозволить підвищити точність вимірювань та достовірність прийняття рішень про стан здоров’я людини.
Аннотация: Предложена структурная схема системы диагностирования функционального состояния человека на основании первичного инвариантного преобразователя параметров БАТ, что даст возможность повысить точность измерений и достоверность принятия решения о состоянии здоровья человека.
Annotation: The flow diagram of the system diagnosing functional state of man is offered on the basis of primary invariant transformer of parameters BAT, which will allow to promote exactness of measurings that authenticity of decision-making about the state of health oman.

Вступ

Протягом останніх десятиліть, у зв’язку із різким збільшенням негативних наслідків традиційного лікування хімічними препаратами, все більшої актуальності набувають методи нетрадиційної медицини, що враховують взаємозв’язок усіх функціональних систем і дають змогу на ранніх стадіях, коли ще немає явних ознак, виявити органічні зміни та здійснити корегувальний вплив по відновленню нормального функціональної стану людини. Незважаючи на широке застосування даних методів у різних діагностичних системах [1-3], залишаються недостатньо дослідженим вплив завад, тестового навантаження, шунтувальних біотканин та інших факторів, які можуть привести до значного спотворення вимірювальних сигналів у БАТ і в кінцевому підсумку прийняття невірного рішення про стан здоров’я людини. В [6] здійснене дослідження похибок окремих структурних компонентів системи для діагностування параметрів БАТ, проте недостатньо досліджений вплив завад та інших негативних чинників на вимірювальний сигнал і засоби їх зменшення.

Актуальність

Із літературних джерел [7, 8] та проведених досліджень [9] відомо, що найбільшу вагу серед складових впливних величин, які спричиняють похибки вимірювання, займають завади. Крім того, спотворення корисного сигналу може бути викликане сукупним впливом таких факторів як виділення або забруднення на поверхні шкіри під час вимірювання, наявністю шунтувального впливу зі сторони інших органів і тканин людини [10] та ін. Аналіз існуючих методів діагностування показав, що вони не враховують об’єктивно існуючих взаємозв’язків між БАТ, що може привести до отримання недостовірних результатів, оскільки під час вимірювань досліджувану БАТ шунтують інші. Тому актуальним є застосування методів інваріантних (незалежних) від впливних величин перетворень параметрів компонентів складних об’єктів для задач медичної діагностики.

Мета

Метою статті є побудова структурної схеми системи діагностування функціонального стану людини на основі інваріантного перетворювача параметрів БАТ, що дозволить підвищити точність результатів вимірювань та достовірність прийняття рішень про стан здоров’я людини.

Задачі

Відповідно до мети досліджень формулюються такі задачі:

1. Реалізація методу інваріантного відносно впливних величин перетворення електричних параметрів БАТ.
2. Побудова структурної схеми системи діагностування функціонального стану людини на основі інваріантного перетворювача параметрів БАТ.

Розв’язання задач

Для забезпечення інваріантного перетворення параметрів досліджуваної *j*-ї репрезентативної БАТ, що представлена на графі взаємозв’язків між БАТ точкою e1, пропонується здійснити виведення усіх інших репрезентативних точок на «віртуальний нуль» (точка e2). Для усунення інших впливних величин всі БАТ з’єднуються з базовим вузлом, який зображений точкою e3.

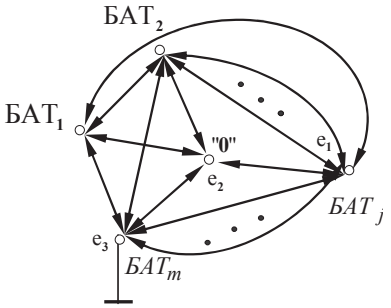


Рисунок 1 – Двонаправлений граф взаємозв'язків між БАТ

Реалізація запропонованого методу інваріантних вимірювальних перетворень електричних параметрів БАТ передбачає створення режиму штучного електричного відокремлення ділянки, параметри якої вимірюються, від інших ділянок, що її шунтують [11]. З цією метою композицію компонентів, що визначають шунтувальний вплив ділять на дві композиції послідовно з'єднаних шунтів $Z_{1\Sigma}$ та $Z_{2\Sigma}$, утворюючи разом із імпедансом Z_x об'єкта дослідження коло типу трикутник. В такому випадку достатньо розірвати шлях протікання струму в одному із шунтів за рахунок забезпечення рівності потенціалів на його полюсах, щоб еталонний струм протікав через ділянку, яка досліджується і не відгалужувався в шунтувальні ділянки.

На рис.2 наведений первинний інваріантний перетворювач (ПІП) параметрів БАТ. Тут режим електричного відокремлення при вимірюванні імпедансу Z_x еквівалентної схеми *j*-ї БАТ (ЕС БАТ$_j$) від шунтувального впливу величин $Z_{1\Sigma}$ та $Z_{2\Sigma}$ забезпечується рівністю потенціалів на полюсах e_2 та e_3 завдяки підключенню до полюсу e_2 операційного підсилювача, який виконує функцію порівняння, при цьому різницевий сигнал підсилюється і через зворотній зв'язок подається на вхід підсилювача. Завдяки цьому в точці e_2 при певному коефіцієнті підсилення з високою точністю встановлюється нульовий потенціал. При цьому в ділянці $Z_{2\Sigma}$ організується режим електричного розриву.

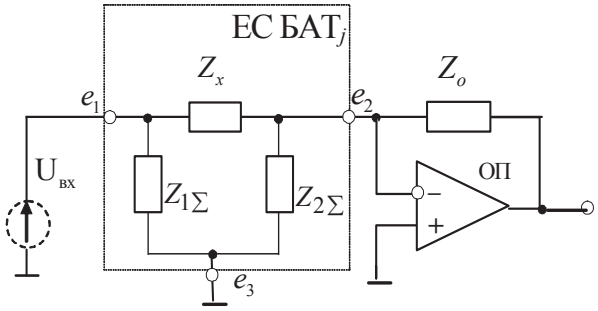


Рисунок 2 – Первинний інваріантний перетворювач параметрів БАТ

Вплив гілки $Z_{1\Sigma}$ виключається за рахунок низького опору джерела тестового сигналу. Тому струм буде протікати тільки через досліджувану і не відгалужуватиметься у шунтувальні ділянки. При цьому струм, протікаючи через зразковий елемент, викликає падіння напруги перетворювача, яке і можна вважати за результат вимірювання

$$U_{вих} = -U_{вх} \cdot \frac{Z_0}{Z_x}, \qquad (1)$$

де $U_{вх}$ - вхідна напруга, яка подається ззовні у методах діагностування із застосуванням тестового впливу [1-5] або власний біопотенціал у методах без застосування тестового впливу [12]; Z_0 - імпеданс еталонного двополюсника; ОП – операційний підсилювач.

Комп'ютерне моделювання роботи ПІП в реальних умовах, коли корисний сигнал реєструється на фоні завад дозволило отримати вихідну напругу, що містить корисну складову сигналу у БАТ та складову завад (рис.3).

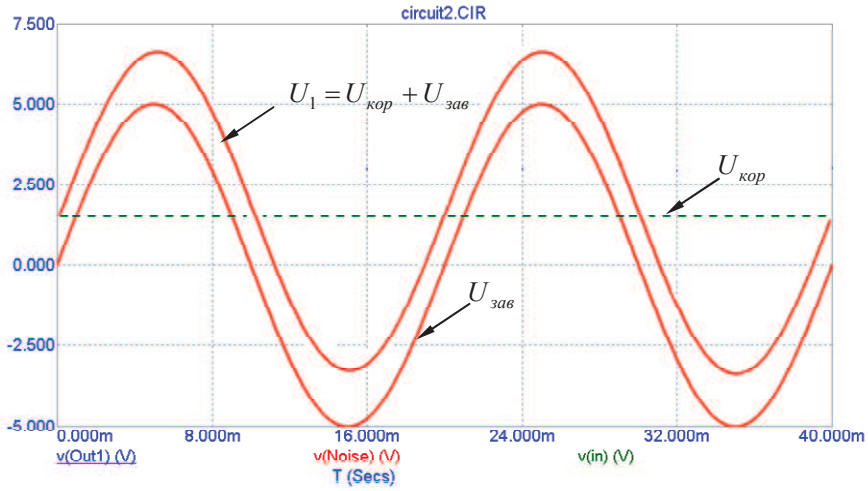


Рисунок 3 – Складові вихідної напруги первинного інваріантного перетворювача БАТ

Виходячи із наведеного вище вираз (1) можна також представити у вигляді

$$U_1 = U_{кор} + U_{зав} \tag{2}$$

де $U_{кор}$ - корисна складова напруги сигналу; $U_{зав}$ - складова завад.

Для виділення корисної складової напруги сигналу пропонується до схеми, наведеної на рис. 2 додати диференційний підсилювач, який реалізований на підсилювачах $П_2$ і $П_3$ з гілками прямого і зворотного зв’язку, що містять резистори $R_1 - R_4$, які мають однакові значення опору (рис. 4).

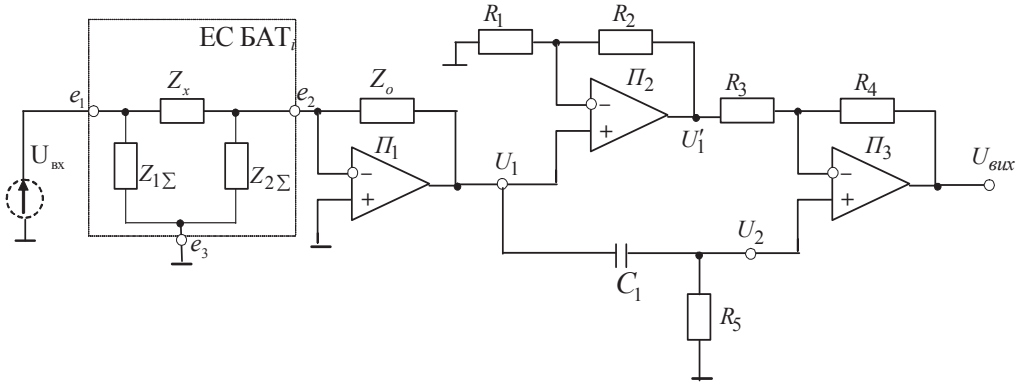


Рисунок 4 – Схема ПІП з відокремленням корисної складової сигналу

При цьому на виході підсилювача $П_2$ отримаємо вираз (3)

$$U_1' = U_1 \frac{R_1 + R_2}{R_1}. \tag{3}$$

При умові, що значення резисторів є однаковими ($R_1 = R_2$), вираз (3) буде мати вигляд

$$U_1' = 2U_1. \tag{4}$$

В іншій гілці вихідна напруга ПІП поступає на C_1R_5, де $R_5 \to \infty$, що дозволить виділити тільки складову сигналу, яка містить завади (5)

$$U_2 = U_{зав}. \tag{5}$$

Напруги, що виражаються через (4), (5) поступають на входи підсилювача $П_3$, який виконує функцію порівняння, в результаті чого напруга на виході, при умові, що $R_3 = R_4$, становитиме

$$U_{вих} = -U_1' + U_2 \frac{R_3 + R_4}{R_3} = -2U_1 + 2U_2 = -2U_{кор} - 2U_{зав} + 2U_{зав} = -2U_{кор} \tag{6}$$

Таким чином, вихідна напруга буде вільною від завад і міститиме корисну складову, що дозволить підвищити точність вимірювання параметрів БАТ.

З урахуванням наведеного вище запропонована вимірювальна система на основі інваріантного перетворення та відокремлення корисної складової вимірювального сигналу у БАТ (рис.5), де ЕС БАТ$_j$ –

еквівалентна схема *j*-тої біологічно активної точки; Z_{xj} - комплексний опір *j*-тої БАТ; $Z_{1\Sigma}$, $Z_{2\Sigma}$ – комплексні опори шунтувальних тканин; e_{1j} – точка вимірювання *j*-тої БАТ; e_2 – точка, у якій створюється режим електричного відокремлення; e_3 -базова точка; $Кл_{1,...,}Кл_m$ – ключі комутатора для перемикання з режиму виведення точки діагностування на e_3 в режим вимірювання; $y_{1,\ ...,}\ y_m$ – сигнали керування перемиканням ключів комутатора $Кл_{1,...,}Кл_m$; $U_{вх}$ – вхідна напруга; $Кл_{m+1}$ – ключ для перемикання з режиму подачі тестової напруги у режим вимірювання без тестового впливу; $z_{1,}\ z_2$ – сигнали керування перемиканням ключа $Кл_{m+1}$; $R_1,\ldots,R_5$ - резистори; ППП– первинний інваріантний перетворювач, БК – блок керування; МП – масштабний підсилювач; ПВЗ – пристрій вибірки та зберігання із керуванням; АЦП – аналогово-цифровий перетворювач; $g_1,\ldots,g_4$ –сигнали керування ППП, МП, ПВЗ та АЦП; ПК – персональний комп'ютер.

В даній системі інваріантне перетворення забезпечується виведенням усіх крім досліджуваної БАТ на шину нульового рівня шляхом керування перемиканнями ключів $y_{1,\ ...,}\ y_m$ комутатора. В підсистемі передбачена можливість безперервного реєстрування вимірювання показників БАТ за допомогою подачі керуючого сигналу з БК, а також проведення вимірювань як із застосуванням зовнішнього тестового впливу, так і без шляхом керування перемиканням ключа z_k.

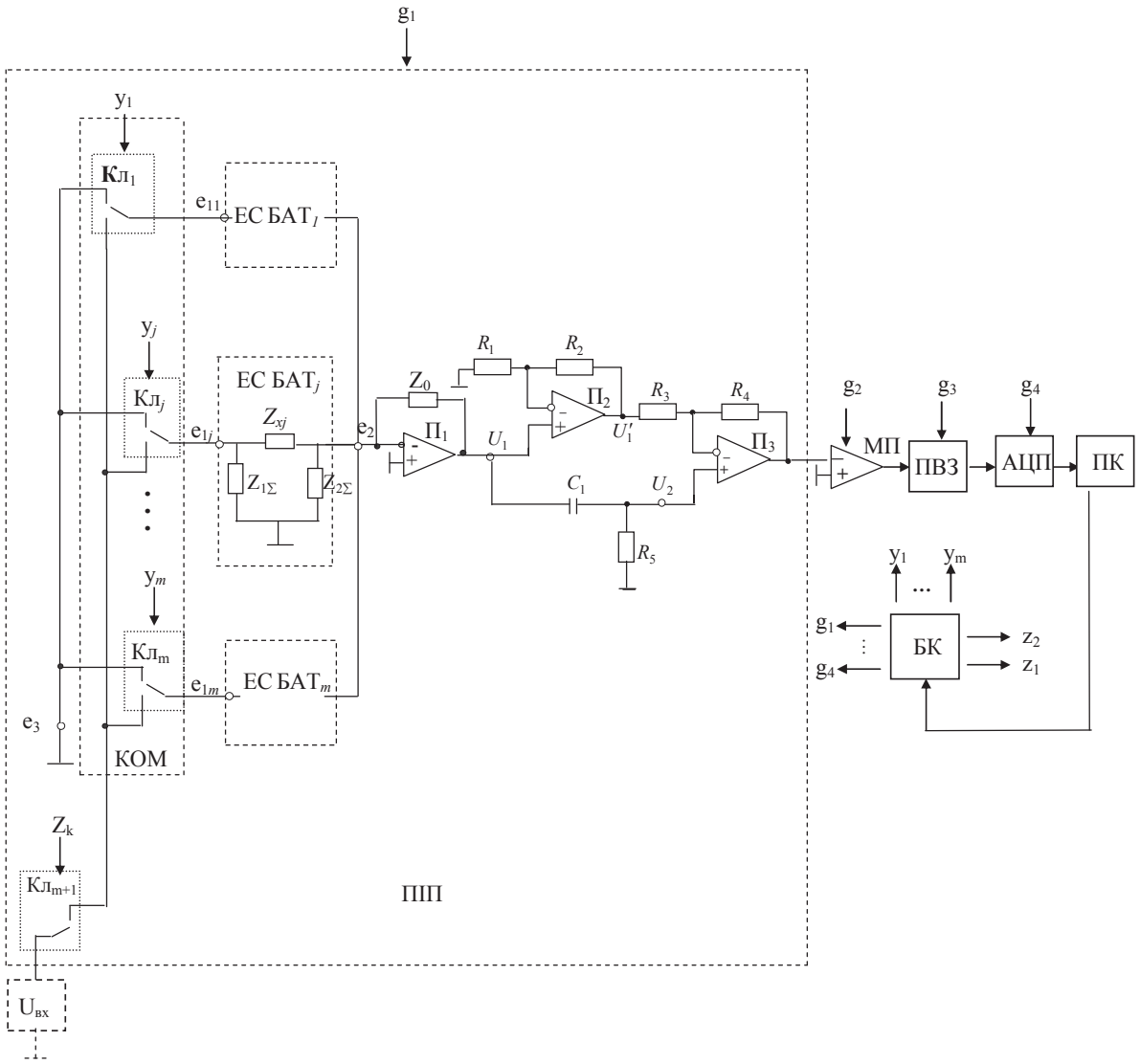


Рисунок 5 – Структурна схема системи діагностування функціонального стану людини на основі інваріантного перетворювача параметрів БАТ

Висновки

1. На основі аналізу впливних величин, що спричиняють похибки вимірювання параметрів БАТ, зокрема таких як завади, а також вплив виділень та забруднення на поверхні шкіри під час вимірювання, наявність шунтувального впливу зі сторони інших органів і тканин людини та ін., було запропоновано і реалізовано метод інваріантного перетворення електричних параметрів БАТ.
2. Для відокремлення корисної складової сигналу у БАТ до складу розробленого первинного інваріантного перетворювача було введено диференційний підсилювач, що дозволило отримати результат вимірювання вільний від складової завад.
3. Побудована структурна схема системи діагностування функціонального стану людини на основі інваріантного перетворювача параметрів БАТ, що дасть змогу підвищити точність результатів вимірювань та достовірність прийняття рішень про стан здоров’я людини.

Список літератури

1. Портнов Ф.Г. Электропунктурная рефлексотерапия. – Рига: Зинатне, 1988. – 352 с.
2. В.Г. Вогралик, М.В. Пунктурная рефлексотерапия: чжень-цзю. – Горький: Волго-Вятское кн. изд-во, 1988 – 335 с.
3. Лучинина Е.В. Изучение диагностической эффективности метода электропунктурной диагностики по Накатани и компьютерного комплекса «Диакомс» при артериальной гипертонии: Автореф. дис. к-та мед. наук: 05.13.01/ Саратовский Государственный медицинский университет. – Москва, 2002.
4. А.А. Малин, Ю.Г. Быстров Многоточечный информационно-измерительный комплекс для снятия показателей с точек акупунктуры // Технические аспекты рефлексотерапии и системы диагностики. Сб. науч. трудов. – Калинин: КГУ, 1984. – С 14 – 15.
5. Dimmick SL, Mustaleski C, Burgiss SG, Welsh T.A. case study of benefits and potential savings in rural home telemedicine. Home Health Nurse. 2000; 18:124-135
6. Азаров О.Д., Снігур А.В. Багатоканальні ІВС опрацювання стрибкоподібних сигналів на базі АЦП із ваговою надлишковістю: Монографія. – Вінниця: УНІВЕРСУМ – Вінниця, 2008. – 138 с.
7. Основи метрології та вимірювальної техніки / [М. Дорожовець, В. Мотало, Б. Стадник, В. Василюк, Р. Борек, А. Ковальчик]; за ред. Б. Стадника. – Львів: Видавництво Національного університету “Львівська політехніка”, 2005. – 655с.
8. Володарський Є.Т., Кухарчук В.В., Поджаренко О.В., Сердюк Г.Б. Метрологічне забезпечення вимірювань і контролю. Навчальний посібник. – Вінниця: Велес, 2001. – 219 с.
9. О.М. Роїк, А.В. Снігур, С.А. Яремко Вибір компонента вимірювального каналу на основі досліджень математичних моделей інформативних сигналів у БАТ // Вісник хмельницького національного університету. – 2009. - №3(132). – С.130-135.
10. Роїк О.М., Власюк А.І. Інваріантні вимірювання параметрів біологічних об’єктів в системах медичної діагностики // Вісник ВПІ, 1999. - №2. – С. 8-11
11. Н.П. Байда, В.И. Месюра, А.М. Роик Самообучающиеся анализаторы производственных дефектов РЭА. – М.: Радио и связь, 1991. – 256 с.: ил.
12. Основи біоактиваційної медицини (відкрита функціонально-енергетична система біологічних об’єктів) / [В.Макац, В.Нагайчук, Д.Макац, Д.Макац]. – Вінниця: Велес, 2001. – 315 с.

Відомості про авторів

Роїк Олександр Митрофанович – д.т.н., професор, завідувач кафедри інформаційного менеджменту, Хмельницьке шосе, 95, м. Вінниця, 21021: тел. 64-53-00.

Вінницький національний технічний університет;

Яремко Світлана Анатоліївна – асистент кафедри економічної кібернетики та інформаційних систем, Соборна, 87, 21100. тел. +380672607014, e-mail: yaremkos@rambler.ru

Вінницький торговельно-економічний інститут Київського національного торговельно-економічного університету.