Національний університет “Львівська політехніка”

Кутельмах Роман Корнелійович

УДК 004.023:519.16

МАТЕМАТИЧНЕ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ДЛЯ РОЗВ’ЯЗУВАННЯ ЗАДАЧІ КОМІВОЯЖЕРА ВЕЛИКИХ РОЗМІРНОСТЕЙ

01.05.03 – Математичне та програмне забезпечення
обчислювальних машин і систем

Автореферат
дисертації на здобуття наукового ступеня
кандидата технічних наук

Львів – 2011

Дисертацією є рукопис.

Робота виконана на кафедрі програмного забезпечення
Національного університету “Львівська політехніка”
Міністерства освіти і науки, молоді та спорту України.

Науковий керівник: доктор технічних наук, професор
Базилевич Роман Петрович,
Національний університет “Львівська політехніка”,
професор кафедри програмного забезпечення.

Науковий консультант: професор **Ремі Дюпа (Rémy Dupas)**,
Університет м. Бордо, Франція
(Université de Bordeaux 1 - LAPS, France).

Офіційні опоненти: доктор технічних наук, професор
Гребеннік Ігор Валерійович,
Харківський національний університет
радіоелектроніки,
професор кафедри системотехніки;

доктор технічних наук, доцент
Романишин Юрій Михайлович,
Національний університет “Львівська політехніка”,
професор кафедри електронних засобів
інформаційно-комп’ютерних технологій.

Захист відбудеться 10 червня 2011 року о $16^{\underline{00}}$ годині
на засіданні спеціалізованої вченої ради Д 35.052.05
в Національному університеті “Львівська політехніка”
за адресою: 79013, м. Львів, вул. С. Бандери, 12.

З дисертацією можна ознайомитися у науково-технічній бібліотеці
Національного університету “Львівська політехніка”
за адресою: 79013, м. Львів, вул. Професорська, 1.

Автореферат розісланий _____ 2011 р.

Вчений секретар спеціалізованої вченої ради
доктор технічних наук, професор Р.А. Бунь

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. Задача комівояжера (ЗК) має широке прикладне застосування в транспортних системах, автоматизованому проектуванні, тестуванні та виготовленні інтегрованих схем, виробництві друкованих плат, лазерній нарізці пластмас і металів, дослідженні структури білка, технологіях вишивання, зварювання та малювання неперервною лінією, рентгенівській кристалографії та інших галузях. Важливою особливістю цих задач є їх велика розмірність, яка для багатьох із них становить більше мільйона точок. ЗК відноситься до задач класу *NP*, оскільки має факторіальну обчислювальну складність. Це не дає змоги отримувати точні розв'язки для задач великих розмірностей за прийнятний час.

Активно задача почала досліджуватися у 50-х роках минулого століття. G. Dantzig, R. Fulkerson та S. Johnson у 1954 році сформулювали ЗК у вигляді задачі дискретної оптимізації та запропонували метод, який забезпечив отримання точного розв'язку. Вони розв'язали задачу розмірністю 49 точок та довели, що не існує коротшого маршруту. У 1970-х роках R. Karp обгрунтував *NP*-повноту задачі, а S. Lin та B. Kernighan розробили один з найбільш ефективних евристичних методів розв'язування ЗК.

Впродовж останніх років K. Helsgaun удосконалив класичну версію методу Lin-Kernighan, запропонувавши метод LKH (Lin-Kernighan-Helsgaun), який вважається найкращим на даний час евристичним методом розв'язування задачі. На основі методу гілок і меж D. Applegate, R. Bixby, V. Chvátal та W. Cook розробили прикладну програмну систему "Concorde", що вважається найкращою на даний час для знаходження точного розв'язку ЗК.

Зазначені методи мають обчислювальну складність не нижче $O(N^2)$, що робить їх малопридатними для задач великих розмірностей. Тому виникла потреба в розробленні методів та прикладної програмної системи для розв'язування таких задач. З метою організації ефективних обчислень, придатних для задач великих розмірностей, у дисертації розроблено набір спеціалізованих декомпозиційних методів, адаптованих до конкретних типів структур вхідних даних. Розроблено прикладну програмну систему, яка дає змогу розв'язувати ЗК великих розмірностей, досліджувати та інтегрувати нові методи.

Зв'язок роботи з науковими програмами, планами, темами. Робота виконувалась відповідно до напрямків наукової діяльності кафедри програмного забезпечення Національного університету "Львівська політехніка", у тому числі за двома спільними українсько-французькими науково-дослідними проектами "Дніпро" (договір М/78-2005 від 04.12.2005 "Застосування еволюційних та декомпозиційних алгоритмів для розв'язування динамічної транспортної задачі", № держреєстрації 0105U007631 та договір М/120-2007 від 27.03.2007 "Транспортні та мережні задачі великої розмірності зі специфічними властивостями: кооперація кластеризаційних та еволюційних алгоритмів для оптимізації розв'язків", № держреєстрації 0107U005115), а також НДР "Важковирішувані комбінаторні задачі високої та надвисокої розмірності"

(№ держреєстрації 0108U0050037), в яких автор був відповідальним виконавцем. У перелічених науково-дослідних проектах автору належать: розроблення методів кластеризації вхідних даних для ЗК великих розмірностей, дослідження відомих методів розв'язування ЗК, розроблення декомпозиційних методів знаходження початкового розв'язку ЗК на основі макромоделювання та нарощення часткового розв'язку, розроблення методів оптимізації початкового розв'язку ЗК великих розмірностей, розроблення прикладної програмної системи для розв'язування ЗК великих розмірностей та експериментальні дослідження запропонованих методів.

Мета і завдання дослідження. Метою дисертації є розроблення математичного та програмного забезпечення для розв'язування ЗК великих розмірностей. Для реалізації цієї мети необхідно вирішити такі завдання:

- класифікувати типи ЗК та дослідити відомі методи її розв'язування;
- обґрунтувати необхідність розроблення ефективних декомпозиційних методів з лінійною чи лінійно-логарифмічною обчислювальною складністю, що забезпечують високу точність розв'язків;
- розробити методи знаходження початкового розв'язку та його оптимізації для ЗК великих розмірностей з адаптацією до структури вхідних даних;
- розробити прикладну програмну систему для розв'язування ЗК великих розмірностей;
- оцінити ефективність розроблених методів з точки зору якості та часу розв'язання задачі на основі проведених експериментальних досліджень та порівняльного аналізу отриманих результатів.

Об'єкт дослідження – процес обчислення розв'язку задачі комівояжера великих розмірностей.

Предмет дослідження – методи, алгоритми та програмне забезпечення для розв'язування задачі комівояжера великих розмірностей.

Методи дослідження. При виконанні дисертаційної роботи для знаходження початкового розв'язку ЗК та його оптимізації використано положення теорії множин та евристичні методи розв'язування складних комбінаторних задач, для розбиття вхідної множини точок на підмножини обмеженої розмірності – методи кластеризації та тріангуляції, при розробленні програмного забезпечення використано положення теорії алгоритмів, методи об'єктно-орієнтованого програмування.

Наукова новизна одержаних результатів. У дисертаційній роботі вирішено актуальну наукову задачу – розроблено ефективні методи розв'язування задачі комівояжера великих розмірностей:

- удосконалено декомпозиційний метод розв'язування ЗК шляхом макромоделювання на основі кластеризації вхідних даних, який включає етапи макромоделювання, знаходження початкового розв'язку та його оптимізації, який забезпечує зменшення довжини маршруту на 0,5-2% у порівнянні з відомим методом 2-opt;
- вперше розроблено декомпозиційний метод знаходження початкового розв'язку ЗК на основі нарощення часткового розв'язку, який включає етапи розбиття вхідної множини точок на підмножини невеликої

розмірності, знаходження часткового розв'язку для деякої початкової множини та його нарощення, забезпечуючи відхилення від розв'язку, отриманого за допомогою методу LKH, не більше 0,5% зі зменшенням обчислювальної складності з $O(N^{2,2})$ до $O(N \log N)$;

- вперше розроблено метод оптимізації початкового розв'язку за маршрутом з лінійною обчислювальною складністю, який полягає в послідовній чи послідовно-паралельній мінімізації довжини виділених ділянок початкового маршруту та забезпечує покращання в середньому на 0,22% початкового розв'язку, довжина якого на 0,2-2% не перевищує довжину оптимального;
- набув подальшого розвитку метод геометричної оптимізації початкового розв'язку, який полягає в мінімізації довжини ділянок початкового маршруту у виділених геометричних областях та забезпечує покращання в середньому на 0,4% початкового розв'язку задачі;
- вперше розроблено метод оптимізації початкового розв'язку з лінійно-логарифмічною обчислювальною складністю, який полягає в мінімізації довжини маршруту послідовно в суміжних областях на тріангуляції Делоне та забезпечує покращання в середньому на 0,27% початкового розв'язку задачі.

Практичне значення одержаних результатів. Розроблене математичне та алгоритмічне забезпечення дало змогу створити прикладну програмну систему "VRP Modeler" для розв'язування та експериментального дослідження ЗК великих розмірностей, в якій з метою організації ефективних обчислень інтегровано в єдиному комплексі відомі та розроблені декомпозиційні методи. Система має такі особливості та переваги:

- отримано виграш в часі розв'язання задач розмірністю 1000-100000 точок в 2,7-164 раз відповідно у порівнянні з методом 2-opt, забезпечивши при цьому зменшення довжини маршруту на 0,5-2%;
- отримано виграш в часі розв'язання задач розмірністю 1000-85900 точок із відомих бібліотек в 19-402618 раз відповідно у порівнянні з точним методом, реалізованим у ППС "Concorde". Для ряду досліджених задач знайдено оптимальні розв'язки, а відхилення від оптимального розв'язку для інших задач не перевищили 0,58%;
- отримано виграш в часі для досліджених задач розмірністю 3000-100000 точок у 2,16 – 15,7 разів відповідно, а втрати якості не перевищили 0,36% у порівнянні з найкращим на сьогодні евристичним методом LKH;
- отримано розв'язки для відомих тестових задач розмірністю 10^6 та 10^7 точок, забезпечивши при цьому відхилення у якості розв'язку не вище 0,11% у порівнянні з найкращим, знайденим на даний час;
- впроваджено розроблені ефективні декомпозиційні методи знаходження початкових розв'язків ЗК та їх оптимізації, які виникають у широкому колі прикладних задач: виготовлення друкованих плат, проектування схем мікроелектроніки, НВІС, маршрутизації автотранспорту, технології малювання неперервною лінією та ін.

Програмне забезпечення розроблене автором самостійно у середовищі Microsoft Visual Studio 2008 на мові C++ із використанням методів об'єктно-орієнтованого програмування. Результати дисертаційної роботи впроваджено в навчальний процес кафедри програмного забезпечення Національного університету "Львівська політехніка", ПП "АР-ТРАНС" (м. Трускавець), ТзОВ "СЕА Електронікс" (м. Київ), а також трьох НДР, в яких автор був відповідальним виконавцем.

Особистий внесок здобувача. Всі положення й результати дисертаційної роботи отримані автором самостійно. У роботах, опублікованих у співавторстві, здобувачеві належать: розроблення та дослідження методів оптимізації розв'язків ЗК [1, 6, 8], розроблення методів знаходження початкового розв'язку задачі [2], дослідження ефективності відомих методів розв'язування ЗК [3], розроблення та дослідження декомпозиційних методів розв'язування ЗК великих розмірностей [4, 5, 9 – 13], розроблення методів кластеризації вхідної множини точок для ЗК [7], розроблення методів та засобів для розпаралелювання задачі, розроблення архітектури програмного забезпечення для паралельного розв'язування ЗК великих розмірностей [14 – 16].

Апробація результатів дисертації. Наукові й практичні результати роботи доповідалися та обговорювалися на: міжнародних наукових семінарах "Сучасні проблеми інформатики в управлінні, економіці, освіті" (Шацьк, 2006, 2007, 2008, 2009), міжнародних конференціях "Computer Science and Information Technologies" (Львів, 2006, 2007, 2010), міжнародній конференції "Комп'ютерні науки та інженерія – 2006" (Львів, 2006), міжнародній конференції "Francoro/Roadef 2007", (Гренобль, Франція, 2007), науковому семінарі "Decomposition Approach for Large Scale TSP", (Бордо, Франція, 2007), "3rd, 4th Indian International Conferences on Artificial Intelligence" (Пуне, Тумкур, Індія, 2007, 2009), "The 2008 International Conference on Theoretical and Mathematical Foundations of Computer Science" (Орландо, США, 2008), міжнародній конференції "The Experience of Designing and Application of CAD Systems in Microelectronics" (Поляна, 2009), Міжнародній науково-технічній конференції "Dependable Systems, Services and Technologies" (Кіровоград, 2009), Міжнародній науково-практичній конференції "Інформаційні технології та комп'ютерна інженерія" (Вінниця, 2010), щорічних наукових семінарах кафедри програмного забезпечення Національного університету "Львівська політехніка" (Львів, 2006-2010).

Публікації. Результати дисертації опубліковано у 16 наукових працях, у тому числі 8 у фахових наукових виданнях (серед яких 7 статей задовільняють вимогу щодо однієї публікації у виданні), та 8 у матеріалах і тезах конференцій.

Структура та обсяг дисертації. Дисертаційна робота складається зі вступу, 4 розділів, висновків по роботі, трьох додатків (21 с.) та списку використаних джерел із 237 найменувань (25 с.). Загальний обсяг дисертації складає 181 сторінку, в тому числі 123 сторінки основного тексту, включаючи 82 рисунки, 22 таблиці.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** сформульовано мету роботи, обґрунтовано актуальність досліджень. Дано характеристику наукової новизни та практичної цінності отриманих результатів. Показано зв'язок роботи з науковими програмами, планами, темами.

У **першому розділі** проведено вивчення літературних джерел за темою дисертації. В Україні значний внесок в розробку методів та алгоритмів для розв'язування задач комбінаторної оптимізації зробили вчені І.В. Сергієнко, В.С. Михалевич, Ю.Г. Стоян, М.І. Гіль, О.А. Павлов, А.В. Панішев та ін. Задачу також досліджували R. Bellman, C. Papadimitriou, N. Christofides, M. Padberg, G. Rinaldi, F. Glover, D. Johnson, G. Reinelt, J. Bentley, І. Сігал, Е. Гімаді, Г. Гутін, Ю. Фінкельштейн та інші.

Класифіковано відомі методи розв'язування задачі, описано приклади застосування ЗК великих розмірностей. Проаналізовано сучасні дослідження для розв'язування ЗК великих розмірностей, виконаних проф. K. Helsgaun (Данія), проф. Y. Nagata (Японія), проф. W. Cook (США), проф. P. Molitor (Німеччина), проф. D. Johnson (США), проф. C. Rego (США) та іншими.

Згідно із класичним формулюванням задачі комівояжера (Traveling Salesman Problem, TSP) заданими вважають:

1) множину точок P, описаних координатами:

$P=\{p_1, p_2, \ldots, p_N\}$, $p_i=(x_i, y_i)$, $i \in \{1, 2, \ldots, N\}$;

2) метрику $dist$: P x $P \to R$ на множині P, наприклад:

$dist_E(p_i, p_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$, $i,j \in \{1, 2, \ldots, N\}$,

або $dist_M(p_i, p_j) = |x_i - x_j| + |y_i - y_j|$, $i,j \in \{1, 2, \ldots, N\}$.

Задача полягає в мінімізації довжини замкнутого маршруту M, який відвідує усі точки з P:

$$len\ M \to \min$$

$$\text{з обмеженням } \forall i\colon p_i \in M, |M| = N,$$

де $len\ M = \sum_{i=1}^{N-1} dist(m_i, m_{i+1}) + dist(m_N, m_1)$ є функцією визначення довжини маршруту $M = \langle m_1, m_2, \ldots, m_N \rangle$, $m_i \in P$.

Введемо функцію оцінки якості розв'язку задачі (маршруту M) $f_{quality}$:

$$f_{quality}(M, M^*) = \frac{len\ M - len\ M^*}{len\ M^*},$$

де M^* – оптимальний розв'язок задачі. Якщо M^* – маршрут мінімальної довжини, то $len\ M \geq len\ M^*$, а також $f_{quality}(M, M^*) \geq 0$. При $f_{quality}(M, M^*) = 0$ маршрут M (розв'язок задачі) вважається *оптимальним*, інакше – *наближеним*. Якщо M^* є невідомим, то порівняння здійснюється з найкращим на даний час відомим розв'язком. В такому разі функція оцінки якості розв'язку може приймати від'ємне значення, що свідчить про отримання кращого розв'язку, ніж відомий.

Відомі точні методи, зокрема метод гілок і меж, реалізований у ППС “Concorde”, можуть бути використані для задач невеликих розмірностей. На ПК з процесором Celeron 2,5 ГГц час знаходження оптимального розв’язку задачі розмірністю 100 точок склав 0,54 с, 500 точок – 205 с, 2000 точок – 35872 с. Відомі евристичні методи мають обчислювальну складність $O(N^2)$ і вище, що є також неефективним для задач великих та надвеликих розмірностей. Найкращий на даний час евристичний метод LKH має обчислювальну складність, близьку до $O(N^{2,2})$. На ПК з процесором Celeron 2,5 ГГц час знаходження розв’язку задачі розмірністю 100 точок методом LKH склав 0,07 с, 500 – 6,99 с, 2000 – 81 с, 10000 – 16646 с, 20000 точок – 124563 с.

У зв’язку з великими часовими витратами відомих методів для розв’язування ЗК великих розмірностей необхідним є розроблення ефективного математичного та програмного забезпечення з близькою до лінійної чи лінійно-логарифмічної обчислювальною складністю, яке забезпечує високу якість отриманих розв’язків. Сформульована задача дисертаційного дослідження — розроблення математичного та програмного забезпечення для розв’язування ЗК великих розмірностей.

У **другому розділі** запропоновано декомпозиційні методи розв’язування ЗК великих розмірностей. Розвинуто відомі та розроблено нові декомпозиційні методи, в яких задача розв’язується за декілька етапів: розбиття всієї множини точок на підмножини обмеженої розмірності (≈ 500-2000 точок), для яких можна отримати високоякісні часткові розв’язки з невеликими часовими витратами; зшивання часткових розв’язків у початковий розв’язок та наступне його покращення розробленими оптимізаційними методами. Обчислювальна складність методів є близькою до лінійно-логарифмічної, що обґрунтовано теоретично та підтверджено експериментально.

Експериментальні дослідження розроблених методів на задачах із відомих бібліотек тестів ППС “Concorde”, конкурсу DIMACS TSP Challenge та TSPLIB проводились на ПК з процесором Intel Celeron 2,5 ГГц та об’ємом оперативної пам’яті 1 Гб.

Розроблено декомпозиційний метод знаходження розв’язку ЗК з *кластерним* розподілом точок**,** який складається з таких етапів:

1) макромоделювання – формування множини кластерів точок та макромаршруту M_{macro};
2) знаходження початкового маршруту M_0;
3) оптимізація початкового маршруту M_0.

На етапі макромоделювання кожен кластер $C_i \in C$ описується точковою моделлю – однією точкою $p_{cond}(C_i)$ з координатами (x_i, y_i):

$$p_{cond}(C_i) = (x_i, y_i),\ i \in \{1, 2, \ldots, K\}.$$

Задача з N точками множини P замінюється задачею з K точками, що утворюють множину P':

$$P' = \{p_{cond}(C_1), p_{cond}(C_2), \ldots, p_{cond}(C_K)\}.$$

Оскільки $K << N$, це суттєво спрощує задачу. За допомогою вибраного базового (ефективного) методу розв'язується ЗК для множини з K точок. Отримується макромаршрут M_{macro}:

$$M_{macro} = <m_1, m_2, \ldots, m_K>, m_i \in P', \text{ для } i \in \{1, 2, \ldots, K\}.$$

Множина кластерів C та макромаршрут M_{macro} є *макромоделлю* задачі (рис. 1).

Рис. 1. Вхідна множина точок та макромодель:
а) вхідна множина точок; б) виділені кластери точок; в) макромаршрут

Початковим розв'язком вважається маршрут M_0, що створюється зшиванням розв'язків у всіх кластерах $C_i \in C$ у порядку їх відвідування в *макромаршруті* M_{macro}:

$$M_0 = M(m_1) * M_{1,2} * M(m_2) * \ldots * M(m_K),$$

де $M(m_i)$ – розв'язок ЗК у кластері C_i, а $M_{i,i+1}$ – ребро, що з'єднує кластери C_i та C_{i+1}.

Наступним етапом є оптимізація маршруту з використанням макромоделі. Для оптимізації маршруту M_0 обирається деяка його ділянка (*елементарна область сканування*), для якої знаходиться окремий розв'язок задачі, і, в разі зменшення довжини, дана ділянка замінюється на нову. *Елементарною областю сканування* (ЕОС) вважається ділянка S_{area} маршруту M_0, що складається з точок деякого кластера C_i та певних точок попереднього кластера C_{i-1} і (або) наступного C_{i+1} у порядку їх відвідування в макромаршруті (рис. 2а). Процес повторюється для інших ЕОС до тих пір, поки не буде переглянуто всіх точок маршруту M_0.

Запропоновано та досліджено низку стратегій оптимізації (рис. 2б), що відрізняються розміром елементарної області сканування: ЕОС1 (алгоритми ЕОС1-1, ЕОС1-2) – скануються один кластер та найближча точка наступного або попереднього кластера; ЕОС2 (алгоритми ЕОС2-1, ЕОС2-2) – скануються один кластер та найближчі точки наступного і попереднього кластерів; ЕОС3 (алгоритми ЕОС3-1, ЕОС3-2) – скануються два суміжні кластери; ЕОС4 (алгоритми ЕОС4-1, ЕОС4-2, ЕОС4-3) – скануються два кластери та найближчі точки наступного і попереднього кластерів.

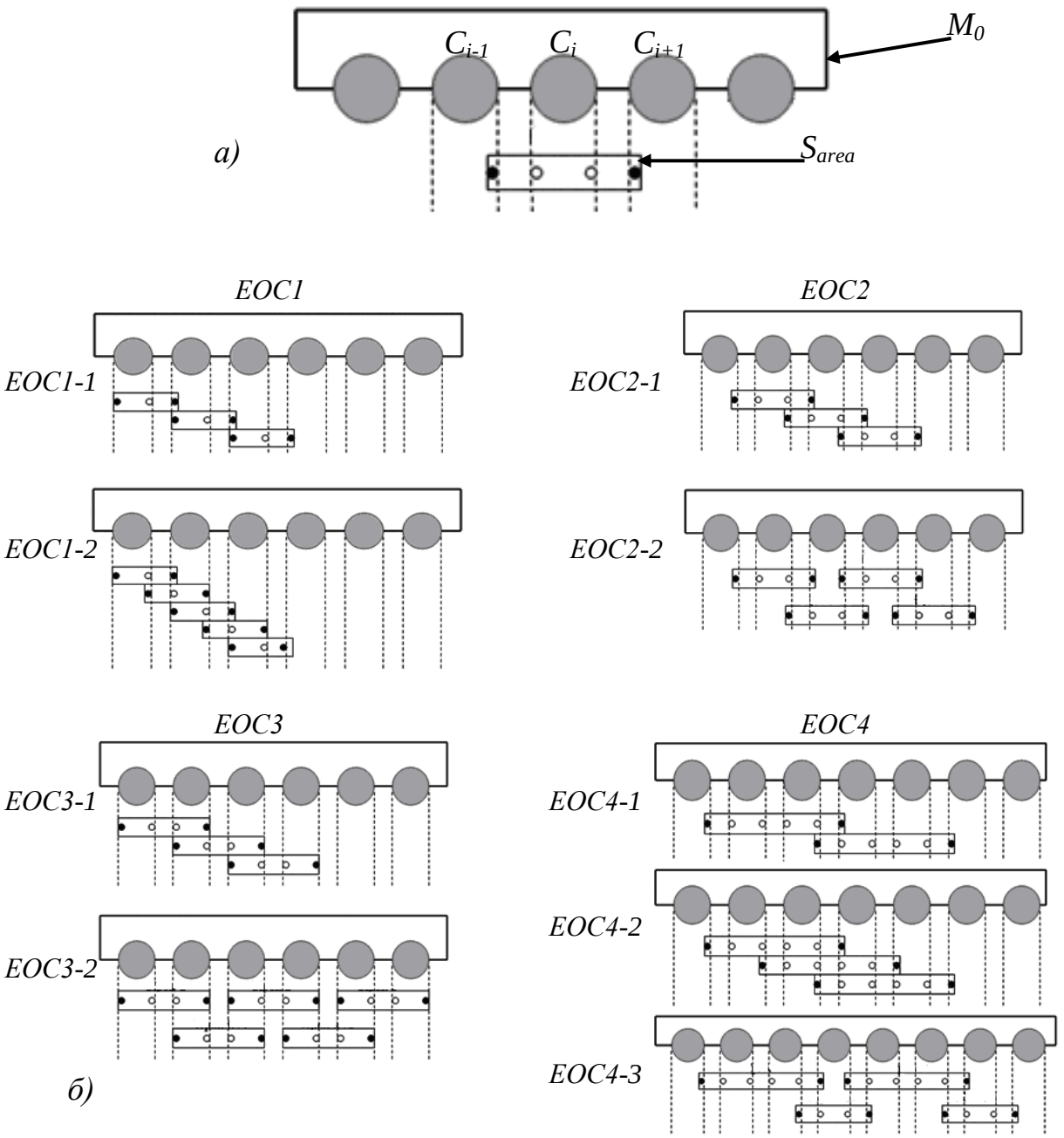

Рис. 2. Оптимізація маршруту з використанням макромоделі: а) елементарна область сканування; б) стратегії оптимізації розв’язку та варіанти їх алгоритмічної реалізації

Використання методу забезпечило зменшення довжини маршруту на 0,5-2% у порівнянні з методом *2-opt* для задачі розмірністю 10000 точок за час, у 28 разів, а для задачі розмірністю 100000 точок – у 164 рази менший.

Розроблено декомпозиційний метод знаходження початкового розв’язку ЗК шляхом нарощення часткового розв’язку, який передбачає етапи:

1) розбиття вхідної множини P, утворюючи множину $U = \{U_1, U_2, \ldots, U_K\}$ з K підмножин: $U_1 \cup U_2 \cup \ldots \cup U_K = P$, $U_i \cap U_j = 0$, $|U_i| < D_{max}$, для $i,j \in \{1, 2, \ldots, K\}$, де D_{max} є параметром методу;
2) вибір початкової підмножини U_i та знаходження розв’язку ЗК у ній (маршруту M_0);

3) послідовне чи послідовно-паралельне зшивання розв'язків у всіх підмножинах (нарощення розв'язку M_0).

Для зшивання розв'язків у підмножинах U_i та U_j формується множина U_j' (рис. 3а), яка міститиме усі точки U_j та частину точок підмножини U_i, тобто утворює з нею перетин $\Delta U_{ij} = U_i \cap U_j'$. Число точок перетину є параметром, який впливає на якість та час розв'язування задачі. Для множини U_j' визначаються *граничні вхідні* та *граничні вихідні* точки (рис. 3б). За допомогою вибраного відомого методу знаходиться розв'язок ЗК для точок, що належать множині U_j' (рис. 3в).

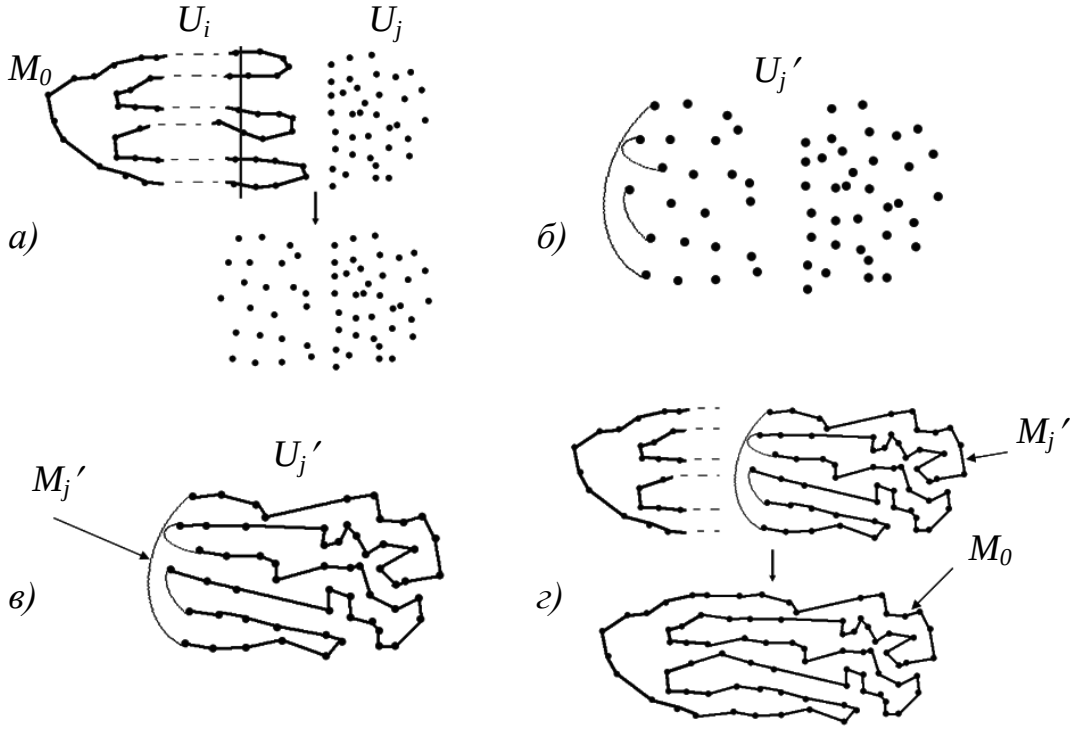

Рис. 3. Основні етапи методу:
а) формування множини U_j'; б) визначення граничних вхідних та граничних вихідних точок; в) розв'язок ЗК для точок множини U_j'; г) зшивання розв'язків у підмножинах U_i та U_j

Одержується маршрут M_j', який зшивається з маршрутом M_0. Таким чином маршрут M_0 нарощується (рис. 3г). Операція зшивання розв'язків у підмножинах повторюється до приєднання усіх підмножин. Отриманий маршрут M_0 охоплює всі точки множини P та вважається *початковим розв'язком* задачі. Метод забезпечив отримання початкового розв'язку для тестових задач розмірністю 1000-10000 точок із бібліотеки DIMACS TSP Challenge за час 7,5 – 78,2 с відповідно з відхиленням від оптимального на 0,021%–0,09%. Час знаходження оптимального розв'язку з використанням ППС "Concorde" для задачі розмірністю 10000 точок становить близько 10 днів на цьому ж процесорі.

Можливим є розпаралелення процесу розв'язування задачі. Відмінною особливістю методу від попереднього є те, що він не передбачає використання *макромаршруту*. Існує низка алгоритмів вибору початкової підмножини та наступних підмножин для нарощення розв'язку, наприклад, зшивання підмножин

зліва направо, або почергово – зліва направо і справа наліво і т.д., спірально та ін. Процес знаходження початкового розв'язку за допомогою спірального алгоритму зшивання показано на рис. 4.

Рис. 4. Процес знаходження початкового розв'язку задачі
за допомогою спірального алгоритму зшивання

Метод забезпечив покращання якості для тестової задачі "Mona Lisa" розмірністю 100000 точок на 0,015% у порівнянні з найкращим на сьогодні евристичним методом LKH, зі зменшенням часу в 4,76 раз. Для тестових задач із бібліотеки TSPLIB розмірністю 1173–33810 точок метод забезпечив знаходження початкових розв'язків з відхиленням від оптимального не більше 1,73% за 18-15661 с відповідно, забезпечивши при цьому виграш у часі в 1,25-74 раз. Для задачі розмірністю 2319 точок було знайдено оптимальний розв'язок за 261 с.

Розроблений декомпозиційний метод знаходження розв'язку ЗК з кластерним розподілом точок та метод знаходження початкового розв'язку ЗК шляхом нарощення часткового розв'язку доцільно застосовувати в задачах великих розмірностей з метою отримання якісних розв'язків зі значним зменшенням часових витрат. Для забезпечення бажаного компромісу між якістю розв'язку та часом обчислень для розроблених методів введені такі параметри: розмір *області перетину*, *потужність підмножин (кластерів) та базовий алгоритм*. На основі проведених експериментів визначена функційна залежність, яка дає змогу вибирати необхідний компроміс між цими двома критеріями при розв'язуванні конкретних задач з врахуванням заданих вимог.

У **третьому розділі** запропоновано методи оптимізації розв'язків. Для оптимізації початкового розв'язку (маршруту M_0), обирається деяка його ділянка (*елементарна область оптимізації*), для якої розв'язується задача мінімізації і, в разі зменшення її довжини, така ділянка замінюється на нову. Одержується маршрут M_1 такий, що $len\ M_1 < len\ M_0$. Процес повторюється для всіх *елементарних областей оптимізації* до завершення перегляду всіх точок маршруту M_0. Одержується послідовність маршрутів $M_0, M_1, M_2, \ldots, M_T$, де $len\ M_{i+1} < len\ M_i$ для $i \in \{0, 1, \ldots, T\text{-}1\}$, а T – кількість замін ділянок маршруту M_0 на коротші.

Розроблено методи:

- оптимізації за маршрутом;

- геометричної оптимізації;
- оптимізації на тріангуляції Делоне.

Обчислювальна складність методів є близькою до лінійної, оскільки визначається числом елементарних областей оптимізації, яке лінійно зростає зі збільшенням розмірності задачі.

Метод *оптимізації за маршрутом* (рис. 5) полягає в мінімізації довжини ділянок маршруту M_0, які аналізуються послідовно вздовж нього.

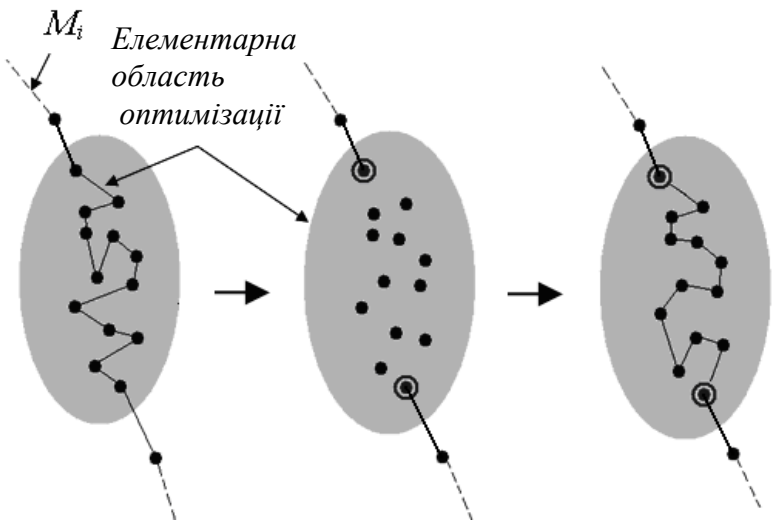

Рис. 5. Схематичне зображення етапу заміни ділянки маршруту на коротшу згідно з методом оптимізації за маршрутом

Метод *геометричної оптимізації* (рис. 6) полягає в мінімізації довжини всіх ділянок маршруту M_0, які входять у вибрану геометричну область. Розміри і форма області оптимізації та перетину двох суміжних ділянок, а також послідовність їх вибору є параметрами методу.

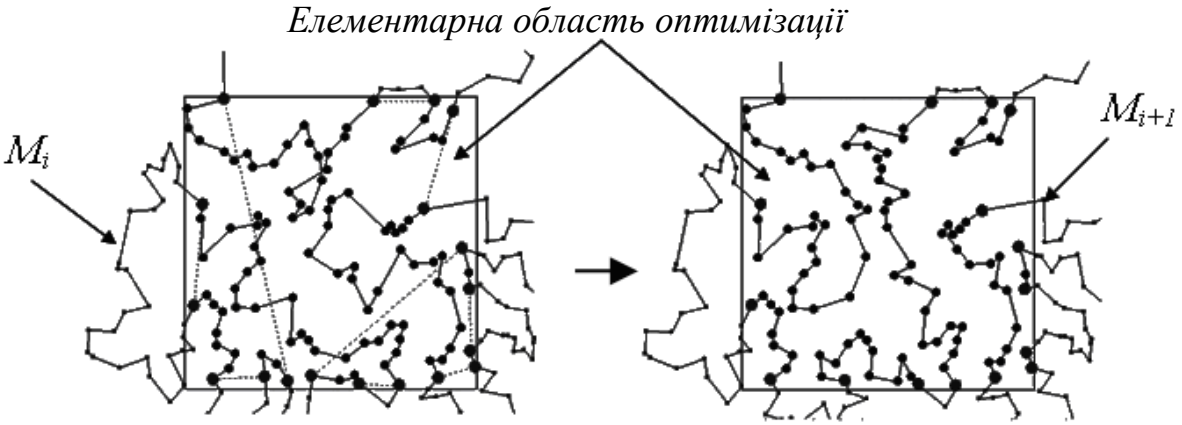

Рис. 6. Схематичне зображення етапу заміни ділянок маршруту на коротші за методом геометричної оптимізації

Метод *оптимізації на тріангуляції Делоне* передбачає зменшення довжини маршруту M_0 послідовно в суміжних областях на такій тріангуляції (рис. 7).

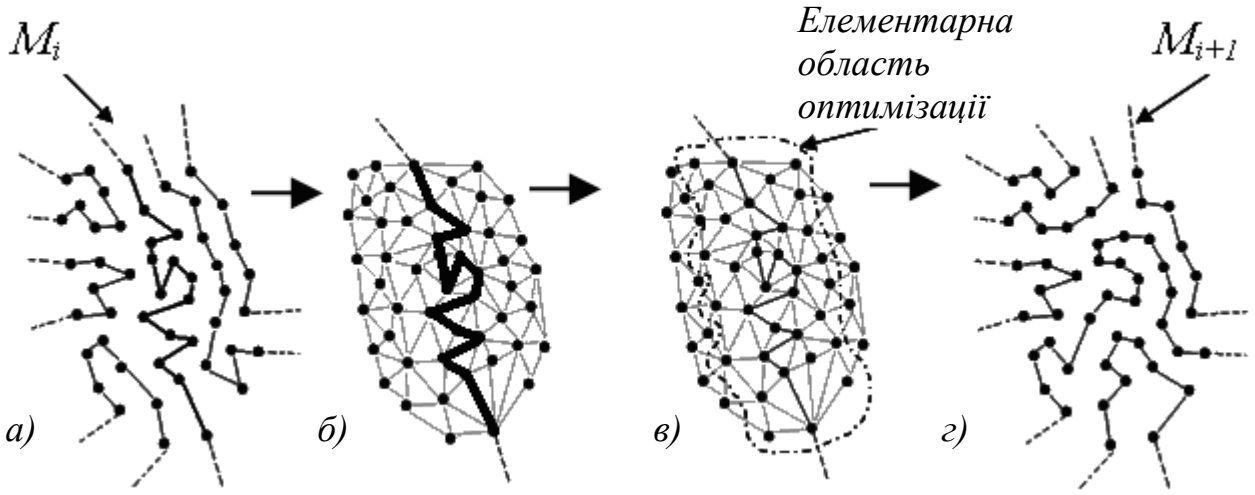

Рис. 7. Етапи заміни ділянок маршруту на коротші за методом оптимізації на тріангуляції Делоне: а, б) вибір ділянки маршруту M_0; в) визначення елементарної області оптимізації та розв'язування ЗК в ній; г) результат заміни

Для досліджених тестових задач із бібліотеки TSPLIB розмірністю 1173–85900 точок розроблені методи оптимізації забезпечили знаходження розв'язків із відхиленням в межах 0,012%-0,97% від оптимального. Час оптимізації розв'язку для тестової задачі розмірністю 1173 точок становив 9,35 с, задачі розмірністю 85900 точок – близько 34 хвилин. Для порівняння час знаходження оптимального розв'язку задачі розмірністю 85900 точок становить близько 136 років процесорного часу згідно з опублікованими матеріалами авторів ППС "Concorde".

Запропоновані методи оптимізації забезпечують можливість вибору належного компромісу між якістю та часом розв'язання. Для цього введено параметри: розміри елементарної області оптимізації та області перетину, базовий метод – вибраний метод, що застосовується для розв'язування ЗК в елементарних областях оптимізації.

У **четвертому розділі** описано розроблену прикладну програмну систему "VRP Modeler" та результати дослідження при комплексному застосуванні запропонованих методів на тестових задачах із відомих бібліотек – ППС "Concorde", конкурсу DIMACS TSP Challenge та TSPLIB.

У ППС "VRP Modeler" реалізовано методи розв'язування ЗК великих розмірностей, запропоновані у розділах 2 та 3. Система працює в ОС Windows. Програмне забезпечення системи розроблене автором самостійно в середовищі Microsoft Visual Studio 2008 на мові C++ із використанням методів об'єктно-орієнтованого програмування. Концептуально система складається з модуля керування та функціональних модулів. Модуль керування відповідає за базові функції для маніпуляції з даними, бібліотеками методів, модулями візуалізації вхідних та вихідних даних та ін. Функціональні модулі реалізують методи, візуалізатори, обробники даних.

У систему "VRP Modeler" входить менеджер проектів, що дозволяє зручно керувати всіма даними у вигляді багатовіконного інтерфейсу. За його допомогою можна завантажувати проекти з диску, створювати нові, закривати їх а також маніпулювати даними всередині проекту. У системі "VRP Modeler" дані

представляються за допомогою модулів візуалізації, тобто алгоритмів, що відмальовують точки, лінії та полігони у відповідності зі своїми налаштуваннями.

Усі основні компоненти системи “VRP Modeler” оформлено як класи, які можна розширювати за функціональністю при допомозі технологій об’єктно-орієнтованого програмування. Є можливість розроблення нових модулів, методів та їх інтеграції в систему. Для цього передбачений стандартизований інтерфейс класів та функцій.

Основними компонентами структури прикладної програмної системи “VRP Modeler” є: модуль керування, модулі візуального перегляду вхідних, вихідних даних та проміжних результатів, модуль генераторів вхідних даних, модуль опрацювання вхідних даних, бібліотека методів розв’язування задачі (рис. 8).

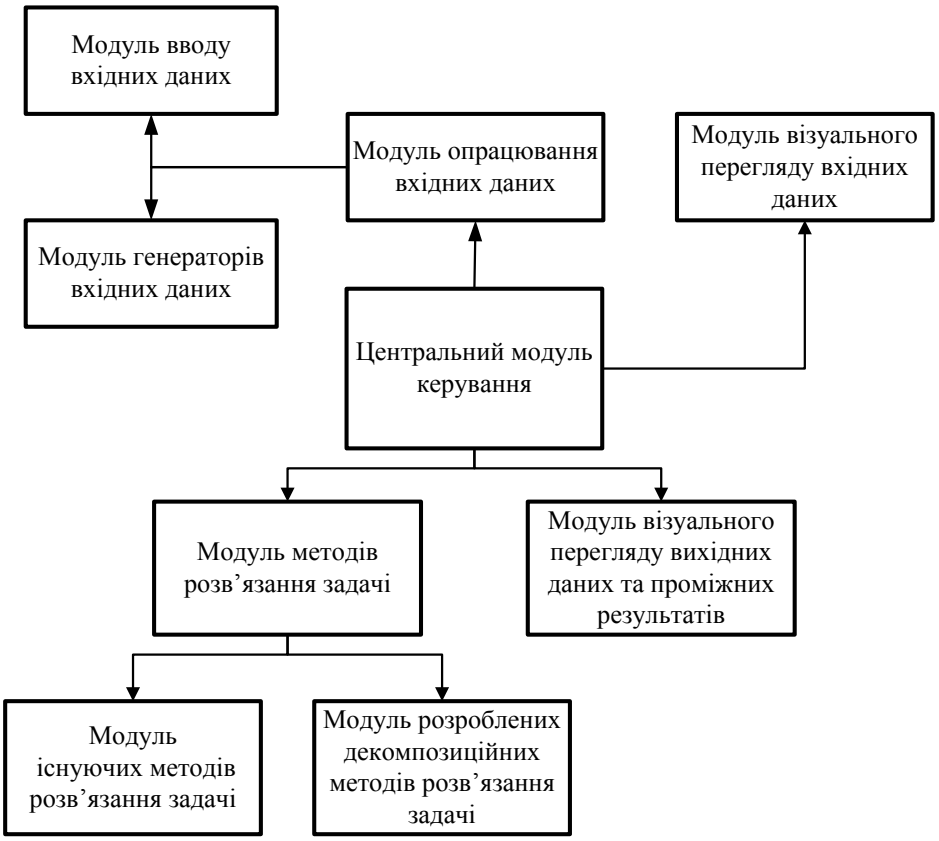

Рис. 8. Структура ППС “VRP Modeler”

З метою організації ефективних обчислень в єдиному комплексі інтегровано розроблені декомпозиційні з відомими методами (2-opt, 3-opt, Lin-Kernighan, LKH та точним методом ППС “Concorde”). На рис. 9 представлено структуру модуля методів розв’язування ЗК програмної системи “VRP Modeler”.

Графічний інтерфейс реалізовано з допомогою бібліотеки класів MFC. Вигляд інтерфейсу користувача ППС “VRP Modeler” показано на рис. 10. Приклади деяких досліджених тестових задач представлено на рис. 11.

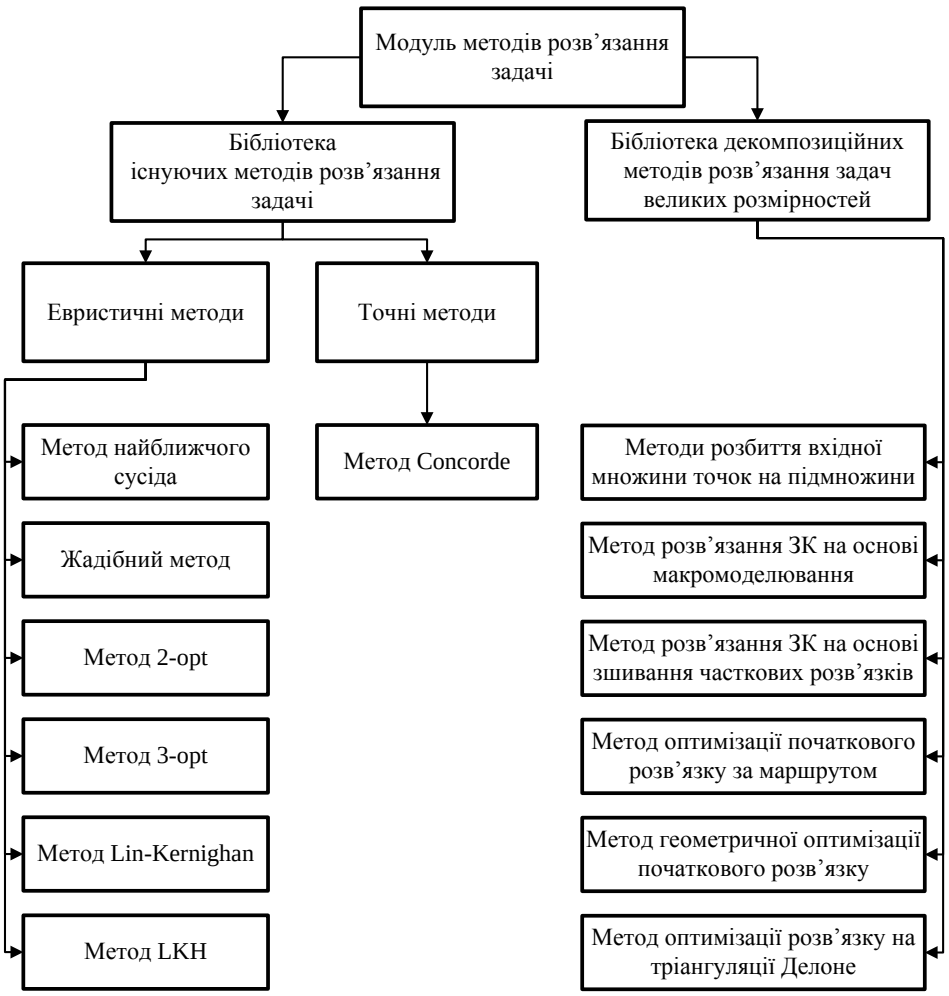

Рис. 9. Структура модуля методів розв’язування ЗК ППС “VRP Modeler”

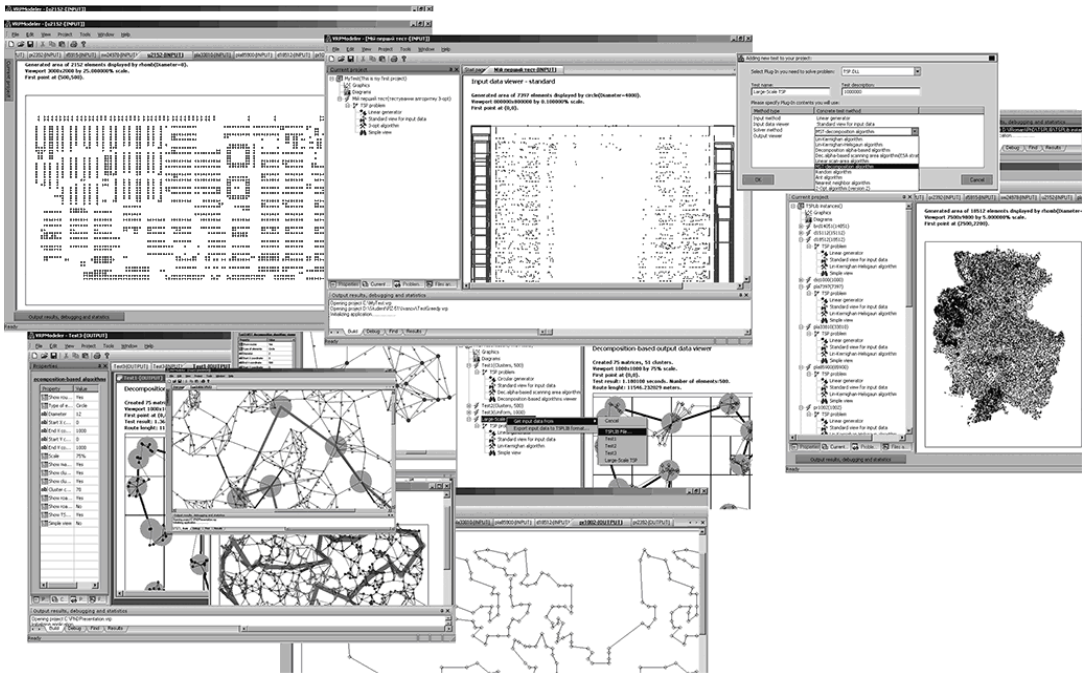

Рис. 10. Вигляд інтерфейсу користувача ППС “VRP Modeler”

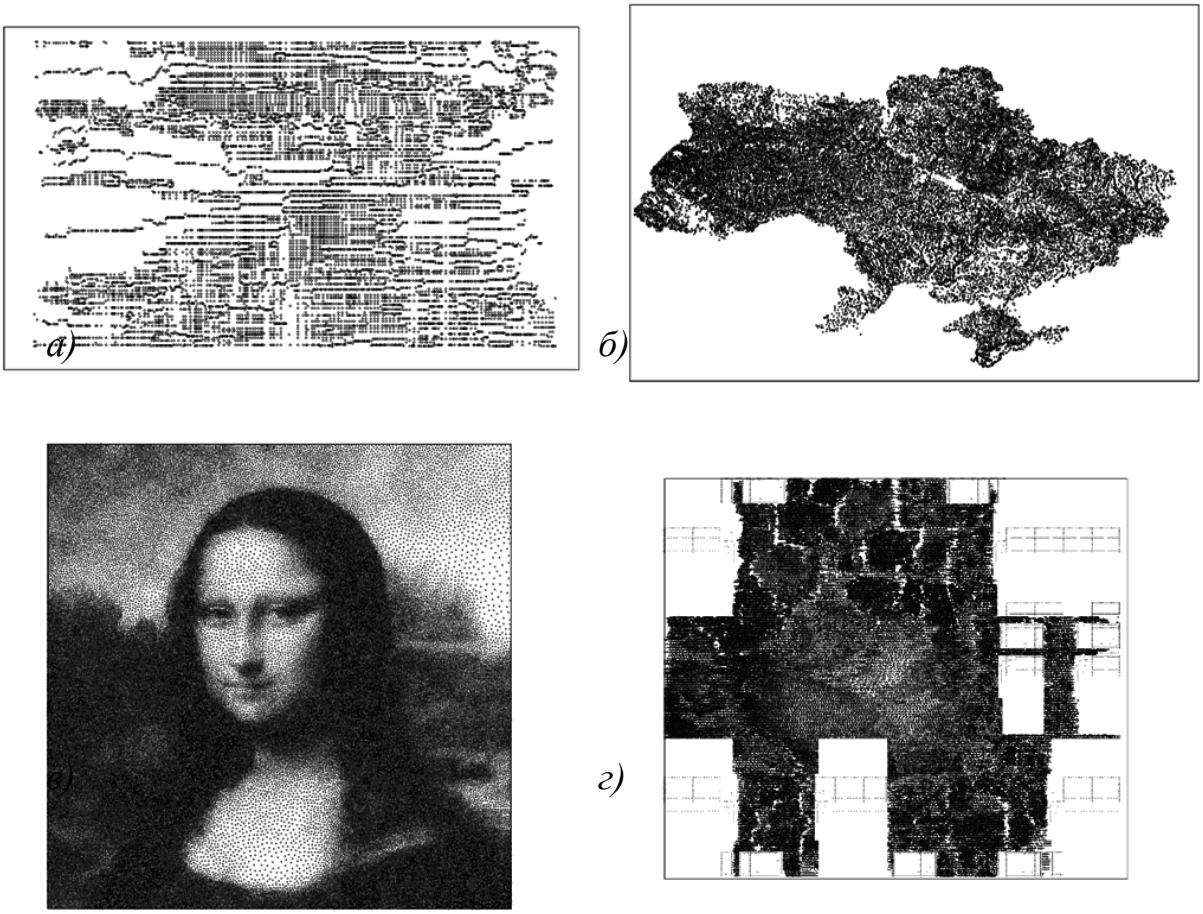

Рис. 11. Приклади деяких досліджених тестових задач: а) тестова задача rl11849 (друкована плата) розмірністю 11849 точок; б) тестова задача ua56583 (населені пункти України) розмірністю 56583 точки; в) тестова задача Mona Lisa (технологія малювання неперервною лінією) розмірністю 100000 точок; г) тестова задача lrb744710 (НВІС) розмірністю 744710 точок

З метою оцінки ефективності розроблених методів проведено експерименти на тестових задачах відомих бібліотек, а саме – бібліотек сайту ППС “Concorde, конкурсу DIMACS TSP Challenge та TSPLIB. Розмірність задач, що досліджувалися, – від 10^3 до 10^7 точок. Як базовий обрано метод LKH.

У порівнянні з методом, реалізованим у ППС “Concorde”, отримано виграш в часі в 19 – 402618 раз для досліджених задач розмірністю 1000-85900 точок відповідно з відхиленням не більше 0,58%. Для ряду досліджених задач (розмірністю 1000-2392 точок) знайдено оптимальні розв’язки. У порівнянні з методом LKH для досліджених задач розмірністю 3000-85900 точок отримано виграш в часі у 2,16 – 15,7 раз відповідно. Отримано розв’язки для відомих тестових задач розмірністю 10^6 та 10^7 точок, забезпечивши при цьому відхилення у якості розв’язку не вище 0,11% у порівнянні з найкращим знайденим на даний час. Для тестових задач із технології малювання неперервною лінією (бібліотека тестів ППС “Concorde”) розмірністю 100000-200000 точок отримано виграш в часі в 4,76 – 9,18 раз при покращенні якості до 0,02%.

ОСНОВНІ РЕЗУЛЬТАТИ ТА ВИСНОВКИ

У дисертаційній роботі вирішено актуальну наукову задачу – розроблено ефективні методи та алгоритми розв’язування задачі комівояжера великих розмірностей з близькою до лінійно-логарифмічної обчислювальною складністю. На основі проведених досліджень розроблено прикладну програмну систему “VRP Modeler” для розв’язування ЗК великих розмірностей. Отримано такі наукові та практичні результати:

1. Для ЗК великих розмірностей розвинуто відомі та розроблено нові декомпозиційні методи, в яких задача розв’язується за декілька етапів: розбиття вхідної множини точок на підмножини обмеженої розмірності (≈ 500-2000 точок), для яких отримуються високоякісні часткові розв’язки з невеликими часовими витратами; нарощення часткових розв’язків у початковий розв’язок, та його покращення розробленими оптимізаційними методами.
2. Розроблено декомпозиційний метод знаходження початкового розв’язку ЗК великих розмірностей на основі нарощення часткового розв’язку, що забезпечує відхилення від розв’язку, отриманого за допомогою методу LKH, не більше 0,5% зі зменшенням обчислювальної складності з $O(N^{2,2})$ до $O(N \log N)$.
3. Розроблено методи оптимізації початкового розв’язку ЗК великих розмірностей з близькою до лінійної та лінійно-логарифмічної обчислювальною складністю, які забезпечують відхилення від розв’язку, отриманого за допомогою методу LKH, не більше 0,36%, а також 0,58% від оптимального (для відомих тестових задач розмірністю 3000-85900 точок): метод оптимізації за маршрутом, геометричної оптимізації та метод оптимізації на тріангуляції Делоне.
4. Розроблено прикладну програмну систему “VRP Modeler” для розв’язування ЗК великих розмірностей, в якій з метою організації ефективних обчислень інтегровано набір розроблених декомпозиційних та відомих методів. Модулі, що входять у ППС “VRP Modeler”, можуть бути інтегровані в різні системи для розв’язування прикладних задач, де є необхідність знаходження розв’язків ЗК великих розмірностей.
5. Розроблені методи та програмні засоби забезпечують отримання розв’язків ЗК великих розмірностей з високою якістю за значно менший час у порівнянні з найкращим відомим евристичним методом LKH (в 2,16-15,7 раз для досліджених задач розмірністю 3000-85900 точок), та можуть застосовуватися в широкому колі прикладних задач, які вимагають знаходження субоптимальних маршрутів: маршрутизації автотранспорту, автоматизованому проектуванні, тестуванні та виготовленні НВІС, виробництві друкованих плат, лазерній нарізці пластмас і металів, дослідженні структури білка, технології малювання неперервною лінією, рентгенівській кристалографії та інших.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. Базилевич Р. П. Алгоритм оптимізації розв'язків задачі комівояжера у локальній області / Р. П. Базилевич, Р. К. Кутельмах // Радіоелектронні і комп'ютерні системи. – Харків. – 2009. – № 7(41). – С. 41-45.
2. Базилевич Р. П. Алгоритм приєднання часткових розв'язків у підмножинах при декомпозиції задачі комівояжера / Р. П. Базилевич, Р. К. Кутельмах // Вісник Національного університету "Львівська політехніка". – 2009. – № 653 : Інформаційні системи та мережі. – С. 3-12.
3. Базилевич Р. П. Дослідження ефективності існуючих алгоритмів для розв'язання задачі комівояжера / Р. П. Базилевич, Р. К. Кутельмах // Вісник Національного університету "Львівська політехніка". – 2009. – № 650 : Комп'ютерні науки та інформаційні технології. – С. 235-245.
4. Базилевич Р. П. Декомпозиційні алгоритми для розв'язування задачі комівояжера / Р. П. Базилевич, Р. К. Кутельмах // Вісник Національного університету "Львівська політехніка". – 2007. – № 598 : Комп'ютерні науки та інформаційні технології. – С. 138-148.
5. Базилевич Р. П. Алгоритми динамічного формування моделі робочого поля для задачі комівояжера з кластерним розподілом точок / Р. П. Базилевич, Р. К. Кутельмах // Вісник Національного університету "Львівська політехніка". – 2006. – № 565 : Комп'ютерні науки та інформаційні технології. – С. 200-207.
6. Базилевич Р. П. Оптимізація розв'язків задачі комівояжера методом послідовного сканування / Р. П. Базилевич, Р. К. Кутельмах // Вісник Національного університету "Львівська політехніка". – 2009. – № 638 : Комп'ютерні науки та інформаційні технології. – С. 254-260.
7. Базилевич Р. П. Методи кластеризації множини точок в задачі комівояжера з обмеженнями / Р. Базилевич, Р. Кутельмах, Б. Кузь // Вісник Національного університету "Львівська політехніка". – 2010. – № 672 : Комп'ютерні науки та інформаційні технології. – С. 87-90.
8. Базилевич Р. П. Використання алгоритмів локальної оптимізації для розв'язування задачі комівояжера з кластерним розподілом точок / Р. П. Базилевич, Р. Дюпа, Р. К. Кутельмах // Вісник Національного університету "Львівська політехніка". – 2006. – № 565 : Комп'ютерні науки та інформаційні технології. – С. 207-212.
9. Bazylevych R. Decomposition and scanning optimization algorithms for TSP / Bazylevych R., Kutelmakh R., Prasad B., Bazylevych L. // Proceedings of the International Conference on Theoretical and Mathematical Foundations of Computer Science. – Orlando, USA, 2008. – P. 110-116.
10. Bazylevych R. Decomposition algorithms for large-scale clustered TSP / Bazylevych R., Kutelmakh R., Dupas R., Bazylevych L. // Proceedings of the 3rd Indian International Conference on Artificial Intelligence. – Pune, India, 2007. – P. 255-267.

11. Bazylevych R. A decomposition algorithm for uniform Traveling Salesman Problem / Bazylevych R., Prasad B., Kutelmakh R., Dupas R., Bazylevych L. // Proceedings of the 4th Indian International Conference on Artificial Intelligence. – Tumkur, India, 2009. – P. 47-59.
12. Bazylevych R. Decomposition of clustered TSP / Bazylevych R., Dupas R., Kutelmakh R. // Conférence Francoro V / Roadef 2007. – Livre des resumes. – Grenoble, France, 2007. – P. 57-58.
13. Bazylevych R. Scanning-area algorithms for clustered TSP / Bazylevych R., Dupas R., Kutelmakh R. // Proceedings of International Conference "Computer Science and Information Technologies". – Lviv, 2006. – P. 148-152.
14. Bazylevych R. Methodology and software architecture for solving large-scale Traveling Salesman Problem / Bazylevych R., Kuz' B., Kutelmakh R. // Proceedings of International Conference "Computer Science and Information Technologies". – Lviv, 2010. – P. 101-102.
15. Bazylevych R. Parallel approach for solving large-scale clustered TSP / Bazylevych R., Dupas R., Kutelmakh R. // Proceedings of the 10-th International Conference "The experience of designing and application of CAD systems in microelectronics". – Polyana, 2009. – P. 452-456.
16. Базилевич Р. Алгоритмічне забезпечення для розпаралелювання задачі комівояжера великої розмірності / Базилевич Р., Кутельмах Р., Кузь Б. // Вісник конференції "Інформаційні технології та комп`ютерна інженерія". – Вінниця, 2010. – С. 88-89.

АНОТАЦІЇ

Кутельмах Р. К. Математичне та програмне забезпечення для розв'язування задачі комівояжера великих розмірностей. – Рукопис.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 01.05.03. – Математичне та програмне забезпечення обчислювальних машин та систем. – Національний університет "Львівська політехніка", Львів, 2011.

У дисертаційній роботі розглянуто задачу комівояжера великих розмірностей, яка має широке прикладне застосування в транспортних системах, автоматизованому проектуванні, тестуванні та виготовленні інтегрованих схем, виробництві друкованих плат та інших галузях. Розвинуто відомі та розроблено нові декомпозиційні методи, в яких задача розв'язується за декілька етапів: розбиття вхідної множини точок на підмножини обмеженої розмірності (≈ 500-2000 точок), для яких отримуються високоякісні часткові розв'язки з невеликими часовими затратами; зшивання часткових розв'язків у початковий розв'язок, та його покращення розробленими методами оптимізації.

Розроблено прикладну програмну систему "VRP Modeler" для розв'язування ЗК великих розмірностей, яка є спеціальним програмним забезпеченням, що дає змогу розв'язувати реальні задачі комівояжера, їх моделювати, досліджувати та інтегрувати нові методи. Розроблені методи та програмні засоби кластеризації вхідних даних, побудови макромаршруту, знаходження початкового розв'язку та

його оптимізації можна застосовувати для широкого кола прикладних задач, де використовується ЗК та близькі до неї.

Ключові слова: задача комівояжера, комбінаторні задачі класу *NP*, декомпозиція, локальна оптимізація, кластеризація.

Кутельмах Р. К. Математическое и программное обеспечение для решения задачи коммивояжера больших размерностей. – Рукопись.

Диссертация на соискание ученой степени кандидата технических наук по специальности 01.05.03. – Математическое и программное обеспечение вычислительных машин и систем. – Национальный университет “Львовская политехника”, Львов, 2011.

В диссертационной работе рассмотрена задача коммивояжера больших размерностей, широко применяемая в транспортных системах, автоматизированном проектировании, тестировании и изготовлении интегрированных схем, производстве печатных плат и других отраслях. Развиты существующие и разработаны новые декомпозиционные методы, в которых задача решается за несколько этапов: разбиение входного множества точек на подмножества ограниченой размерности (≈ 500-2000 точек), для которых получаются высококачественные частичные решения с небольшими временными затратами; сшивания частичных решений в начальное решение, и его улучшение разработанными методами локальной оптимизации.

Разработана прикладная программная система "VRP Modeler" для решения ЗК больших размерностей, которая является специальным программным обеспечением, что позволяет решать реальные задачи коммивояжера, их моделировать, исследовать и интегрировать новые методы. Разработанные методы и программные средства кластеризации входных данных, построения макромаршрута, нахождения начального решения и его оптимизации можно применять для широкого круга прикладных задач, где используется ЗК и близкие к ней.

Ключевые слова: задача коммивояжера, комбинаторные задачи класса NP, декомпозиция, локальная оптимизация, кластеризация.

Kutelmakh R.K. Mathematical methods and software for solving large scale traveling salesman problem. – Manuscript.

Thesis for a Ph.D. degree in technical sciences by speciality 01.05.03 – Mathematical support and software of computational machines and systems. – Lviv Polytechnic National University, Lviv, 2011.

The Traveling Salesman Problem (TSP) is extensively applied in transportation systems, automated design, testing and manufacturing of integrated circuits, printed circuit boards, laser cutting of plastics and metals, as well as in protein structure research, continuous line drawings, X-ray crystallography and a number of other fields. A significant feature of such problems is their large-scale complexity amounting to over a million points for many of them. The TSP is viewed as an intractable combinatorial problem (discrete optimization), referred to the class of *NP*-hard problems due to its factorial computational complexity.

The computational complexity of existing effective methods is not lower than $O(N^2)$, which makes them inappropriate for large scale and very large scale problems. Therefore there is a need for developing methods and an application software system for solving such problems. In order to achieve high computing performance, appropriate for complicated large-scale problems, this dissertation presents a set of specialized decomposition methods adapted to specific types of input data structures and an application software system that enables solving real TSPs, as well as modelling, exploring and integrating new methods.

The following tasks were done: the different types of TSP such as uniform and clustered were classified, high quality heuristic methods were chosen as basic, methods of finding the initial solution as well as it's optimization for large scale TSP with their adaptation to the structure of input data were developed. Software for solving large scale TSP was developed. Experimental investigations of the developed methods were done as well as the comparative analysis of the results and estimated their effectiveness in terms of quality and solving time.

Existing methods has been improved as well as new decomposition methods for finding initial solution and solution optimization have been developed. The problem is solved in several steps: partitioning of the input set of points into smaller subsets (≈ 500-2000 points), finding partial high-quality solutions for them within small time frames, merging them into the whole initial solution, and further solution optimization. Developed methods have computational complexity very close to linear.

According to the experimental results, the proposed decomposition approaches provide substantial reduction in the run time in comparison with the currently best heuristic Lin-Kernighan-Helsgaun algorithm. Quality loss is negligible. Problem instances up to 10^7 points were investigated.

Developed mathematical methods provided the possibility to develop special software system "VRP Modeler" for large scale TSP solving and experimental research, in which with the purpose of organization of effective calculations it is computer-integrated in the unique complex known existing and developed decomposition methods.

Software is developed by author independently in Microsoft Visual Studio 2008 IDE on C++ programming language with the use of methods of the object-oriented programming. Dissertation job performances were applied in the educational process of software department of the Lviv National Polytechnic University and also in the "AR-TRANS" and "SEA Electronics" companies. Developed software system "VRP Modeler" for solving large scale TSP, which is a special software that allows to solve real travelling salesman problems, model, explore and integrate new methods. The methods and tools for clustering of input data, building macroroute, finding the initial solution and its optimization can be applied to a wide range of applications, which used the TSP and those close to it: VLSI optimization, systems and networks on a crystal (SoC and NoC), optimizations of production processes, vehicle routing problems, continuous line drawings et al.

Keywords: traveling salesman problem, combinatorial NP-type problems, decomposition, local optimization, clustering.

Підписано до друку 28.04.2011.
Формат 60х84/16. Папір офсетний. Друк на різографі.
Умовн. друк. арк.0,9.
Наклад 100 прим.
Замовлення № 0405/1

Друк СПД “Синчук В.В.”
м. Львів, вул. Чупринки, 38
тел./факс (032) 297-04-74