

Каграманянц Виктор Александрович
E-mail: vict@mopevm.tsure.ru.

Kravchenko Pavel Pavlovich
Taganrog Institute of Technology - Federal State-Owned Educational Establishment of Higher Vocational Education "Southern Federal University".
E-mail: kravch@tsure.ru.
44, Nekrasovskiy, Taganrog, 347928, Russia.
Phone: 88634314945.

Kagramanyants Victor Alexandrovich
E-mail: vict@mopevm.tsure.ru.

УДК 528

В.С. Василенко

**ОПТИМИЗАЦИЯ ПОИСКА АЛГОРИТМА РЕШЕНИЯ ЗАДАЧИ
В ПРОСТРАНСТВЕ ВЕРСИЙ В ЭКСПЕРТНЫХ ГИС**

*Рассмотрены достоинства и недостатки алгоритма автоматического поиска решения задач в ГИС. Предложен способ оптимизации этого алгоритма.
Геоинформационная система; картография; пространство версий.*

V.S. Vasilenko

**OPTIMIZATION OF THE TASK SOLUTION ALGORITHM SEARCH
IN THE VERSION SPACE OF EXPERT GIS**

*Merits and demerits of automatic scan of problem solving algorithm in GIS are considered. The way of optimisation of this algorithm is offered.
A geointelligence system; cartography; version space.*

Современные геоинформационные системы (ГИС) являются автоматизированными информационными системами, предназначенными для обработки пространственно-временных данных, основой интеграции которых служит географическая информация [1]. При этом большинство существующих геоинформационных систем, таких, как система Geo Base, продукт ArcGIS компании ESRI, семейство продуктов GeoMedia корпорации Intergraph, MapInfo компании Pitney Bowes MapInfo, сводят процесс автоматизации к предоставлению конечному пользователю возможностей СУБД, редакторов растровой и векторной графики и набора высокоуровневых аналитических средств [2]. Решение некоторой картографической задачи в таких ГИС, а именно построение тематической карты, отвечающей требованиям предметной области, представляет собой достаточно трудоемкий процесс, требующий экспертных знаний в области геоинформатики.

В общем случае решение картографической задачи в современных ГИС состоит из формулировки задачи, анализа набора инструментальных средств, предоставляемых конкретной системой, анализа картографических данных, представленных в исследуемой области, выборки полезных картографических объектов на основе определенных цензов и формирования из них тематических слоев, выполнении над полученными слоями ряда аналитических операций и группировании их в тематическую карту [3]. Такой подход обладает рядом принципиальных недостатков – требует от пользователя наличия экспертных знаний в области решения картографических задач, требует широкой полосы пропускания информационного

канала, связывающего пользователя с сервером ГИС, занимает относительно много времени. Эти недостатки значительно ограничивают область применения таких геоинформационных систем и приводят к необходимости построения автоматизированных систем более высокого уровня – экспертных ГИС.

Оптимальным решением проблемы является освобождение пользователя от необходимости создания алгоритма решения картографической задачи (перенос логики ее решения на сервер ГИС) и предоставление ему возможности сформулировать задачу с помощью близкого к естественному языку декларативного описания требований к итоговой тематической карте. При этом сервер ГИС должен сам сгенерировать необходимый алгоритм решения (или определить, что такого алгоритма не существует) по заданной декларации и применить его к исходным данным. Это возможно, так как решение любой задачи в ГИС пользователем-экспертом представляет собой набор предоставляемых данной системой элементарных операций, выполняемых над исходными данными в определенной последовательности.

При этом класс картографических объектов $c_i \in C$, $i \in [0, N]$ рассматривается как множество свойств $p_j \in P$, $j \in [0, M]$:

$$c_i \equiv \{p_k, p_{k+1}, \dots, p_l, \dots, p_{n_i}\},$$

где $k \leq l \leq n_i$, $i \in [0, N]$, $k, l, n_i \in [0, M]$ и каждое свойство обладает типом $t \in T$, соответствующим типу принимаемых свойством значений, доменом, определяющим диапазон допустимых значений свойства, и связностью с другими свойствами в классе. Классы, свойства, из которых они состоят, типы и домены свойств составляют метаданные, определяющие структуру картографических данных. Сами данные представляются как экземпляры классов, т.е. как определенный набор значений свойств, принадлежащих данному классу. Предоставляемые геоинформационной системой функции $f_i \in F$, $i \in [0, L]$ по преобразованию данных определяются множествами картографических свойств своей области определения и области значения:

$$f_i(p'_k, p'_{k+1}, \dots, p'_m, \dots, p''_l, p''_{l+1}, \dots, p''_h) = \{p'_k, p'_{k+1}, \dots, p'_m, \dots, p'''_s, p'''_{s+1}, \dots, p'''_u\},$$

где $k \leq m$, $l \leq h$, $s \leq u$, $k, m, l, h, s, u \in [0, M]$, $i \in [0, L]$.

Процесс автоматического построения алгоритма решения картографической задачи основывается на отделении картографических данных от метаданных, описывающих структуру данных (выделения в предметной области классов объектов и составляющих их свойств), и определении множества элементарных преобразований, переводящих метаданные из одного состояния в другое (определении множества функций F), множество последовательностей S которых с согласованными по типам областей определения и значения членами составляют пространство версий экспертной ГИС. Картографическая задача представляется в виде исходного A и целевого B множества элементарных состояний метаданных, а алгоритм ее решения ищется как последовательность элементарных преобразований, переводящих исходное множество состояний в целевое:

$$s_i(A) = B, \quad s_i \in S.$$

Эта последовательность строится с помощью восходящего или нисходящего поиска – формирования нового множества метаданных как множества типов значений элементарных операций на основе, соответственно, исходного или целевого множества с помощью применения к элементарным состояниям метаданных этих множеств, подходящих по типу принимаемых значений элементарных операций, и продолжения такого преобразования до тех пор, пока не будет получено необходимое множество элементарных состояний метаданных или не будет обнаружено, что никакая последовательность элементарных преобразований не может перевести исходное множество элементарных состояний метаданных в целевое. Алгоритмом решения задачи будет последовательность найденных элементарных преобразований $s_{target}(A) = B$, $s_{target} \in S$, а окончательное решение задачи получится после применения этого алгоритма к исходным данным.

Такой подход для поиска алгоритма решения картографической задачи, несомненно, эффективнее простого перебора, однако при непосредственном использовании в современных геоинформационных системах он обладает рядом принципиальных недостатков – потенциальными неоптимальностью и несходимостью.

Картографическая информация в современных ГИС представляется в виде небольшого числа классов объектов, где каждый класс характеризуется определенным типом представления географической информации и уникальным набором атрибутов. Например, в продукте ArcInfo тип географической составляющей определяется типом примитива, используемого для ее представления, таким, как точка, линия, множество точек, полилиния, полигон, а набор атрибутов класса соответствует столбцам связанной атрибутивной таблицы. Метаданными для автоматического поиска решения картографической задачи при такой организации будут соответственно тип географической составляющей и типы атрибутов, а элементарные операции будут определены на множестве этих типов. Так как пользователь формулирует картографическую задачу в терминах, свойственных естественному языку и опирающихся на семантику картографических данных, а не на их представление в базе данных ГИС, то такая формулировка должна быть предварительно преобразована в форму, совместимую со способом представления данных в ГИС. Таким образом, представляемые картографические данные не связаны с семантикой объектов и одним и тем же способом могут представлять разные объекты предметной области. Очевидно, формулировки различных картографических задач при этом будут похожи, а возможные решения будут либо не соответствовать смыслу задачи, либо отличаться друг от друга только перестановкой элементарных функций.

В процессе поиска алгоритма решения возможны циклы и рекурсии, когда несколько элементарных операций имеют обратно совместимые типы принимаемых и возвращаемых значений. То есть на некотором шаге i полученное множество свойств включает множество свойств, полученное на шаге $j < i$. При этом поиск никогда не закончится. Поиск не дает гарантии нахождения оптимального алгоритма, а перебор всех возможных последовательностей элементарных преобразований с целью поиска оптимального алгоритма не эффективен, так как при ограниченном количестве типов данных и потенциальной необходимости конечных циклов преобразований число таких последовательностей может быть очень велико.

Эти проблемы могут быть решены путем введения нового класса метаданных и их элементарных преобразований, основывающихся не на физическом представлении картографических данных в базе данных ГИС, а на их семантическом содержании. Метаданные и операции их преобразования при таком подходе будут надстройкой над базовыми метаданными и операциями, но смогут быть легко масштабированы

ны и организованы в иерархическую структуру в соответствии со своей семантикой. При этом поиск алгоритма решения картографической задачи не требует предварительной формализации задачи в терминах физического представления данных в ГИС (что создает дополнительный уровень инкапсуляции метода представления данных) – задача может быть легко сформулирована в терминах, ориентированных на семантику метаданных. Процесс поиска алгоритма решения теперь представляет собой многоступенчатый процесс, где сначала ищется решение в терминах (наборах свойств) метаданных, находящихся на высоких уровнях иерархии. Практически решение будет получено в виде последовательности картографических задач, каждая из которых должна решаться в терминах метаданных, находящихся на более низком уровне иерархии. Такой процесс следует продолжать до получения алгоритма решения исходной картографической задачи в терминах, связанных с представлением картографических данных в базе данных ГИС.

При таком подходе обеспечивается относительная оптимальность искомого алгоритма и сходимости алгоритма поиска за конечное время, что связано с простотой задач и малым количеством возможных преобразований (в терминах текущего уровня иерархии метаданных), решаемых при проходе по уровням иерархии. Дополнительным преимуществом такого алгоритма является легкость его распараллеливания для выполнения в многопроцессорных системах – каждая промежуточная задача определяется только множествами исходных и целевых элементарных состояний и может быть решена независимо от других промежуточных задач.

Метод автоматического решения картографических задач, основанный на построении алгоритма решения таких задач на основе последовательности преобразований, связанных с ними метаданных, позволяет существенно упростить использование геоинформационных систем и расширить область их использования. Потенциальная неоптимальность и несходимость метода могут быть исправлены путем введения иерархической системы метаданных и операций, основывающихся на семантике картографических данных.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Slocum H., Terry A.* Thematic Cartography and Visualization. – NY: Diasoft, 2006. – 427 p.
2. *ДеМерс М.Н.* Географические информационные системы. Основы: Пер. с англ. – М.: Дата+, 1999. – 491с.
3. *Тихунов В.С.* Моделирование в картографии. – М.: Изд-во Моск. ун-та, 1997. – 405 с.

Василенко Виталий Сергеевич

Технологический институт Федерального государственного образовательного учреждения высшего профессионального образования «Южный федеральный университет» в г. Таганроге.

E-mail: vittech@mail.ru

347923, Россия, г. Таганрог, пер. Красногвардейский, 13.

Тел.: 89044439340.

Vasilenko Vitaly Sergeevich

Taganrog Institute of Technology – Federal State-Owned Educational Establishment of Higher Vocational Education “Southern Federal University”.

E-mail: vittech@mail.ru.

13, Krasnogvardeyskiy street, Taganrog, 347923, Russia.

Phone: 89044439340.