

## OPTIMAL ALGORITHM OF HYPOTHESES TESTING AT THE SHIP USE RESEARCH BASED ON THE GEOCHRONOLOGICAL TRACKING

**Y. A. Ivakin<sup>1,2</sup>, S. N. Potapychev<sup>1,2</sup>, R. Y. Ivakin<sup>3</sup>**

<sup>1</sup> — St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, St. Petersburg, Russian Federation

<sup>2</sup> — Concern OCEANPRIBOR JSC, St. Petersburg, Russian Federation

<sup>3</sup> — The Bonch-Bruевич Saint Petersburg State University of Telecommunications, St. Petersburg, Russian Federation

*Information technology of the geochronological tracking is an assembly of processes that accumulate and integrate data about geographic relocation of ships, figures for a given time interval and represent the results as a generalizing graph in GIS. Hypotheses on the stable tendencies in migration could be represented as the above graph's sub-graphs. Such tendencies testing is reduced to the search and evaluation of the statistical significance of the matching graphs isomorphism. Full-featured development of computer interpretation of the graph theory methods based on geochronological tracking provides new quality of research using modern GIS-tools. Namely, researcher can use the quantitative methods of the corresponding logical-analytical apparatus. The proposed paper deals with a consideration of qualitatively new possibilities of such approach and the corresponding algorithmic apparatus. Dispatching for geospatial processes of maritime transport with implementation of geochronological tracking tools is tightly related with completion of more developed model of data processing in the automated vessel traffic control system. The proposed model is focused not only on each vessel's data processing but also on spatial situations tracking based on the data of geographic theatre. This is the main feature for implementation of software component of geochronological tracking for dispatching of geospatial processes for maritime transport. The article is dedicated to the thorough consideration of this feature.*

*Keywords: Geographic information systems, GIS-technologies for historic research, geochronological track and tracking, graphs isomorphism, optimal algorithm, GIS-based interdisciplinary research.*

**For citation:**

Ivakin, Yan A., Sergei N. Potapychev, and Roman Y. Ivakin. "Optimal algorithm of hypotheses testing at the ship use research based on the geochronological tracking." *Vestnik Gosudarstvennogo universiteta morskogo i rechnogo flota imeni admirala S. O. Makarova* 11.3 (2019): 448–460. DOI: 10.21821/2309-5180-2019-11-3-448-460.

**УДК 681.1.003**

## РАЦИОНАЛЬНЫЙ АЛГОРИТМ ПРОВЕРКИ ГИПОТЕЗ РЕТРОСПЕКТИВНЫХ ИССЛЕДОВАНИЙ ИСПОЛЬЗОВАНИЯ ВОДНОГО ТРАНСПОРТА НА БАЗЕ ГЕОХРОНОЛОГИЧЕСКОГО ТРЕКИНГА

**Я. А. Ивакин<sup>1,2</sup>, С. Н. Потапычев<sup>1,2</sup>, Р. Я. Ивакин<sup>3</sup>**

<sup>1</sup> — Санкт-Петербургский институт информатики и автоматизации Российской академии наук, Санкт-Петербург, Российская Федерация

<sup>2</sup> — АО «Концерн «ОКЕАНПРИБОР», Санкт-Петербург, Российская Федерация

<sup>3</sup> — Санкт-Петербургский государственный университет телекоммуникаций имени профессора М. А. Бонч-Бруевича, Санкт-Петербург, Российская Федерация

*Рассмотрена информационная технология геохронологического трекинга как совокупность процессов накопления и интеграции данных о географическом перемещении судов, должностных лиц, артефактов за установленный период времени с представлением результатов в виде обобщающего графа в географической информационной системе или в ГИС-приложении соответствующей информационно-логистической системы. Гипотезы об устойчивых тенденциях в миграции указанных сущностей представлены как подграфы указанного графа. Показано, что проверка таких гипотез сведется к поиску и оценке*

статистической значимости изоморфизма соответствующих графов. Исследованы различные аспекты того, что полнофункциональное развитие компьютерной интерпретации методов теории графов на базе геохронологического трекинга способно обеспечить новое качество ретроспективных исследований вопросов специфики организации эксплуатации водного транспорта с использованием современного ГИС-инструментария. Такие аспекты выражаются в предоставлении возможности исследователю использовать количественные методы соответствующего логико-аналитического аппарата в своей предметной области — области организации и реализации перевозок пассажиров и грузов средствами водного транспорта как на внутренних водных путях страны, так и на международных морских линиях перевозок. Такая возможность открывает новые пути для анализа логистики современных транспортно-поставочных сетей, оптимизации транспортных потоков, систем диспетчеризации водного транспорта. Рассмотрению качественно новых возможностей такого подхода, его сильных и слабых сторон, а также рационализации соответствующего алгоритмического аппарата посвящена данная статья. Определены также направления рационализации алгоритмов проверки гипотез исследования на базе геохронологического трекинга в ГИС, которые связаны с постановкой и решением оптимизационной задачи определения временной сложности указанных алгоритмов, а также определения строгих граничных условий такой оптимизации.

*Ключевые слова:* Географические информационные системы, ГИС-технологии для ретроспективных исследований, геохронологический трек и трекинг, изоморфизм графов, рациональный алгоритм, междисциплинарные исследования на базе ГИС.

**Для цитирования:**

Ивакин Я. А. Рациональный алгоритм проверки гипотез ретроспективных исследований использования водного транспорта на базе геохронологического трекинга / Я. А. Ивакин, С. Н. Потапычев, Р. Я. Ивакин // Вестник Государственного университета морского и речного флота имени адмирала С. О. Макарова. — 2019. — Т. 11. — № 3. — С.448–460. DOI: 10.21821/2309-5180-2019-11-3-448-460.

### **Введение (Introduction)**

Достижение высокой эффективности, безопасности эксплуатации судов и других объектов водного транспорта невозможно без ретроспективного анализа организации и результатов их использования. Именно результаты такого анализа являются основой для обоснования стратегических управленческих решений по развитию маршрутной сети, открытию новых линий перевозок и др. Мощным средством такого анализа на базе геоинформационных технологий является *геохронологический трекинг*. Процесс геохронологического трекинга представляет собой совокупность способов сбора первичной ретроспективной, хронологической информации и последовательности приемов обобщения геохронологических треков судов и кораблей (объектов, артефактов или их совокупности) на электронной карте (в ГИС). Соответственно геохронологический трек есть интеграция хронологических и географических данных в виде графа, соединяющего географические точки нахождения указанных исторических сущностей (вершины трека имеют строгое историко-географическое определение, дуги носят характер условно-логической связи). В статьях [1], [2] дано комплексное описание узкоспециализированной программной технологии геохронологического трекинга применительно к проблематике водного транспорта, в работе [3] показаны возможности применения аналитического аппарата теории графов и статистических исследований на базе геохронологического трекинга.

Выполнение разработки апробационных примеров построения геохронологических треков по данным из архивных источников для различных групп исторических личностей позволило сделать вывод о том, что финишная версия графа для достаточно представительной выборки исторических сущностей (личностей, объектов и др.), как правило, имеет высокосложную и полно- либо высокосвязную структуру. Такая структура может быть строго упорядочена, о чем свидетельствует пример трека, приведенный на рис. 1. Именно на основе такого итогового графа геохронологического трекинга становится возможным исследование различных миграционных процессов, выявления некоторых частных исторических закономерностей в перемещении социальных групп, статистически подтвержденная проверка исследовательских гипотез о характере миграций. Существование, концептуальная модель и методологический аппарат таких исследований детально описаны в статье [3].

Концептуальная идея проверки исследовательских гипотез заключается в следующем: итоговый граф геохронотрекинга представляется как граф-базис, в структуре которого выявляется подграф изоморфный заданному, т. е. устанавливается наличие взаимно однозначного отображения одного графа на подграф другого, при котором сохраняется отношение инцидентности [4,5]. Граф, на изоморфность к которому в составе базового графа геохронологического трекинга определяется подграф, топологически описывает ту или иную определенную гипотезу исследования об устойчивой особенности в перемещениях исторических личностей, объектов или других существ в географическом пространстве. Далее определяется степень устойчивости в признании гипотезы исследования о выявляемой особенности в перемещениях с использованием статистического аппарата доверительной вероятности и доверительных интервалов.

Программно-алгоритмическая реализация проверки исследовательских гипотез на базе геохронологического трекинга представляет собой сложную итеративную вычислительную процедуру, практическое воплощение которой может иметь экспоненциальную сложность и приводить к трансвычислительному характеру решения задачи при определенных входных данных и граничных условиях. Именно этот факт диктует необходимость обоснования и разработки рационального или оптимального алгоритма проверки исследовательских гипотез на базе геохронологического трекинга, т. е. такой локализации вычислительного алгоритма решения задачи определения всех изоморфных вложений в граф геохронотрека, которая позволяет за конечное число подстановок (итераций) определить все комбинации вложений, изоморфных заданному графу-гипотезе, и не сделать решение трансвычислительным.

Синтез указанного обобщенного алгоритма поиска в графе-базисе подграфа-вложения, изоморфного заданному, применительно к специфике входных данных и граничных условий проверки исследовательских гипотез на базе геохронологического трекинга в ГИС, а также обоснование его рационального характера является предметом рассмотрения данной статьи.

### Методы и материалы (Methods and Materials)

Базовым методом описываемого исследования выступает графовая модель проверки исследовательских гипотез на базе геохронотрекинга. Итогом реализации трекинга как специализированного программного процесса в ГИС является географически привязанный граф, обобщающий геохронологические треки отдельных исторических существностей или артефактов, информация о миграциях которых занесена в соответствующую базу данных [6]. Именно этот граф является базисом проверки исследовательских гипотез, т. е. в его составе выявляются подграфы, изоморфные графам-гипотезам (рис. 1).

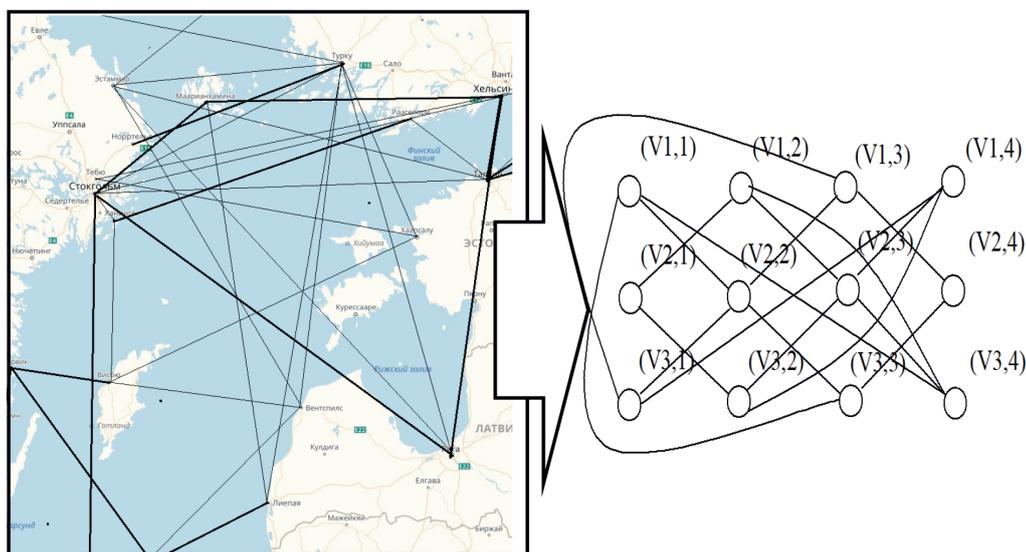


Рис. 1. Рассмотрение геохронологического трека в ГИС в качестве упорядоченного графа

Представленный на рис. 1 пример обобщающего геохронотрека дает возможность понять всю сложность и комбинаторную вариабельность решения частной задачи рационального выделения подграфа, изоморфного заданному. Существо указанного выделения заключается в следующем: два графа  $G_1 = (X, U)$  и  $G_2 = (X', U')$  изоморфны, если между парами множеств их вершин, ребер и дуг существуют взаимно однозначные соответствия, сохраняющие смежность и ориентацию для дуг [7]. Значение попарно изоморфных графов с заданным значением вершин и заданным значением ребер конечно. Изоморфное отображение  $\varphi$  графа  $G_1$  на граф  $G_2$  задается перестановкой

$$\varphi = \begin{pmatrix} x_1 & x_2 & \dots & x_i & \dots & x_p \\ \varphi(x_1) & \varphi(x_2) & \dots & \varphi(x_i) & \dots & \varphi(x_p) \end{pmatrix},$$

называемой *изоморфной*, т. е. при распознавании изоморфизма графов  $G_1 = (X_1, R_1)$  и  $G_2 = (X_2, R_2)$  необходимо определить факт их изоморфности. В случае установления изоморфизма необходимо указать изоморфную подстановку. Ввиду того, что в теории графов не определены граничные условия и условия оптимальности решения задачи определения строгого соответствия графов, правомерно применение целого ряда математических методов установления изоморфизма в качестве методологического базиса разработки искомого алгоритма в случае проверки гипотез исследования на базе геохронологического трекинга.

Условно все множество математических методов установления изоморфизма графов (методы решения задачи распознавания изоморфизма графа заданному) можно классифицировать как это показано на рис. 2 [7]. Очевидно, что доминирующими математическими методами в определении изоморфизма графов являются методы, использующие инварианты (инвариант графа — это некоторая функция, сопоставляющая графу  $L$  — соответствующий элемент  $f(L)$  из множества  $M$ , элементами которого выступают числа или системы чисел, векторы, многочлены, матрицы и другие алгебраические структуры. Для изоморфных графов значения этой функции совпадают [8]. В соответствии с разнообразием выбора однотипных фрагментов графа различают три класса инвариантов: локальные, квазиглобальные и глобальные.

Условно все множество математических методов установления изоморфизма графов (методы решения задачи распознавания изоморфизма графа заданному) можно классифицировать как это показано на рис. 2 [7]. Очевидно, что доминирующими математическими методами в определении изоморфизма графов являются методы, использующие инварианты (инвариант графа — это некоторая функция, сопоставляющая графу  $L$  — соответствующий элемент  $f(L)$  из множества  $M$ , элементами которого выступают числа или системы чисел, векторы, многочлены, матрицы и другие алгебраические структуры. Для изоморфных графов значения этой функции совпадают [8]. В соответствии с разнообразием выбора однотипных фрагментов графа различают три класса инвариантов: локальные, квазиглобальные и глобальные.



Рис. 2. Классификации математических методов установления изоморфизма графов

При решении задачи поиска изоморфизма графов при различных условиях (размерность графов, их регулярность, однородность и др.) определяется функция временной сложности самой задачи. Именно она позволяет конкретизировать математический метод решения задачи до прикладного алгоритма, адаптированного к граничным условиям геохронологического трекинга (упорядоченный граф, количество вершин  $n$  от 20 до 100 и др.). Как правило, это полиномиальный или экспоненциальный алгоритм поиска изоморфизма. Различия между двумя указанными типами алгоритмов особенно заметны при решении задач большой размерности.

Приведенные в таблице данные позволяют понять причины, по которым полиномиальные алгоритмы более предпочтительны для поиска изоморфизма на геохронотреке по сравнению с экспоненциальными: большинство экспоненциальных алгоритмов — это варианты полного перебора, в то время как полиномиальные алгоритмы можно построить лишь тогда, когда удастся строго формализовать предметную суть решаемой задачи. Иными словами, задача строго формализована, если для ее решения получен полиномиальный алгоритм [9].

### Временная сложность экспоненциальных алгоритмов решения задачи поиска изоморфизма графов

Функция временной сложности	Размер графа-базиса (количество вершин $n$ )				
	10	20	30	40	50
$n$	0,00001 с	0,00002 с	0,00003 с	0,00004 с	0,00005 с
$n_2$	0,0001 с	0,0004 с	0,0009 с	0,0016 с	0,0025 с
$n_3$	0,001 с	0,008 с	0,027 с	0,064 с	0,125 с
$n_5$	0,1 с	3,2 с	24,3 с	1,7 мин	5,2 мин
$2n$	0,001 с	1,0 с	17,9 мин	12,7 дней	35,7 лет
$3n$	0,059 с	58 с	6,5 лет	3855 столетий	2·108 столетий

В свою очередь, необходимо показать, что линейное установление изоморфизма графов алгоритмически отличается от более сложной задачи, т. е. распознавания изоморфного вложения в составе графа-базиса, изоморфного вложения или изоморфизма подграфа: граф  $G_2$  изоморфно вложим в граф  $G_1$ , если в графе  $G_1$  существует подграф, изоморфный графу  $G_2$  [4]. Эта задача отличается от задачи распознавания графов: в частности, чтобы решить задачу изоморфизма подграфа с использованием известных алгоритмов распознавания изоморфизма графов, необходимо реализовать процедуру выявления в графе  $G_1$  подмножества вершин  $X \subset X_1$ , равномошного множеству вершин  $X_2$  графа  $G_2$ . Данная процедура включает  $k_1$  действий, где  $k_1 = \binom{p_1}{p_2}$ ,  $p_1 = |X_1|$ ,  $p_2 = |X_2|$ . Следовательно,  $k_1$  раз необходимо применить алгоритм распознавания изоморфизма графов как некоторую частную процедуру в составе более общего алгоритма.

Таким образом, конкретизация графовой модели проверки гипотез исторического исследования в ГИС заключается в определении наилучшего математического алгоритма решения задачи изоморфного вложения графов, соответствующего ограничениям и условиям предметной области процедуры геохронотрекинга.

### Результаты (Results)

Выполнена рационализация процедуры распознавания изоморфного вложения подграфа-гипотезы в составе графа-базиса, обобщающего данные по перемещению судов и других объектов водного транспорта.

В целях разработки конкретизированного алгоритма распознавания изоморфного вложения подграфа-гипотезы в составе графа-базиса, применительно к задаче проверки исследовательских гипотез в ГИС структуре сводного геохронотрека, сопоставлен  $N$ -вершинный граф  $L$ . Пронумеровав его вершины натуральными числами, можно осуществить синтез изоморфного к  $L$  графа, вершинами которого служат эти числа. Вариантов такого графа будет  $n!$ , т. е. столько, сколько имеется перестановок из  $n$  элементов. Далее задается квадратная матрица:

$$A(L) = \|a_{ij}\|_n \quad (1)$$

с  $n$  строками и  $n$  столбцами, элементы которой определяются в виде:

$$a_{ij} = \begin{cases} 1, & \text{если } i\text{-я и } j\text{-я вершины в } L \text{ смежны;} \\ 0, & \text{если эти вершины не смежны.} \end{cases} \quad (2)$$

Очевидно, что  $a_{ii} = 0$  и матрица  $A(L)$  симметрична:  $a_{ij} = a_{ji}$ . Такая матрица будет матрицей смежности графа  $L$  с заданной нумерацией вершин, т. е. эта квадратная матрица  $A = \{a_{ij}\}$  будет являться матрицей смежности графа  $L$ , если при  $a_{ij} = l$  в графе  $L$  вершины  $x_i$  и  $x_j$  соединены  $l$  ребрами, при  $a_{ij} = 0$  вершины  $x_i$  и  $x_j$  в  $L$  несмежны.

Степенью  $s(L, x)$  вершины  $x$  графа  $L$  называется число его вершин, смежных с  $x$ , или, что то же, число ребер, инцидентных этой вершине. При всяком изоморфизме  $\leftrightarrow$  графов  $L$  и  $L'$  соответствующие друг другу вершины должны иметь одинаковую степень, т. е. для любой  $x \in X$  из  $x \leftrightarrow x'$  ( $x' \in X'$ ) должно выполняться равенство:  $s(L, x) = s(L', x')$ .

В самом деле, если для какой-то вершины  $x$  и соответствующей  $x'$  окажется  $s(L, x) > s(L', x')$ , то среди тех  $s(L, x)$  вершин графа  $L'$ , которые отвечают смежным с  $x$  вершинам  $L$ , хотя бы одна не будет смежной с  $x'$ , т. е. соответствие  $\leftrightarrow$  не будет изоморфизмом, но это условие может оказаться приемлемым, в случае, если граф  $L'$  изоморфно вложен в граф  $L$ , обратное утверждение ложно. Также для ориентированных графов различают полустепени исхода и захода. *Полустепень исхода вершины* — это число инцидентных исходящих дуг,  $S^-(x)$ . *Полустепень захода вершины* — число инцидентных заходящих дуг,  $S^+(x)$ . Для удобства алгоритмизации полустепени захода вершин записываются в  $n + 1$  строке, а полустепени исхода вершин — в  $n + 1$  столбце матрицы смежности  $n$ -вершинного ориентированного графа геохронотрека или его подграфов.

Перестановочная матрица (матрица подстановок) представляет собой матрицу, в каждой строке и столбце которой находится по одной единице на пересечениях координат соответствующих друг другу вершин изоморфных графов. При умножении матрицы смежности одного графа слева и справа на перестановочную матрицу можно получить матрицу смежности изоморфного графа [8]. Тогда становится возможным предложить конкретизированный алгоритм распознавания изоморфного вложения подграфа-гипотезы в составе графа-базиса, соответствующий ограничениям и условиям предметной области геохронотрекинга.

В соответствии с приведенной ранее классификацией методов решения такой задачи, математический метод, реализованный в алгоритме, является комбинированным методом направленного перебора. Он объединяет в себе основные преимущества, которые дают методы направленного перебора, использующие локальные, квазиглобальные и глобальные инварианты. В разработанном в ходе исследования алгоритме используются следующие инварианты:

количество вершин  $n(L)$ ;

количество ребер  $m(L)$ ;

вектор степеней  $S(L) = (s_1, s_2, \dots, s_n)$ , который, в частности, дает числовые инварианты:  $s(L) = \min S(L, x)$  и  $s'(L) = \max S(L, x)$ ;

полустепень исхода  $S^-(x)$ ;

полустепень захода  $S^+(x)$ ;

матрица смежности  $A(L)$ .

При синтезе указанного алгоритма были приняты следующие допущения и ограничения:

- предполагается, что вершины (ребра) графов имеют одинаковые свойства, т. е. в алгоритме не учитываются весовые коэффициенты вершин (ребер), поскольку для различных видов графов будут и различные требования к весовым коэффициентам;
- все вершины должны быть пронумерованы;
- рассматриваются только ориентированные графы, т. е. при анализе неориентированных графов необходимо задавать одно ребро как два, связывающих вершины в обоих направлениях;
- не рассматриваются «несвязанные» вершины, т. е. каждое географическое местоположение должно иметь хотя бы одно отношение с другим местоположением (географической точкой).

Методологическим базисом предлагаемого алгоритма решения задачи определения изоморфного вложения графа в граф-базис является тезис о том, что логические схемы  $S_1 = \{A_1, R_1\}$  и  $S_2 = \{A_2, R_2\}$  создают в совокупности систему ограничений, которая воздействует на множество гипотетически возможных вариантов решения, существенно его сокращая. Т. е. производится не перебор вариантов решения, что привело бы к  $N$ -факториальным переборам, а наложение системы ограничений по определенному алгоритму на специально созданное поле и на этом поле в результате последовательности действий, направленных на удовлетворение требований этой системы ограничений, формируется уже готовый вариант решения. Этот вариант может выглядеть для графов одинаковой размерности как матрица подстановок. Является допустимым поле с множеством гипотетически возможных подстановок представить в виде матрицы размерностью  $n \times m$  для  $n$  и  $m$  — число вершин графов соответственно. Такую матрицу далее корректно называть *матрицей возможных подстановок*.

Логико-математическое существо матрицы возможных подстановок заключается в том, что на ней отражено все поле возможных решений текущей задачи изоморфизма графов. Так, при решении этой задачи для графов одинаковой размерности путем прямого перебора считается, что каждой вершине  $x_i \in X_1$  графа 1 может быть сопоставлена любая из  $x_j \in X_2$  графа 2. Количество возможных подстановок на матрице размерностью  $n \times n$  будет равно  $n!$ . Эта матрица является удобным средством для фильтрации всех невозможных подстановок. Такая фильтрация имеет два этапа. На первом этапе производится устранение с поля возможных решений тех вариантов подстановок, которые невозможны принципиально по условию задачи (используя как фильтры глобальные, квазиглобальные и локальные инварианты, так и веса дуг или вершин и др.). На втором этапе фильтрация вариантов имеет место для выдвигаемых гипотез о той или иной подстановке.

### Обсуждение (Discussion)

Суть работы предлагаемого алгоритма распознавания изоморфного вложения подграфа-гипотезы в составе графа-базиса, применительно к задаче проверки исследовательских гипотез в ГИС можно эффективно проиллюстрировать на примере решения рассматриваемой задачи для графов, приведенных на рис. 3. Матрицы смежности и матрица возможных подстановок для этой пары графов приведены на рис. 4. В матрице возможных подстановок  $C$  в первом столбце указаны все вершины графа  $B$ , в первой строке — все вершины графа  $A$ . В столбце  $N + 1$  записывается количество возможных подстановок.

Первоначальное заполнение матрицы возможных подстановок осуществляется путем анализа полустепней исхода и захода исходных графов по правилу: вершина  $B_i$  может соответствовать вершине  $A_j$  только в том случае, если полустепени исхода и захода этой вершины больше, чем у вершины  $A_j$ , или равны. Тем самым производится первая фильтрация вариантов решения. Математически это формализуется и записывается в виде:

$$C_{ij} = \begin{cases} 1, & \text{если } s(B, b_i) \leq s(A, a_i) \text{ и } s(B, b_j) \leq s(A, a_j); \\ 0 & \text{— в остальных случаях.} \end{cases} \quad (3)$$

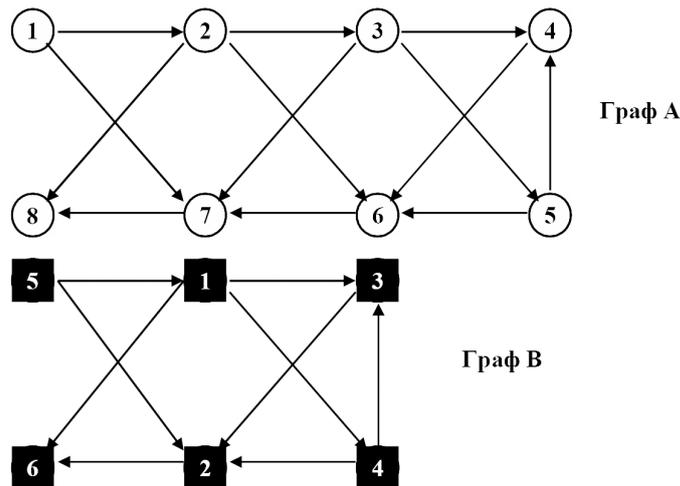


Рис. 3. Пример распознавания изоморфного вложения графа-гипотезы в составе графа-базиса

а)

**Матрица смежности графа А**

	1	2	3	4	5	6	7	8	$S^-(x)$	
1	0	1	0	0	0	0	0	1	0	2
2	0	0	1	0	0	0	1	0	1	3
3	0	0	0	1	1	0	0	1	0	3
4	0	0	0	0	0	0	1	0	0	1
5	0	0	0	1	0	1	0	0	0	2
6	0	0	0	0	0	0	0	1	0	1
7	0	0	0	0	0	0	0	0	1	1
8	0	0	0	0	0	0	0	0	0	0
$S^+(x)$	0	1	1	2	1	3	3	2		

б)

**Матрица смежности графа В**

	1	2	3	4	5	6	$S^-(x)$
1	0	0	1	1	0	1	3
2	0	0	0	0	0	1	1
3	0	1	0	0	0	0	1
4	0	1	1	0	0	0	2
5	1	1	0	0	0	0	2
6	0	0	0	0	0	0	0
$S^+(x)$	1	3	2	1	0	2	

в)

**Матрица возможных подстановок (матрица С)**

	A1	A2	A3	A4	A5	A6	A7	A8	Количество подстановок
B1	0	1	1	0	0	0	0	0	2
B2	0	0	0	0	0	1	1	0	2
B3	0	0	0	1	0	1	1	0	3
B4	0	1	1	0	1	0	0	0	3
B5	1	1	1	0	1	0	0	0	4
B6	0	0	0	1	0	1	1	1	4

Рис. 4. Матрицы смежности графов и возможных подстановок:  
 а — для графа А; б — для графа В;  
 в — вид результирующей матрица подстановок графа В в граф А

Дальнейшая корректировка матрицы производится путем наложения на нее следующей системы ограничений: графы не могут быть изоморфными, если в столбце матрицы «Количество возможных подстановок» имеется хотя бы одна нулевая строка. При однозначном соответствии вершины  $V_i$  вершине  $A_j$  ( $V_i \leftrightarrow A_j$ ), в столбце  $j$  матрицы возможных подстановок ( $C$ ) не должно быть других символов «1». Для удовлетворения требований этого ограничения необходимо исключить из матрицы в столбце  $j$  так называемые *лишние символы* «1». В связи с тем, что найденное соответствие для какой-либо вершины может оказаться ложным, исключение символов «1» из матрицы необходимо производить таким образом, чтобы оставалась возможность восстановления матрицы на определенном шаге. С этой целью введена переменная *Mirage*, значение которой изменяется после каждого цикла наложения системы ограничений на матрицу возможных подстановок. Из ранее изложенного следуют действия, которые математически можно представить так: если  $C[N+1, i] = 1$  и  $C[j, i] = \langle 1 \rangle$  то элементу  $C[k, i]$ , имеющему значение единица (для  $k = 1 \dots m; k \neq i$ ), присвоить текущее значение переменной *Mirage*. Если все значения столбца матрицы «количество возможных подстановок» больше единицы, то для скорейшего нахождения решения, очевидно, необходимо взять для работы строку с наименьшим значением, а первой из неиспользованных ранее ячеек соответствующей строки присваивается так называемый *фокус*, т. е. назначается активная ячейка (определяются координаты вершин подстановки  $C[\text{FocB}, \text{FocA}]$ ). Остальные символы «1» необходимо «закрыть» переменной *Mirage*.

При  $V_i \leftrightarrow A_j$ , вершинам  $B_k$  ( $k = 1 \dots m$ ), смежным с  $B_j$ , могут соответствовать только смежные вершины  $A_l$  ( $l = 1 \dots n$ ) с вершиной  $A_i$ . Для удовлетворения этого требования необходимо так называемые *лишние символы* «1» в матрице возможных подстановок «закрыть» переменной *Mirage*. Математически это можно записать следующим образом:

если  $C[j, i] = \langle 1 \rangle$  и  $B[j, \text{FocB}] = \langle 1 \rangle$  и  $A[j, \text{FocA}] \neq \langle 1 \rangle$  то  $C[j, i] = \text{Mirage}$ ;

если  $C[j, i] = \langle 1 \rangle$  и  $B[\text{FocB}, i] = \langle 1 \rangle$  и  $A[\text{FocA}, i] \neq \langle 1 \rangle$  то  $C[j, i] = \text{Mirage}$ ,

где  $A, B$  — исходные матрицы смежности;

$C$  — матрица возможных подстановок.

Другие частные ограничения учитывают специфику распознавания подграфа-гипотезы в составе графа-базиса применительно к задаче проверки исследовательских гипотез в ГИС.

Приведенный перечень ограничений не является полным и закрытым. В данном случае ограничения выполняют роль логического фильтра. В зависимости от специфики структуры подграфа, описывающего гипотезу, в структуру графа-базиса могут вводиться так называемые *дополнительные фильтры-требования*. Тогда текстуальное описание работы алгоритма распознавания изоморфного вложения подграфа-гипотезы в составе графа-базиса применительно к задаче проверки исследовательских гипотез в ГИС, как работы программы решения задачи нахождения изоморфного вложения графов, можно представить в следующем виде (для примера графов, приведенных на рис. 3):

=> **Начало**

1. Выполнена процедура *Prov\_0\_str*.  $Pr_0 = 0$  переход на *Prov\_1\_str*.

2. *Prov\_1\_str* =>  $Pr_1 = 0$  переход на *Prioritet*.

3. *Prioritet* (с наименьшим количеством подстановок две строки (1-я и 2-я). Выбор первой строки и присвоение ей первого приоритета).  $PrTab[1] = 1$ ;  $Mirage = '2'$ ;  $PrEnd = 0$ ; переход на *Mirage1*.

4. *Mirage1*. Определена активная ячейка  $C[1, 2]$ , т. е. выдвинута гипотеза, что  $B1 \leftrightarrow A2$ . На основании этой гипотезы получается:  $C[1, 3], C[4, 2], C[5, 2] = '2'$ ;  $PrExit = 0$ ;  $Pr_0 = 0$ . Переход на *Mirage3*.

5. *Mirage3*. В связи с тем, что  $B1$  имеет исходы в  $B3, B4, B6$  ( $B1 \rightarrow B3, B4, B6$ ), а  $A2 \rightarrow A3, A6, A8$ , то, следовательно, и вершинам  $B3, B4, B6$  могут соответствовать только вершины из множества  $A3, A6, A8$ . В таком случае в матрице  $C$  получается:  $C[3, 4], C[3, 7], C[4, 2], C[4, 5], C[6, 4], C[6, 7] = '2'$ ; аналогично для  $B \leftarrow B5$  и для  $A2 \leftarrow A1$ , т. е.  $B5 \leftrightarrow A1$ , а  $C[5, 3], C[5, 5] = '2'$ ; переход на *Mirage2*;

6. *Mirage2*.  $C[2, 6], C[6, 6] = '2'$ . Переход на *Prov\_1\_str* => *ZapolnMatrCM* => *ProvEnd*.  $PrExit = 1$  (найденная перестановочная матрица оказалась неверной) =>  $Nvar = 0$  => *ExitToHome* (восстановление матрицы  $C$  в прежнем виде) => *Mirage1*.

7. *Mirage1*. Определена активная ячейка  $C[1, 3]$ , т. е. выдвинута гипотеза, что  $B1 \leftrightarrow A3$ . На основании этой гипотезы получается:  $C[1, 2], C[4,3], C[5,3] = '2'$ ;  $PrExit = 0$ ;  $Pr\_0 = 0$ . Переход на *Mirage3*.

8. *Mirage3*. В связи с тем, что  $B1 \rightarrow B3, B4, B6$ , а  $A3 \rightarrow A4, A5, A7$ , то, следовательно, и вершинам  $B3, B4, B6$  могут соответствовать только вершины из множества  $A4, A5, A7$ . В таком случае в матрице  $C$  получается:  $C[3, 6], C[4, 2], C[6, 6], C[6, 8] = '2'$ ; аналогично для  $B1 \leftarrow B5$  и для  $A3 \leftarrow A2$ , т. е.  $B5 \leftrightarrow A2$ , а  $C[5, 1], C[5, 5] = '2'$ ; переход на *Mirage2*.

9. *Mirage2*. Изменений матрицы  $C$  не происходит, переход на *Prioritet*.

10. *Prioritet*. Активной выбирается строка 4.  $Pr = 2$ ;  $(PrTab[4]=2)$ ;  $Mirage = '3'$ . Переход на *Mirage1*.

11. *Mirage1*. Определена активная ячейка  $C[4,5]$ , т. е. выдвинута гипотеза, что  $B4 \leftrightarrow A5$ . Переход на *Mirage3*.

12. *Mirage3*. В связи с тем, что  $B4 \rightarrow B2, B3$ , а  $A5 \rightarrow A4, A6$ , то, следовательно, и вершинам  $B2, B3$  могут соответствовать только вершины из множества  $A4, A6$ . В таком случае в матрице  $C$  получается:  $C[2, 7], C[3, 7] = '3'$ ; аналогично для  $B4 \leftarrow B1$  и для  $A5 \leftarrow A3$ , т. е.  $B1 \leftrightarrow A3$ , изменений матрицы не происходит; переход на *Mirage2*.

13. *Mirage2*.  $C[6,4] = '3'$ . Переход на *Prov\_1\_str* => *ZapolnMatrCM* => *ProvEnd*.  $PrExit = 1$ . (решение найдено) =>  $Nvar = 1$  => *Prioritet* (попытка поиска автоморфизмов) =>  $ProvEnd = 1$ . => **Конец**.

При этом матрица возможных подстановок будет в процессе выполнения программы в соответствии с описанным алгоритмом принимать вид, последовательно представленный на рис. 5.

а)

	A1	A2	A3	A4	A5	A6	A7	A8	Количество подстановок
B1		1	2						1
B2						2	1		1
B3				2		1	2		1
B4		2	1		2				1
B5	1	2	2		2				1
B6				2		2	2	1	1

б)

	A1	A2	A3	A4	A5	A6	A7	A8	Количество подстановок
B1		2	1						1
B2						1	1		2
B3				1		2	1		2
B4		2	2		1				1
B5	2	1	2		2				1
B6				1		2	1	2	2

в)

	A1	A2	A3	A4	A5	A6	A7	A8	Количество подстановок
B1		2	1						1
B2						1	3		1
B3				1		2	3		1
B4		2	2		1				1
B5	2	1	2		2				1
B6				3		2	1	2	1

Рис. 5. Динамика изменений вида матрицы возможных подстановок: а — на первой итерации; б — на второй итерации; в — в результате выполнения алгоритма

Из приведенной динамики изменений вида матрицы возможных подстановок видно, что корректировки выполнялись на основе выдвинутых трех гипотез соответствия пар вершин и за-

ключались, по сути дела, лишь в так называемой *подгонке матрицы возможных подстановок под требования системы ограничений для конкретного вида графа-гипотезы.*

### Заключение (Conclusion)

Использование алгоритмов и вычислительных процедур теории графов применительно к аппарату геохронологического трекинга дает принципиально новые потенциальные возможности для внедрения строгих математических и расчетных методик в сфере ГИС-задач для ретроспективных исследований. Также очевидна и перспективность дальнейшей адаптации расчетных методов, моделей и методик «мягких» вычислений (использование нечетких множеств и нечетких чисел, функций лингвистической переменной и др.) для ГИС-приложений, применяемых в ходе реализации социологических, статистических и других исследовательских ГИС-проектов. Данный подход уже сегодня является предметом интереса специалистов в области вычислительной математики, современной алгоритмики и геоинформатики, что подтверждается публикациями [10]–[12]. Он может быть интересен специалистам по автоматизации управления водным транспортом, а также исследователям, использующим теорию графов и ее приложения в решении задач водного транспорта, о чем свидетельствуют работы [13]–[15]. Также возможен интерес к предлагаемым приложениям теории графов в смежных областях, что наглядно видно из статей [16], [17].

Дальнейшие направления рационализации алгоритмов проверки гипотез исследования на базе геохронологического трекинга в ГИС связаны с постановкой и решением оптимизационной задачи определения временной сложности указанных алгоритмов, определения строгих граничных условий такой оптимизации и др. Констатация перспективности описанных направлений позволяет прогнозировать широкое внедрение оптимизированных расчетных алгоритмов в ГИС-инструменты поддержки прикладных исследований использования различных объектов водного транспорта.

*Работа выполнена при поддержке РФФИ (проект № 19-07-00006).*

### СПИСОК ЛИТЕРАТУРЫ

1. *Ивакин Ян А.* Информационная технология геохронологического трекинга для проверки гипотез ретроспективных исследований использования водного транспорта / Ян А. Ивакин, С. Н. Потапычев // Вестник Государственного университета морского и речного флота имени адмирала С. О. Макарова. — 2018. — Т. 10. — № 2. — С. 452–461. DOI: 10.21821/2309-5180-2018-10-2-452-461.
2. *Ивакин Я. А.* Геохронологический трекинг – специализированный ГИС-инструментарий исторического исследования / Ян А. Ивакин, С. Н. Потапычев // Историческая информатика. Информационные технологии и математические методы в исторических исследованиях и образовании. — 2016. — № 1-2 (15-16). — С. 3–11.
3. *Нечепуренко М. И.* Алгоритмы и программы решения задач на графах и сетях / М. И. Нечепуренко [и др.]. — Новосибирск: Наука (Сиб. отд-ние), 1990. — 515 с.
4. *Зыков А. А.* Основы теории графов / А. А. Зыков. — М.: Вузовская книга, 2004. — 664 с.
5. *Печенкин В. В.* Прикладные аспекты использования алгоритмов ранжирования для ориентированных взвешенных графов (на примере графов социальных сетей) / В. В. Печенкин [и др.] // Труды СПИИРАН. — 2018. — № 6 (61). — С. 4. DOI: 10.15622/sp.61.4.
6. *Ивакин Ян А.* Проверка гипотез исторического исследования на базе геохронологического трекинга / Ян А. Ивакин, С. Н. Потапычев, В. Я. Ивакин // Историческая информатика. — 2018. — № 1 (23). — С. 86–93. DOI: 10.7256/2585-7797.2018.1.25344.
7. *Воротников В. И.* Метод линеаризующей обратной связи в задаче управления по части переменных при неконтролируемых помехах / В. И. Воротников, А. В. Вохмянина // Труды СПИИРАН. — 2018. — № 6 (61). — С.3. DOI: 10.15622/sp.61.3.
8. *Дюваль П. М.* Непрерывная интеграция. Улучшение качества программного обеспечения и снижение риска / П. М. Дюваль, С. Матиас, Э. Гловер. — СПб.: Символ, 2016. — 240 с.
9. *Курейчик В. В.* Муравьиный алгоритм для решения оптимизационных задач с явно выраженной целевой функцией / В. В. Курейчик, М. А. Жиленков // Информатика, вычислительная техника и инженерное образование. — 2015. — № 2 (22). — С. 10–12.

10. Ajwani D. Average-Case Analysis of Incremental Topological Ordering / D. Ajwani, T. Friedrich // *Discrete Applied Mathematics*. — 2010. — Vol. 158. — Is. 4. — Pp. 240–250. DOI: 10.1016/j.dam.2009.07.006.
11. Ammar A. B. Query optimization techniques in graph Databases // *International Journal of Database Management Systems (IJDMs)*. — 2016. — Vol. 8. — No. 4. DOI: 10.5121/ijdm.2016.8401
12. Sarma A. D. Fast distributed pagerank computation / A.D. Sarma et al // *International Conference on Distributed Computing and Networking*. — Springer, Berlin, Heidelberg, 2013. — Pp. 11–26. DOI: 10.1007/978-3-642-35668-1\_2.
13. Иванов Д. А. Анализ тенденций изменения принципов управления предприятиями в условиях развития технологий индустрии 4.0 / Д. А. Иванов, М. А. Иванова, Б. В. Соколов // *Труды СПИИРАН*. — 2018. — № 5 (60). — С. 97–127. DOI: 10.15622/sp.60.4.
14. Dobrowski J. Real-time Web-based GIS for Analysis, Visualization, and Integration of Marine Transport Environment Data / J. Dobrowski et al // *Information Fusion and Geographic Information Systems*. Springer, Berlin, Heidelberg, 2009. — Pp. 277–288. DOI: 10.1007/978-3-642-00304-2\_19.
15. Codescu M. DO-ROAM: Activity-Oriented Search and Navigation with OpenStreetMap / M. Codescu et al // *International Conference on GeoSpatial Semantics*. — Springer, Berlin, Heidelberg, 2011. — Pp. 88–107. DOI: 10.1007/978-3-642-20630-6\_6.
16. Мичурин С. В. Результативность и качество программных комплексов ситуационного управления для автоматизированных систем диспетчеризации пространственных процессов авиатранспорта / С. В. Мичурин, Я. А. Ивакин // *Информационно-управляющие системы*. — 2016. — № 4 (83). — С. 19–25. DOI: 10.15217/issn1684-8853.2016.4.19.
17. Ивакин Я. А. Иерархия показателей оценки качества программно-аппаратных комплексов центров обработки и хранения данных / Я. А. Ивакин [и др.] // *Вестник Санкт-Петербургского государственного университета технологии и дизайна. Серия I: Естественные и технические науки*. — 2017. — № 5. — С. 65–68.

## REFERENCES

1. Ivakin, Yan A., and Sergei N. Potapychev. “Information technology of geochronological tracking for hypotheses testing in research of ship use.” *Vestnik Gosudarstvennogo universiteta morskogo i rechnogo flota imeni admirala S. O. Makarova* 10.2 (2018): 452–461. DOI: 10.21821/2309-5180-2018-10-2-452-461.
2. Ivakin, Yan A., and Sergey N. Potapychev. “Geochronology tracking - specialized GIS-tool for history researches.” *Istoricheskaya informatika. Informatsionnye tekhnologii i matematicheskie metody v istoricheskikh issledovaniyakh i obrazovanii* 1-2 (15-16) (2016): 3–11.
3. Nechepurenko, M. I., V. K. Popkov, S. M. Mainagashev, S. B. Kaul', V. A. Proskuryakov, V. A. Kokhov, and A. B. Gryzunov. *Algoritmy i programmy resheniya zadach na grafakh i setyakh*. Novosibirsk: Nauka. Sib. otdnie, 1990.
4. Zykov, A. A. *Osnovy teorii grafov*. M: Vuzovskaya kniga, 2004.
5. Pechenkin, Vitaly Vladimirovich, Mikhail Sergeevich Korolev, and Lyubomir Vankov Dimitrov. “Applied aspects of ranking algorithms for oriented weighted graphs (on the example of social network graphs).” *SPIIRAS Proceedings* 6(61) (2018): 4. DOI: 10.15622/sp.61.4.
6. Ivakin, Yan. A., S. N. Potapychev, and V. Ya. Ivakin. “Proverka gipotez istoricheskogo issledovaniya na baze geokhronologicheskogo trekinga.” *Istoricheskaya informatika* 1(23) (2018): 86–93. DOI: 10.7256/2585-7797.2018.1.25344.
7. Vorotnikov, Vladimir Il'ich, and Anastasiya Vladimirovna Vokhmyanina. “Feedback linearization method for problem of control of a part of variables in uncontrolled disturbances.” *SPIIRAS Proceedings* 6(61) (2018): 3. DOI: 10.15622/sp.61.3.
8. Dyugal', P. M., S. Matias, and E. Glover. *Nepriyemlyaya integratsiya. Uluchshenie kachestva programmnogo obespecheniya i snizhenie riska*. SPb.: Simvol, 2016.
9. Kureichik, Vladimir Victorovich, and Mikhail Aleksandrovich Zhilenkov. “ANT COLONY algorithms for solving optimization problems with the explicit objective function.” *Informatika, vychislitel'naya tekhnika i inzhenernoe obrazovanie* 2(22) (2015): 10–21.
10. Ajwani, Deepak, and Tobias Friedrich. “Average-case analysis of incremental topological ordering.” *Discrete Applied Mathematics* 158.4 (2010): 240–250. DOI: 10.1016/j.dam.2009.07.006.
11. Ammar, Ali Ben. “Query optimization techniques in graph Databases.” *International Journal of Database Management Systems (IJDMs)* 8.4 (2016). DOI: 10.5121/ijdm.2016.8401.

12. Sarma, Atish Das, Anisur Rahaman Molla, Gopal Pandurangan, and Eli Upfal. “Fast distributed pagerank computation.” *International Conference on Distributed Computing and Networking*. Springer, Berlin, Heidelberg, 2013. 11–26. DOI: 10.1007/978-3-642-35668-1\_2.

13. Ivanov, Dmitry Alexandrovich, Marina Alexandrovna Ivanova, and Boris Vladimirovich Sokolov. “Analysis of transformation trends in enterprise management principles in the era of industry 4.0 technology.” *SPIIRAS Proceedings* 5(60) (2018): 97–127.

14. Dąbrowski, Jacek, Marcin Kulawiak, Marek Moszyński, Krzysztof Bruniecki, Łukasz Kamiński, Andrzej Chybicki, and Andrzej Stepnowski. “Real-time web-based GIS for analysis, visualization and integration of marine environment data.” *Information Fusion and Geographic Information Systems*. Springer, Berlin, Heidelberg, 2009. 277–288. DOI: 10.1007/978-3-642-00304-2\_19.

15. Codescu, Mihai, Gregor Horsinka, Oliver Kutz, Till Mossakowski, and Rafaela Rau. “DO-ROAM: Activity-oriented search and navigation with OpenStreetMap.” *International Conference on GeoSpatial Semantics*. Springer, Berlin, Heidelberg, 2011. 88–107. DOI: 10.1007/978-3-642-20630-6\_6.

16. Michurin, S. V., and Ya. A. Ivakin. “The Effectiveness and Quality of Situational Management Software for Automated Dispatching of Air Transport Spatial Processes.” *Information and Control Systems* 4(83) (2016): 19–25. DOI: 10.15217/issn1684-8853.2016.4.19.

17. Ivakin, Ya. A., E. G. Semenova, S. A. Morozov, and M. G. Smirnova. “Hierarchy of indicators for assessing the quality of software and hardware complexes of data processing and storage centers.” *Vestnik of St. Petersburg State University of Technology and Design. Series 1. Natural and technical science* 5 (2017): 65–68.

#### ИНФОРМАЦИЯ ОБ АВТОРАХ

#### INFORMATION ABOUT THE AUTHORS

**Ивакин Ян Альбертович —**

доктор технических наук, профессор, ведущий научный сотрудник  
Санкт-Петербургский институт информатики и автоматизации Российской академии наук (СПИИРАН)  
199178, Российская Федерация, г. Санкт-Петербург, 14-я линия В. О., 39  
АО «Концерн «ОКЕАНПРИБОР»  
198226, Российская Федерация, г. Санкт-Петербург, Чкаловский пр., 46  
e-mail: [ivakin@oogis.ru](mailto:ivakin@oogis.ru)

**Потапычев Сергей Николаевич —**

кандидат технических наук, старший научный сотрудник  
Санкт-Петербургский институт информатики и автоматизации Российской академии наук (СПИИРАН)  
199178, Российская Федерация, г. Санкт-Петербург, 14-я линия В. О., 39  
АО «Концерн «ОКЕАНПРИБОР»  
198226, Российская Федерация, г. Санкт-Петербург, Чкаловский пр., 46  
e-mail: [potapychev@mail.ru](mailto:potapychev@mail.ru)

**Ивакин Роман Янович —**

инженер-исследователь НИИ «Технологии связи» Санкт-Петербургский государственный университет телекоммуникаций имени профессора М. А. Бонч-Бруевича  
193232, Российская Федерация, г. Санкт-Петербург, пр. Большевиков, д. 22, корп. 1  
e-mail: [romanivakin2018@yandex.ru](mailto:romanivakin2018@yandex.ru)

**Ivakin, Yan A. —**

Dr. of Technical Sciences, professor, leading researcher  
St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences  
39 14<sup>th</sup> liniya V.O., St. Petersburg, 199178, Russian Federation  
Concern OCEANPRIBOR JSC  
46 Chekalovskiy Av., St. Petersburg, 198226, Russian Federation  
e-mail: [ivakin@oogis.ru](mailto:ivakin@oogis.ru)

**Potapychev, Sergei N. —**

PhD, senior researcher  
St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences  
39 14<sup>th</sup> liniya V.O., St. Petersburg, 199178, Russian Federation  
Concern OCEANPRIBOR JSC  
46 Chekalovskiy Av., St. Petersburg, 198226, Russian Federation  
e-mail: [potapychev@mail.ru](mailto:potapychev@mail.ru)

**Ivakin, Roman Ya. —**

Engineer-Researcher of SRI “Technologies of Telecommunication”  
The Bonch-Bruевич Saint Petersburg State University of Telecommunications  
22/1 Bolshevnikov Av., St. Petersburg, 193232, Russian Federation  
e-mail: [romanivakin2018@yandex.ru](mailto:romanivakin2018@yandex.ru)

Статья поступила в редакцию 21 января 2019 г.  
Received: January 21, 2019.