

УДК 004.6

Р. А. БАДАМШИН, А. С. ПАВЛОВ**МНОГОПОЛЬЗОВАТЕЛЬСКАЯ ОБРАБОТКА
РАСПРЕДЕЛЕННО-ХРАНЯЩЕЙСЯ
ПРОСТРАНСТВЕННОЙ ИНФОРМАЦИИ
В НАУЧНО-ОБРАЗОВАТЕЛЬНОЙ ГИС РБ**

Описывается проблема организации многопользовательской обработки распределенно-хранящейся пространственной информации. *Геоинформационная система ; инфраструктура пространственных данных ; геопортал ; метаданные ; пространственные данные*

Своевременное обеспечение органов государственной власти, научных исследований, потребностей промышленности и населения достоверными и непротиворечивыми пространственными данными (ПД) о своей территории позволяет организовать эффективное управление процессами и явлениями, происходящими в большинстве сфер жизнедеятельности.

Одним из эффективных решений этой проблемы в мировой практике является создание инфраструктуры пространственных данных (ИПД) как элемента единого информационного пространства страны для формирования эффективной системы производства и предоставления пользователям информационных ресурсов пространственных данных. Начиная с середины 90-х гг. прошлого века уже более чем в 120 странах создаются национальные инфраструктуры пространственных данных [3, 4].

Российская Федерация также приступила к созданию ИПД – иерархической территориально-распределенной системы сбора, обработки, хранения пространственных данных на основе ГИС-технологий, которая призвана объединить и обеспечить пользователям регламентированный удаленный доступ к пространственным данным на трех уровнях: федеральном, территориальном (уровень субъектов РФ) и муниципальном [5]. Естественным образом из этого вытекает актуальность создания ИПД в Республике Башкортостан (РБ) как подсистемы уровня субъекта Российской Федерации.

1. СОСТОЯНИЕ ВОПРОСА

Как правило, при создании ИПД выделяют ряд основных компонентов, наиболее существенными из которых являются пространственные информационные ресурсы: базовые пространственные данные (БПД), описывающие базовые пространственные объекты и их наборы; специализированные пространственные данные (СПД), описывающие пространственные объекты и их наборы, представляющие интерес для одного конкретного предприятия или органа государственной власти и необходимые для их деятельности; метаданные – данные о специализированных и базовых пространственных данных, предназначенные для поиска, оценки, назначения, пригодности и возможности обработки пространственных данных. Для определения информации, входящей в состав ИПД РБ, был проведен анализ пространственных данных, используемых органами государственной власти и крупными предприятиями в повседневной деятельности [1, 2], который позволил построить обобщенную структуру пространственных данных (рис. 1), состоящую из блока общегеографических пространственных данных и блоков пространственных данных по основным видам жизнедеятельности.

Большая часть специализированных пространственных данных и метаданные о них зачастую являются собственностью коммерческих предприятий, их создание связано с существенными материальными затратами, а использование ограничено определенным уровнем регламентированного доступа.

Исходя из вышесказанного, становится ясно, что возможность получения даже части специализированных данных, принадлежащих

коммерческим предприятиям, или организация доступа к ним в рамках ИПД РБ представляются, по крайней мере, на ранней стадии, весьма затруднительными и дорогостоящими.

2. ПОСТАНОВКА ЗАДАЧИ

Для отработки научных и технологических принципов создания ИПД количество баз данных (БД) СПД и их ведомственная принадлежность не играют существенной роли, в то же время у ученых, изучающих территорию РБ и протекающие на ней различные процессы (природные, техногенные, социальные, экономические и др.), существует актуальная потребность, и даже необходимость, в активном обмене и совместном использовании пространственных данных (в виде различных тематических карт РБ).

Поэтому предлагается в качестве прототипа ИПД РБ разработать научно-образовательную

геоинформационную систему Республики Башкортостан (ГИС РБ), позволяющую организовать многопользовательский доступ большого числа пользователей (ученых, аспирантов, студентов и др.) к распределенно-хранящейся в различных научных организациях и образовательных учреждениях пространственной информации.

На рис. 2 представлена структура предлагаемой научно-образовательной ГИС РБ, включающая:

- БД СПД научных и образовательных учреждений РБ, которые содержат информацию, полученную в ходе выполнения научно-исследовательской, образовательной и хозяйственной деятельности;
- БД о территории РБ некоторых масштабов;
- метаданные о БД и СПД.

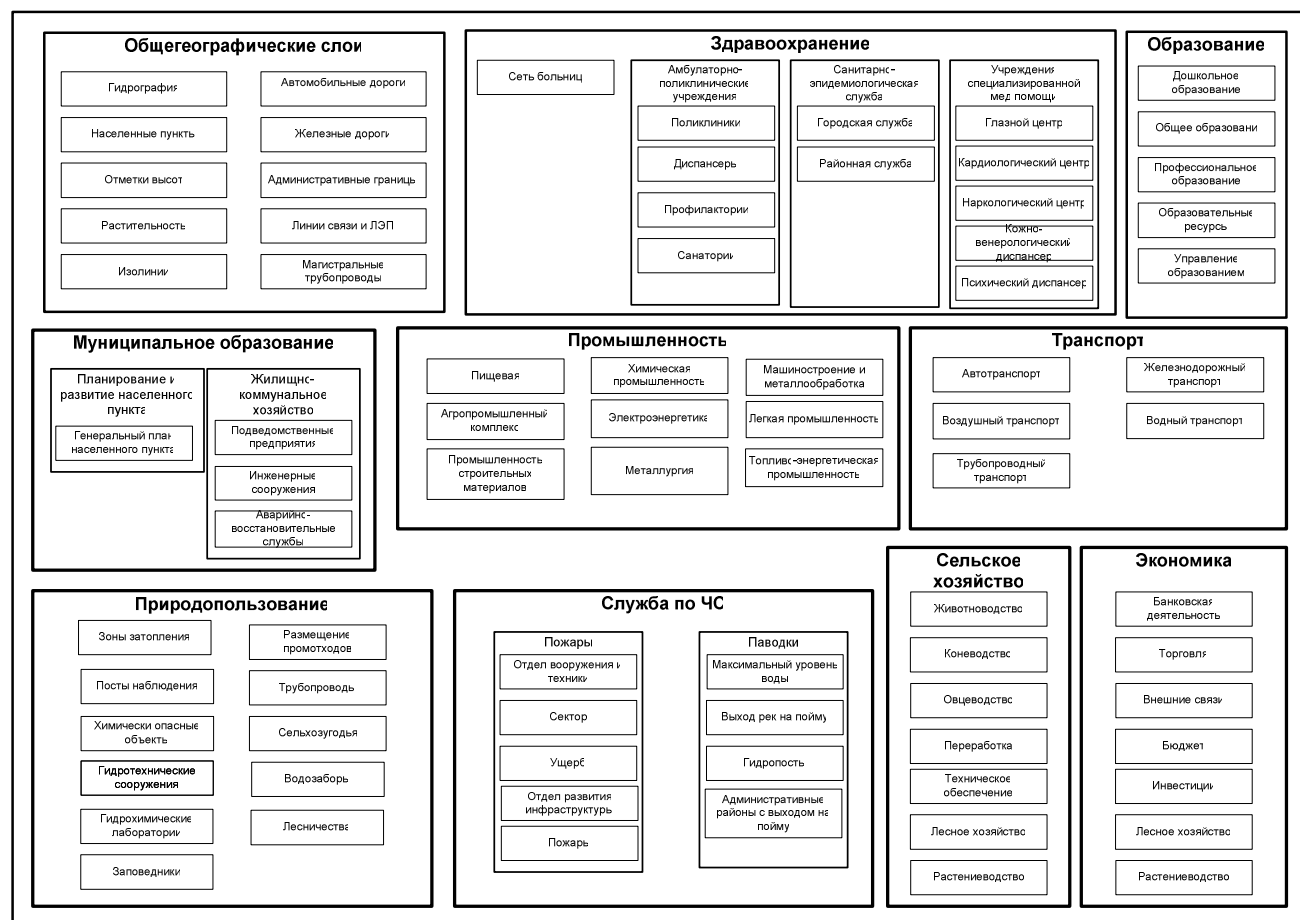


Рис. 1. Обобщенная структура пространственных данных, необходимых для деятельности органов государственной власти и крупных предприятий РБ

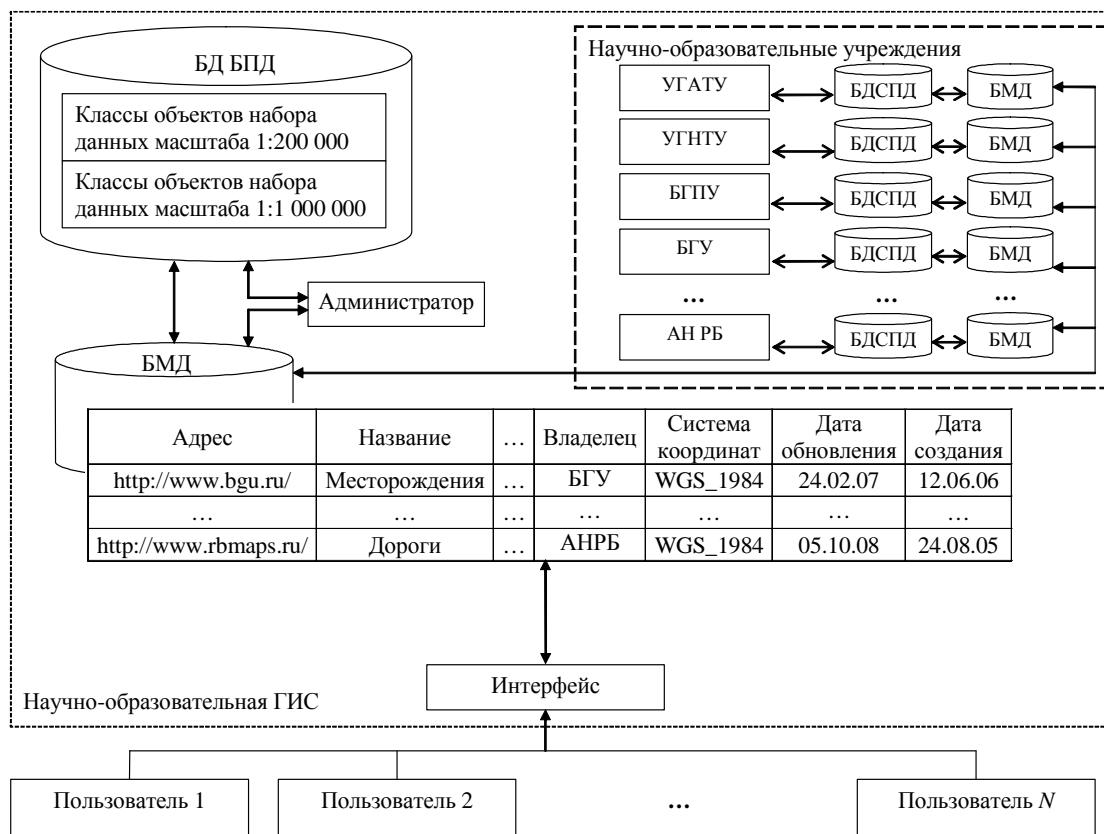


Рис. 2. Структура научно-образовательной ГИС РБ

Поскольку часть информации, содержащейся в БД СПД, может носить конфиденциальный характер или интересовать узкий круг специалистов, предлагается организация удаленного многопользовательского доступа следующим образом (рис. 3). Возможность полного доступа к пространственным данным, хранящимся в БД СПД научно-образовательного учреждения, предоставляется пользователям, имеющим доступ к соответствующей локальной сети. Возможность многопользовательского удаленного доступа к базе метаданных (БМД), БД БПД, а также некоторой части пространственных данных, хранящихся в БД СПД, предоставляется пользователям, обращающимся через интернет к геопорталу.

Геопортал – географический информационный узел, предназначенный для поиска, просмотра и обмена пространственными данными. Одной из основных составляющих геопорталов является база метаданных, которая позволяет пользователям осуществлять поиск пространственных данных по необходимым критериям, определять место расположения их источника, получать доступ к этим источникам.

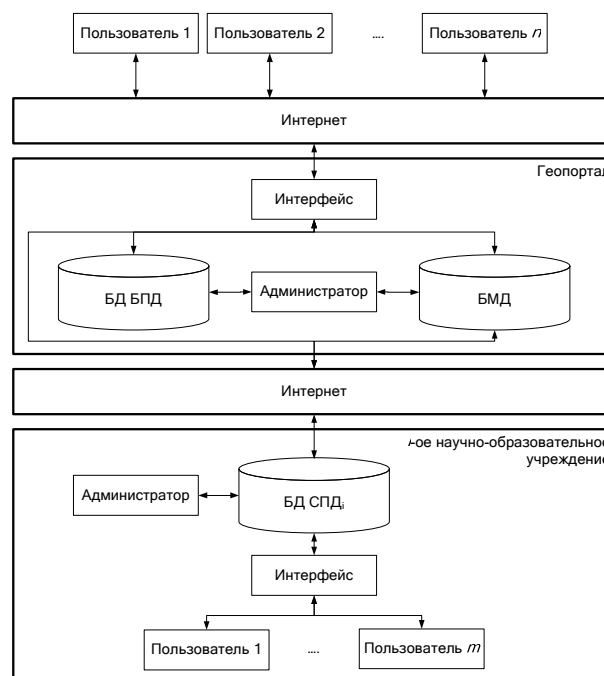


Рис. 3. Схема организации многопользовательского доступа к пространственным данным в научно-образовательной ГИС РБ

Для создания базы метаданных необходимо выявить общие для всех БД пространственные данные, провести анализ требований к этим

данным всех пользователей научно-образовательной ГИС РБ и исходя из этого определиться с местом хранения той или иной части данных. При этом в качестве основных требований выдвигаются: непротиворечивость данных, что может быть достигнуто исключением (или хотя бы минимизацией) дублирования данных (с точки зрения хранения), время получения пользователем информации о пространственных данных, а также объем занимаемой памяти.

3. ПОДХОД К РЕШЕНИЮ

Для формализованного описания всех частей пространственной информации и связей между ними введем обозначение – базу данных конкретного i -го научного или образовательного учреждения РБ обозначим через $БД_i$, где $i = \overline{1, n}$, n – число $БД_i$, входящих в состав ГИС РБ. В $БД_i$ выделим две части: $БД_i^1$ – та часть $БД_i$, которая входит в состав БД БПД, $БД_i^2$ – часть $БД_i$, входящая в состав БД СПД. Таким образом, получаем:

$$БД_i = БД_i^1 \cup БД_i^2, i = \overline{1, n}. \quad (1)$$

Следовательно, можно записать, что база данных БПД (которую обозначим $БД_0$) на всю республику является совокупностью $БД_i^1$:

$$БД_0 = \bigcup_{i=\overline{1, n}} БД_i^1, \quad (2)$$

Пусть $БД_i^1$ для всех $БД_i$ представлены в виде совокупности отношений

$$БД_i^1 = \{R_i^1, R_i^2, \dots, R_i^{m_i}\}, i = \overline{1, n}, \quad (3)$$

m_i – количество отношений $БД_i$. Тогда $БД_0$ можно представить как совокупность отношений, входящих хотя бы в две из $БД_i^1$,

$$БД_0 = \bigcup_{\substack{i=\overline{1, n} \\ j=\overline{1, m_i}}} R_i^j. \quad (4)$$

Предположим, что информация об одних и тех же сущностях, а также одни и те же связи между одинаковыми сущностями в различных БД входят с одним и тем же значением индекса. Учитывая это предположение, переобозначим отношения из (3), входящие в состав $БД_0$, следующим образом:

$$БД_0 = \{R_*^1, R_*^2, \dots, R_*^M\} \quad (5)$$

где $R_*^i, i = \overline{1, M}$ – полный неповторяющийся перечень, а M – количество различных отношений из $\{БД_i^1\}_{i=\overline{1, n}}$, входящих в $БД_0$,

$$R_*^i \neq R_*^j, \text{ при } i \neq j \quad (6)$$

$$\bigcup_{i=\overline{1, M}} R_*^i = \bigcup_{\substack{i=\overline{1, n} \\ j=\overline{1, m_i}}} R_i^j \quad (7)$$

Пусть схема каждого k -го отношения (3) представлена в виде совокупности атрибутов

$$S(R_*^k) = \{a_{k*}^1, a_{k*}^2, \dots, a_{k*}^K\} \quad (8)$$

где $a_{k*}^j, k = \overline{1, M}, j = \overline{1, K}$ – множество упорядоченных некоторым образом атрибутов всех отношений $R_i^k, i = \overline{1, m}$, так что

$$\bigcup_{j=\overline{1, M}} a_{k*}^j = \bigcup_{\substack{j=\overline{1, m_i} \\ i=\overline{1, n}}} a_{ik}^j, k = \overline{1, M}. \quad (9)$$

Заметим, что

$$m_i \leq M \text{ для всех } i = \overline{1, n}, \quad (10)$$

поэтому пробегание индекса k от 1 до M в (10) корректно. Выражение (2) и способ построения БПД, описываемый соотношениями (2) – (9), задают соответствующее преобразование баз данных

$$F_i : БД_i^1 \rightarrow БД_0, i = \overline{1, n}, \quad (11)$$

позволяющее получить информацию о составе и структуре БД БПД и БД СПД, входящих в состав научно-образовательной ГИС РБ, на основании которой будет создана БМД.

Поскольку со временем число источников информации, входящих в состав научно-образовательной ГИС РБ, может увеличиться, предложен алгоритм добавление q -го источника ПД (рис. 4), реализующий соотношения (2–9), предназначенный для автоматизации выявления общей для всех или нескольких БД пространственной информации, и оперативного изменения состава и структуры БД БПД, а также занесения информации о соответствующих изменениях в БМД. При появлении нового q -го источника информации осуществляется сравнение отношений, входящих в $БД_q$ и $БД_0$. Если некоторая часть отношений $БД_q$ уже присутствует в $БД_0$, она принимается за $БД_q^1$ в соотношении (1), а оставшаяся часть принимается за $БД_q^2$. После этого сравниваются отношения из $БД_q^2$ со всеми отношениями из $БД_i, i = \overline{1, n}$ и определяются совпадающие отношения, которые также переносятся в $БД_0$, а информация о соответствующих изменениях состава и структуры $БД_0, БД_q, БД_i$ заносится в БМД. Если же совпадающий отношений в $БД_q$ и $БД_0$ нет, то $БД_q^2 = БД_q$, и количество БД СПД увеличивается на единицу $n = n + 1$.

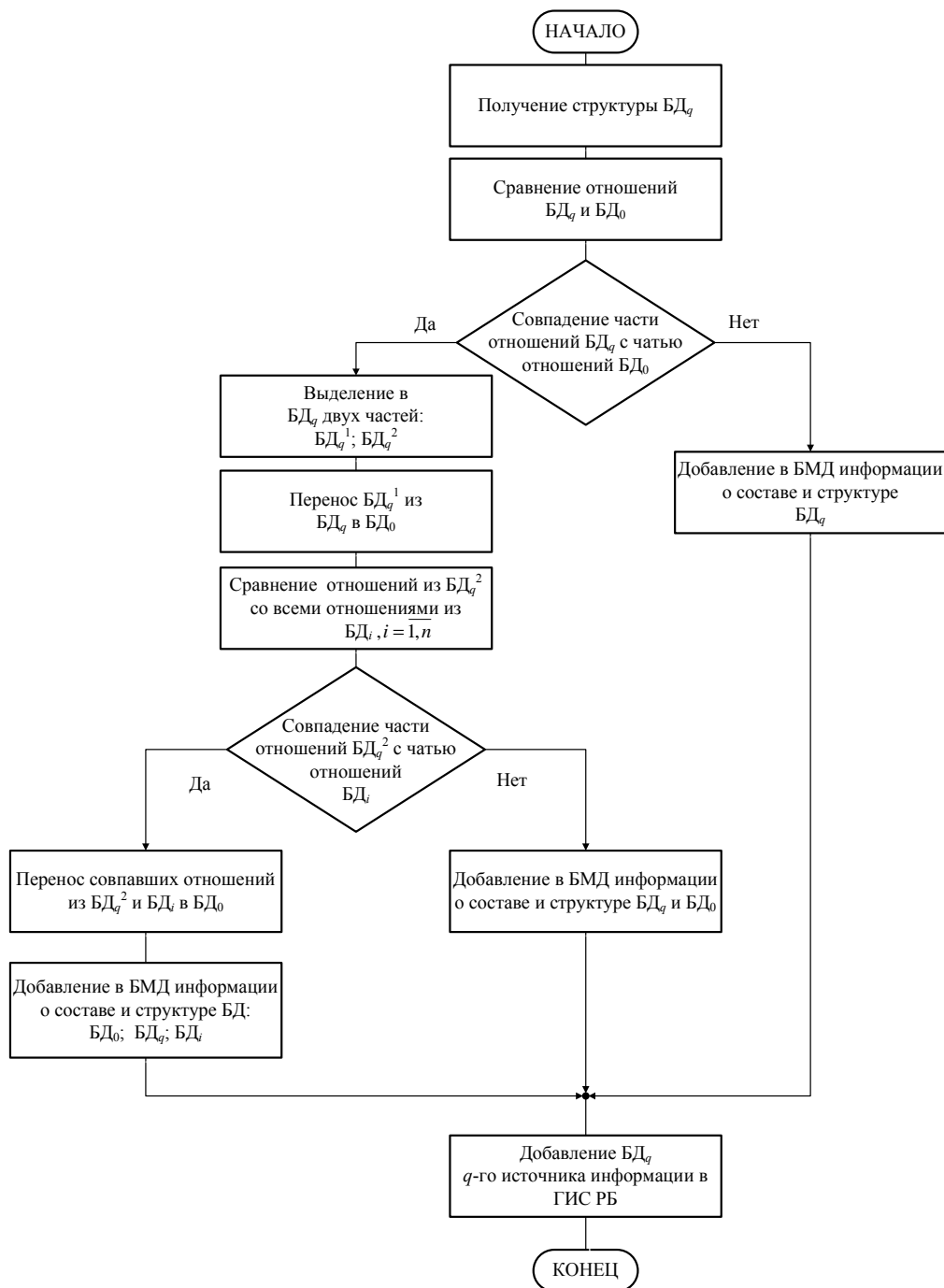


Рис. 4. Алгоритм добавления q -го источника ПД в научно-образовательную ГИС РБ

ЗАКЛЮЧЕНИЕ

Предложен алгоритм формирования и обновления БД БПД основанный на формальном анализе структур БД СПД, входящих в состав научно-образовательной ГИС РБ, позволяющий исключить дублирование и организовать многопользовательский доступ к распределенно-хранящимся пространственным данным.

СПИСОК ЛИТЕРАТУРЫ

1. Бахтизин, Р. Н. Информационная модель нефтегазового комплекса в составе инфраструктуры пространственных данных Республики Башкортостан / Р. Н. Бахтизин, С. В. Павлов, Г. М. Сайфутдинова, А. С. Павлов // Нефтегазовое дело : науч.-техн. журнал. Уфа : Нефтегазовое дело, 2007. Т. 5. С. 25–30.
2. Бадамшин, Р. А. Применение геоинформационных технологий для создания инфраструктуры пространственных данных Республики Башкортостан / Р. А. Бадамшин, Р. Н. Бахтизин, С. В. Павлов,

Г. М. Сайфутдинова, А. С. Павлов // Межвуз. науч. сб. «Геоинформационные технологии в проектировании и создании корпоративных информационных систем». Уфа : УГАТУ, 2008. С. 56–66.

3. **Кошкарев, А. В.** США – национальная инфраструктура пространственных данных / А.В. Кошкарев // ГИСинфо. 2005. № 4(10). С.27–30.

4. **Кошкарев, А.В.** Инфраструктура пространственных данных Финляндии / А. В. Кошкарев // Пространственные данные. 2008. № 1. С.7–17.

5. **Концепция** создания и развития инфраструктуры пространственных данных Российской Федерации, одобренная распоряжением Правительства Российской Федерации от 21 августа 2006 г. № 1157-р.

ОБ АВТОРАХ



Бадамшин Рустам Ахмарович, проф. каф. техн. киберн., проректор по науч. и иннов. деят-ти УГАТУ. Дипл. инж.-электромех. (УГАТУ, 1973). Д-р техн. наук (УГАТУ, 2000). Действ. чл. Междунар. акад. наук высш. шк.



Павлов Александр Сергеевич, асп. той же каф. Дипл. магистр техники и технологии (УГАТУ, 2006). Иссл. в обл. геоинформационных систем.