

Методы контроля данных при обновлении границ лесхозов и авиаотделений для решения задач аэрокосмического мониторинга лесных пожаров

А.С. Подольская, Д.В. Ершов

*Центр по проблемам экологии и продуктивности лесов РАН,
117997, Москва, Профсоюзная ул., 84/32
E-mail: alexandra@ifi.rssi.ru*

В геоинформационной технологии ежегодного обновления границ лесхозов и авиаотделений предусмотрен блок контроля данных. После внесения пользователем изменений в атрибутивную и справочную информацию цифровых покрытий производственно-территориального деления лесного фонда осуществляется контроль введенных данных для предотвращения возможных ошибок. Например, появление разных подразделений с одинаковыми атрибутами, несогласованность в длине и структуре кода, которая может привести к некорректному определению административной принадлежности подразделения.

Моделирование вероятных ошибок при внесении изменений позволило разработать методы контроля данных и комплекс программных мер, предназначенный для выявления ложной информации до формирования новых и архивирования существующих цифровых покрытий.

Поля атрибутивных таблиц и справочников имеют различные типы в зависимости от содержащихся в них данных - числовые и строковые. Учитывая этот факт, было выделено два вида контроля данных – контроль кодировки и названий.

Контроль кодировки происходит на 3-х уровнях: ограничение длины кода, контроль состава кода и проверка характера символов. Для проведения контроля названий осуществляется их сравнение с несколькими наборами уникальных словосочетаний, содержащих собственные названия заповедников, национальных парков и принятые в лесной отрасли сокращения. Дополнительно реализована проверка разрешенных в названии символов. Например, названия подразделений должны содержать только русские буквенные символы, а также «», -, ., /, () и пробел.

Введение

Для осуществления геоинформационной поддержки работ по обнаружению и тушению лесных пожаров, предоставления информации и технологий для анализа последствий лесных пожаров в авиационную службу охраны лесов от пожаров ФГУ «Авиалесоохрана» создана информационная система дистанционного мониторинга лесных пожаров (ИСДМ МПР РФ). Обнаруженные по спутниковым данным лесные пожары пространственно идентифицируются в соответствии с границами административных районов, лесхозов, авиаподразделений и зон мониторинга, которые должны быть актуальными по состоянию на начало пожароопасного сезона.

Среди основных функций геоинформационной технологии обновления границ лесхозов и авиаотделений, таких как сбор информации об изменениях, внесение текущих изменений, формирование производных цифровых покрытий после внесения изменений, автоматизированное формирование цифровых карт выделяется контроль данных (рис. 1).

Как видно из представленной блок-схемы технологии контроль данных применяется к атрибутивной и нормативно-справочной информации. Таблицы справочников авиаподразделений и лесхозов являются производными от атрибутивной таблицы цифрового покрытия границ [1]. В них сохраняются названия, типы и размерности полей, поэтому используются те же методы контроля данных, что и для атрибутивной таблицы.



Рис. 1. Блок-схема технологии

Виды контроля данных

Проанализировав структуру атрибутивной таблицы (Табл. 1) были выделены поля двух типов – числовые и строковые [2]. Однако несколько строковых полей (Fe_code, Ao_code и Ab_code) содержат числовую информацию. Данный факт связан с тем, что кодировка лесхозов, авиаподразделений и авиабаз до десятого номера начинается с нулевого символа. Например, код лесхоза 012310. Таким образом, контроль данных разделяется на два вида – контроль кодов и названий.

Таблица 1. Структура атрибутивной таблицы

Название поля	Тип поля	Размерность поля	Описание
Obl_name*	строковый	38	Название субъекта РФ
Name_rus*	строковый	38	Название административных районов субъекта
Adm_code	числовой	4	Код субъекта
Fe_name	строковый	50	Название лесхоза
Fe_code	строковый	6	Код лесхоза
Fe_gul	числовой	8	Код лесхоза по государственному учету леса
Ao_name	строковый	50	Название авиаподразделения
Ao_code	строковый	6	Код авиаподразделения
Ab_code	строковый	2	Код авиабазы

* данные этих полей не контролируются

Контроль кодов

1. Контроль состава кода

- Проверка правильности кодирования лесхозов в соответствии с административными кодами субъектов;

Любому лесхозу поставлено в соответствие два кода: код ведомства Рослесхоза и код-ГУЛ (государственного учета лесного фонда). ГУЛ-код имеет восьмизначную кодировку, первые четыре цифры которой взяты из административного кода субъекта РФ. На данном этапе контроля проверяется совпадение первых четырех цифр ГУЛ-кода и административного кода субъекта, на территории которого расположен лесхоз [3].

- Проверка правильности кодирования авиаподразделений в соответствии с номером авиабазы;

Авиаподразделения кодируются составным кодом. Принцип его формирования: первые две цифры – номер авиабазы, следующие две – номер авиазвена (если на территории авиабазы выделены авиазвенья, в противном случае присваивается 00), последние две – собственный номер авиаотделения. Контроль данных - две первые цифры кода авиаподразделения должны быть равны номеру авиабазы из поля Ab_code.

2. Ограничение длины кода

Лесхозы, авиаподразделения и авиабазы имеют фиксированную длину кода. А именно, лесхозы – код министерства имеет 6 символов и ГУЛ-код 8 символов, авиаподразделения – 6 символов, авиабазы – 2 символа. Длина введенного пользователем кода должна быть не более и не менее указанной длины (Табл. 1) [4].

3. Проверка характера символов при кодировании

При осуществлении проверки характера символов критерием правильности кодировки служит тип поля. Числовые поля, содержащие коды, могут состоять только из цифр.

Для дополнительного ограничения пользователя при редактировании кодов во избежание ложной информации запрещен ввод любых символов кроме цифр.

На рис. 2 представлена блок-схема проведения контроля данных для нормативно-справочной информации. Условные обозначения: I = 1 – справочник лесхозов, I = 2 – справочник авиаотделений; J – обозначения кодов, для лесхозов J = 1 - “Adm_code”, J = 2 - “Fe_code”, J = 3 - “Fe_gul”, для авиаотделений J = 1 - “Adm_code”, J = 2 - “Ao_code”, J = 3 - “Ab_code”.

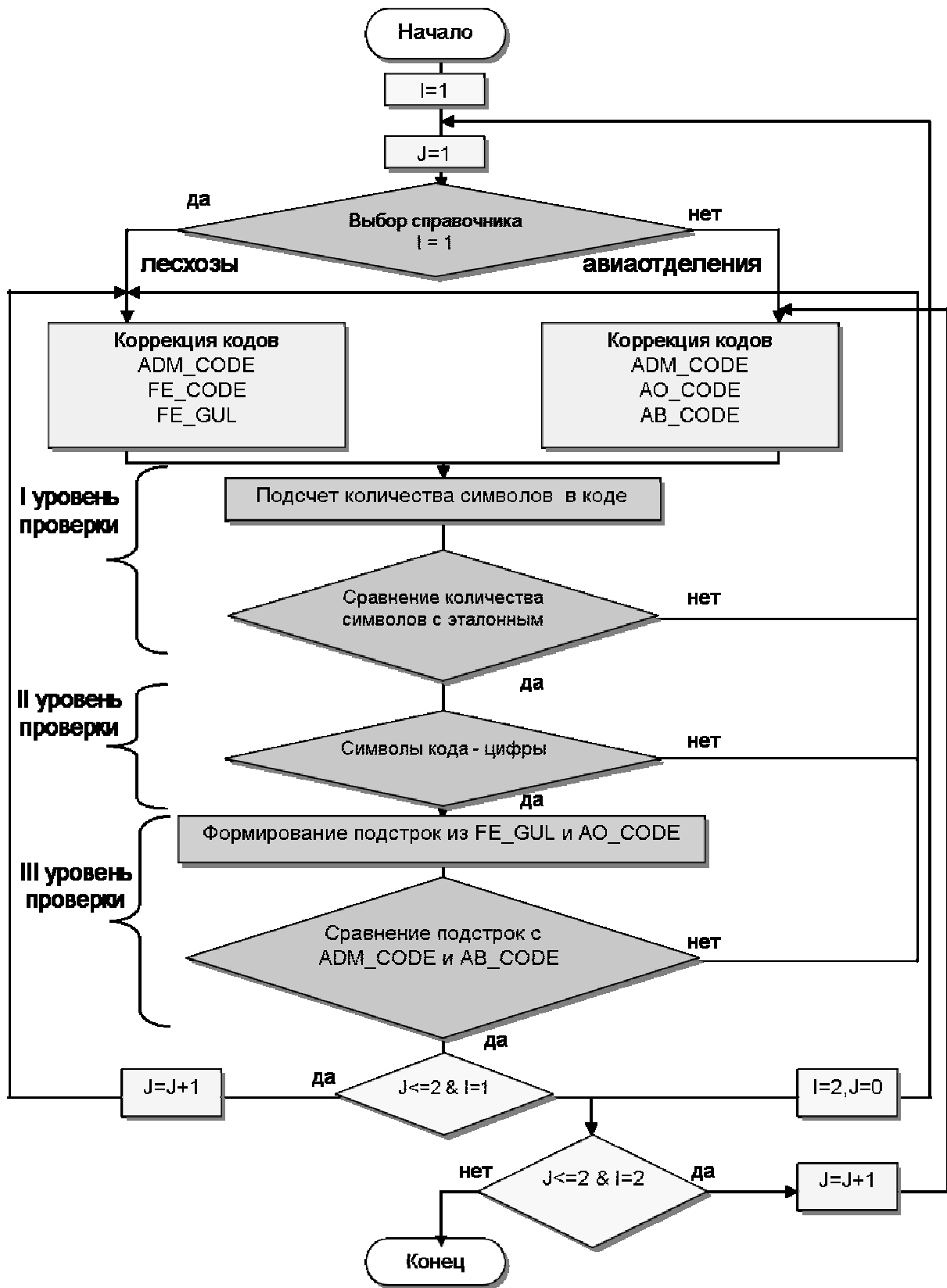


Рис. 2. Блок-схема контроля кодов

Контроль названий

1. Контроль разрешенных символов

Разрешенными символами для полей, содержащих названия, являются буквы русского алфавита, «», -, ., /, () и пробел.

2. Сравнение исходного названия с набором уникальных словосочетаний

Проанализировав названия лесхозов и авиаподразделений в справочниках, было принято решение о создании трех наборов уникальных словосочетаний. Первый набор состоит из десяти названий, таких как "Леспаркхоз 'Кумысная Поляна'", "ОПЛХ 'Могутовский Лес'", "ОЛПО 'Бузулукский Бор'" и т.д. Второй набор, в состав которого вошли принятые в лесной отрасли сокращения: "ЗП", "ОПЛХ", "ОЛПО" и "НЦП". Третий набор: "ЛХ", "лесхоз-техникум", "опытное лесное хозяйство" и т.д.

При контроле названий происходит обнаружение (первого) символа пробел « » и формирование 2-х подстрок - до пробела и после пробела [5]. Далее подстроки сравниваются со вторым и третьим наборами уникальных словосочетаний.

Дополнительный контроль - выявление дублированных записей

Каждое подразделение (лесхоз, авиаотделение) имеет уникальный код, с помощью которого возможна его идентификация. Помимо этого, на территории одного субъекта не могут находиться лесхозы или авиаподразделения с одинаковыми названиями. Поэтому проводимый контроль данных должен препятствовать появлению дублированных кодов и названий для различных подразделений.

Примеры реализации контроля кодов

Индикатором проведенного контроля данных в случае обнаружения ошибки является заливка красным цветом соответствующего поля (рис. 3), при отсутствии ошибочных данных поля заливаются зеленым цветом (рис. 4). При обнаружении ошибки в данных возникает сообщение с описанием характера ошибки и указанием имя поля (рис. 3, 4, 5). Интерфейс редактирования спроектирован таким образом, что контроль данных обязателен после внесения изменений в каждую запись.

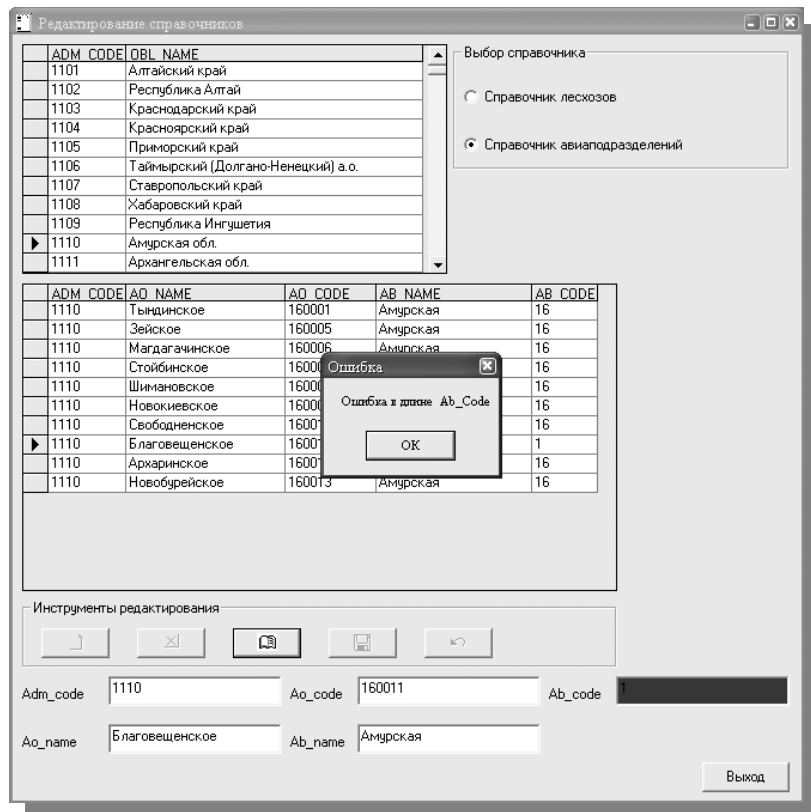


Рис. 3. Уровень: контроль количества символов в коде – номер авиабазы представляет собой двузначный код

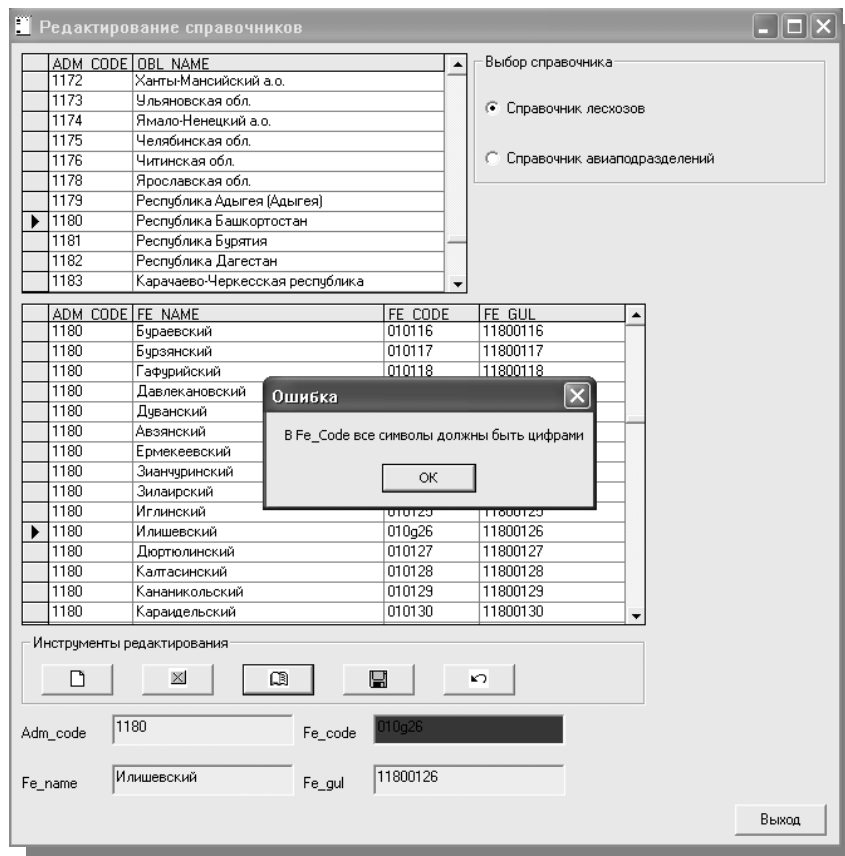


Рис. 4. II уровень: контроль характера символов в коде
Недопущение ввода никаких символов кроме цифр

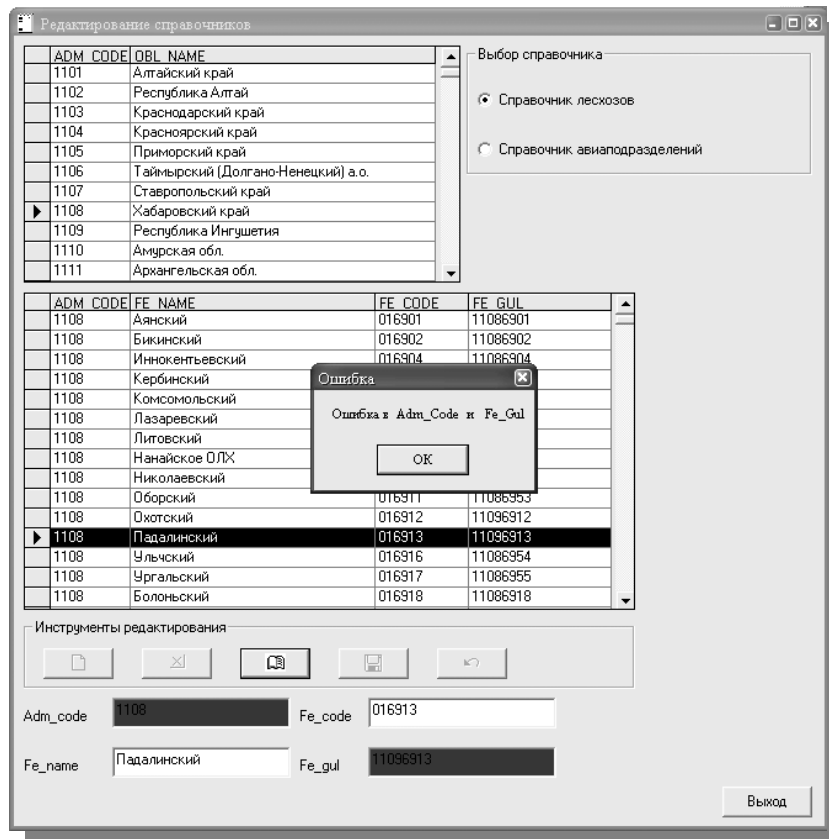


Рис. 5. III уровень: контроль состава кода

Заключение

Результатом проведения контроля данных является устранение ошибочных данных при пользовательском редактировании атрибутивной и справочной информации. Например, неправильная кодировка при создании нового лесхоза может повлечь за собой его ошибочную принадлежность к другому субъекту. Данная процедура носит предупредительный характер, выводя сообщения для пользователя об обнаруженных ошибках.

Литература

1. Подольская А.С., Ершов Д.В. Геоинформационная технология обновления границ лесхозов и авиаотделений для решения задач аэрокосмического мониторинга лесных пожаров // Современные проблемы дистанционного зондирования Земли из космоса: Физические основы, методы и технологии мониторинга окружающей среды, потенциально опасных явлений и объектов. Сборник научных статей.– М.: ООО «Азбука-2000», 2006. Выпуск 3. Том I. - С. 366-372.
2. Кузьменко В.Г. Базы данных в Visual Basic и VBA. Самоучитель. – М.: ООО «Бином-Пресс», 2004 г. – 416 с.
3. Microsoft Visual Basic 6.0 для профессионалов. Шаг за шагом.: Практик. пособие / Пер. с англ. М.: Изд-во ЭКОМ, 2004. 720 с.
4. Visual Basic 6.0 / Пер. с нем. К.: Издат. группа ВНУ, 2000. 288 с.
5. Visual Basic 6.0 / Пер. с англ. СПб.: БХВ-Петербург, 2003. 992 с.