

Информационная электронная орнитологическая система.

К. В. Большаков

Биологическая станция "Рыбачий" Зоологического института РАН, rybachy@bioryb.koenig.su

М. Б. Дианов

Зоологический институт РАН, mix@zin.ru

Н. П. Зеленова

Биологическая станция "Рыбачий" Зоологического института РАН, rybachy@bioryb.koenig.su

А. Л. Лобанов

Зоологический институт РАН, all@zin.ru

1 Введение

Биологическая станция "Рыбачий" ЗИН РАН на Куршской косе Балтийского моря со времени своего основания в 1956 году проводит широкий спектр орнитологических исследований. За 43 года своего существования она накопила огромный фактический материал по экологии, миграциям, гнездованию, физиологии и многим другим аспектам биологии более двухсот видов птиц. Сотрудниками станции опубликовано более семисот научных работ, затрагивающих различные аспекты орнитологических исследований.

К сожалению, существование таких уникальных по объему и широте охвата данных до сих пор является мало доступным мировому научному сообществу как из-за недостатка информации, так и из-за языковых барьеров, не позволяющих многим нашим зарубежным коллегам найти литературу по заданной теме на русском языке и познакомиться с ней.

Создаваемая нами в настоящее время информационная система предназначена для того, чтобы:

1) предоставить заинтересованным пользователям возможность оценки объема и характера имеющегося фактического материала по более чем двумстам видам птиц;

2) позволить за краткое время найти публикации, затрагивающие конкретные виды птиц и/или аспекты орнитологических исследований.

2 Современное состояние

В основе создаваемой информационной системы лежат четыре уже имеющиеся на станции базы данных, созданные независимо друг от друга, в различное время и на разной основе.

База данных кольцевания, находок окольцованных птиц и данных их прижизненной обработки.

Создание ее начато в 1973 году [2]. В настоящее время данная база насчитывает более 2,3 млн. записей, имеет объем более 800 Мб, и продолжает пополняться. В целом

Первая Всероссийская научная конференция
ЭЛЕКТРОННЫЕ БИБЛИОТЕКИ:
ПЕРСПЕКТИВНЫЕ МЕТОДЫ И ТЕХНОЛОГИИ,
ЭЛЕКТРОННЫЕ КОЛЛЕКЦИИ
19 - 21 октября 1999 г., Санкт-Петербург

она включает информацию о районе и времени кольцевания или находки птицы, ее виде, поле и возрасте, а так же до 40 морфологических и физиологических характеристик, регистрируемых во время сопутствующего кольцеванию прижизненного обследования птиц.

Информацию, поступающую в данную базу, можно по смыслу разделить на четыре основных блока.

1) Данные кольцевания птиц. Сюда относятся сведения о том, когда, где и на какую птицу надето данное кольцо, а так же результаты сопровождающего кольцевание прижизненного обследования птицы. Такое обследование включает в себя данные о районе, времени и методе отлова, размерах и весе тела птицы, состоянии оперения, уровне жировых резервов, накапливаемых птицами во время перелетов, обнаруженных при обследовании болезнях, травмах и паразитах, а так же ряд других морфологических и физиологических параметров. Эта информация поступает как с расположенных на Куршской косе стационарных пунктов кольцевания, так и от исследователей, отлавливающих птиц в различных районах косы в рамках кратковременных научных программ, или же кольцующих птенцов на гнездах.

Ежегодно поступает информация о 50 000 - 110 000 окольцованных птиц. В общей сложности накоплено уже более 2,2 млн. данных.

Именно эти данные легли в основу СУБД "ЗООЛОГИЯ-80" (позже - "АРХИМЕД") созданной в 1973 г., и до 1993 года были единственными, входящими в нее [1]. Затем, в связи с участием станции в международных программах по изучению миграций птиц и открытием второго стационара для отлова и кольцевания, требования к базе данных существенно изменились: а) расширился список регистрируемых параметров прижизненного обследования; б) резко увеличился обмен данными с зарубежными участниками программ, возникла потребность быстрой подготовки выборок с разными наборами параметров в форматах, легко воспринимаемых распространенными в мире программами для обработки данных; в) сильно возросло количество повторных отловов уже окольцованных птиц в пределах Куршской косы, возникла возможность сравнения и анализа изменений состояния отдельных птиц с течением времени; г) увеличилось количество запросов о кольцевании и находках за пределами Куршской косы птиц, которые отлавливались на косе. Возмозности СУБД "АРХИМЕД" [2] не позволили справиться с этими задачами, поэтому были предприняты перевод данных в среду FoxPro и несколько существенных реорганизаций структуры базы. При этом в базу

были включены следующие три дополнительных блока информации:

2) Данные повторных отловов птиц, окольцованных ранее на Куршской косе. Птицы, отловленные с кольцами, обследуются вновь, при этом регистрируются те же параметры, что и при кольцевании. Количество таких отловов колеблется в разные годы от 10 000 до 30 000.

Сравнение полученных результатов с уже имеющимися данными кольцевания позволяет проследить динамику изменения различных параметров в зависимости от сезона и возраста птицы, стратегии поведения птиц в зависимости от сезона, а, зная местонахождение птицы между отловами, и выявить характеристики окружающей среды, оказывающие влияние на те или иные физиологические процессы.

3) Данные о находках окольцованных птиц за пределами Куршской косы. Они могут поступать как непосредственно от других станций кольцевания, где наши птицы были отловлены, так и от граждан, обнаруживших ослабевших или мертвых птиц или нашедших кольцо, и сообщивших об этом в Центр Кольцевания, который затем по номеру кольца выясняет его происхождение и переадресует эти данные Биологической станции. Количество таких сообщений составляет от 100 до 500 в год. В настоящее время накоплены данные о более чем 6000 птиц.

4) Данные о кольцевании птиц, отловленных на Куршской косе с иностранными кольцами. Данные о каждой из таких находок направляются в Центр Кольцевания, который затем сообщает информацию о месте и времени кольцевания каждой из этих птиц, а дальнейшие контакты с кольцевателями позволяют выяснить детали прижизненного обследования каждой из этих птиц. Каждый год происходит от 100 до 500 таких находок. К настоящему времени накоплены данные о более чем 5,5 тысяч птиц.

Информация, перечисленная в пунктах (3) и (4), позволяет выяснить расположение районов гнездования, миграционных путей и зимовок различных видов птиц, а также изучить скорость перемещения различных видов.

При создании современной структуры данной базы более целесообразным оказалось разбиение ее на файлы по другому принципу: в зависимости от наличия тех или иных параметров. Это реляционная база данных на FoxPro2.5, объединяющая 5 основных таблиц (SUM, FRIN, RYB, ESF, RECOVERS) и 26 словарей.

Самые основные данные, присутствующие в каждом сообщении об отлове или находке птицы, и входящие в большинство запросов, а так же содержащие основную информацию о перемещениях птицы и ее судьбе, включены в таблицу SUM. Это данные о кольце, виде, поле и возрасте птицы, дате, времени, месте и обстоятельствах каждого отлова или находки (включая метод отлова и номер орудия лова). Таблица SUM содержит в настоящее время более 2,4 млн. записей.

При помощи уникального кода данная таблица связана с таблицей FRIN, содержащей минимальный набор данных прижизненного обследования, регистрируемый при большинстве отловов: данные о длине крыла и хвоста, массе тела птицы, состоянии оперения, уровне жировых резервов, степени пневматизации черепа, признаках размножения, болезнях. В настоящее время в таблицу FRIN включено 1,3 млн. записей.

Таблица RYB содержит параметры, регистрируемые главным образом в рамках международных программ на полевом стационаре "Рыбачий". Это детальные данные

об отлове (номер, сторона, высота паутиной сети, где была поймана птица), уровень жировых резервов птицы по девятибалльной шкале, несколько дополнительных характеристик линьки, включая указание на наличие карточки линьки, длина третьего первостепенного махового пера, длина вырезки на втором первостепенном маховом (для камышевок). По уникальному коду данная таблица связана с таблицами SUM и FRIN, а так же с таблицей данных карточек линьки, а по номеру паутиной сети - с дополнительными таблицами, содержащими описание биотопа и детальные описания растительности вокруг каждой сети в разные годы и сезоны. В таблицу в настоящее время включено около 220 тыс. записей.

Таблица ESF содержит данные о длине цевки и длинах десяти маховых перьев, регистрируемые для определенных видов птиц в рамках международной программы по изучению миграций птиц из Европы в Африку [3]. Таблица в настоящее время включает более 40 тысяч записей.

В таблицу RECOVERS включены параметры, характерные птиц, отловленных или найденных не только на косе, но и за ее пределами: время между кольцеванием и находкой птицы, расстояние между местом кольцевания и местом находки, азимут из точки кольцевания в точку находки, средняя скорость перемещения птиц по данным кольцевания, источник информации о находке птицы, дата получения сообщения, информация об иностранном кольце, а так же два уникальных кода, позволяющих связать каждую запись с записями в SUM, касающимися 1) последнего отлова птицы в районе кольцевания, и 2) первого отлова в точке находки. Два последних поля необходимы так же для выбора данных при построении электронных карт миграций птиц.

В управление базой вовлечены более ста программ и служебных файлов, обеспечивающих ввод и контроль данных и быстрый поиск необходимой информации.

Специализированные программы ввода новых данных, в зависимости от поступающей информации, могут производить ввод а) базового набора параметров (таблицы SUM и FRIN); б) расширенного набора параметров (таблицы SUM, FRIN, RYB); в) дополнительных измерений (таблица ESF) для птиц, данные которых уже имеются в таблице SUM; г) дальних находок и данных кольцевания птиц с иностранными кольцами (таблицы SUM и RECOVERS). В последнем случае программа оснащена механизмом автоматического вычисления направления, скорости, расстояния и времени перелета.

Контроль вводимых данных производится в три этапа. В процессе ввода данные поступают во временный файл, заполняемый при каждом сеансе работы, и очищаемый по его окончании. Когда в ходе ввода данных курсор перемещается на поле, для которого имеется словарь, то словарь по данному полю становится доступным оператору, что делает возможным выбор или проверку кода данного поля. Для числовых параметров и дат на этом этапе предусмотрен грубый контроль, не позволяющий вводить заведомо невозможные значения.

По окончании сеанса ввода (или, по желанию оператора, после ввода некоторого количества данных) осуществляется более тщательная проверка. При этом учитываются: 1) соответствие длины крыла и массы птицы пределам значений, заданных для данного вида (а в ряде случаев - и пола) птицы; 2) соответствие кодов возраста сезону; 3) соответствие кодов пункта кольцевания/находки и метода и номера применяемого орудия лова.

После этого этапа данные переносятся в специальный

файл накопления. Включение их в базу данных, как и третий этап контроля, производится администратором базы данных. На третьем этапе проверяется соответствие данных первых и повторных отловов для вида, пола и возраста птицы с учетом дат повторных отловов. При этом а) исключаются ошибки, допущенные при прочтении номеров колец и записи их в журнал регистрации; б) выявляются и устраняются противоречия в определении пола и возраста птицы при кольцевании и повторных отловах; в) при возможности - уточняются сведения о поле и возрасте птицы на основе сопоставления всей имеющейся информации; г) проверяется наличие дублирующихся записей, которые могут возникнуть результате поступления одних или тех же данных из разных источников, или хранения их на нескольких формах бумажных носителей.

Дополнительная система контролирует наличие в базе данных информации по всем используемым номерам колец, и сообщает а) об отсутствии данных кольцевания, если номер кольца помечен, как использовавшийся б) о наличии данных кольцевания для номеров колец, помеченных как брак/потерянное/не использовавшееся. Это позволяет провести дополнительный контроль на неправомерное прочтение или запись номеров колец при отловах окольцованных птиц, а так же выявить потерянные или не предоставленные вовремя кольцевателями первичные данные.

Поиск информации может производиться по латинскому, русскому, английскому или немецкому названию вида птицы, а так же по любому значению или пределам значений вводимых параметров. Проводящееся в настоящее время дополнение словарей позволит осуществлять вывод любых параметров запрашиваемой информации не только на русском, но и на английском и немецком языках. Завершение разработки блока информации о находках птиц, окольцованных на косе, за ее пределами, и отловах птиц с иностранными кольцами, позволит создать на его основе электронный атлас путей миграций различных видов птиц, встречающихся на Куршской косе.

Данные регистрации гнезд на Куршской косе Балтийского моря. Регистрация гнезд, в основном воробьиных птиц, проводится на Куршской косе в течение уже 40 лет. Каждое обнаруженное гнездо тщательно обследуется. Затем проводятся регулярные наблюдения обнаруженных гнезд, данные которых так же регистрируются. Компьютеризация этих данных начата в 1998 году, в настоящее время наряду с вводом информации производится выработка оптимальной структуры данной базы, разработка словарей и служебных программ. База так же выполнена на FoxPro2.5. В случае кольцевания птиц на гнезде информация из данной базы может быть связана с базой данных кольцевания, повторных находок птиц и данных прижизненной обработки по номеру гнезда. На конец августа 1999 в ней содержатся данные о более чем 1500 гнезд.

База данных о научных публикациях Биологической станции создана в 1996, в настоящее время существует в формате Access7.0 и содержит данные о более чем семистах публикациях сотрудников станции, вышедших начиная с 1956 года. Библиографическая информация на русском и английском языках в ряде случаев снабжена так же кратким содержанием статей на русском и английском языках, а в некоторых случаях так

же имеет ссылки на содержащиеся в отдельных файлах тексты или сканированные изображения статей. В данной базе предусмотрен поиск библиографических ссылок на основе информации об авторе, годе издания и словам из заголовка и краткого содержания. В тех случаях, когда статьи снабжены ключевыми словами, возможно осуществление поиска по ним. Параллельно разработана система управления этой базой на Delphi4.0, позволяющая не только просматривать содержимое базы и сканированные изображения статей, но и при необходимости вводить новые изображения непосредственно со сканера. В настоящее время проводится сбор недостающих текстов статей для дополнения базы. Планируется размещение данной информации на CD.

База метеорологических данных существует с 1998 г. В настоящее время она включает данные по температуре, влажности и давлению воздуха на Куршской косе за 1994-98 гг. Система управления этими данными в настоящее время разрабатывается.

3 Перспективы развития

Разрабатываемая в настоящее время на основе Delphi4.0 информационная система предусматривает возможность получения из единого интерфейса сведений из перечисленных выше различных по структуре и содержанию баз данных. По любому из зарегистрированных на Куршской косе видов птиц предусматривается возможность вывода следующих блоков информации:

Общая информация о виде (на основе литературных данных): будет включать название, систематическое положение, описание и внешний вид птицы, область распространения вида, его статус на Куршской косе.

Данные кольцевания В этот блок войдут данные о количестве окольцованных птиц каждого вида (как суммарном, так и с разбивкой по годам), количестве контрольных отловов этих птиц в районе кольцевания, количестве дальних находок окольцованных птиц, а также, количестве отловов и находок на косе птиц с иностранными кольцами. После завершения работы по вводу и проверке данных о дальних находках птиц и отловах птиц с иностранными кольцами возможно включение в этот блок карты миграций данного вида по данным кольцевания. Кроме того, предусмотрена выдача информации об общем количестве птиц, для которых имеются данные прижизненной обработки, с указанием количества измерений по каждому из зарегистрированных параметров. Это позволит пользователю примерно оценить количество потенциально доступных для анализа данных по конкретному виду.

Данные о гнездовании на Куршской косе Количество обнаруженных гнезд (как суммарное, так и с разбивкой по годам), средние сроки гнездования. После завершения работ по вводу и анализу данных возможно включение в этот блок карт, показывающих распределение гнезд по территории косы.

Публикации станции. Список публикаций станции, затрагивающих данный вид или группу птиц. Знакомство с кратким содержанием любой из выбранных статей на русском или английском языках, а после дополне-

ния имеющихся данных - и с полным содержанием имеющихся публикаций на языке оригинала. Описываемая система может оказать значительную помощь специалисту при планировании работ с новым видом или группой видов птиц, в то же время она может представлять значительный интерес для любого посетителя Биологической станции в Рыбачьем, стремящегося расширить свой кругозор.

Планируется размещение основной информации, предоставляемой системой, в Интернете на сервере Зоологического института РАН (<http://www.zin.ru>).

Библиография

- [1] Морозов Ю. Г., Ефремов В. Д. 1997: *О некоторых проблемах обмена информацией баз данных кольцевания и прижизненного обследования птиц*. Базы данных и компьютерная графика в зоологических исследованиях. Труды ЗИН РАН, том 269, С.88–90. Санкт-Петербург.
- [2] Морозов Ю. Г., Ефремов В. Д. 1997: *Программное обеспечение для поддержки и эксплуатации орнитологических баз данных*. Базы данных и компьютерная графика в зоологических исследованиях. Труды ЗИН РАН, том 269, С.91–98. Санкт-Петербург.
- [3] Bairlein F. 1997: *Spatio-Temporal Course, Ecology and Energetics of Western Palaearctic - African Songbirds Migration*. Institut fuer Vogelforschung 'Vogelwarte Helgoland', Wilhelmshaven.