

С. С. КОЗЛОВ, ассистент, НТУ “ХПИ”

ХРАНЕНИЕ ДАННЫХ РАДАРА НЕКОГЕРЕНТНОГО РАССЕЯНИЯ В ФОРМАТЕ XML

Розглянуто алгоритм процедури обробки іоносферної інформації за допомогою радара некогерентного розсіяння та обчислювальних систем. Запропоновано впровадження локальної бази даних радіолокаційних вимірювань та збереження даних у форматі XML. Визначено переваги та недоліки запропонованого формату. Запропоновано структуру та спосіб кодування даних для формату XML.

The procedure algorithm of ionospheric information treatment by incoherent scatter radar and computer systems is considered. Introduction of local base of these radio-location measuring and storage of data in the XML format is offered. Dignities and lacks of the offered format are certain. A structure and methods of information code for the XML format is offered.

Постановка задачі. Необхідність автоматизації процесу зондування і обробки в методі некогерентного розсіяння (НР) потребує впровадження пакетів прикладних програм, орієнтованих не тільки на рішення спеціалізованих задач в умовах все зростаючого об'єму вхідної інформації, але і на забезпечення взаємодії радіотехнічних систем і засобів обробки даних. В наші часи іоносферні дані, отримані в Інституті іоносфери з допомогою радара НР і станції вертикального зондування, в основному представлені в вигляді двобитних файлів, які обробляються програмами на базі мови програмування TurboBasic. Для розширення можливостей управління і обміну цими даними нинішня програмна оболонка і формат зберігання повинні бути модифіковані.

Аналіз літератури. Програмна оболонка, що реалізує процедури вторинної обробки сигналу НР з допомогою персонального комп'ютера (ПК), створена на базі мови програмування TurboBasic [1]. Так як мова має обмеження на текстовий об'єм створюваної програми (не більше 64 Кбайт), то вона розділена на короткі bas-модулі. Для компоновки програми в цілому ці модулі компілюються в exe-модулі, які поочередно запускаються з допомогою стартової програми, організуючої обмін масивами в вигляді автокореляційних функцій (АКФ) сигналу НР [2].

Процес настільки складний, що при виконанні визначеного етапу вимагається участь кількох спеціалістів. На рис. 1 видно, що процес обробки відбувається поступово. Автором свідомо не показані деякі зворотні зв'язки, оскільки вони слабо інформативні.

Ціль статті – аналіз можливості управління даними, отриманими з допомогою радара НР, з метою покращення інформаційного забезпечення і підвищення швидкості процесу обробки іоносферних даних.

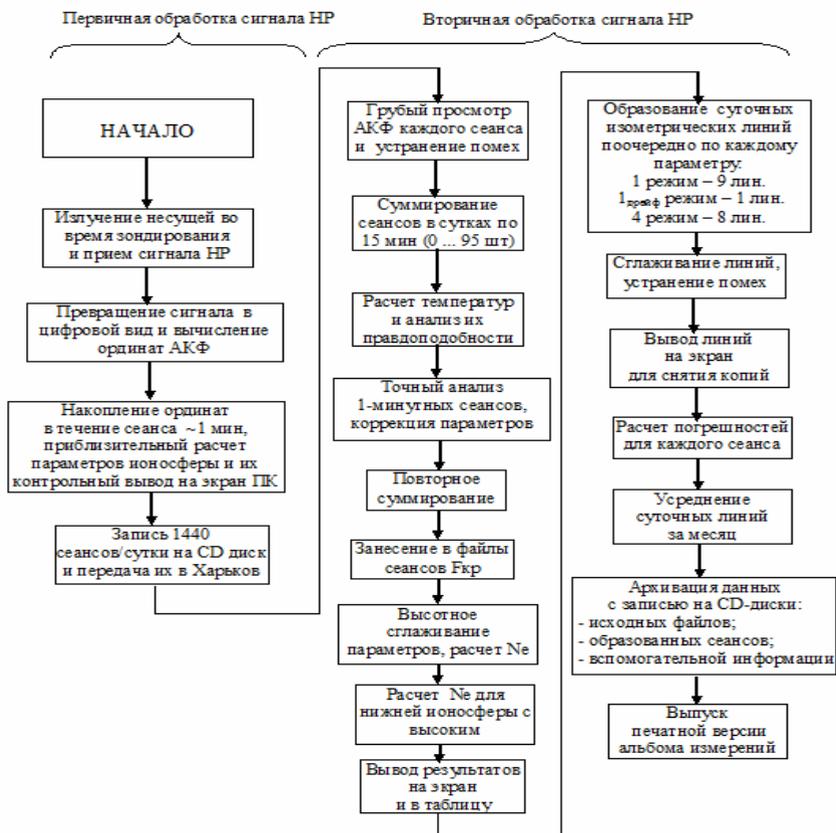


Рис. 1. Алгоритм процедуры обработки ионосферной информации с помощью радара НР и вычислительных систем

База данных радара некогерентного рассеяния. Для улучшения управления и повышения быстродействия обработки предлагается создание локальной базы для хранения экспериментальных данных харьковского радара НР, учитывая большую, до 10 Гбайт/сутки, плотность информационных потоков данной системы [3].

В целом, структура базы данных радара должна учитывать особенности использования метода НР (см. рис. 2). База предназначена для хранения радиотехнической информации, получаемой в процессе геофизических экспериментов, хранения измеренных высотно-временных зависимостей АКФ и спектров сигнала НР, вычисленных высотно-временных вариаций параметров ионосферы и результатов их анализа.

Организация базы данных подразумевает разработку компьютеризированной системы управления. Процесс формирования компонентов и их запись обычно является наиболее трудоемким и наименее автоматизированным из всех операций. Предлагается создать архив из файлов универсального формата, например, XML.

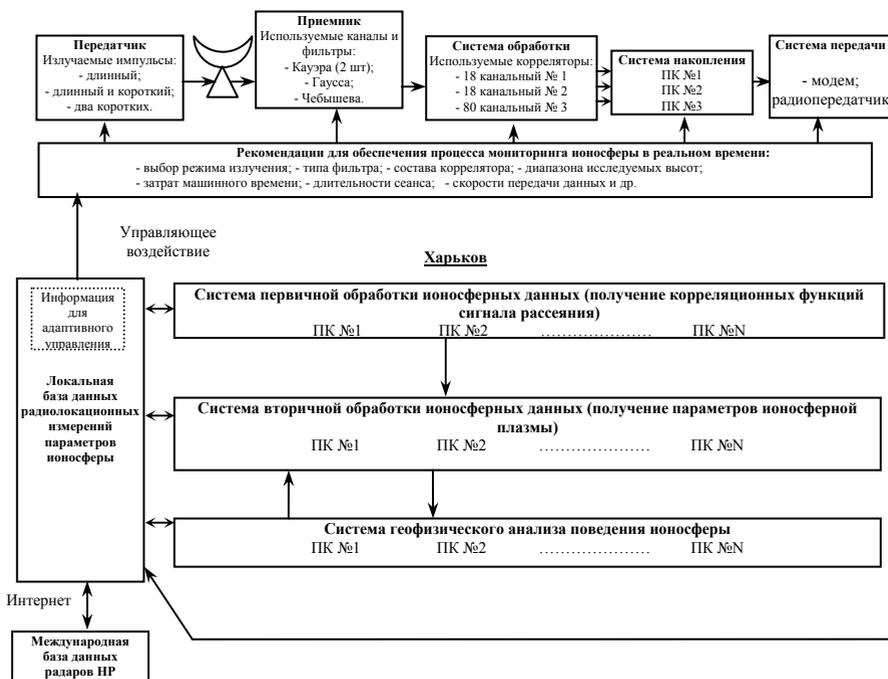


Рис. 2. Особенности использования локальной базы данных радиолокационных измерений

Формат XML. XML – текстовый формат, который используется для хранения структурированных данных и обмена информацией между программами. XML – это иерархическая структура, предназначенная для хранения любых данных. Визуально структура может быть представлена как дерево. Наиболее важное и обязательное синтаксическое требование заключается в том, чтобы документ имеет только один корневой элемент (альтернативно называемый элементом документа). Это означает, что текст или другие данные всего документа должны быть расположены между

единственным начальным корневым тегом и соответствующим ему конечным тегом [4].

Под XML подразумевается совокупность трех тесно связанных стандартов. Сюда входит сам XML как средство описания структуры документов, XSL как средство преобразования XML-документов для отображения, а также XLL – расширяемый язык связывания документов. Последний позволяет устанавливать многонаправленные ссылки и ссылаться не на весь документ, а на его конкретные его элементы.

XML, как основной формат, используется в офисных приложениях OpenOffice и MS Office 2007 и применяется для описания векторной графики SVG, формул MathML. Практически все современные системы управления базами данных поддерживают формат XML, а многие частично или полностью ориентированы на него.

Достоинства:

- 1) XML – язык разметки, позволяющий отобразить двоичные данные в текст, читаемый человеком и анализируемый компьютером;
- 2) XML поддерживает Юникод;
- 3) В формате XML могут быть описаны такие структуры данных как записи, списки и деревья;
- 4) XML – это самодокументируемый формат, который описывает структуру и имена полей так же, как и значения полей;
- 5) XML имеет строго определённый синтаксис и требования к анализу, что позволяет ему оставаться простым, эффективным и непротиворечивым. Одновременно с этим, разные разработчики не ограничены в выборе экспрессивных методов (например, можно моделировать данные, помещая значения в параметры тегов или в тело тегов, можно использовать различные языки и нотации для именования тегов и т. д.);
- 6) XML – формат, основанный на международных стандартах;
- 7) Иерархическая структура XML подходит для описания практически любых типов документов, кроме аудио и видео мультимедийных потоков, растровых изображений, сетевых структур данных и двоичных данных;
- 8) XML представляет собой простой текст, свободный от лицензирования и каких-либо ограничений;
- 9) XML не зависит от платформы;
- 10) XML является подмножеством SGML (который используется с 1986 г.). Накоплен большой опыт работы с языком и созданы специализированные приложения;
- 11) XML не накладывает требований на расположение символов в строке [5];
- 12) В отличие от бинарных форматов, XML содержит метаданные об именах, типах и классах описываемых объектов, по которым приложение может обработать документ неизвестной структуры (например, для динамического построения интерфейсов [4]);

13) XML имеет реализации парсеров для всех современных языков программирования [5];

14) XML поддерживается на низком аппаратном, микропрограммном и программном уровнях в современных аппаратных решениях. Основные достоинства формата XML следующие: он позволяет отобразить двоичные данные в текст, читаемый человеком и анализируемый компьютером; основан на международных стандартах; не зависит от платформы.

Недостатки:

1) Синтаксис XML избыточен;

2) Размер XML документа существенно больше бинарного представления тех же данных;

3) XML содержит мета-данные (об именах полей, классов, вложенности структур), и, одновременно, XML позиционируется как язык взаимодействия открытых систем. При передаче между системами большого количества объектов одного типа (одной структуры), передавать метаданные повторно нет смысла, хотя они содержатся в каждом экземпляре XML-описания.

Радиолокационные данные в формате XML. Пример ионосферных данных, для которых предлагается создать архив – выходные файлы системы четырёхканального коррелятора радара НР [2]. Эти файлы состоят из заголовка (дата, время и номер сеанса) и данных в формате четырёхбайтных целых чисел. Корреляционные функции НР сигнала хранятся в четырёх двумерных массивах (каждый 19 точек с 680 высотных участков в виде косинусной и синусной составляющих для двух каналов).

При создании архива ионосферных данных имеет смысл использовать интерфейс XML-RPC (XML Remote Procedure Call). XML-RPC – протокол вызова удалённых процедур, основанный на XML, отличается исключительной простотой применения. XML-RPC, как и любой другой интерфейс RPC, определяет набор стандартных типов данных и команд, которые программист может использовать для доступа к функциональности другой программы, находящейся на другом компьютере в сети [6].

Двоичные данные при включении в файл XML кодируются при помощи схемы base64. Base64 буквально означает – позиционная система счисления с основанием 64. Здесь 64 – это наибольшая степень двойки (2^6), которая может быть представлена с использованием печатных символов ASCII. Эта система широко используется в электронной почте для представления бинарных файлов в тексте письма (транспортное кодирование). Все широко известные варианты, известные под названием Base64, используют символы A-Z, a-z и 0-9, что составляет 62 знака, для остальных двух знаков в разных системах используются различные символы.

Результатирующие закодированные по base64 данные имеют длину, большую, чем исходный текст в соотношении 4:3, и напоминают по виду случайные символы. В частности можно использовать Radix-64 – разновидность кодирования Base64 двоичных данных в текстовый формат,

используемая в PGP. От Base64 отличается тем, что в конец добавляется контрольная сумма в 24 бита.

Дата и время сеанса хранятся в формате по стандарту ISO 8601. Формат ISO 8601 представляет собой международный стандарт, выданный организацией ISO, описывающий формат даты и времени.

Значения даты и времени расположены в порядке от более к менее значимому. Каждое число (год, месяц, день, время) записывается с фиксированным числом знаков и, при необходимости, должно быть дополнено до него ведущими нулями.

Кроме того, стандарт поддерживает добавление десятичной точки к наименьшей единице времени, когда требуется большая точность.

Существенный недостаток формата XML компенсируется ростом объёмов носителей информации.

Выводы. Предложен способ хранения данных радара НР в формате XML. Представление файлов ионосферных данных в формате XML позволит улучшить информационное обеспечение, повысить быстродействие процесса обработки ионосферных данных и создать функциональную базу данных, которая предоставит быстрый и удобный доступ к информации.

Список литературы: 1. Пуляев В.А. Инструкция и описание процедуры обработки данных радиолокатора НР с помощью компьютерных программ вторичной обработки. – Х.: Институт ионосферы, 2007. – 16 с. 2. Лысенко В.Н., Кононенко А.Ф., Черняк Ю.В. Корреляционная обработка сигнала некогерентного рассеяния // Вестник НТУ «ХПИ». Сборник научных трудов. Тематический выпуск «Радиофизика и ионосфера». – 2004. – № 23. – С. 49 – 62. 3. Козлов С.С. Организация базы данных харьковского радара НР // Технічна електродинаміка. Тематичний випуск “Силова електроніка та енергоефективність”. – 2009. – Ч. 2. – С. 113 – 114. 4. <http://www.xml.org/> 5. Хантер Д., Рафтер Дж. и др. XML. Базовый курс. – М.: Вильямс, 2009. – С. 1344. 6. Козлов С.С., Богомаз А.В., Пуляев В.А. О возможности применения формата XML для хранения ионосферных данных // Конференция молодых учёных “Дистанционное радиозондирование ионосферы (ИОН-2010)” (Харьков, Украина, 6 – 9 апреля 2010 г.). – Сборник тезисов. – 2010. – С. 38.

Поступила в редколлегию 31.05.2010