

БАЗА ДАННИ ЗА "ИКБ-1300" - ПРОБЛЕМИ И РЕШЕНИЯ

Николай Банков¹, Стефан Чапкънов¹, Людмила Тодориева¹, Михаел Касчиев²

¹⁻ ИКИ – БАН, nbankov@bas.bg

²⁻ ИМ – БАН

* работата е спосорирана от договор НЗ 1309/03 с МОН

АНОТАЦИЯ

Представени са организацията на базата и данните, които съдържа към текущия момент. Описание на съпътстващата геофизична и орбитална информация и нейната организация също е включена. Специално внимание е обърнато на принципите, използвани за подобряване на точността на “привязката” по време.

ИЗЛОЖЕНИЕ

В предхождащи доклади и публикации [1,2], една предварителна и вероятно подлежаща на преоценка стратегия, избрана от авторите за изграждане на база данни от “ИКБ 1300”, бе описана в най-общ вид. Предлаганата сега работа се предполагаше да съдържа описание на текущото състояние, и по-точно, достъпната към момента информация в базата. Това ще бъде направено по-долу, но преди това това ще опишем някои принципни положения, залегнали в базата, и някои проблеми, възникнали при нейната реализация. Трябва да отбележим, че създаването на такава база не изисква никакви специални умения или опит в тази област, достатъчно е да се определят съответствията между основните теоретични принципи и конкретната реализация. Така например, принципа “един обект – една таблица” в дадения случай може да се интерпретира като “един прибор – една таблица”, и в повечето случаи това изглежда като един много добър избор. Сравнително еднозначно е и определянето на ключови полета – очевидно универсалното време UT, пакетирано във формат UNIX, е подходящ кандидат за първичен ключ (primary key), “номер на сеанса” е възможен избор за вторичен ключ, и т.н.. От друга страна, при подготовката на данните за въвеждане в базата, информацията от обекта преминава различни нива на обработка, и в зависимост от качеството на изходните данни, използваните алгоритми и т.н. се получават данни с различна степен на достоверност, което, освен че е нетипично за конвенционалните бази, води до необходимостта от въвеждане на допълнителни параметри, характеризиращи точността. Естествено, първият въпрос, на който трябва да се отговори в процеса на оценка на достоверността, е този за точността на така наречената “привязка по време”, въпреки че, както ще стане ясно по-нататък, под това име ще разбираме по-широк кръг проблеми. Тук ще се спрем по-подробно на този въпрос не само поради първостепенното му значение, но и поради факта, че до този

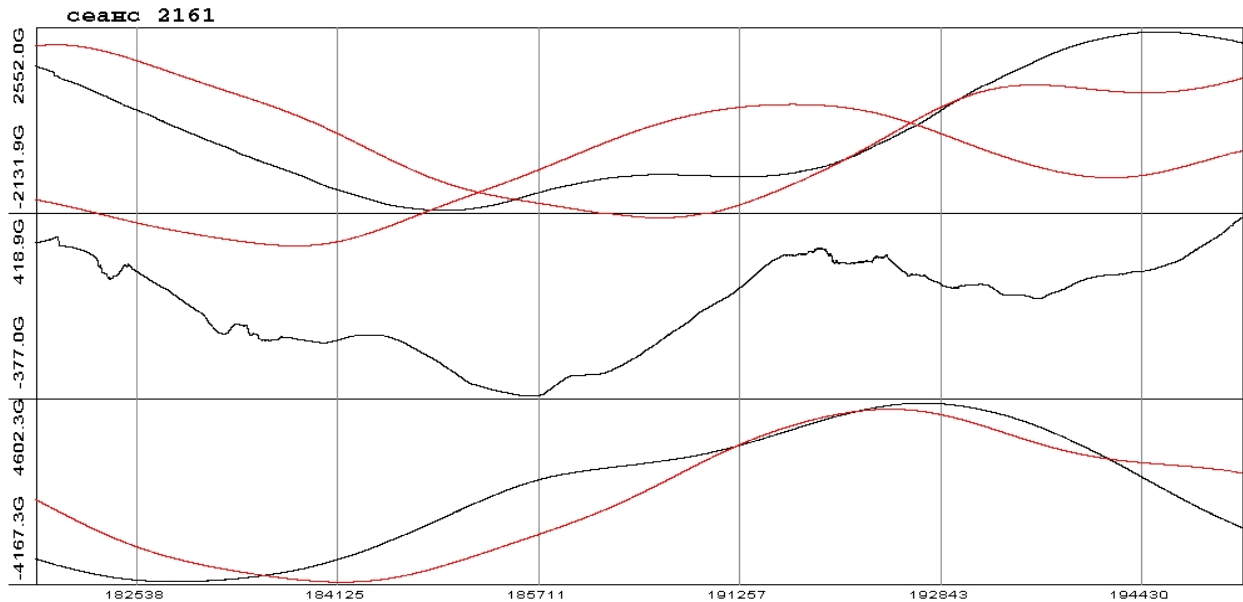
момент не са анонсирани проблеми с привязката, въпреки че са правени и други опити да се разработи база данни за този обект, т.е. напълно възможно е да се окаже, че в даден момент в публичното пространство съжителствуват данни с едни и същи номера на сеанси или орбити, но с различно време и съответно местоположение. За по-пълното разбиране на проблема са ни необходими някои кратки сведения за телеметричната система. На борда на "ИКБ-1300" бе използвана аналогова ТС с два основни режима на работа с дискретизация на сигнала 80 и 320 милисекунди, наричани съответно ЗАП2 и ЗАП3, т.е. 256 информационни канали от приборния състав и служебните системи, т.н. телеметричен кадър, се сканират циклично и равномерно във времето на всеки 80 или 320 мс. Тази информация се записва на лентата на бордовия магнетофон, като на всеки 15 сек, по сигнал на бордовия часовник, върху полезния сигнал се записват, приоритетно, секундни и минутни маркери, с чиято помощ на земята се възстановява времето, асоциирано с началото на всеки телеметричен кадър. Истинското време се добавя към записа по време на етапа "предварителна" обработка, като оператора въвежда ръчно стойностите на календарните дата и време, отнесени към началото на сеанса (това е процесът наричан "привязка") в съответствие с времето на подадената команда за стартиране на измерванията. Именно това ръчно въвеждане (посредством забравените сега перфокарти) е главният източник на преобладаващата част от грешките в привязката по време. Издирването на този тип грешки, някои от тях очевидни, позволи откриването и на някои неочаквани ситуации, като в крайна сметка възможностите бяха групирани така:

- погрешно въведени цифра/и във някое от полетата за време, дата или номер на сеанс
- един сеанс представен като два различни
- телеметричните кадри подредени в обратен ред.

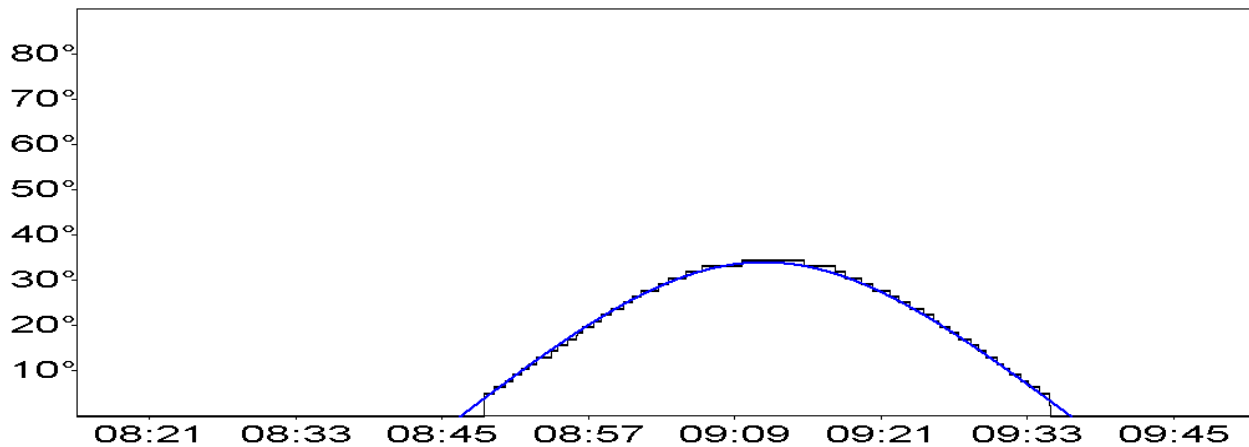
Последният случай е отделен в отделна група, тъй като е забелязан един случай (за съжаление използван в публикации), сеанс 231 - ЗАП4, каквито са само няколко сеанса, и за който лесно се вижда, че времето, определено като начало на сеанса, може да бъде евентуално времето на последния телеметричен кадър.

Грешките от втория тип бяха открити след като бе определена стратегия за откриване и евентуално корегиране, при определени условия, на възможните грешки по вина на оператора, т.е. тези от първи тип. Бе избрана стратегия, която се основава на сравнение на данните от измерените на борда от прибора ИМАП координати на вектора на постоянното магнитно поле с моделните данни за полето (в координатната система, свързана със спътника). При установяване на подозрително големи разлики в сравняваните величини, като допълнителни критерии се използват данните за височина на Слънцето над хоризонта на спътника, измерени и моделни (естествено когато орбитата има участък на слънце), както и наличните сведения за научната програма, по-точно записите за времето на включване на приборите относително времето на пресичане на еkvатора. При така избраната стратегия, за "поправими" сеанси приемаме тези, за които описаните по-горе критерии се удовлетворяват с възможно най-малък брой корекции в началното време. Такъв критерий е естествен, тъй като, предполагайки като източник на несъответствията грешка на оператора при въвеждането на началните данни за времето, то не би трябвало да се очаква този оператор да е допускал "много" грешки при единично въвеждане. Процедурата за идентифициране на сеанси с проблемно време ще илюстрираме на примера на оригиналните сеанси 2161 и 2171. Казусът на пръв поглед изглежда елементарен, и в известна степен е такъв, но е добра илюстрация на различните възможности. На фиг.1 са показани трите компоненти на магнитното поле, измерено и моделно, за сеанса с номер 2161. Вижда се, че има

сериозно разминаване в графиките. Веднага се установява, че датата на начало на сеанса е 07.01.1981, т.е. очевидна грешка (от първи тип) в цифрата на единиците за годината – трябва да е 1982. Поправяме грешката и получаваме приемливо поведение на полето, при това се получава и добро съответствие на данните от датчика за височина на Слънцето с пресметнатите такива (фиг.2). Сега обаче се оказва, че със същата дата и практически същото време съществува още един сеанс – 2171 (Московското време на 2161 е 11:16:29, а на 2171 11:15:43), така че е налице ситуация с много грешки, която би трябвало да е неразрешима.



фиг.1 червената линия – моделното магнитно поле, черната – измереното
сеанс 2161



фиг.2 синята линия – пресметнатата височине на Слънцето,
черната – показанията на датчика

При внимателно разглеждане на случая се установи, че данните от измерванията на приборите съвпадат, така че става дума за един сеанс с различни дата-време, или грешка от втория наблюдаван тип. Не можем да гадаем как се е стигнало до тази ситуация, но бяха установени още няколко такива случая, и няма гаранция, че няма и други.

Всички тези бележки бяха необходими за по-доброто разбиране на въпроса с достоверността на информацията.

Трябва да отбележим и една друга принципна разлика между разглежданата тук база и възприетите стандарти. Става дума за принципа за изчисляемост, или схващането, че една добре планирана база не трябва да съдържа данни, пресметнати с помощта на други данни от базата. Тук този принцип е неприложим, тъй като се предполага, че основните извадки, които потребителите ще заявяват, ще са свързани с пространственото положение на спътника и/или някои моделни величини, описващи геофизичната обстановка, които в случая са изчисляеми параметри. Подробно за тях и групирането им в таблици ще говорим по-долу, а тук ще споменем за някои идеи за съдържание на таблици, касаещи приборите, и нарушаващи принципа.

Така например, за някои от вълновите измервания, и по-специално електрическото и магнитно полета в диапазона 0.2-8.0 херца, е решено в базата да се съхранят спектралните им коефициенти по секунди вместо изходния сигнал, въпреки че тези коефициенти са изчисляеми, но пък са значително по-информативни. От друга страна, за данните от корелатора вероятно по-доброто решение е да се запазят корелограмите, а не изчисляваният с тяхна помощ сигнал, така че в този случай принципът ще е спазен.

Данните за съпътстващата геофизична и навигационна информация са пресметнати със стъпка 1 секунда и са групирани тематично в 6 таблици: mainkd (ALT, LAT, LONG, L, ILAT, Bmag, Xmag, Ymag, Zmag, MLT, Tloc, GST, Dmag, Imag, Ldip, COSPEQ), coor (Xabs, Yabs, Zabs, Vxabs, Vyabs, Vzabs, Xgr, Ygr, Zgr, VXgr, Vygr, VZgr), sunmoon (SXabs, SYabs, Szabs, UnitVectorToSunInsat, Mxabs, Myabs, Mzabs), suns (Zsun, Asun, HSHADE, Salt, BSANGL, Tsun, AULONG, Taur, SUNLAT, SUNLON, ASLONG), angles (BXANGL, BYANGL, BZANGL, X VELM, Y VELM, Z VELM, VXANGL, VYANGL, VZANGL, D VELM, I VELM, HMIN), trace (LAT Lo, LON Lo, LATcLo, LONcLo).

Всички тези таблици съдържат времето (PT) и № на сеанс. Съкращенията в първите пет таблици са взети директно от [3], а тези в шестата таблица от [4].

Към настоящия момент са обработени всички сеанси с данни от приборите:

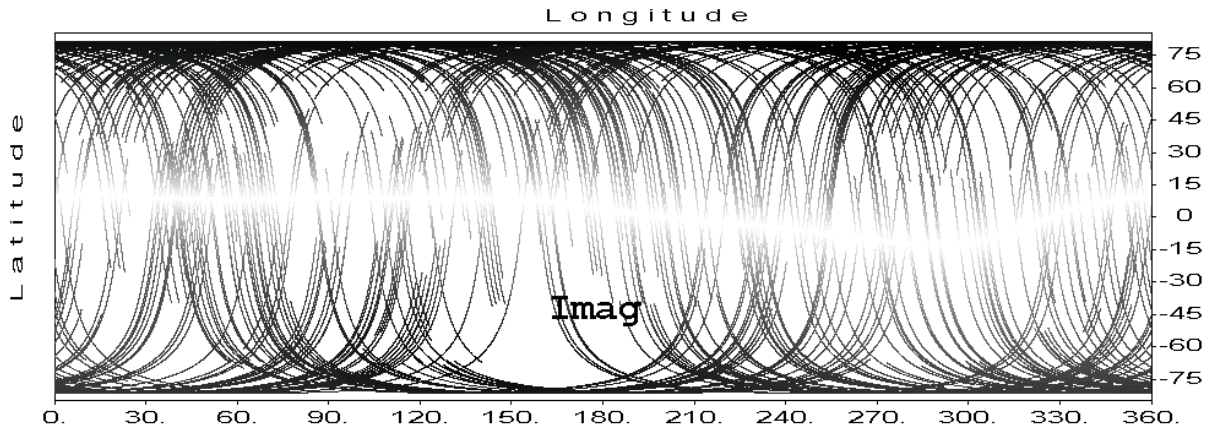
- ИМАП– измереното магнитно поле,
- ИЕСП– измервания на квазипостоянното електрично поле,
- П6-ЗЕЛ – измерване на пълната йонна концентрация,
- АНЕПЕ – анализатор на частици с енергия 1кев,
- ОНЧ – филтри,
- ЕМО5 – светене.

Основно това са приборите, чиито измервания са синхронизирани с честотата на телеметричната система, така че по същество са най-информативни и като следствие обемът на информацията от тях е най-голям (с изключение на корелатора на ИЕСП).

За илюстрация на бързодействието на базата бе проведен следния експеримент:

-извличат се данните за "Imag" заедно с "LAT" и "LONG" от таблица "mainkd" за всички записи, чиито времена съвпадат с времената на измерванията от таблицата за данните от прибор Р63ЕЛ, т.е. едно не много сложно множество, но с претърсване на цялата база. За 9.437 сек (компютър PIV 2.4G, 512RAM, SATA) бяха открити и извлечени 680876 записа. Селектираните данни са изобразени в "gray" скала на фиг.3.

В заключение ще отбележим, че направеното към настоящия момент позволява изграждането на една напълно използвана база с данните от приборите, които работиха с най-голяма информативност.



фиг.3

ЛИТЕРАТУРА:

1. Банков Н. Г., Чапкънов Ст., Тодориева Л. (ИКИ-БАН), Касчиев М. (ИМИ-БАН) БАЗИ ДАННИ ЗА "ИНТЕРКОСМОС БЪЛГАРИЯ – 1300" — Юбилейна научна сесия 2005 "120 години от Съединението", 21-22 април 2005г. сс 253-256.
2. Банков Н. Г., Чапкънов Ст., Тодориева Л., Касчиев М. - БАЗА ДАННИ ЗА "ИКБ-1300", Юб. Научна сесия 2006 "130г. от Априлското Въстание", 20-21 април 2006,(под печат)
3. Гальперин Ю.И., Ю.Н.Пономарев, В.М.Синицин, НЕКОТОРЫЕ АЛГОРИТМЫ РАСЧЕТА СПРАВОЧНОЙ ГЕОФИЗИЧЕСКОЙ ИНФОРМАЦИИ ВДОЛЬ ОРБИТ ОКОЛОЗЕМНЫХ СПУТНИКОВ, Пр. № 544, Институт Космических Исследований Акад. Наук СССР.
4. TSYGANENKO N.A. - A SET OF FORTRAN SUBROUTINES FOR COMPUTATIONS OF THE GEOMAGNETIC FIELD IN THE EARTH'S MAGNETOSPHERE (GEOPACK; release of January 5, 2001, corrected April 3, 2001), UNIVERSITIES SPACE RESEARCH ASSOCIATION (USRA), CODE 690.2, NASA GSFC, GREENBELT, MD 20771, USA