

СИСТЕМА СБОРА И ОБРАБОТКИ ГЕОМАГНИТНЫХ ДАННЫХ ДЛЯ ИНТЕРАКТИВНОГО РЕСУРСА ДАННЫХ ПО СОЛНЕЧНО-ЗЕМНОЙ ФИЗИКЕ SPIDR НА БАЗЕ СУЩЕСТВУЮЩЕЙ СЕЙСМОЛОГИЧЕСКОЙ СЕТИ ПЕРЕДАЧИ ДАННЫХ В ИНТЕРНЕТ

А.В. Андреев, М.Н. Жижин, А.А. Бурцев, А.Н. Поляков (Центр изучения геофизических данных ИФЗ РАН)

Основной задачей проекта является попытка минимизировать временные и финансовые затраты на организацию (восстановление) академической сети передачи цифровых геомагнитных данных с обсерваторий до коллективных центров компьютерной обработки и представления данных.

В качестве коллективных центров обработки и хранения данных используются распределенные узлы Сетевого интерактивного ресурса данных по солнечно-земной физике SPIDR. Первый узел SPIDR реализован в МЦД по солнечно-земной физике в Москве на базе параллельного кластера из 12 линукс-компьютеров с общим объемом дискового пространства 1 Тбайт и оптическим высокоскоростным выходом в Интернет 1 Гбит/с. Имеются результаты по сбору геомагнитных данных в режиме, близком к реальному времени, со станций Интермагнет (2002-2003 гг.). С 2002 г. SPIDR также используется для сбора и обмена данных с цифровых ионозондов США, Австралии и Южной Африки в реальном времени.

В качестве сети передачи данных предлагается в первую очередь использовать каналы Интернет совместно с Геофизической Службой РАН. В ЦОМЭ ГС РАН уже создана и расширяется сеть передачи мультидисциплинарных геофизических данных в реальном времени (в основном сейсмика и GPS) и имеется подготовленная инфраструктура, обслуживающая аппаратную часть, к которой могут быть подключены большинство среднеширотных геомагнитных и ионосферных обсерваторий. Пропускная способность этой сети позволит добавление потоков геомагнитных и ионосферных данных без снижения качества передачи сейсмических волновых форм.

В настоящее время отрабатывается система цифровой регистрации геомагнитных вариаций на базе кварцевых датчиков Боброва, разработанная в ИЗМИРАНе. Приборы достаточно просты в установке, обеспечивают необходимую точность измерений для магнитных обсерваторий, устойчивы к воздействиям окружающей среды, экономически выгодны по сравнению с зарубежными аналогами, и имеют интерфейс для подключения к системе передачи данных.

Интерактивный ресурс данных по солнечно-земной физике SPIDR

Интерактивный ресурс данных по солнечно-земной физике (Space Physics Interactive Data Resource, SPIDR) – это распределенная сеть синхронных баз данных и серверов приложений, позволяющая выбирать, визуализировать и моделировать исторические данные по космической погоде в сети Интернет. Система допускает интеграцию с вычислительными моделями околоземной среды и космической погоды в реальном времени. Распределенная архитектура системы является ключевым фактором для сокращения времени ожидания при визуализации мультимедийных данных и быстрой доставки данных на компьютеры пользователей для последующей обработки и анализа.

Проект SPIDR развивался в основном кооперативными усилиями двух научных центров: Национальным Центром Геофизических Данных (NGDC NOAA) в Болдере, США, и в Центром Изучения Геофизических Данных (ЦИГЕД РАН) в Москве. Международные организации для обмена данными по системе SPIDR – это мировые центры данных по солнечно-земной физике WDC STP. Они также являются сетевыми узлами SPIDR. Центру ионосферных прогнозов (Ionosphere Prediction Service, IPS) в Сиднее, Австралия, был присвоен статус узла SPIDR и в июле 1999 в нем были установлены сервер SPIDR и базы данных. В декабре 2000 сервер SPIDR и базы данных были установлены в Родезийском университете, Грехемстаун, Южная Африка, в январе 2001 сервер SPIDR и базы данных были установлены в Университете г. Нагойя, Япония, и в апреле 2003 узел SPIDR был установлен в Центре космической погоды Пекинского университета, Китай.

Основное преимущество такого распределения серверов SPIDR состоит в том, что региональные подключения к сети Интернет являются обычно более быстрыми, чем межконтинентальные связи. Таким образом, региональный web- сервер будет "более интерактивен" при запросах к данным и визуализации, хотя при этом может потребоваться войти в контакт с удаленными базами данных, если региональный узел SPIDR не имеет достаточной производительности на сервере приложений или просто не заинтересован в хранении локальной копии определенной базы данных.

Узел системы SPIDR обычно реализуется на компьютерном кластере и обеспечивает параллельный поиск, обработку и визуализацию больших объемов данных с применением методов искусственного интеллекта и нечеткой логики. Каждый сервер SPIDR имеет интерфейс управления базами данных, который позволяет выполнять модификацию данных удаленным пользователям с правами администратора по сети Интернет. Узлы системы SPIDR автоматически обмениваются обновлениями баз данных и программного обеспечения.

После того, как клиент с помощью браузера соединяется с web-сервером SPIDR, клиентской машине передаются несколько встраиваемых в браузер Java-приложений, которые там и работают (рис. 1). Они помогают исполнению операций и запросов. Делая запрос к системе, пользователь выбирает интервал времени, все необходимые параметры, возможно список станций или спутников, и заполненная форма запроса данных посылается через web-сервер на Java-сервлет, выполняющийся на сервере приложений. В свою очередь Java-сервлет распределяет JDBC-запросы на серверы базы данных и, после получения и обработки результатов, посылает их назад клиентской машине. Далее пользователь может оперировать с полученными данными с помощью загруженных ранее Java-апплетов и функций JavaScript, выполняющихся в пределах его браузера.

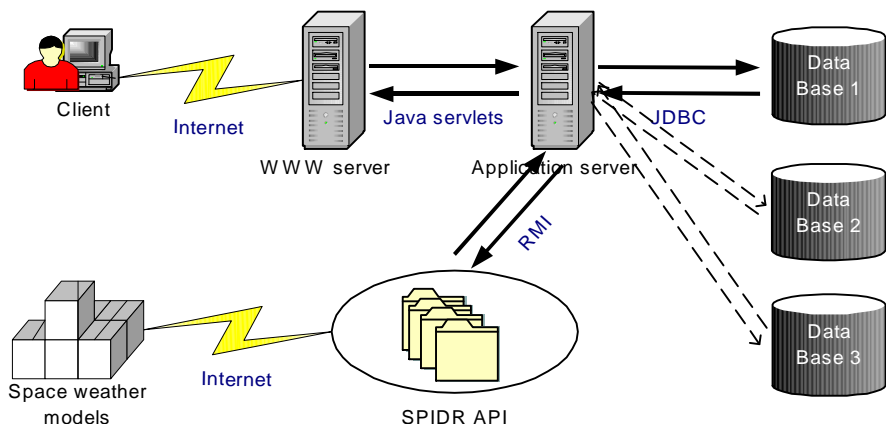


Рис. 1. Узел SPIDR

Наряду с интерактивной работой в системе SPIDR через web-сервер, данные из системы доступны и для запросов из других вычислительных систем (рис. 1). Вычислительные модели космической погоды могут посылать запросы данных и получать отобранные данные от SPIDR по сети Интернет, используя специальный прикладной программный интерфейс (API). API сформирован набором Java-классов с методами, вызываемыми через протокол удаленного обращения к методам (RMI и веб-сервисы). При запуске сервера приложений, классы самостоятельно регистрируются в специальном именованном пространстве, так называемом системном реестре RMI. Чтобы получать данные от SPIDR, вычислительная модель должна найти соответствующий класс API в системном реестре RMI сервера приложений и вызывать метод запроса к данным с параметрами, отражающими интервал времени, базу данных, таблицу базы, и столбец(цы) с параметрами, и, возможно, спутник или станцию наблюдений (все доступные классы, методы, и значения их параметров изложены в документации API).

Сервер приложений собирает данные из кластера баз данных, заносит их в общий класс-контейнер, называемый моделью данных SPIDR (data model), и посылает этот класс-контейнер в ответ на запрос по сети Internet.

Минимальные аппаратные требования для оптимальной работы сервера приложений системы SPIDR: мобильный компьютер типа notebook с процессором типа Pentium 300 МГц, оперативной памятью 32 Мбайт, дисковым пространством порядка 3 Гбайт, подключенный к сети Интернет по протоколу TCP/IP с помощью модема или выделенной линии. Типовая конфигурация основного комплекта оборудования, используемая узлами SPIDR в настоящее время: два процессора Pentium-III 1 ГГц, оперативная память 512 Мбайт, 100 Гбайт дискового пространства, подключение к Интернет со скоростью 100 Мбит/с.

Московский узел SPIDR расположен в помещении объединенного вычислительного центра ЦИГЕД ИФЗ и ГЦ РАН. Базы данных и программное обеспечение системы размещены на вычислительном кластере из 12 параллельных компьютеров типа IBM PC под управлением операционной системы Linux, соединенных между собой по протоколу Gigabit Ethernet с высокоскоростным оптоволоконным 1 Гбит выходом в Интернет. Типовой компьютер кластера - IBM PC с процессором Intel Celeron 733 МГц, оперативной памятью 512 Мбайт. Управление и внешний доступ к кластеру осуществляются через сервер приложений с двумя процессорами Intel Pentium III 1 ГГц с оперативной памятью 1 Гбайт. Пиковая производительность кластера подобной конфигурации достигает 10 Гфлоп/сек.

SPIDR объединяет распределенные тематические базы данных по различным направлениям солнечно-земной физики. Она также содержит дополнительные базы данных для административных целей. Схема подключений баз данных в SPIDR изображена на рис. 2. Общий конфигурационный файл содержит адреса распределенных серверов баз данных (*configuration data*). Административная база данных пользователей регионального узла (*Users*) связана со встроенной базой данных отзывов и предложений по работе системы (*StpBugrat*). База мета-данных (*Metadata*) содержит имена Java-классов для загрузки и экспорта данных из тематических баз, интервалы времени наличия этих данных, физические единицы измерений, заголовки и способы размещения графиков временных рядов, имена и координаты геомагнитных, ионосферных, солнечных обсерваторий и станций космических лучей, а также параметры орбиты и времени жизни научных спутников, данные от которых хранятся в системе. Ядром SPIDR является база данных *StpCore*, которая содержит основные геомагнитные и солнечные индексы, параметры межпланетного магнитного поля (IMF), и обычно устанавливается на каждый узел SPIDR.

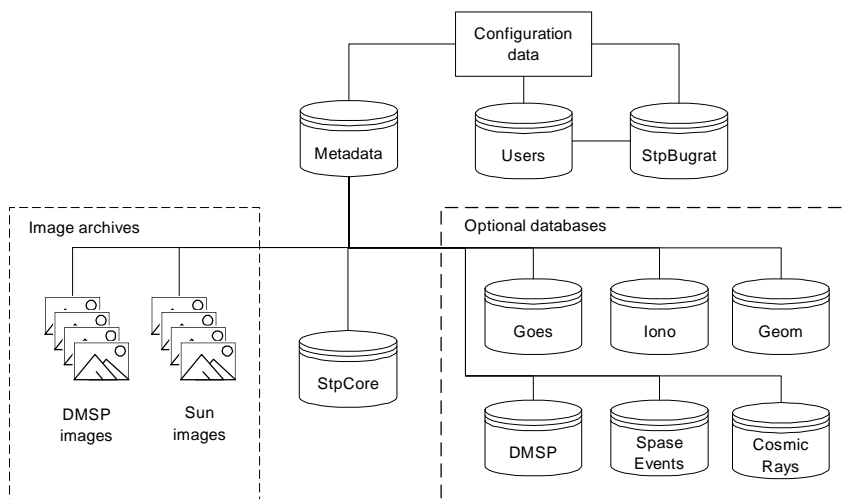


Рис. 2. Конфигурация баз данных SPIDR

Дополнительные базы данных (optional databases, рис. 2) могут быть установлены локально на узле SPIDR или запрашиваться с удаленного узла сервером приложений, в зависимости от конкретной конфигурации. Они включают базы данных изображений Солнца с различных спутников и обсерваторий (*Sun images*), базу данных по параметрам околоземной среды со спутников *GOES*, базу данных с изображениями Земли и орбитальными данными спутников *DMSP*, временные ряды наблюдений ионосферных и геомагнитных станций (базы данных *Iono* и *Geom*), временные ряды наблюдений со станций космических лучей (база данных *CosmicRays*), и каталог событий космической погоды (база данных *SpaceEvents*).

В таблице ниже приводятся краткие описания и временные интервалы наблюдений для данных, доступных в системе SPIDR:

База данных	Таблицы	Вид наблюдений	Года
Users		Регистрационные данные о пользователях SPIDR, место постоянного хранения пользовательских «корзин данных»	-
Stp Bugrat		Отчеты пользователей об ошибках системы и запросы о новых свойствах и функциях SPIDR	-
GOES satellites		Данные спутников GOES, 1 мин и 5 мин вариации и параметры орбит	1986 – 2002
Stp Core	IMF	Параметры межпланетного магнитного поля и солнечного ветра, 1 час вариации	1973 – 2002
	geomInd	Геомагнитные и солнечные индексы DST, Kp, ..	1932 – 2002
	SSN	Число и группы солнечных пятен	1800 – 2002

Ionoshere	-	Параметры ионосферы и ионосферных станций	1900 – 2002
Geomagnetic field	-	Минутные и часовые вариации магнитного поля Земли и параметры геомагнитных станций	1957 – 2002
Cosmic rays	-	Наблюдения космических лучей и параметры станций	1998 – 2002
Space weather events	-	События космической погоды (солнечные вспышки и выбросы частиц)	1975 – 2002
Sun images	-	Ежедневные изображения Солнца в различных диапазонах с обсерваторий и спутников	1995 – 2002
DMSP images	-	Изображения Земли в инфракрасном и видимом диапазонах со спутников	1991 – 2002
Night time lights	-	Изображения Земли с указанием положения стабильных ночных огней, полученные в результате совместного проекта NGDC/DMSP	-

Новая база данных может быть относительно просто добавлена в систему посредством объединенного интерфейса без остановки работы сервера приложений. Чтобы добавить базу данных в SPIDR, нужно указать сетевой адрес новой базы данных в конфигурационном файле, подготовить и разместить на web-сервере SPIDR HTML-форму для запроса к данным, и добавить мета-данные для нового типа данных (имя Java-класса для доступа к данным, размер окна графика, логарифмические/линейные масштабы и метки осей на графиках, и т.п.). После очередного автоматического обновления конфигурации сервера приложений (обычно раз в 10 мин), новая база данных станет доступна из системы SPIDR.

Для понимания архитектуры SPIDR важно определить термин "корзина данных". Это – специальный объект-контейнер, связанный с каждым зарегистрированным пользователем, который содержит список всех наборов данных, отобранных пользователем для графического анализа и экспорта из системы. Эта информация сохраняется от сессии к сессии для каждого зарегистрированного пользователя так, чтобы всякий раз, когда пользователь вновь входит в систему, он уже имел список наборов данных и интервалы дат, с которыми он работал в предыдущей сессии.

Корзина данных позволяет одновременную выборку и визуализацию многодисциплинарных данных, существенно расширяя возможности анализа и доступа к данным в целом. Например, добавив в корзину данные по солнечной активности и состоянию ионосферы (различные формы запроса данных и различные базы данных SPIDR), пользователь получает возможность построить все графики на одной странице в едином масштабе времени и загрузить все запрошенные данные с сервера в одном файле.

Визуализация временных рядов в SPIDR, главным образом, основана на технологии Java-апплетов, динамически встраиваемых в web-страницу. После того, как web-сервер получает запрос на визуализацию данных с формы тематической базы данных или со страницы с содержимым «корзины данных», специализированный графический модуль на сервере приложений

SPIDR соединяется с нужными базами данных (рис. 3), заносит выборку в объект-контейнер Data SequenceSet, сохраняет этот объект на жестком диске сервера в директории, доступной клиенту по протоколу HTTP, и посылает клиенту динамически сформированную HTML-страницу с внедренными ссылками на Java-апплеты, строящие графики (возможно несколько приложений-графиков на одной странице, рис. 4) и на сериализованные объекты-контейнеры с выборками данных (взаимно-однозначное соответствие между контейнерами и приложениями-графиками).

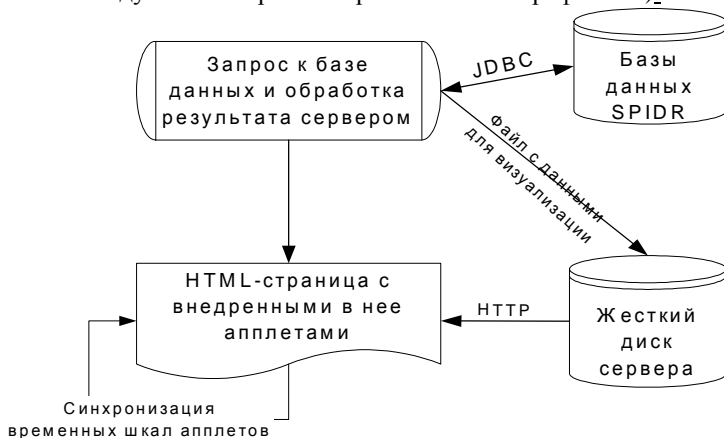


Рис. 3. Принцип визуализации данных в SPIDR

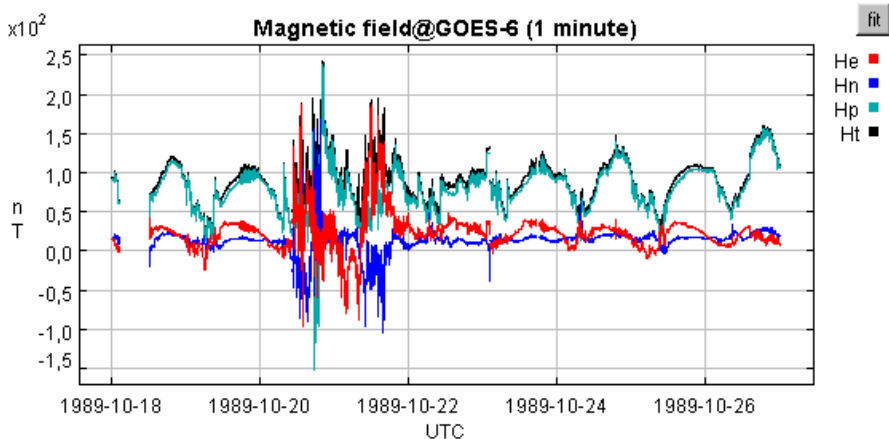


Рис. 4. Пример окна Java-апплета с графиком вариаций магнитного поля, зарегистрированных спутником GOES-6

Если на странице имеется больше чем один график, Java-апплеты «договариваются» между собой путем обмена событиями через объект-синглетон, и синхронно масштабируют временные оси всех графиков. В случае, если пользователь хочет изменить масштаб изображения в интересный ему период времени на одном из графиков, все остальные графики на этой странице автоматически масштабируются к тому же временному окну.

Информация обо всех графиках на странице приводится в обобщающей таблице, где указаны интервалы времени, для которых построены графики (могут различаться из-за пропусков в данных), физические единицы измерений, имена и координаты регистрирующих станций (см. табл. ниже). Для распечатки графиков на принтере и для включения их в статические электронные документы, сервер приложений может создавать графики временных рядов в виде статических изображений в формате GIF (в этом случае масштабирование графиков будет недоступно). Система автоматизированной синхронизации данных между региональными узлами SPIDR основана на обмене данными по подписке по электронной почте и удаленном вызове процедур загрузки файлов с данными на региональных узлах, подписавшихся на соответствующие обновления. Например, если администратор ионосферной базы данных на региональном узле в Москве загружает новый файл с данными, и региональный узел в Болдере подписан на обновления ионосферных данных из Москвы, то после проверки формата, *интерактивной* визуализации и успешной локальной загрузки данных на московском узле, сжатая копия файла данных с приложением инструкций по загрузке, выполненных в Москве, будет отправлена по электронной почте на специальный адрес в Болдере. Там это сообщение будет проверено на соответствие электронных подписей, файл с данными разархивирован и *автоматически* загружен в локальную базу данных. SPIDR позволяет загружать данные автоматически по электронной почте (так работает загрузчик геомагнитных данных сети Интермагнет) или с использованием «тонкого» клиента, который вызывается из командной строки скриптом операционной системы и передает файл данных и управляющие команды по протоколу SOAP на специальный веб-сервис на узле SPIDR (в настоящее время так загружаются ионосферные данные).

Первоначально SPIDR планировался как система для выборки и анализа данных из научных архивов в системе Мировых центров данных по солнечно-земной физике. Однако зачастую данные поступают в эти центры уже в реальном времени или с задержками порядка нескольких часов. Уже сегодня обмен ионосферными данными с обсерваторий между узлами SPIDR достигает 10000 сообщений в день. Возникла новая задача расширения системы SPIDR для работы с данными в реальном времени. В первом приближении это достигается периодическими обновлениями метаданных. Более сложными становятся проверка качества загружаемых данных, замена

предварительной версии данных, полученной в реальном времени, на окончательную версию, выработанную Мировыми центрами данных, и синхронизация региональных узлов.

Сеть передачи данных

В качестве сети передачи данных для автоматической загрузки в SPIDR предлагается в первую очередь использовать каналы Интернет совместно с Геофизической Службой РАН. В ЦОМЭ ГС РАН уже создана и расширяется сеть (рис. 5) передачи мультимедийных геофизических данных в реальном времени (в основном сейсмика и GPS) и имеется подготовленная инфраструктура, обслуживающая аппаратную часть, к которой могут быть подключены большинство среднеширотных геомагнитных и ионосферных обсерваторий. Пропускная способность этой сети позволяет добавление потоков геомагнитных и ионосферных данных без снижения качества передачи сейсмических волновых форм

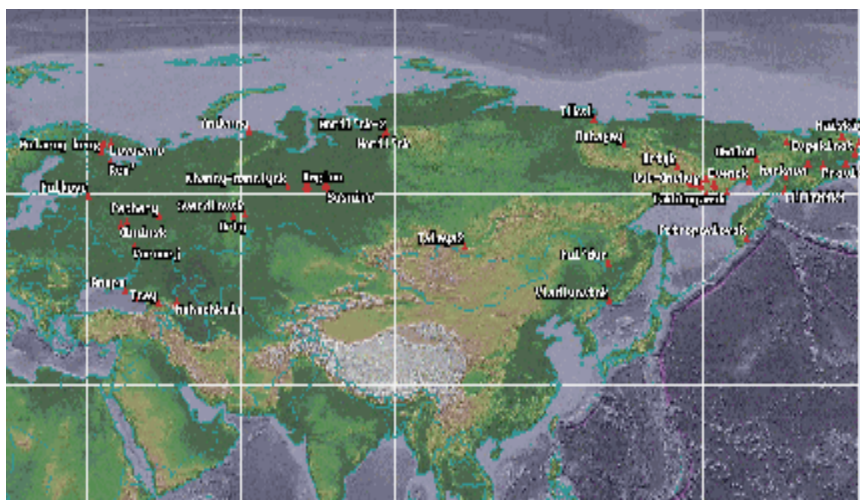


Рис. 5. Сеть станций ЦОМЭ ГС РАН, с которых возможен сбор цифровых геомагнитных данных.

В таблице ниже приводится список сейсмических станций, передающих данные в реальном времени и которые в первую очередь планируется использовать для сбора геомагнитных данных.

п/п	Название станции	Код	Широта, град.	Долгота, град.	Н, м
1	<u>Арги</u>	ARU	56.430N	58.562E	250
2	<u>Кисловодск</u>	KIV	43.955N	42.686E	1054
3	<u>Ловозеро</u>	LVZ	67.898N	34.651E	630

4	<u>Норильск</u>	NRIL	69.505N	88.441E	92
5	<u>Обнинск</u>	OBN	55.116N	36.569E	130
6	<u>Талая</u>	TLY	51.681N	103.644E	579

Цифровые станции

Цифровая магнитовариационная станция является высокочувствительным магнитометром, выполненным на основе кварцевых магнитных датчиков системы В.Н.Боброва в виде переносного прибора, и предназначена для непрерывной регистрации вариаций трех составляющих X (H), Y (D), Z вектора магнитной индукции (ВМИ) поля Земли, а также для исследования полей, создаваемых искусственными источниками и для специальных целей.

Внешний вид одного из вариантов ЦМВСБ показан на рис. 6.

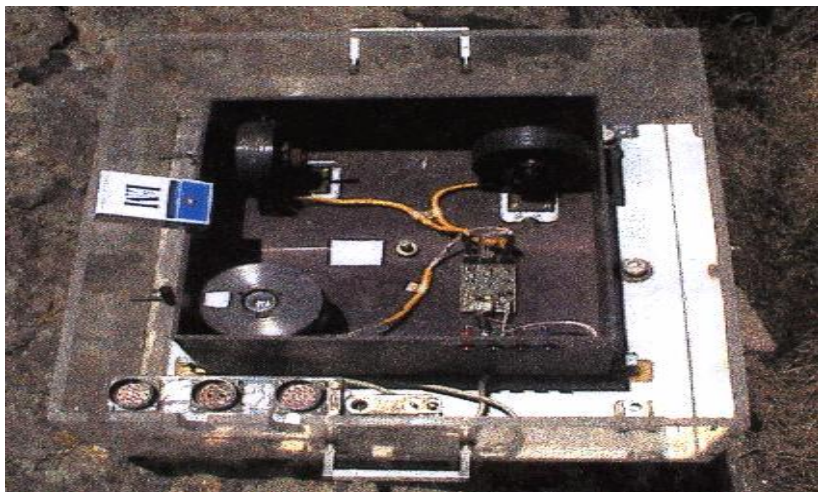


Рис. 6. Внешний вид цифровой магнитовариационной станции

Прибор включает в себя: корпус с тремя регулировочными ножками и двумя жидкостными уровнями, внутри которого на массивном основании расположен трехканальный магнитный измерительный преобразователь (МИП), выполненный на основе кварцевых датчиков (КД) с электрическим выходом оригинальной конструкции, и два соединительных кабеля для подключения источника питания и цифрового регистратора. МИП содержит три КД, отличительной особенностью которых является то, что фотоэлектрический преобразователь установлен непосредственно на кварцевой рамке внутри корпуса датчика, плату усилителей (УС) и плату цифрового регистратора (ЦР) данных. На корпусе ЦМВСБ расположены органы управления и контроля станции и соединительные разъемы,

позволяющие производить ориентацию и установку КД, питание ее от источника постоянного тока (аккумуляторной батареи или стандартного сетевого адаптера) и передачу данных через последовательный канал связи в специальное устройство сбора и накопления информации (УСНИ) или в персональный компьютер (ПК).

На боковых стенках корпуса прибора около каждого КД установлено по два компенсационных магнита, позволяющих производить настройку вариометров в измерительный диапазон в заданном районе проведения работ. КД канала Z-вариометра установлен на кронштейне, а КД каналов D и H, – на угольниках, укрепленных на корпусе и ориентированных взаимно перпендикулярно с точностью до ± 5 угл. мин. Юстировка всех КД выполнена так, что оси вращения подвижных магнитов и их магнитные оси ориентированы параллельно посадочным поверхностям КД с точностью до ± 3 угл. мин.

Плата УС содержит три идентичных канала усиления, позволяющих обеспечивать необходимую крутизну преобразования поле-напряжение МИП, датчик измерения температуры, два линейных стабилизатора напряжения питания усилителей и источник питания для свето- фотодиодов. Плата ЦР содержит АЦП с управляющим микропроцессором, энергонезависимую память, объем которой позволяет накапливать измеренную информацию о вариациях магнитного поля на заданном промежутке времени, последовательный канал передачи данных во внешний регистратор и преобразователь напряжения DC-DC.

Питание ЦМВСБ осуществляется от аккумуляторной батареи или от стандартного сетевого адаптера напряжением постоянного тока 12 ± 3 В, при этом потребляемая мощность не более 0,5 Вт.

Основные технические характеристики :

Динамический диапазон измерения вариаций, составляющих ВМИ по каждому измерительному каналу	± 1000 нТл
Цена единицы счета наименьшего разряда цифрового отсчетного устройства	$\pm 0,1$ нТл
Цикл автоматических измерений	0,1...3600 с
Относительная нестабильность показаний во времени	< 1 нТл/сут.
Диапазон рабочих температур	от -40 до $+60^{\circ}\text{C}$
Относительная нестабильность от влияния температуры	$< \pm 0,1$ нТл/ $^{\circ}\text{C}$
Зависимость показаний от угла наклона	$< \pm 0,2$ нТл/угл.мин.
Объем энергонезависимой памяти	1 (8) Мбайт
Скорость передачи данных по последовательному каналу RS232	9600 бод
Напряжение питания от источника постоянного тока	12 ± 3 В
Мощность потребления	$< 0,5$ Вт
Габаритные размеры	320x320x150 мм
Масса	< 5 кг