

doi: 10.11720/wtyht.2020.1485

张伟,胡蕾,张钊搏.LEMI-417I 型地球深部电磁场观测系统的数据格式解析[J].物探与化探,2020,44(4):810-815.http://doi.org/10.11720/wtyht.2020.1485

Zhang W, Hu L, Zhang Z B. Raw data format analysis of LEMI-417 earth deep electromagnetic field observation system[J]. Geophysical and Geochemical Exploration, 2020, 44(4): 810-815. http://doi.org/10.11720/wtyht.2020.1485

# LEMI-417 型地球深部电磁场观测系统的数据格式解析

张伟<sup>1,2</sup>, 胡蕾<sup>3</sup>, 张钊搏<sup>3</sup>

(1. 中国地质调查局 成都地质调查中心, 四川 成都 610082; 2. 电子科技大学 资源与环境学院, 四川 成都 611731; 3. 中科院成都信息技术股份有限公司, 四川 成都 610041)

**摘 要:** LEMI-417 长周期大地电磁测深仪是目前国内普及程度较高的一套通过观测超低频段天然电磁场信号来探测地球深部壳幔电性结构的观测系统。本文对该系统采集的天然大地电磁场数据的原始二进制数据结构进行了深入研究, 剖析出了其对应的计算机数据结构, 并列出了核心的程序代码, 以作为下一步的后期数据解释软件开发之基础。

**关键词:** LEMI-417 长周期大地电磁测深仪; 原始数据格式; C++

**中图分类号:** P631.3      **文献标识码:** A      **文章编号:** 1000-8918(2020)04-0810-06

## 0 引言

长周期大地电磁法目前是探测地球深部(>150 km)上地幔尺度电性结构特征的主要方法, 由乌克兰 Lviv Centre of Institute of Space Research 研究所研制的 LEMI-417 仪器是最早引入到我国用于长时间、稳定、连续观测超低频大地电磁场的一套成熟装备。2007 年起, 中国科学院地质与地球物理研究所<sup>[1]</sup>、成都理工大学<sup>[2]</sup>、中国地质大学(北京)<sup>[3-4]</sup>以及中国地震局地质研究所<sup>[5]</sup>等科研院所开始使用该仪器探测青藏高原及邻区、龙门山构造带以及川滇黔相邻区地震活动带的深部地质结构, 用于揭示其关键地质带的深部动力学特征。与该仪器配套的数据处理软件是由俄罗斯 Ivan M. Varantsov 等编写的 PRC\_MTMV 商业软件, 用于实现将大地电磁场从时间域到频率域的转换, 通过傅里叶变换计算得到功率谱, 并通过功率谱估计得到大地电磁阻抗要素。

众所周知, 计算机对于二进制格式数据的读取效率远远大于文本格式数据, 因此在进行傅里叶变

换时直接读取二进制数据进行计算将大大提高程序的运行效率。但由于乌克兰仪器生产厂家没有公开给出其原始数据的二进制格式, 因此在处理 LEMI-417 数据时, 必须首先使用厂家提供的转换程序将仪器二进制记录数据转换为文本格式, 再使用 PRC\_MTMV 软件的 LEMI2PRC 子模块将 ASCII 文本格式数据再次转换为二进制数据后方能进行后续的傅里叶变换处理<sup>[3]</sup>, 这无疑明显地增加了一步不必要的处理操作。为此, 作者对照 V5-2000 仪器的二进制数据格式的特点<sup>[6]</sup>, 对 LEMI-417 仪器的原始二进制数据格式进行了深入研究, 在不断试错后, 成功剖析出了该仪器二进制记录数据的数据结构。

## 1 LEMI-417 型系统的原始数据格式

LEMI-417 是专门针对长时间采集超低频段天然电磁场而研制的一款观测系统, 其主要特点是: ①采用磁通门传感器来采集磁场数据, 观测范围为  $\pm 70\,000\text{ nT}$ , 可以记录到几十万秒的长周期信号; ②电道采集盒子通过专门的电路设计能够保证在长

收稿日期: 2019-10-21; 修回日期: 2019-12-05

基金项目: 中国地质调查局地质调查项目 (DD20190033); 国家自然科学基金项目 (41804144)

作者简介: 张伟 (1983-), 高级工程师, 在职博士后。主要从事金属矿产地球物理勘查与物化探软件开发工作。

时间范围内保持很低的漂移,提供了 4 组电道接口,可以同时采集 4 组电极的电场数据;③采用高精度 GPS 同步授时能够保证在长时间范围内具有很小的时移;④整体功耗很低,能够保证在 1 块 12V45ah 的铅酸电池供电下连续工作一周以上。

该仪器以 1 Hz 的采样率来连续观测天然电磁场,采集的数据以二进制格式实时储存在仪器内部的 CF 卡中,文件名以 L417\_×××.B××的规则来进行命名,其中:L417 为仪器的型号名称;\_×××为单点的数据存放文件序号,第一个数据文件编号为 001,

当 GPS 时间为格林威治时间的零点时,仪器系统自动新建一个新文件来存放数据,文件序号自动在前一个文件序号的基础上累加 1;B××的首字母 B 是二进制(binary)的缩写,××是采集仪器的出厂编号。

1.1 头文件格式

通常与测点数据采集相关的记录信息(如仪器型号、采集日期、经纬度、高程、电极距等)一般以文件头的形式存放于二进制文件的首端。LEMI-417I 型仪器共分配了 32 个字节的大小空间来存放信息(表 1),各字节位置信息阐述如下。

表 1 LEMI-417 原始二进制记录数据的头文件格式(tag)  
Table 1 LEMI-417 header file format for raw binary data (tag)

字节位置	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16
信息内容	4	1	7	型号	编号	年	月	日	时	分	秒	纬度(度分)				N S
字节位置	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
信息内容	经度(度分)					E W	空 字 符	高程	采样 设置	VIN	VBAT	L1	L2	L3	L4	

1~4 字节:4 个字节长度,存放采集仪器的型号信息,对于 LEMI-417I 型仪器来说,第 1 个字节为 ASCII 字符‘4’(16 进制为 34),第 2 字节为 ASCII 字符‘1’(16 进制为 31),第 3 字节为 ASCII 字符‘7’(16 进制为 37),第 4 字节为 ASCII 字符‘1’(16 进制为 49)。

5 字节:1 个字节长度,存放采集仪器的出厂编号。

6 字节:1 个字节长度,存放开始采集数据时的当前年信息。

7 字节:1 个字节长度,存放开始采集数据时的当前月信息。

8 字节:1 个字节长度,存放开始采集数据时的当前日信息。

9 字节:1 个字节长度,存放开始采集数据时的当前小时信息(24 小时制)。

10 字节:1 个字节长度,存放开始采集数据时的当前分信息。

11 字节:1 个字节长度,存放开始采集数据时的当前秒信息。

12~15 字节:4 个字节长度,存放当前测点 GPS 同步的纬度。

16 字节:1 个字节长度,存放当前测点坐标是位于北纬还是南纬,以我国位于北半球为例,在我国测量数据时应为 ASCII 字符‘N’(16 进制为 4E)。

17~21 字节:5 个字节长度,存放当前测点 GPS 同步记录的经度信息。

22 字节:1 个字节长度,存放当前测点坐标是位于东经还是西经,以我国位于东半球为例,在我国测量数据时应为 ASCII 字符‘E’(16 进制为 45)。

23 字节:1 个字节长度,为空字符(对应的 16 进制为 00)。

24~25 字节:2 个字节长度,存放当前测点 GPS 同步的高程(单位:m)。

26 字节:1 个字节长度,存放仪器的采样频率。

27 字节:1 个字节长度,存放仪器内部的 VIN(外部供电电源电压)信息。

28 字节:1 个字节长度,存放仪器内部的 VBAT(内部纽扣电池电压)信息。

29 字节:1 个字节长度,存放电道盒子上第 1 对电极的极距长度(单位:m)。

30 字节:1 个字节长度,存放电道盒子上第 2 对电极的极距长度(单位:m)。

31 字节:1 个字节长度,存放电道盒子上第 3 对电极的极距长度(单位:m)。

32 字节:1 个字节长度,存放电道盒子上第 4 对电极的极距长度(单位:m)。

1.2 时间序列格式

如图 1 所示,在头文件(Tag)后紧接着的是时间序列数据,LEMI-417 以 30 个字节的大小空间来

存放一个采样时间片采集到的电磁场数据段 (segment, 表 2 所示), 由 17 个时间片组成一个记录块 (record), 一个数据段结束后再接下一个文件头数

据, 依次循环、直到采集结束。其中, 单个时间片数据段中各字节位置信息阐述如下。

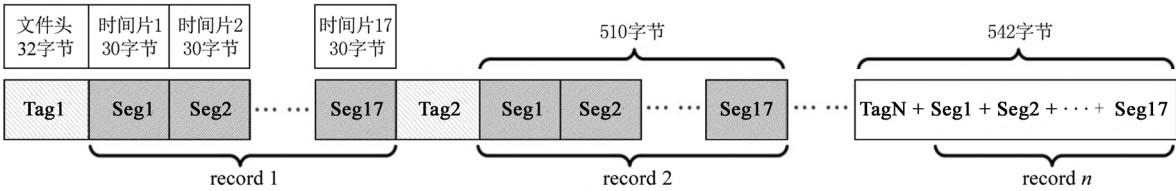


图 1 LEMI-417 原始数据结构分解示意

Fig.1 Schematic diagram of original data structure decomposition

表 2 LEMI-417 原始二进制记录数据的时间序列格式 (segment)

Table 2 LEMI-417 time series format of raw binary recorded data

字节位置	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
信息内容	$H_x$			$H_y$			$H_z$			$E_1$				$E_2$	
字节位置	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
信息内容	$E_2$		$E_3$				$E_4$				$T_f$		$T_e$		$No$

1~3 字节: 3 个字节长度, 存放磁通门传感器  $X$  方向的测量值 (单位: nT)。

4~6 字节: 3 个字节长度, 存放磁通门传感器  $Y$  方向的测量值 (单位: nT)。

7~9 字节: 3 个字节长度, 存放磁通门传感器  $Z$  方向的测量值 (单位: nT)。

10~13 字节: 4 个字节长度, 存放第 1 道电极间的测量值 (单位: mkV/m)。

14~17 字节: 4 个字节长度, 存放第 2 道电极间的测量值 (单位: mkV/m)。

18~21 字节: 4 个字节长度, 存放第 3 道电极间的测量值 (单位: mkV/m)。

22~25 字节: 4 个字节长度, 存放第 4 道电极间的测量值 (单位: mkV/m)。

26~27 字节: 2 个字节长度, 存放磁通门传感器温度 (单位:  $^{\circ}\text{C}$ )。

28~29 字节: 2 个字节长度, 存放电道采集盒子的温度 (单位:  $^{\circ}\text{C}$ )。

30 字节: 1 个字节长度, 存放时间片采样间隔 (单位: m)。

2 计算机数据类型与程序实现

2.1 头文件数据结构与核心代码

在 C++ 语言中可利用 struct 类型变量来结构化

地读取和存储头文件信息, 再按表 1 所示各信息的字节位置和字节长度, 定义其内部的子变量数据类型, 其核心代码如下:

```
// 头文件数据结构定义
struct TagData
{
    unsigned char str_1; //4
    unsigned char str_2; //1
    unsigned char str_3; //7
    unsigned char str_4; //仪器型号
    unsigned char no; //仪器编号
    unsigned char year; //年
    unsigned char month; //月
    unsigned char day; //日
        unsigned char hour; //时
        unsigned char minute; //分
    unsigned char second; //秒
    unsigned char latitude[4]; //纬度
    char N_S; //北纬还是南纬
    unsigned char longitude[5]; //经度
    char E_W; //东经还是西经
    char separator; //空字符
    unsigned char altitude[2]; //高程
    unsigned char average; //采样率
    unsigned char vin; // 外部电压
    unsigned char vbat; // 内部电压
    unsigned char L1; //第一对极距长度
```

```
unsigned char L2; //第二对极距长度
unsigned char L3; //第三对极距长度
unsigned char L4; //第四对极距长度
}
// ... ..由于篇幅有限省略部分非关键代码
// 读取仪器原始数据文件
TagData tag;
FILE * fp = fopen(strFilePath,"rb");
char dataBuf[32] = "";
fread(dataBuf,sizeof(dataBuf),1,fp);
fclose(fp);
// 结构体各子变量赋值
tag.str_1=dataBuf[0];
tag.str_2=dataBuf[1];
// ... ..由于篇幅有限省略部分代码
tag.latitude[0]=dataBuf[11];
tag.latitude[1]=dataBuf[12];
tag.latitude[2]=dataBuf[13];
tag.latitude[3]=dataBuf[14];
// ... ..由于篇幅有限省略部分代码
tag.L4=dataBuf[31];
```

2.2 时间序列数据结构与核心代码

同理,利用 struct 类型变量来结构化地读取和存储时间片数据段信息,再按表 2 所示各信息的字节位置和字节长度,定义其内部的子变量数据类型,其核心代码如下:

```
// 单个时间片数据段结构定义
struct SegmentData
{
    char Hx[3]; //Hx
    char Hy[3]; //Hy
    char Hz[3]; //Hz
    char E1[4]; //E1
    char E2[4]; //E2
    char E3[4]; //E3
    char E4[4]; //E4
    char TF[2]; //TF
    char TE[2]; //TE
    char No; //
}
// ... ..由于篇幅有限省略部分非关键代码
// 定义变量
int Hx=0,Hy=0,Hz=0, E1=0, E2=0, E3=0, E4=0;
short int Tf=0,Te=0;
char No;
TagData tag;
SegmentData segment;
// 读取仪器原始数据文件
fseek(infile_origin,0L,SEEK_SET);
```

```
cur_rc = 0;
int No = 0;
// 如果当前文件指针读取位置不是结尾
while( cur_rc < fileSize)
{
    if(i % 17 == 0)
    {
        fread(tag,sizeof(unsigned char),32L,infile_origin);
        cur_rc = ftell(infile_origin);
        i++;
        continue;
    }
    fread(segment,sizeof(unsigned char),30L,infile_origin);
    cur_rc = ftell(infile_origin);
    i++;
    j++;
    // 将二进制数值转换为整型变量,详见下文
    convertSegmentData(segment,&Hx,&Hy,&Hz,&E1,&E2,&E3,&E4,&Tf,&Te,&No);
    pre_rc = cur_rc;
}
```

2.3 二进制数值到整型变量的转换

LEMI-417 原始数据是在仪器采集时存储的二进制数据,将它赋值到计算程序语言的整型变量时必须进行二进制位操作转换。以磁通门传感器的  $x$  方向磁场数据( $H_x$ )为例,由于整型变量(int)无论是在 32 位还是 64 位操作系统其字节长度均为 32 位,幸运地避开了不同操作系统程序代码的特殊处理问题,位转换的实质就是将原始数据中各字节位置的数值完整地赋值到整型变量的各字节位置上,其核心代码如下:

```
int temp1=0,temp2=0,temp3=0;
temp1 = segment.Hx[2];
temp1= temp1 << 16;
temp1 = temp1 & 0x00FF0000;
temp2 = segment.Hx[1];
temp2 = temp2 << 8;
temp2 = temp2 & 0x0000FF00;
temp3 = segment.Hx[0];
temp3 = temp3 & 0x000000FF;
Hx = temp1|temp2|temp3;
if(Hx >=0)
{
    Hx = Hx & 0x00FFFFFF;
}
else
{
```



```
Hx = Hx | 0xFF000000;  
}
```

3 应用开发实例

在对原始数据格式成功剖析的基础上,笔者针对该 LEMI-417 仪器在实际工作中存在的一些问题,补充性地开发出一套基于 MFC 的、可以在任何 Windows 系统运行的界面程序,辅助一线野外工作人员对个别偶然会出现的数据问题进行人为纠正。

实际应用中,LEMI-417 型仪器在无 GPS 信号时会出现记录紊乱的现象<sup>[4]</sup>。在峡谷地貌或林木茂密区开展工作时可能会出现个别时段 GPS 卫星信号偶然解锁的现象,使得仪器无法同步当前的 GPS 时钟和测点坐标,在对应的文件头信息中记录下多组差别迥异的信息。例如,野外工作中仪器部署在深切山谷地貌区时,在部分时段会出现仪器 GPS 无法锁定到≥3 颗卫星的情况,此时仪器记录的测点经纬度和高程均变为 0,个别情况甚至会出现 1 颗卫星都无法锁定的情况,此时便会出现卫星失锁问题,无法同步到正确的 GPS 时钟,使采集数据时序出现混乱,从而导致在后期数据导出时无法顺利形成连续的时间序列数据,需要进行 2 次返工,造成不必要的成本支出。

针对该常见问题,开发出了 Lemi-417 数据修改器(图2),可以将当前测点采集的连续多天的数据

加载、合并到一起,遍历每个文件头中的 GPS 时钟、测点经度、纬度、高程信息,判断每段数据对应的 GPS 时钟是否是以设置的采样率进行连续观测,会缺失的、不连续的时间序列进行三次样条插值修复,此外判断每段数据对应的经度、纬度和高程信息是否出现较大的数据偏差,通过取众数方法,识别出正常状态下和故障状态下的记录,对故障状态下的记录进行自动修复。最后通过最小二乘法对正常状态下的 GPS 经度、纬度和高程进行线性回归拟合得到高精度的测点定位信息。由此,通过计算机程序解决了以往由于 GPS 偶然失锁而需要重复返工的问题。

4 结论

随着我国广大学者对大地电磁时频转换算法研究的不断深入<sup>[7-9]</sup>,国产化数据处理软件<sup>[10]</sup>、仪器设备<sup>[11]</sup>的研制工作也已取得长足进步。本文深入剖析了 LEMI-417I 型地球深部探测主流观测系统的原始数据格式,较详细地给出了对应的核心程序代码,为广大从事天然电磁场现代数字信号处理方面研究的学者提供一定的前期基础支撑,对今后同类型国产仪器、软件的研发工作具有一定的借鉴意义。此外,也针对 LEMI-417I 型仪器存在的 GPS 解锁时的记录紊乱问题,提出了一个经济有效的计算机辅助纠正技术方案。

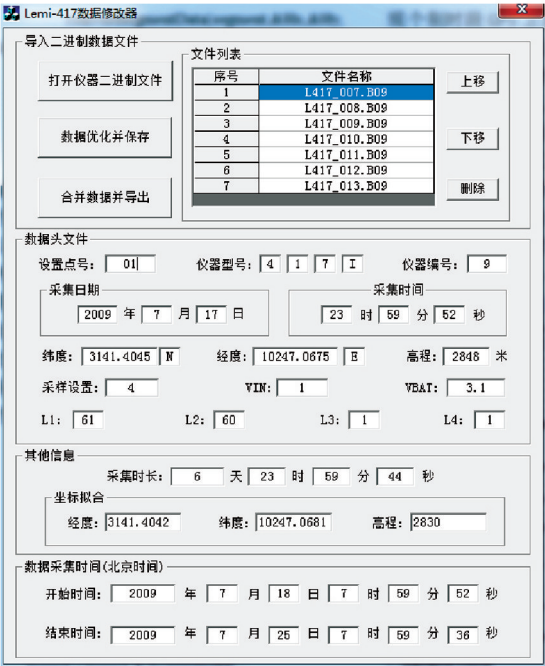


图2 窗口程序界面截图

参考文献(References):

[1] 肖鹏飞,白登海,Varentsov I M,等.长周期大地电磁测深研究——青藏高原东部 LMT 响应函数及应用[J].地震地质,2010,32(1):38-50.  
Xiao P F, Bai D H, Ivan M.Varentsov, et al. Study on long-period magnetotelluric sounding—The LMT transefer function in eastern Tibetan plateau[J].Seismology and Geology,2010,32(1):38-50.  
[2] 张伟.长周期大地电磁法信号处理关键技术研究与应用[D].成都:成都理工大学,2011.  
Zhang W. Research and application on key signal processing technologies of Long-period Magnetotelluric[D].Chengdu:Chengdu University of Technology,2011.  
[3] 郭泽秋,叶高峰,王辉,等.PRC\_MTMV 和 EMTF 在长周期大地电磁数据处理中的应用[J].地球物理学进展,2013,28(3):1192-1198.  
Guo Z Q, Ye G F, Wang H, et al. The application of PRC\_MTMV and EMTF in long period magnetotelluric data processing[J].Progress in Geophys,2013,28(3):1192-1198.  
[4] 巩秀钢,魏文博,金胜,等.CS3301 在长周期大地电磁测深仪中的应用研究[J].地球物理学报,2012,55(12):4051-4057.  
Gong X G, Wei W B, Jin S, et al.Application of CS3301 to the

Fig.2 Screenshot of the window program interface

long-period MT instrument [J]. Chinese Journal of Geophysics, 2012, 55(12): 4051–4057.

[5] 程远志, 汤吉, 陈小斌, 等. 南北地震带南段川滇黔接壤区电性结构特征和孕震环境[J]. 地球物理学报, 2015, 58(11): 3965–3981.

Cheng Y Z, Tang J, Chen X B, et al. Electrical structure and seismogenic environment along the border region of Yunan, Sichuan and Guizhou in the south of the North-South seismic belt [J]. Chinese Journal of Geophysics, 2015, 58(11): 3965–3981.

[6] 柳建新, 刘春明, 马捷, 等. V5-2000 大地电磁测深仪文件头数据格式研究[J]. 物探化探计算技术, 2007, 29(4): 359–362.

Liu J X, Liu C M, Ma J, et al. The data format of MT sounding instrument V5-2000 [J]. Computing Techniques for Geophysical and Geochemical Exploration, 2007, 29(4): 359–362.

[7] 王辉, 姚郁松, 许滔滔, 等. 利用地磁台站数据和站间传递函数估算长周期大地电磁测深的本地阻抗[J]. 地球物理学进展, 2019, 34(1): 200–207.

Wang H, Yao Y S, Xu T T, et al. Estimation of local impedance for long period magnetotelluric data using geomagnetic data and interstation transfer functions[J]. Progress in Geophysics, 2019, 34(1): 200–207.

[8] 李晋, 燕欢, 汤井田, 等. 基于匹配追踪和遗传算法的大地电磁噪声压制[J]. 地球物理学报, 2018, 61(7): 3086–3101.

Li J, Yan H, Tang J T, et al. Magnetotelluric noise suppression based on matching pursuit and genetic algorithm [J]. Chinese Journal of Geophysics, 2018, 61(7): 3086–3101.

[9] 王辉, 程久龙, 姚郁松, 等. 基于站间天然电磁场单位脉冲响应的大地电磁时间序列去噪方法[J]. 地球物理学报, 2019, 62(3): 1057–1070.

Wang H, Cheng J L, Yao Y S, et al. A new method of noise deletion in magnetotelluric time-series based on impulse response of inter-station transfer function [J]. Chinese Journal of Geophysics, 2019, 62(3): 1057–1070.

[10] 代小威. 基于 V5-2000 格式 MT 时间序列处理与功率谱估计及软件开发[D]. 西安: 长安大学, 2014.

Dai X W. Processing and power spectrum estimation of MT time serieses in V5-2000 format development [D]. Xian: Chang'an University, 2014.

[11] 高嵩, 汪莹莹, 李志鹏, 等. 三轴磁通门探头球形反馈线圈的设计方法研究[J]. 地球物理学报, 2019, 62(10): 3785–3793.

Gao S, Wang Y Y, Li Z P, et al. Design of a spherical feedback coil for the three-axis fluxgate probe [J]. Chinese Journal of Geophysics, 2019, 62(10): 3785–3793.

Raw data format analysis of LEMI-417 earth deep  
electromagnetic field observation system

ZHANG Wei<sup>1,2</sup>, HU Lei<sup>3</sup>, ZHANG Zhao-Bo<sup>3</sup>

(1. Chengdu Center of China Geological Survey, Chengdu 610082, China; 2. School of Resources and Environment, University of Electronic Science and Technology of China, Chengdu 611731, China; 3. Chengdu Information Technology Co., Ltd., Chinese Academy of Sciences, Chengdu 610041, China)

**Abstract:** When exploration of the deep electrical structure of the earth’s crust and mantle scale is conducted. LEMI-417 long-period magnetotelluric sounding instrument is a prevailing system for continuously observing natural electromagnetic field. In this paper, through in-depth study of the original binary data structure of the system, the authors analyzed the corresponding data structure in C++ computer language and also listed the core program code so as to provide necessary support for researchers who are engaged in the research on deep earth exploration and signal processing in electromagnetic field.

**Key words:** LEMI-417; raw data format; C++ program

(本文编辑: 沈效群)