

mergeTbn30miDataFile说明

软件当前版本: v20.01.020

文档更新时间: 2023-02-09

作者 : fu.sky

版本变化历史记录:

2023-02-09	: 因河北尚义石井电场发现bug(不能删除多个结果文件夹过期文件) 由 v20.01.000 升级到 v20.01.010
2022-06-17	: 因河北尚义石井电场需求(输出多个文件)由 v20.01.000 升级到 v20.01.010
2021-03-17	: 初始版本 : v20.01.000

目录

- -----1.功能说明
- -----2.软件目录结构说明
- -----3.格式要求
 - -----3.1 1分钟数据源文件
 - -----3.2 生成的结果文件
- -----4.软件部署
- -----5.其他需要说明的情况
- -----6.文档变更历史记录

1 功能说明

简述: 根据已经生成的30个单风机1分钟数据文件生成单风机30分钟数据文件

详细情况: 此程序根据配置文件cfg/cfg.cfg的实际配置值，再根据已经生成好的30个 1分钟数据文件（此文件由其他满足要求的软件生成，不是此软件功能）合成 30分钟的上传文件。

程序根据运行时间找对应数据文件的规则如下：

（为了举例说明方便假定1分钟文件名类似:genwnd_1_20210317_1330.cime
生成的30分钟结果文件名类似:JIMENHE_HT_20210317_1430.DJ
）

（1）如果程序运行时刻对应的分钟数值在0-29之间（包括0和29）则找前一
小时30-59分钟对应的文件；结果文件名的时间小时值为前一小时分钟
值为30
eg：程序运行时刻为：2021-03-17 15:01:**
对应要找的1分钟数据文件名为:genwnd_1_20210317_1430.cime
到genwnd_1_20210317_1459.cime 这30个文件
生成的目标文件名为:JIMENHE_HT_20210317_1430.DJ

（2）如果程序运行时刻对应的分钟数值在30-59之间（包括30和59）则找当前
小时0-29分钟对应的文件；结果文件名的时间小时值为当前小时分钟
值为00
eg：程序运行时刻为：2021-03-17 15:31:**
对应要找的1分钟数据文件名为:genwnd_1_20210317_1500.cime

到genwnd_1_20210317_1529.cime 这30个文件

生成的目标文件名为:JIMENHE_HT_20210317_1500.DJ

说明: 生成的目标文件名除时间之外, 可以在配置文件中配置

返回目录

2 软件目录结构说明

软件所在目录的结构和相关说明如下:

```
../merge_tbn_30miDataFile      <----- #程序所在目录 (包) 名
├─ mergeTbn30miDataFile.sh      <----- #运行脚本的名称
├─ myDiyShFunction.sh           <----- #脚本用到的一些函数文件
├─ manageCron                   <----- #在终端运行此程序来配置mergeTbn30miDataFile.sh的cron设置
├─ version.txt                  <----- #程序运行后此文件记录当前脚本的版本号
├─ readme.md                    <----- #MarkDown格式的说明文档
├─ .tRecorded                   <----- #用于记录当天已经生成文件 (防止重复生成)
├─ cfg                          <----- #程序配置文件目录
│   └─ cfg.cfg                  <----- #程序的配置文件
├─ gen_tmp_srcfile              <----- #程序自测用于生成数据源文件工具 (在没有生成1分钟程序时方便测试用)
│   └─ gen.sh                   <----- #自测试时生成文件脚本
│       └─ result               <----- #gen.sh生成的结果文件目录
├─ log                          <----- #软件运行日志目录
│   └─ cron.txt
│       └─ mergeTbn30miDataFile_20210318.log <--软件运行日志 (年月日根据实际运行时间变动)
├─ result                       <----- #软件自测时的临时目录
├─ result1                      <----- #软件自测时的临时目录
├─ sample_output_file           <----- #开发此需求时的样例文件目录
│   └─ JIMENHE_HT_20210311_1430.DJ
│   └─ SHIJING_SQ_20220505_0200.DJ
│   └─ SHIJING_JF_20220505_0200.DJ
└─ tmp                          <----- #软件运行时产生的临时文件目录
    └─ do                       <----- #软件运行时将1分钟源文件拷贝到此目录再处理
        └─ genwnd_1_20210318_0958.cime <--- #软件运行时临时拷贝的处理数据源文件
            └─ genwnd_1_20210318_0959.cime
        └─ tmp_fj_EC_9.bak       <----- #对某个风机EC处理的临时文件
            └─ tmp_fj_EC_9.txt   <----- #对某个风机EC处理的临时文件
```

返回目录

3 格式要求

要用此软件合成30分钟文件, 需要有如下格式要求, 否则不支持合成功能

3_1 1分钟数据源文件

文件名格式:

xxxxxxx_20210318_0958.xxxxxx #其中"xxxx"可以任意, 其中的数字代表年月日, 小时分钟

文件内容除了文件头和尾其他内容类似如下格式:

@	EC	PP_AVG	PQ_AVG	WS_AVG	SF_OGN	FAULT
#	0	35.861	2.123	9.856	15	16^17
@	EC	PPAVG	PQ_AVG	WS_AVG	SF_OGN	FAULT
#	1	15.369	21.230	8.624	45	16^17
@	EC	PPAVG	PQ_AVG	WS_AVG	SF_OGN	FAULT
#	2	15.369	6.369	9.856	44	16^17

- 注意：
- 1.其中@行的字段名对脚本来说不重要，脚本不解析，脚本只解析#行
 - 2.文件格式中的每列代表的函义不能变，变了此脚本生成的逻辑就要更改了；
 - 3.每一行的FAULT列可以有多列多列用空格隔开不影响脚本正常生成结果文件。

3_2 生成的结果文件

注意:从脚本的v20.01.010 开始可以配置生成多个文件

文件名格式:

```
JIMENHE_HT_20210311_1430.DJ
#JIMENHE:风场名; HT:风机厂家编码;
#20210311_1430 年月日时分
# 文件命名时间部分采用向前取时的方式，例如
# 风机2016年09月26日0点0分至0点29共三十分钟的单机运行数据，
# 其命名应为“JIMENHE_HT_20160926_0000.DJ”
```

文件内容类似如下格式:

```
<DANJI::JIMENHE DATE='2021-03-11'>
@INDEX ID TYPE TIME PWRAT PWRREACT SPD STATE FAULT
#1 1#FJ HT01 14:30 250.76 -156.00 4.67 2 '(0)'
#2 1#FJ HT01 14:31 103.62 -149.00 3.80 2 '(0)'
#3 1#FJ HT01 14:32 126.41 -319.00 3.86 2 '(0)'
...
#31 2#FJ HT01 14:30 251.87 -155.00 4.30 2 '(0)'
#32 2#FJ HT01 14:31 310.37 -150.00 5.93 2 '(0)'
#33 2#FJ HT01 14:32 232.59 -319.00 4.57 2 '(0)'
...
</DANJI::JIMENHE>
```

- 注意：
- 1.文件内容中的JIMENHE及DATE=后的引号(单引号、双引号)及“@INDEX ID TYPE TIME PWRAT PWRREACT SPD STATE FAULT”是可以在配置文件中修改
 - 2.不能修改文件中每一列原来所代表的含义(每一列的顺序也是不可修改的)
结果文件每一列的含义为: 序号、风机ID、风机型号、源数据生成时间、有功平均值、无功平均值、风速平均值、风机状态、风机故障代码
 - 3.结果文件内容的生成逻辑简单描述如下（此逻辑可以方便维护和研发人员判断是否可以根据源文件格式生成结果文件的格式）
(1). INDEX列只是一个顺序号自脚本运动顺序添加
(2). ID列是通过源文件的EC编码+1并配合配置文件的g_turbn_ID_suffix配置项的值生成
(3). TYPE列是通过源文件EC编码为下标配合配置文件的g_turbn_TTYPE[x]的值生成
(4). TIME列是通过源文件名"xxxxxxx_20210318_0958.xxxxxx"中的0958所在的域生成
(5). PWRAT PWRREACT SPD 直接取的源文件的PP_AVG PQ_AVG WS_AVG 列原样（顺序都不变）输出，但精度(保留小数位数)可以在配置文件的g_PP_scale等配置项配置

- (6). STATE列是通过源文件的SF_OGN列的值为下标配合配置文件的g_turbn_TSTATE[x]和g_default_TSTATE值生成
- (7). FAULT列是将源文件的多个列(如果一个列有多个值默认用^分隔,如果分隔符不是^可以通过配置文件的g_1mi_fixCnt_FaultJinChar配置项修改)合并成结果文件的FAULT列,结果文件不修改源文件的值只是把源文件中的多个值用()分隔,所有的值用单引号括起来

返回目录

4 软件部署

1. 确定生成1分钟数据文件的软件运行在哪个操作系统用户下，则此程序也部署在相同的用户下
 2. 将此软件包解压后（解压请在linux下进行）放在与公司常用软件目录；例如放在/zfmd/wpfs20目录下
 3. 在软件包下找到配置文件cfg/cfg.cfg，根据实际情况和配置文件注释说明进行配置(调试时g_debugL_value可设置成49)
- 注意:
1. 修改文件配置时需要仔细阅读配置文件中每一个配置项对应的注释说明，有的不需要配置的可以用默认值即可
 2. 配置文件中各配置项名称不能变（修改配置文件前把配置文件备份一下）
4. 在软件包目录下打开终端执行 `chmod u+x *.sh` 和 `chmod u+x manageCron` 并用命令 `./mergeTbn30miDataFile.sh` 手动运行程序，查看配置文件中g_dst_result_dir项配置的目录下是否有正常结果文件生成
 5. 查看log目录下日志文件是否有ERROR字样的报错信息
 6. 在第4，5步确认正常后将此软件配置到crontab定时任务中去
- 在脚本的同级目录下打开终端执行 `./manageCron` 然后会出现类似如下的提示

```
[提示]:如果发现当前用户下已经有当前脚本的配置
        但路径或定时频率不符合要求,需要先选择[2]
        进行卸载然后再进行其他操作
```

```
请输入如下数字，选择相应的操作：
    [0].查看当前用户[fusky]所有cron设置
    [1].安装脚本的cron设置
    [2].卸载脚本的cron设置
    [3].暂停脚本的cron设置
    [4].重新启用脚本的cron设置
    [5].退出，什么都不做
```

你的选择是：

根据提示操作即可：例如初装选择1即可

7. 观察软件运行1小时左右，确认日志文件是否正常，确认生成的结果文件是否正常（前提是生成1分钟数据文件的软件在正常生成数据）
8. 将配置文件中的g_debugL_value值修改为17
9. 部署结束.

返回目录

5 其他需要说明的情况

1. 配置文件中有配置过期则把相应过期文件删除的配置项，即使是配置成删除且有符合条件的文件需要删除，程序不是立即执行删除，当前系统时间对应分钟数在15到25之间才执行删除操作（有此条件是以量避免影响程序的正常功能）

2. 此软件在生成30分钟结果文件过程中（即正在生成结果文件还没生成完毕，但有文件出现了），结果文件的临时目录是tmp/do，当文件生成完毕后才对其进行mv操作，移动到最终的目标文件夹。

此举是为了避免其他处理结果文件的程序在文件没生成完毕就对其进行处理(而得到非想要的结果)。

6 文档变更历史记录

1. 2021-04-06: 对文档添加”5.其他需要说明的情况“中的第二条说明,并对文档进行目录结构化

2. 2022-06-17: 对文档进行v20.01.010相应版本的变化

3. 2023-02-09: 对v20.01.020进行说明(修改bug：有多个结果文件目录时,不能删除第1个之后的过程文件)

[返回目录](#)
