地信网
WWW.3S001.COM

MapGIS 向 ArcGIS 数据转换过程中若干问题



杨晓超
(河南省地图院，河南 郑州 450008)

摘　要：空间数据转换是提高空间数据利用率及减少重复采集的有效办法，研究空间数据格式转换有助于显著提高工作效率。从 MapGIS，ArcGIS 软件特点的分析入手，重点探索了转换过程中遇到的问题，尤其对于文件对应转换关系、属性字段变化、图廓点三方面出现的问题进行深入分析，找到解决问题的具体方案，充分利用 2 个软件的优势互补，使数据资源最大化地共享。

关键词：MapGIS；ArcGIS；数据格式转换
中图分类号：P208　　**文献标志码**：B　　**文章编号**：1672-4623 (2012) 01-0121-03

目前，MapGIS 和 ArcGIS 是应用较为广泛的 2 种地理信息系统软件。为了提高空间数据的利用率及减少重复采集，经常需要在 ArcGIS 和 MapGIS 之间进行空间数据转换；然而在转换过程中经常存在投影信息丢失、图元丢失、属性字段遗漏、属性乱码，甚至文件无法转换，或转换后 ArcGIS 软件无法识别等问题，严重影响了地理数据的共享与应用。国内已有不少刊物发表了有关 MapGIS 和 ArcGIS 空间数据转换方面的研究成果，许多研究者均对二者之间的数据转换问题进行了论述。本文主要从文件对应转换、属性字段、图廓点三方面出现的问题进行研究分析，找到解决问题的办法。

1　MapGIS 和 ArcGIS 特点分析

MapGIS 是由武汉中地公司开发的地理信息系统平台。它的特点是图形编辑功能很强，简单易用，符合中国人的习惯，出图色彩漂亮，能够输出精美的纸质地图，对系统资源占用低、操作方便，服务方便，价格较便宜。利用 MapGIS 的点线面文件不仅可以存储空间数据，而且可以存储与空间信息相关的属性信息。由于 MapGIS 强大的图形表达能力和优秀的图形打印功能，其已经在我国的很多行业得到广泛的应用，大量数据是以 MapGIS 文件形成存储、交换。但是 MapGIS 在后期空间数据的管理和分析利用上相对较弱，对空间数据库的支持性能不太理想 [1]，多年来一直强调图面的表达，非可视化的属性信息往往被忽视 [2]。

ArcGIS 是美国环境系统研究所 (ESRI) 开发的旗舰产品。ArcGIS 的特点是大而全，功能要比 MapGIS 强大的多，数据格式比 MapGIS 更加丰富，可以构造复杂的数据模型用于不同的分析目的。它是最早开发出具有拓扑关系的空间数据模型，在对空间数据库开发、一体化管理方面独具优势 [3]。它的数据格式日渐成为行业中的通用格式，ArcGIS 是把具有共同属性项的一类要素放到一层中，每个数据库记录对应的一个实际要素，它的数据格式与表示特征的类型没有关系 [4]。然而 ArcGIS 在数据的采集编辑上效率不是太高，同时喷绘或出版印刷地图功能较弱，往往需要转换到其他格式进行图形的美化及输出满足印刷要求的数据格式。

很多用户在 MapGIS 下采集数据、制作地图，然后通过中间格式转换将数据转换到 ArcGIS 支持的数据格式。实现了 MapGIS 和 ArcGIS 的数据共享，从而使用户能更好地利用 MapGIS 优良的图形编辑功能和 ArcGIS 强大的空间数据库分析及模型分析功能，使地图数据制作和空间数据的维护更加方便，大大提高工作效率。

2　主要问题分析及解决方案

本文以河南省 1∶50 000 境界要素数据更新项目为例，探讨了 MapGIS 向 ArcGIS 数据转换中出现的若干问题及解决办法。

2.1　文件对应转换关系

MapGIS 默认的文件类型分点、线、区文件 3 种，分别用来描述现实世界中的点状、线状和面状实体对象。ArcGIS 把具有共同属性项的一类要素放到一层中，每个数据库记录对应的一个实际要素，它的数据格式与表示特征的类型没有关系。

按照一般的思维习惯，MapGIS 的各类文件与 ArcGIS 层应该有正常的对应转换关系，即 MapGIS 的点线面文件分别转换为 ArcGIS 的点线面层。但由于数据结构有本质的不同，实际上两者数据转换时的对应关系比较复杂。经过实验和研究，总结出以下规律：MapGIS 的点文件与 ArcGIS 的点层之间转换有正确的对应关系，

收稿日期：2011-01-12

万方数据

地信网
WWW.3S001.COM

即 ArcGIS 的点层都是由 MapGIS 的点文件直接转换而来；MapGIS 的线文件与 ArcGIS 的线层一般不进行对应转换，即 ArcGIS 的线层不是由 MapGIS 的线文件转换而来，它是由 MapGIS 的面文件的弧段转换而来。但是 MapGIS 的面文件的弧段是由 MapGIS 的线文件自动生成的，当发现面要素的形状有变化时，通过修改 MapGIS 的线文件无法达到目的，只有通过修改面文件的弧段才能修改面要素的形状。

针对该项目，成果中要求的数据分层及属性项与 MapGIS 默认的点、线、面文件之间不存在对应转换关系。假设在 MapGIS 中采集的点、线、面文件名称依次为 boupt.wt、bountl.wl、bount.wp，输出的 E00 文件依次为 boupt. E00、bountl. E00、bount. E00，则 ArcGIS 中的 boupt 层（界桩、界碑）是由 bount.wt 直接转换生成的，ArcGIS 中的 bount 层中的县级行政区域（面）由 bount.wp 中的面生成，行政界线（线）由 bount.wp 中的弧段生成，实际转出的 bountl. E00 没有太大作用，但是 bountl.wl 又是必须的，因为 bount.wp 中的弧段是由 bountl.wl 生成的。

当 MapGIS 文件向 ArcGIS 数据转换完成后，如果要修改行政界线信息，则只能在 ArcGIS 中修改 bount 层文件或在 MapGIS 中修改 bount.wp 的弧段信息，不能通过修改 bountl.wl 文件来进行行政界线的更新。然而，由于 MapGIS 中编辑弧段的功能远远比不上编辑线文件的功能，在 MapGIS 中修改 bount.wp 的弧段信息往往比较麻烦，并且很容易出错，因此，在进行 bountl.wl 的线转弧段操作前一定要把 bountl.wl 处理非常好，尽量减少 bountl.wl 文件中的问题，以减少后期数据处理的工作量。要保证拓扑关系严格按照地理信息数据的要求处理好，如矢量化线文件时，注意用捕捉点方法来处理结点关系，各要素无线段自相交、两线相交、线段打折、碎片多边形、悬挂点或伪节点等图形错误；以河或路等线状地物为界的具有公共边线性质、特征时，应用从 1∶50 000 核心要素数据库中拷贝相应要素方法生成，以保证位置的一致性；要保证属性项定义严格按照规定要求进行设置，以保证在生成 bount.wp 时，弧段属性结构不出问题；同时还要认真核对 1∶50 000 勘界协议书附图有关专题信息，准确处理行政界线的走向问题。

2.2 属性字段类型不一致

MapGIS 中有关属性的字段类型主要包括字符串、短整型、长整型、浮点型、双精度型等，字段长度只定义一个长度，不区分输入宽度和输出宽度。在实验数据中定义的数据项如表 1 所示。

表 1 定义数据项

属性项名称	名称描述	字段宽度	字段类型
GB	国标码	6	长整型
PAC	县级行政区划代码	6	长整型
NAME	名称	60	字符型
BNO	界桩	12	字符型

将 MapGIS 的数据在 MapGIS 中转换为 E00 格式，然后在 ArcGIS 的工具包 ArcToolBox 中选择转换工具中的 Import From Interchange File，将 E00 数据转换为 Coverage 格式，之后在 ArcGIS 中打开图形文件，属性浏览时可以发现 GB 和 PAC 字段的输入宽度为 4，输出宽度为 5，与要求的输入宽度和输出宽度均为 6 不符合，且数据属性有严重丢失现象。当我们用“记事本”打开 J50E024002 图幅在 MapGIS 中转换的 E00 文件后，发现有关属性字段的设置信息如下：

```
bount. AAT          XX    8   8   36   16
FNODE#    4-1  14-1  5-1  50-1  -1  -1-1   1
……
bount-ID  4-1  294-1 5-1  50-1  -1  -1-1   7
GB        4-1  334-1 5-1  50-1  -1  -1-1   8-

bount. PAT          XX    6   6   88    5
AREA      8-1  14-1  18 5  60-1  -1  -1-1   1-
……
PAC       4-1  254-1 5-1  50-1  -1  -1-1   5-
NAME      60-1 294-1 60-1 20-1  -1  -1-1   6-
```

我们可以看出，虽然在 MapGIS 数据中 GB 和 PAC 字段的属性字段宽度已经设为 6，类型是 I，然而导出后的 E00 变成输入宽度是 4，输出宽度是 5，类型是 B，发生了较大变化。按照上交成果要求，最终文件的输入宽度和输出宽度均应为 6，字段类型为 I，因此要对导出后的 E00 进行修改，使之符合最终数据上交的要求。

通过对 E00 文件的内部格式的分析，可以看出 E00 文件的 BOUNT. AAT 部分中，字段 GB 后的 4-1 及 5-1 是反映输入宽度和输出宽度的地方，需要把该部分改为 6-1 及 6-1；字段 GB 后的 50-1 是反映字段类型的地方，需要把 50-1 改为 30-1。在 E00 文件的 BOUNT. PAT 部分中，字段 PAC 后的 4-1 及 5-1 是反映输入宽度和输出宽度的地方，需要把该部分改为 6-1 及 6-1；字段 PAC 后的 50-1 是反映字段类型的地方，需要把 50-1 改为 30-1；字段 NAME 后的 294-1 是反映字段从什么地方开始取值的地方（从 25 + 4 = 29 位开始取值作为字段 NAME 的值），需要把 294-1 改为 314-1（从 25 + 6 = 31 位开始取值作为字段 NAME 的值）。

万方数据

地信网
WWW.35001.COM

这里我们通过编程的方法直接以修改文本形式来修改 MapGIS 中导出的 E00 文件，用编程的方法实现所有图幅 E00 文件中有关字段类型的内容是非常容易的，实验证明，软件开发人员用较短时间内就完成了编程任务，并在几分钟内完成了全省 464 幅 1∶50 000 图的修改工作。

在 ArcGIS 中打开 E00 数据转换后的 Coverage 格式图形文件，我们可以发现字段 GB 及 PAC 的輸入宽度和輸出宽度均改为 6，字段类型为 I，字段 NAME 也能够正确显示，满足最终成果要求。

2.3　图廓点问题

按照《1: 50 000 地形要素数据库境界更新技术规定》要求，更新后的境界数据分幅全部按照 1980 西安坐标系标准分幅存储，对原 1954 北京坐标系的图幅，应重新拼接和裁切。境界层（BOUNT）应存放完整内图廓线，并与界线构面。

在将 MapGIS 的线数据和面数据在 MapGIS 中转换为 E00 格式后，然后在 ArcGIS 的工具包 ArcToolBox 中选择转换工具中的 Import From Interchange File，将 E00 数据转换为 Coverage 格式，之后在 ArcGIS 中打开图形文件，我们可以发现图廓线和图廓点并不完全重合，图廓线显示的是一个不规则的四边形，而 4 个图廓点构成的是一个矩形区域，这显然不符合最终要求。当我们用“记事本”打开 J50E024002 图幅在 MapGIS 中转换的 E00 文件后，发现有关图廓点的信息如下：

```
bount. TIC              XX     3  3   20      4
IDTIC       4-1  14-1  5-1  50-1  -1  -1-1          1
XTIC        8-1  54-1  18 5  60-1  -1  -1-1          2
YTIC        8-1  134-1 18 5  60-1  -1  -1-1          3
1 5.04040137607805350e+03 7.97687283283433500e+04
2 5.50088143075206970e+03 7.97687283283433500e+04
3 5.50088143075206970e+03 8.01508813481311810e+04
4 5.04040137607805350e+03 8.01508813481311810e+04
```

从以上可以看出，4 个图廓点构成的确实是一个矩形，并且是图廓线的外接矩形。造成这一结果的原因在于，MapGIS 在采集线文件时，并不能采集图廓点信息，在生成 E00 文件时软件又不是自动生成图廓线 4 个顶点的坐标文件，而是自动生成了图廓线的外接矩形 4 个顶点的坐标文件。显然，要想解决这一问题，必须把 MapGIS 中转换的 E00 文件中的 4 个图廓点坐标改为正确的坐标信息。在国家下发的 1∶50 000 DLG 数据中，元数据文件详细记录了各图幅的图廓点坐标信息，用写字板打开 J50E024002 图幅的元数据文件，可以找到如下的图廓点坐标信息：

```
42　西南图廓角点 X 坐标          3989044.06
43　西南图廓角点 Y 坐标          252020.07
44　西北图廓角点 X 坐标          4007544.07
45　西北图廓角点 Y 坐标          252543.43
46　东北图廓角点 X 坐标          4006935.29
47　东北图廓角点 Y 坐标          275044.07
48　东南图廓角点 X 坐标          3988436.42
49　东南图廓角点 Y 坐标          274568.38
```

我们只要从国家下发的 1∶50 000 DLG 元数据文件中提取出图廓点坐标信息，并替换掉在 MapGIS 中转换的 E00 文件的对应图廓点坐标信息就可以解决这一问题。而用编程的方法实现所有图幅图廓点坐标信息自动替换也是非常容易的。实验证明，软件开发人员用 0.5 d 的时间就完成了编程任务，并在 1 min 内完成了全省 464 幅 1∶50 000 图廓点坐标信息的替换工作。替换工作完成后，我们在 ArcGIS 中打开 E00 数据转换后的 Coverage 格式图形文件，可以发现图廓线和图廓点已经完全重合，实现了既定目标。

3　结　语

实验结果表明，本文通过研究分析，提出的一套数据转换过程中解决问题的办法是行之有效的，通过一系列的处理，使用户能更好地利用 MapGIS 便捷的地图编辑功能和 ArcGIS 强大的空间数据库处理功能，使地图数据制作和空间数据的处理维护更加方便，节省了巨大的人力物力，大大提高了工作效率。实际上数据格式转换过程中出现的问题较多，本文只是从遇到的实际问题出发进行了部分研究，其中必定存在许多没有遇到的问题，这里希望起到抛砖引玉的作用，将这方面的研究引入更广、更深的层次，期待今后有更多的研究人员进一步的探讨和总结，为广大数据处理人员发现并解决问题提供一定的参考。

参考文献

[1] 张建军，张静波. MapGIS 6.1 向 ArcGIS 8.1 的数据转换[J]. 微机发展，2003，13(11)：14-16

[2] 郝瑞卿，薛林福，刘杰，等. MapGIS 6.X 与 ArcGIS 9.2 空间数据转换的方法及问题探讨[J]. 东北师大学报：自然科学版，2008，40(4)：142-146

[3] 张于，武健强，吴夏懿，等. MapGIS 数据向 ArcGIS 数据格式转换方法的实践与探讨[J]. 江苏地质，2007，31(3)：258-261

[4] 叶国华. ArcGIS 向 MapGIS 数据格式转换[J]. 测绘与空间地理信息，2008，31(6)：43-45

[5] 吴信才. MapGIS 地理信息系统[M]. 北京：电子工业出版社，2004

[6] 樊红，詹小国. Arc/Info 应用与开发技术(修订版)[M]. 武汉：武汉大学出版社，2002

作者简介：杨晓超，工程师，主要从事地图制图、地图学与地理信息系统应用方面的工作。

万方数据