Managing a Cadastral SDI Framework Built from Boundary Dimensions

Michael ELFICK, Australia, Tim HODSON, South Africa and Curt WILKINSON, USA

Key words: cadastral survey, GIS, land records, cadastral fabric, boundary network

SUMMARY

A multi-purpose cadastre is the goal of most countries today. Cadastral information is also a key layer in most multi-purpose geographic information systems (GIS), and these systems, in turn, are included within the broader Spatial Data Infrastructure (SDI). The utility of spatial data is greatly improved if it is accurate and consistent between layers, especially given the advent of inexpensive GPS devices. This paper introduces an approach to improve and maintain spatial accuracy of cadastral boundary geometry, and concurrently improve geometry of other layers constructed with reference to the cadastral boundaries. This approach has been used in Australia for several years, while the full integration with a commercial GIS is currently under development.

In the described system, cadastral geometry provides a network of boundaries defined by their dimensions and connectivity, thus making use of pre-existing survey information. Least squares analysis is applied to the network using survey accuracy as a means to weight the network elements. As many cadastres have used a digitizing process to create their digital maps, the system is designed to begin with such data and incrementally improve accuracy as more survey information is added. Adjustments to the cadastral geometry create a field of displacement vectors that drive adjustment of other spatial data layers. The use of the dimensioned boundary network and least squares analysis yields highly accurate geometry with a minimum of control.

Managing a Cadastral SDI Framework Built from Boundary Dimensions

Michael ELFICK, Australia, Tim HODSON, South Africa and Curt WILKINSON, USA

1. INTRODUCTION

Land ownership is a fundamental layer in the Spatial Data Infrastructure (SDI) and many GIS systems use the cadastral boundary layer as a base framework for other layers. Using a formalized surveyor's mathematical approach, the dimensional information available on land boundaries can be used to make a spatially inaccurate cadastral layer much more accurate. More significantly to managers of GIS systems, the spatial updates can be applied to other layers that use the cadastral layer as their base.

In many countries land boundaries are defined by "metes and bounds". This system has evolved during a period when it was relatively easy to measure lines but very difficult to find the exact coordinate location of a point on the earth's surface.

The boundary dimensions and reference to physical objects on or near the boundaries are shown on cadastral survey documents and these may be in the form of plans or written descriptions. The accuracy of data in these documents is a reflection of the measurement technology at each point in time. For example, prior to 1880, theodolites were not commonly in use and angles were measured to the nearest half degree. Over the years, developments in the technology for measuring distances and angles have steadily improved and the modern instruments used in boundary definition measure angles to within five seconds of arc and distances of 1000 meters to a precision of better than five millimetres. Precise GPS systems can also locate points to centimetre accuracy in real time.

Currently, the most common method of building the land ownership layer in GIS systems is by digitizing boundaries from cadastral maps. The accuracy of this data depends on the maps and the quality of the digitizing process. Many of these maps were designed simply to show the relationship of the various attributes to each other rather than being compiled to an accurate coordinate base. Consequently, the accuracy of position varies from place to place and any mistakes in the original map compilation are carried forward into the digitized records, and other spatial layers.

There are many systems in use to improve the accuracy of this type of data and these include "rubber sheeting" or adjusting to control from GPS or photogrammetric sources. These systems improve positional accuracy near each control point but overall, they do nothing to correct the inaccuracies in the base data. To overcome that problem it would be necessary to have a control point at nearly every corner.

While most GIS managers appreciate the limitations of this data set, it has generally been considered too complicated and costly to build the property ownership layer directly from source documents. The reasons for this relate partly to the lack of systems for processing this type of data efficiently and partly to the lack of understanding of the structure and characteristics of the source data.

This paper describes a solution that addresses these shortcomings and that is proving successful in commercially available GIS software. The paper presents the concepts, workflow, and design motivations behind the solution.

2. BOUNDARY DEFINITION DATA

One of the primary intents of a cadastral survey document is to convey instructions from the surveyor on how to locate, in the field, the physical corners that define the boundaries of a property. Cadastral survey documents most typically describe each parcel of land by the bearings and distances around its perimeter. The basic module for the data is therefore a closed polygon and this polygon may have additional "connections" to corners of other parcels or survey reference marks.

The bearing and distance dimensions are derived from the measurements taken by the surveyor in the field and they may not "close" the polygon exactly. Consequently, if you compute around the boundaries using the dimensions, you will not come back exactly to the starting point. This difference is often referred to as the "misclose". The size of the misclose in each parcel is a measure of the accuracy of the data, and is also used by surveyors to check their work. Note that this is just a measure of the consistency in their work for a particular survey and there may be scale differences between different surveys from different periods of time. So the dimensions on a line may be different on adjacent or overlapping plans reflecting separate measurements of the line.

The data can be considered as a series of "observations" grouped in sets with each set being a closed polygon. These polygons can be linked at the connection points indicated by the survey data to form a boundary network. Parcels are connected by common points or by a point on one parcel lying on a line in another parcel.

Because of overlaps in the surveys, there is redundant data. Coordinates can be generated for each parcel corner by weighting the data according to the measurement accuracy available at the date of each survey, and processing the network through a least squares adjustment. The adjustment process must take into account the ways that the surveys were carried out and the fact that the data is a series of polygons with connected traverses and not individual lines. It must also include all of the data and not just the latest information.

Since the point models the physical location, while the coordinate models the current representation of that point, the software system treats coordinates as derived quantities that are held as attributes of a point rather than as a definition of the point itself.

A cadastral fabric built from survey data can be made very accurate using limited control and this is accomplished by using a least squares adjustment that uses all of the survey information (including historic) to distribute error through the fabric.

The boundary definition layer is constantly changing as land is subdivided and consolidated. Each new document will change the boundary network and may also influence the coordinates of points outside of the document, even though this influence is limited. In practice, the size of change becomes small as the proportion of modern plans is increased, or if the area is stabilized with additional GPS control.

Each time changes occur in portions of the cadastral fabric, they are added to an historical series of vectors collected for each point; these define the changes in the coordinate locations, through time, and may be used to update the other GIS layers that are dependent on the cadastral fabric.

3. COORDINATE BASED PROPERTY BOUNDARIES

Many of the physical services represented in the spatial data infrastructure, such as pipelines, underground cables, and so on, are physically positioned at known offsets from property lines. Property lines are usually represented by the physical markers placed or found by the surveyor. Once accurate coordinates are available as evidence in locating these physical boundary markers, locating other physical services in the field, like buried pipelines, becomes that much easier.

Different periods in the history of survey technology have prescribed different methods of recording a surveyor's description of a property. As more fully described in the preceding section, the most traditional method has been the use of bearing and distance dimensions on the boundary lines. Though coordinates have traditionally found limited use for recording boundary locations, with the advent of high accuracy GPS, it is now significantly easier for surveyors to employ them in this capacity.

Over the last fifteen years, various systems have been developed to generate coordinate based property boundaries from cadastral survey documents. New Zealand has converted most of its urban areas and in Australia, states such as the Northern Territory and Queensland are progressively coordinating their data.

Depending on the regulations in place for a specific cadastral system, the traditional cadastral survey documents used for re-locating property boundary corners may be interpreted in a few different ways. Hence, boundary location disputes can arise when different surveyors use different data to re-establish the location of a boundary. A coordinate can provide a unique and unambiguous record of a point, and GPS now provides the necessary survey technology to quickly and accurately re-locate that point.

To gain maximum benefit from existing data, the building process should not only extract data from the documents and build the boundary network, but it should also analyse the data and provide a measure as to the reliability and accuracy of the computed coordinates. This opens the way for coordinates to be used more widely as the primary way for surveyors to convey instructions on how to locate the physical boundaries of a property.

4. CONCEPTUAL MODEL

In GIS software, the geographic layer is used to represent a theme of mapped information. Often, these layers are constructed in relation to base map layers that outline the property ownership rights.

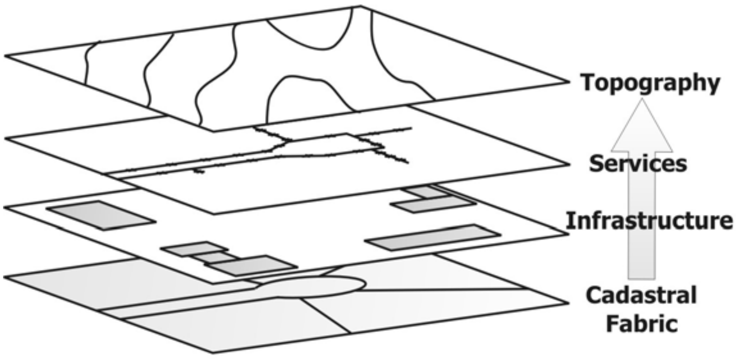

Figure 1: Cadastral Layer as Basemap

The real-world objects modelled in cadastral surveys include: parcels, parcel boundaries, and the physical points that define the parcel boundaries. Surveyors also use the less tangible notion of coordinates and boundary dimensions to represent the findings of their field work; and these findings are recorded in cadastral survey documents. Such documents may take the form of deed titles, registered title diagrams, or subdivision plans; we refer to them as *plans* inclusively.

A network of connected boundaries, often referred to as a *cadastral fabric*, defines the conceptual layer for land ownership within the spatial data infrastructure. The *parcels*, *boundaries*, *points*, *coordinates*, *boundary network*, and *plans* are all elements that must be emulated. Our use of *parcel* simply means a unit of land. The fabric may have any number of different types of parcels, variously described as lots, blocks, sections, and so on, as needed. Thus, hierarchical models of land units may be employed.

A cadastral fabric needs to represent overlapping parcels to support the hierarchical models described above. Also, since new parcels will supersede their parents, and these parent parcels are retained as historical, overlaps will occur naturally between current and historic parcels. Within any one type of parcel in the current fabric, non-overlapping parcels are enforced by the explicit topology of the fabric.

Fabric history is unique in that it does not follow the archive model. It is a living history that actively contributes measurement information in the processes that improve cadastral fabric coordinates, including those that define historical parcels. This maintains spatial relationships between old boundaries and new boundaries through any number of adjustments, and is often visualized as maintaining the verticality between a stack of parcels. Of course, original dimensions and related attributes are not edited; only the coordinates are updated to the current best estimates generated by the least squares analysis.

Coordinates in traditional GIS systems have existed in a different realm of functionality and importance when compared with coordinates amassed by surveyors in a cadastral surveying environment. This difference occurs because of the emphasis placed in making topological representations at mapping scales in a GIS, versus the emphasis placed in making representations of absolute location by the surveyors in a digital cadastral system.

In software where both requirements are equally important, an improved representation is required. The key concept is the flow of information from dimensions and measures, to best estimates of coordinates, to shape geometry used in the GIS. Dimensions and coordinates are maintained in a rigorous survey model, while the shape geometry is allowed to participate in more relaxed GIS analyses that require some movement of coordinates in the shape. A good example of this would be a GIS topology that must create breaks in boundary lines to build a composite representation of the cadastral boundary, utility lines and other related GIS features which intersect each other. As the flow of information is always from the dimensions to the shapes, the correct shape can be reconstituted at any time.

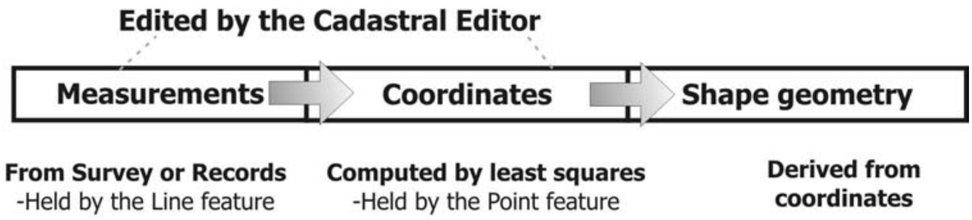

Figure 2: Key architecture concept

Updates to the cadastre are generally modelled as job-workflow transactions against the cadastral fabric. Each cadastral job is a collection of parcels which can be modified, adjusted or extended. Often, the parcels in the cadastral fabric are locked to prevent others from altering them until the job is completed. The job extracts these parcels from the cadastral fabric, and they are edited ’outside’ the database.

A job is edited by a special set of tools that properly updates the cadastral fabric, using the parcel as the unit of work. Specifically, the goals of this cadastral editor are to: capture all possible dimensional information recorded on legal documents; build a boundary network using common points between record sources; and use control points and all available records within an area to provide the best possible solution for the coordinates by a least squares adjustment.

Once a job is completed, it can be used to update the cadastral fabric. The current cadastral fabric is the up-to-date representation of the land status, and is therefore all parcels, excluding parcels marked as historic. The current cadastral fabric is used for updating other layers, and for publishing land record information. At any time, there may be jobs that define parcels that have not been inserted into the cadastral fabric. This occurs in cases where the job is still being processed by an operator, is awaiting approval, or is one of several optional designs for development.

5. DATA MODEL

A geographic database (geodatabase) is the foundation of a GIS. The geodatabase supports many different applications of geographic analysis, modelling and design. Geodatabase tables hold representations of mapped objects called *features*. Tables that store features for mapping are called *feature classes*. Each row in the feature class table stores an item that can be placed on a map. Feature classes therefore contain all the mapping information for a particular GIS map layer. Each item placed on a map has additional attribute information that can be found in the attribute columns of the feature class table.

It follows that the cadastral fabric in the geodatabase is modelled with parcels represented by *parcel line* features, *parcel point* features and *parcel polygon* features, referred to in aggregate as *parcel features*. Along with its mapping shape geometry, each parcel stores other textual information relevant to parcel data maintenance including, for example, attributes in the parcel point feature class for representing its coordinates, and coordinate geometry (COGO) attributes in the parcel line feature class, which represent its dimensional information and accuracy level. Another key attribute is a parcel identification number in the parcel polygon feature class.

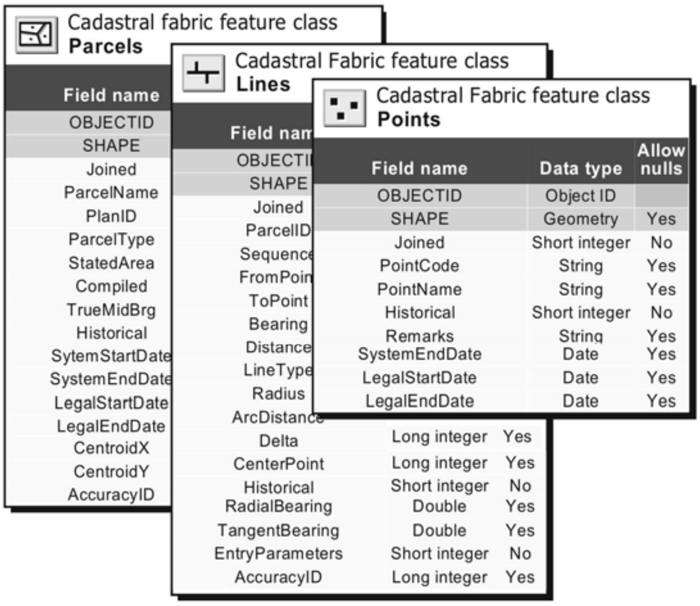

Figure 3: Data Model

These parcel, point and line feature classes are the core representation of the cadastral fabric.

Within the geodatabase system, the fabric is a continuous surface of connected parcels with an explicit topology, defined by common parcel corners and neighbours. This topology is inherent in the model, and is defined and enforced during data entry.

Ideally, digital parcel data sources have attributes on the boundary lines that circumscribe each parcel. These attributes are usually intended to represent the record data from the cadastral documents. Hence, it is important when converting from other digital sources that COGO attributes be converted during transfer into the cadastral fabric. However, not all sources have these COGO attributes. This can be accommodated by calculating a vector geometry representation for the COGO attributes. In addition, the cadastral fabric will record that these source lines are derived from calculated geometry data, versus original record data, and are assigned a lower weight.

The cadastral fabric provides a way to organize the parcel data based on the form in which the data is originally recorded. This is defined through the plan structure. When data is entered, it is usually being entered from a recorded subdivision plan, or from a digital submission of a plan. The *plan* in the cadastral fabric is used to hold information about the subdivision plan record, such as the date, surveyor, entry units, scale factor, and so on. The cadastral fabric has a table to represent plans to store this type of information.

6. COORDINATE SYSTEMS AND JOINED VERSES UNJOINED PARCELS

Cadastral field surveys conducted in Australia and the United States have traditionally proceeded in a local coordinate system, which is sufficient for the cadastral records required, and for the large scale that is typical for the recorded plans.

By contrast, a fabric contains many contiguous plans that, cumulatively, are used for mapping larger areas of the fabric at small scales. The fabric therefore requires a well-defined geographic or projected coordinate system.

The notion of using a local coordinate system carries forward when this data needs to be added into a fabric, since there are often no coordinates, nor any reference to a coordinate system in the source data. Instead the local-coordinate parcels exist in an unjoined space until they are joined into the rest of the cadastral fabric.

A job holds new parcels in an unjoined space, and through the course of completing the job these unjoined parcels are connected into the rest of the fabric.

7. AN INTEGRATED APPROACH

Cadastral systems are subject to continual change and many authorities already have built GIS systems based on digitized cadastral boundaries. The main factor which has limited progress to general coordination of cadastral data has been the limitations in design of GIS systems and the lack of software which can efficiently process and analyse cadastral survey data. Key to the integration of these systems, is the development of processes for updating the geometry of the other layers in a GIS.

The described system can begin with existing data, regardless its quality. Thereafter, survey methods are used for adding and managing cadastral survey data. The system supports ‘remote’ editing of cadastral survey data for updating and upgrading of large blocks of data and can be scaled for very large datasets and keep history for changes.

8. IMPLEMENTATION AND QUALITY ASSURANCE TECHNIQUES

The quality of source data will vary from place to place and will generally depend on the age of surveys and the degree of regulation of the cadastral survey process. In countries where the state guarantees title (such as Australia and New Zealand), there is a higher standard of data and less likelihood that there will be gross mathematical errors in the documents.

The *cadastral editor* is used to enter data directly from survey documents, join the parcels together, add survey control as necessary and then apply a least squares adjustment to the completed cadastral network. The job is then loaded to the geodatabase and the system will merge the new data so that the geodatabase can store and present a seamless fabric.

The following section describes subsystems that manage these steps. There are progressive quality assurance (Q/A) checks at each stage to manage problems which may become apparent in the data.

There are a number of ways to formally assess the quality of a cadastral fabric, such as assessment of the misclose for each parcel, problems in identifying the shared point connections between neighbouring parcels, and analysing the residuals resulting from a least squares adjustment.

The key to success is to progressively test the data as it is entered and assembled so that the operator can detect and resolve any problems early on. Data entry operators must focus on directly recording the geometric data for a parcel, and rely on the software to interpret, manage, and report any problems in the geometry as the parcels are entered. The cadastral editor provides a simple interface for entering this data with immediate generation of misclose information.

Parcel — Properties

Plan	1010187
Name	6
Type	Warranty Deed
Plan Date	01 March 1989
Stated Area	2 hectares
Accuracy	2-After 1980
True-mid bearing	No
Compiled	No
Historical	No

Misclose, Bearing: 133° 28′ 05″ Distance: 0.001

Parcel — Lines

From	Bearing	Distance	To	Type	Radius
1	43° 27′ 45″	60.000	2	Standard	
2	133° 27′ 45″	24.840	3	Standard	
3	189° 22′ 50″	1.795	4	Road	
4	206° 25′ 20″	17.585	5	Road	30.000
5	223° 27′ 45″	35.700	7	Road	
7	268° 27′ 50″	8.485	8	Road	
8	313° 27′ 45″	25.000	1	Road	
8	133° 27′ 45″	32.000	9	Connection	

Misclose, Bearing: 133° 28′ 05″ Distance: 0.001

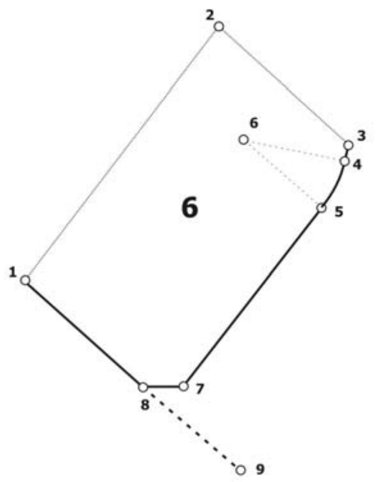

Figure 4: Parcel data entry

Bearings and distances are entered in sequence around the parcel, and the lines and their local reference numbers are displayed on the screen as the data is being entered. Bearings can be input as “whole circle” or “quadrant bearings” and either mode can be displayed as required. Distances are entered in the units used in the plan i.e., feet and inches, or links or meters etc.

Connections shown on survey plats can be included to assist in the joining process and these may take the form of a traverse. At the completion of each parcel, the misclose vector, area, and “order of accuracy” are given. Other editing tools, not shown here, also display quality measures.

In a GIS environment, making maps with topological correctness is important, and a common practice has been to enter parcels to match portions of a pre-existing cadastral fabric base map. Even when sets of newly entered parcels have good geometric consistency, many GIS parcel maintenance workflows distort newly entered data to fit the existing (and often less accurate) map representation.

Later attempts to match the fabric to ortho-photography and other control sources using methods like “rubber-sheeting”, have resulted in additional shifts that, although have provided some spatial improvement, do not use any of the surveyors’ recorded measurement information, and hence do not optimally represent the surveyors’ coordinates for boundary point locations. In effect, these methods cannot provide reliable error estimates for the boundary network coordinates.

The traditional surveying approach to building a property boundary layer from survey documents has been to first set in place a control network, and to then progressively add data to the control. Such an approach can be very costly to implement in the large regions managed by cadastral organizations, and requires staff with considerable expertise in the management of boundary data.

A better approach is to enter all the data, assemble and analyse the boundary network and only then apply control according to the results of the analysis. While this is almost a reversal of “standard surveying practice”, it allows for checks to be made on the internal consistency of the data before the complications imposed by the addition of geodetic control. This approach has the benefit of showing where additional control is needed, reducing costs when collection of new control in the field is required.

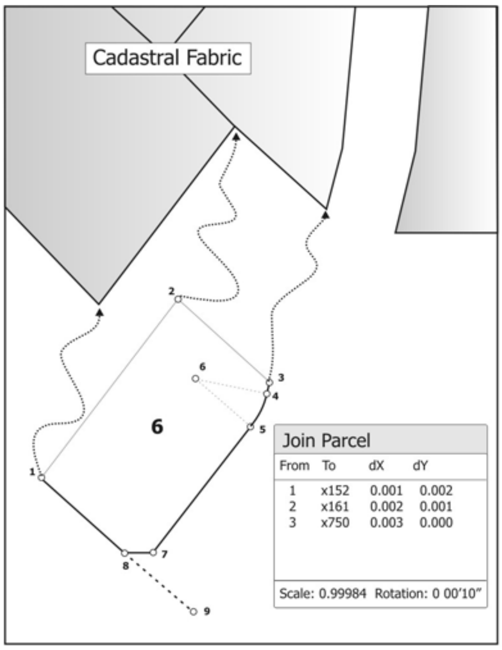

Figure 5: Joining a parcel to the fabric

The parcel joining procedure is designed to define the "connectivity" between parcels in the property boundary layer. Joining is carried out interactively. After a parcel has been selected for joining, it is displayed on the screen and can be “dragged” close to its final location with a mouse. Points connecting the new parcel to network points, and points connected to lines are selected with the mouse. As each connection is made, transformation parameters are computed by a least squares procedure, and the residuals are displayed. The operator can accept the result, or reject it and try other points for the join. If accepted, the parcel is added to the network and the parcel point numbers are changed to network point numbers.

The joining process establishes the topology of the network and it has checks to ensure the integrity of the system. For instance, reports are made where points are close to each other but not joined and where points are close to lines.

The network adjustment procedure is a rigorous least squares procedure to generate coordinates using plan data (bearings and distances) and selected control points. It uses the topology generated by the joining process and treats the dimensions from the survey data as "observations". Closed parcels are treated like direction sets in a geodetic network adjustment in that they have a common orienting parameter.

A report is then generated showing all adjusted co-ordinates and comparisons made with each line in each parcel. Lines with significant differences are tabulated and a statistical estimate is made of the precision of the coordinate data. Analysis of the coordinate data within the larger spatial regions allowed by the cadastral fabric provides explanations of local anomalies that would otherwise not be possible.

9. GEOGRAPHIC LAYER UPDATE

Other GIS layers are edited using the cadastral fabric as the background reference. Since the cadastral fabric may be frequently adjusted with shifts of a large number of coordinates, discrepancies will appear between the cadastral fabric and these layers. It is important for these layers to be adjusted similarly, thereby maintaining relative positioning.

An edit of the cadastral fabric may include an adjustment of the coordinates of a group of parcels. The size of the area adjusted will vary greatly: simple parcel updates will generally adjust only a local group of parcels, regional adjustments will encompass a large number of parcels. When an edit is inserted in the cadastral fabric, the new coordinates for existing points are compared against the old coordinates for the point, and a set of displacement vectors are stored in a table. Thus, the cadastral fabric has a sequential record of the adjustments for each point. These displacement vectors can then be used to adjust any other feature class.

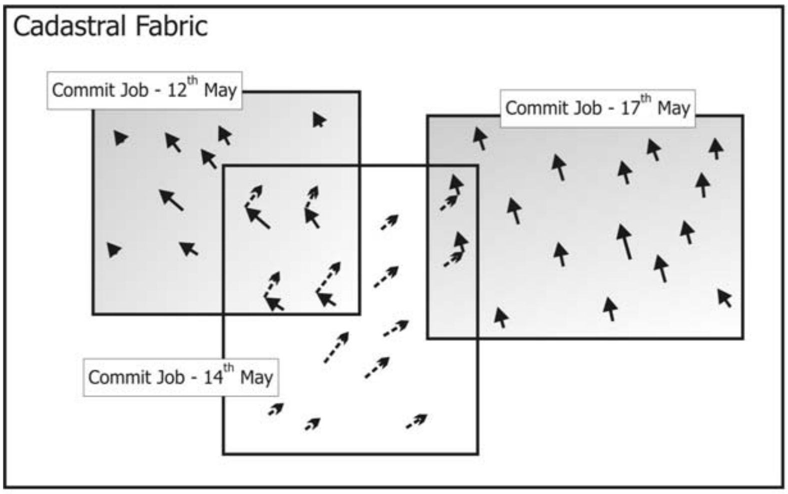

Figure 6: Displacement vectors created by a series of jobs

Figure 6 shows a series of "check-ins" which create sets of displacement links representing changes in the underlying cadastral fabric. The adjustment areas can overlap, so it is important that adjustment order is applied correctly. Displacement vectors for each adjustment are tagged with a sequence ID. The logical grouping of these vectors allows feature classes to have adjustments applied *ad hoc*.

10. CONCLUSIONS

We are now moving into a new era. With GPS, it is now possible to quickly coordinate data points to centimetre accuracy, and the software systems which take advantage of the new 64 bit computer technology will allow millimetre precision for coordinates regardless of the extent of coverage. This means that GIS can hold the master data sets for information with engineering, modelling and design systems alongside traditional GIS data. All that is required is recognition of the specific requirements of survey and cadastral data, and suitable subsystems which properly manage them. As the cadastral data set is a fundamental layer on which so many others are based, it is important that it be as accurate and complete as possible.

In countries where the cadastral boundaries are defined by dimensional data (metes and bounds), the only way to generate a reliable and accurate cadastral boundary network is by direct use of this dimensional data. This requires extending standard GIS systems for the management and storage of the dimensional data, including special software tools for data entry, analysis, and processing. This includes a method for updating GIS feature layers from the changes in the boundary geometry of the cadastral fabric.

REFERENCES

Elfick, M.H., 2001, *Managing the Records which underpin the Land Tenure System*, International Symposium on Spatial Data Infrastructure, Melbourne.

Elfick, M.H., 2003, *Building & Managing a Survey-Accurate Cadastre*, Survey Summit, Twenty Third Annual ESRI User Conference, San Diego.

Kaufmann, J. & Steudler, D. 1998. "Cadastre 2014 A Vision for the Future". Switzerland : FIG.

West, G.S.R., 2002, *Cadastral Reform in the Northern Territory*, Trans Tasman Survey Congress, Adelaide.

Zeiler, M "Modeling Our World – The ESRI Guide to Database Design" ESRI Press.

CONTACTS

Michael Elfick
Geodata Information Systems
P.O. Box 11
Morpeth, New South Wales, 2321
AUSTRALIA
Tel. + 61 2 4930 5381
Email: Michael.Elfick@newcastle.edu.au

Tim Hodson
Environmental Systems Research Institute
380 New York Street
Redlands
USA
Tel. +1 909 793 2853
Email: thodson@esri.com
Web site: www.esri.com

Curt Wilkinson
Environmental Systems Research Institute
380 New York Street
Redlands
USA
Tel. +1 909 793 2853
Email: cwilkinson@esri.com
Web site: www.esri.com