

GE Energy

Smallworld Core Spatial Technology™ 4

Presenting the globe in a flat space –
managing projections and coordinate systems



Abstract

Many of the first Geospatial Information Systems (GIS) products were deployed in the mapping or engineering departments of utilities. At that time most utilities serviced comparatively small territorial regions based around the business' historic customer base. Over time consolidation and globalization resulted in the merging of many of these smaller regional utilities into much larger entities with service territories often covering vast (and often disjointed) areas. As with many new technologies, GIS quickly migrated from established markets in North America and Europe to South America, Asia and the Middle East where national utilities were responsible for very large, sometimes sparsely populated territories. This trend represented a real problem for many conventional GIS architectures since they had been designed to store spatial data in a single coordinate system (often that of the host nation or region). For enterprises servicing customers in more than one coordinate system integrating this data became a complex challenge. Also, many early GISs did not take into account the curvature of the Earth when calculating location or distance leading to inaccurate geometrical calculations.

The Smallworld architecture, designed by GE Energy, has been devised to support this growing trend by allowing data to be stored in one coordinate system and then be projected, on-the-fly, to another coordinate system that can easily be viewed or processed. This advanced integrated technology allows precise calculations to be performed transparently and seamlessly on data stored in multiple coordinate systems without the need for time consuming and costly bulk conversions. It also allows applications written on top of the Smallworld architecture to take advantage of this powerful functionality in a seamless way without the costly re-engineering of code.

In the beginning the world was flat

Many utilities started life as publicly owned or regulated businesses. Each utility was regionally or nationally based with little overlap with adjacent utilities. The spatial data in a region was likely to be in one coordinate system and often geographically small enough to ignore the effects of the curved surface of the Earth on which it was set. For all intents and purposes, the world, at this point, could be regarded as essentially quite small and flat.

However, two major business trends emerged that questioned that view of the world.

Firstly, many countries embarked on a program of privatizing their national utilities, allowing them to merge or be taken over by rivals at home or from abroad. At the same time countries also deregulated their power, water and gas markets, accelerating competition. Both of these drivers led to the creation of very large utilities servicing many regions of one country or indeed regions of many different countries. These service areas were no longer small contiguous

pieces of land. Many of these new multinational fueled by globalization had now grown to service customers over vast stretches of land.

Secondly, gis also began to be adopted by fast growing, developing nations such as India, Brazil, China and so on. These countries had vast territories covering millions of square kilometers (Brazil, for example, is over thirty times larger than Britain).

As well as large service territories, many utilities found themselves servicing a patchwork of customers spread over several countries. Each country often had its own national coordinate system and standards to map spatial data.

This raised a number issues for GIS users:

- It was difficult to have a global, all encompassing view of the data.
- Important business analysis could only be performed regionally rather than company-wide.
- Integration of data was often difficult.

or impossible (unless data was expensively exported and then translated from one coordinate system to another common one).

- For very large areas, the curvature of the earth required much more complex calculations to accurately compute distances, areas and so on.
- The practicalities of mapping meant that it was often impossible to use a single coordinate system to represent several national regions.

Geometric calculations based on a flat view of the world result in significant inaccuracies over large areas. Since these calculations often drive business decisions (such as procurement), having an accurate value of, for example, the length of pipe to be purchased becomes an important business requirement.

Local coordinate systems

Within a local coordinate system, a location is represented by a pair of numbers where each number represents a distance in specified units along an axis from an origin. Given the axes and the units, it is then possible to construct a grid for the coordinate system and the origin specifies where to measure distances from. These are all part of the definition of the local coordinate system. For example, the British National Grid has its local coordinate system origin just west of the Scilly Isles with coordinates expressed as a pair of eastings (distance east of the origin) and northings (distance north of the origin) graduated at one metre intervals.

Geodetic coordinate systems

Simple coordinate systems (for example, a road atlas) often assume the world is flat and for small regions the curvature of the Earth is insignificant enough to make this a valid approximation. However, for large land masses such as the Great Plains of North America, the curvature of the Earth has a pronounced impact introducing significant discrepancies in geometric calculations (for example, computed lengths are shorter).

From space a casual observer might conclude that the Earth is a beautifully perfect sphere. However, closer inspection reveals an uneven planet whose surface is dominated by impressive mountain ranges, deep valleys and vast oceans. Its true geometry is that of an irregular ellipsoid or an ellipsoid of revolution. This irregular surface is too complicated to be modeled mathematically. One way forward is to model this shape using an approximation known as the geoid. This is considered to be the plane of intersection of the earth's gravitational field and its mean sea level. However, this model still yields an irregular surface inconvenient for describing locations. As a result, the Earth's surface is often approximated using a relatively simple mathematical shape, similar to a rugby ball (or American football), called an ellipsoid. Quite often, to improve accuracy further, an ellipsoid is chosen that models only part of the Earth's surface that is called a locally fitting ellipsoid. Within this best-fit area (for example, a state) the ellipsoid gives a good approximation of the Earth's surface, outside, the accuracy usually degenerates significantly due to distortions.

The shape and size of an ellipsoid is not sufficient on its own to form a basis of a coordinate system. The location of the ellipsoid also has to be defined using a datum that fixes the positions of primary reference sites, which can include topographical features such as triangulation points on hilltops or permanent man-made features such as church steeples and relates them to the chosen ellipsoid. Datums are usually defined by national agencies and it is quite frequent for two countries to share a common ellipsoid, but to use different datums (for example, the Netherlands and Germany both base mapping on the Bessel 1841 ellipsoid, but use different datums: the RD/NAP and Potsdam datums, respectively). Global datums reference all parts of the world, for example, WGS84.

From the round to the flat

For the purposes of displaying, printing, processing and manipulating spatial data, a flat representation is required that allows map information to be conveniently presented on a computer screen or printed as a paper map. The mechanism that provides this transformation from the round world to the flat world is called a map projection and forms part of the definition of a coordinate system. This process projects locations on the ellipsoid on to a flat plane representing the computer screen or a piece of paper.

The mathematical principles behind the projection of coordinate data are beyond the scope of this paper, but it might be of interest to the reader to highlight some of the wide range of different projections supported by the Smallworld platform:

Cylindrical	Cassinin, Mercator, Miller, Oblique Mercator, Tranverse Mercator.
Conic	Albers Equal Area, Equidistant Conic, Lambert Conformal Conic, Polyconic.
Aimuthal	Azimuthal Equidistant Gnomonic, Lambert Azimuthal Equal Area, Orthographic, Stereographic.

Something has to give

There is obviously not a one-to-one mapping between something that is round and something that is flat. Projecting a round surface on to a flat paper inevitably results in a certain amount of distortion (due to stretching) in the projected map. A good example is the relative sizes of Greenland and South America in the Mercator projection. These two continents appear on a projected map to be the same size, but, in fact, South America is about eight times the size of Greenland.

A projection might distort area, shape, scale, direction or distance. No projection can be free of all distortions, but most projections will maintain at least one of these features accurately. Some projections

preserve local shape and direction at the expense of area. Others maintain area, but distort shape and scale. Keeping this in mind, the selection of the coordinate system requires careful consideration as the wrong choice might be unsuitable for certain applications. The kind of distortion depends on how the projection is derived. Most projections have one of the following characteristics:

- They are conformal (preserve shape and scale)
- They are equal area (preserve area)
- They are equidistant (preserve distance)
- Or they are compromised between these three

Conformal	Lambert Conformal Conic, Mercator, New Zealand, Oblique Mercator, Stereographic, Transverse Mercator.
Equal area	Albers Equal Area, Lambert Azimuthal Equal Area, Sinusoidal.
Equidistant	Azimuthal Equidistant, Equidistant Conic.
Compromise	Miller, Polyconic.

The projection name often includes its characteristic. Some examples are:

- Albers Equal Area
- Azimuthal Equidistant

Another point to note is that the map distortion is not even: it will be at its least where the projection surface touches the Earth and larger farther away from this point.

One view of the world

Support for coordinate system is an integral part of the Smallworld architecture. Unlike a conventional GIS, this functionality is implemented at a low level allowing the provision of spatial data in a coordinate system without the need for additional application specific code. From the application developer's point of view, spatial data appears to be in an appropriate coordinate system even though, in fact, it might be stored in a different coordinate system. This

transparency is an important advantage when developing enterprise GIS applications.

A consistent application programming interface (API) is provided that allows data from different coordinate systems to be easily processed, combined, shared and manipulated.

Smallworld's advanced coordinate system technology is also integrated directly into its user interface. For example, support is provided for inputting geographic data such as linear features in one coordinate system while actually storing it in another.

Powerful GIS functionality such as spatial queries, polygon analysis and geometric functions are all seamlessly integrated with the underlying coordinate system functionality.

The mathematical processing required to translate coordinates from one coordinate system to another is often not insignificant. For the infrequent calculation this processing overhead is generally not too much of a concern. However, this approach is impractical for large amounts of on-the-fly calculations required when, for example, redrawing a map. So how does Smallworld's technology offer such functionality without requiring such heavyweight computing power? The simple answer is: by design. Using an integrated approach allows Smallworld's coordinate system technology to provide an ingenious mechanism for providing fast conversion of spatial data between coordinate systems. It does this by cleverly trading off accuracy and performance when it is most appropriate to do so. For example, when drawing a map the resolution of the screen is vastly inferior to that of the source spatial data. This allows a set of novel computer algorithms to compute coordinate system conversions with a reduced accuracy that is good enough for display and is also fast enough to provide a pleasing experience to the end user. This approach is particularly impressive when rendering grid data such as scanned paper maps or terrain models that require a large numbers of calculations.

This advanced technology allows data that might be stored in disparate datasets (each with its own native coordinate system) to be cleanly integrated providing a consistent, easy to manage, single view of the world.

The Smallworld platform also supplies the essential administration tools that allow enterprise gis customers to configure new coordinate systems manually or to reuse one of a large set of pre-configured coordinate systems bundled with the Smallworld product.

Conclusion

The Smallworld architecture provides a comprehensive, powerful and highly integrated infrastructure for supporting multiple coordinate systems.

This mechanism provides the flexibility to support the often disjointed operations of today's utilities.

Organizations with spatial data spread over several territories can accurately map and exploit this data in ways that do not require costly data translations or piecemeal analysis.

