The Cartographic Journal *Vol. 49 No. 2 pp. 164–175 May 2012*

REFEREED PAPER

Developing a Flow Mapping Module in a GIS Environment

Kamyoung Kim[1], Sang-Il Lee[2], Jungyeop Shin[2] and Eunyoung Choi[3]

[1]Department of Geography Education, Kyungpook National University, 80 Daehak-ro, Buk-gu, Daegu 702-701, Korea.[2] Department of Geography Education, Seoul National University, 599 Gwanak-ro, Gwanak-gu, Seoul 151-742, Korea.[3] Statistical Research Institute, 282-1 Wolpyung-dong, Seo-gu, Daejeon 302-280, Korea
Email: kamyoungkim@knu.ac.kr

The serious drawback of traditional flow mapping commonly representing geographical movement with straight lines of varying widths is a visual cluttering problem. To resolve these issues, several alternatives reducing the problem or enhancing pattern detection have been suggested. Despite being successful in some aspects, they have their shortcomings such as partial representation or misrepresentation of geographical movement and difficulty in interpreting visualized results. In addition, flow mapping functionality has yet to be fully implemented in modern geographic information systems (GIS) applications. Given these problems, this paper proposes the new designed symbols for improving the traditional flow mapping and develops a flow mapping module implementing them within a commercial GIS software package. Application results demonstrated that the developed module can represent gross, net, two-way and location-specific flows from a single dataset efficiently and can make better flow maps with the alternative symbols. The alternative symbols successfully reduced the visual overlap of symbols and improved the estimation of values.

Keywords: geographical movement, traditional flow mapping, Tobler's *Flow Mapper*, alternative symbols, geographic information systems (GIS)

INTRODUCTION

Spatial interactions, which are location-based movements of humans, information, funds, goods and so on, are one of the essential forces driving many socioeconomic processes and changes (Tobler, 2003; Guo, 2009). Among them, migration is very important for understanding regional changes because it leads to social and economic changes as well as demographic changes in origin and destination. One of the first steps to understand and explain the migration phenomena is to visualize spatial relations by using geographical maps.

There are several ways of rendering geographical movement by using maps, such as choropleth, flow or proportional symbol maps (Figure 1). Common techniques are to make proportional symbol maps or choropleth maps for visualizing in- and out-migrations or in- and out-migration rates measured at individual locations. However, they present only a change of state rather than actual movement (Tobler, 2003). The most simple and traditional method depicting the movement of phenomena between geographic regions is flow maps that typically use lines of varying widths. In traditional flow mapping, origins and destinations are connected with a straight line with a width proportional to flow magnitude with an arrowhead added to indicate the direction of flow. This technique is straightforward and familiar to most people and, consequently, easily understandable.

The most serious drawback of the traditional flow mapping approach is a visual cluttering problem (Tobler, 1987; Phan *et al.*, 2005; Andrienko *et al.*, 2008; Cui *et al.*, 2008). The number of flow lines grows geometrically, and they overlap each other severely as the number of geographic units increase. As a result, it is difficult to read the flow maps, and their effectiveness as a visualisation tool decreases. This visual cluttering problem arises due to some characteristics inherent to geographical movement. First of all, a flow map showing interactions between n regions must visualize $n(n-1)$ flow lines in the worst case. Besides the number of flow lines, the problem stems from spatial dependency of geographical movement. One of the 'laws of migration' suggested by Ravenstein (1885) is that '…the great body of migrants only proceed a short distance…' (p. 198). The short-distance movement tendency reflects the first law of geography (Tobler, 1970), that is, 'everything is related to everything else, but near things are more related than distant things' (p. 236). Tobler (1987) restated this law to fit for geographical movement: 'near places interact more than distant places'. This statement implies that spatial interaction, specifically geographical movement, has

DOI: 10.1179/174327711X13166800242356

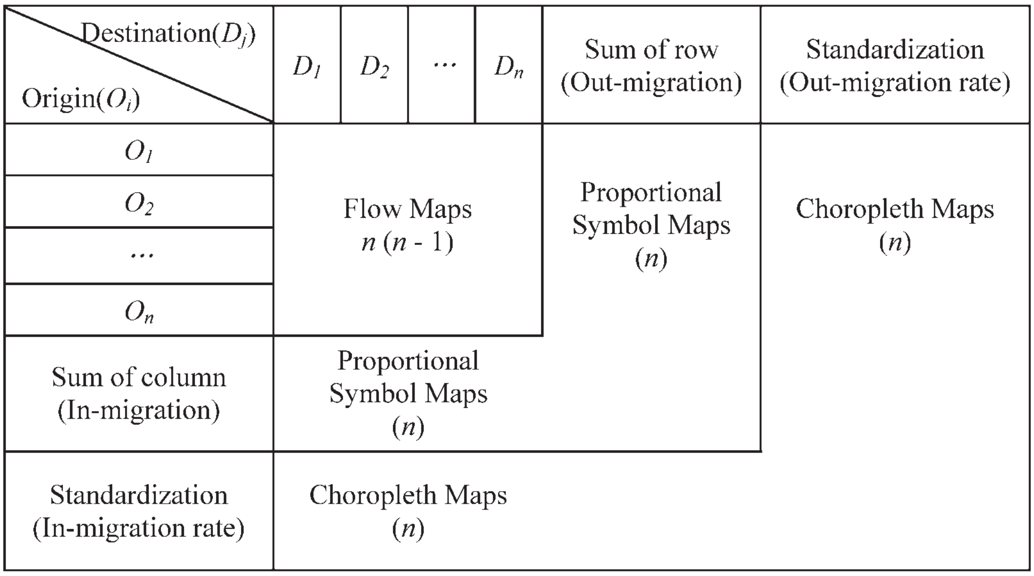

Figure 1. Domains of mapping migration data

spatial dependency. In other words, because most of the flows arise in places in adjacent to one another, most of flow lines are locally clustered. The occlusion of flow lines is intensified with spatial dependency. In addition, sometimes contextual information, such as locations of origin and destination and their boundaries, needs to be visualized for understanding geographical movement in company with flow lines (Tobler, 1987). This situation complicates the problem of reading traditional flow maps.

Several alternatives to the traditional flow mapping have been suggested to express the large volume of flow lines and to reduce visual clutter. They can be classified into three categories according to their strategies. The first strategy is to select a small subset of movement data satisfying some threshold quantity using queries and to visualize them only (Tobler, 1987, 2003; Holland and Plane, 2001; Breukelman *et al.*, 2009). The selection strategy is generally combined with the traditional flow mapping approach. The second strategy is to aggregate flow lines (Phan *et al.*, 2005; Cui *et al.*, 2008) or geographic units (Guo, 2009). The third strategy is to utilize new representation ways such as raster (Nielsen and Hovgesen, 2008; Rae, 2009), tessellation (Wood *et al.*, 2010) or circle (Xiao and Chun, 2009) instead of a straight line vector. Despite being effective in finding answers to certain questions, all of these strategies have their shortcomings. Particularly, the second and third strategies require numerous computational efforts, and their visualized results are sometimes difficult to interpret. Although the alternatives that particularly adopted the second and third strategies could deepen understanding of geographical movement in another dimension, they cannot replace the traditional approach and are inadequate to become a basic flow mapping method within a contemporary GIS environment.

Although there is desktop software like Tobler's *Flow Mapper* (2003), *Flowmap* (Breukelman *et al.*, 2009), and commercial GIS packages allowing for some form of flow mapping, as Holland and Plane (2001) and Rae (2009) pointed out, flow mapping functionality has yet to be fully represented in modern GIS applications compared with other thematic mapping ones such as dot density, choropleth and proportional symbol maps. Given this, the purpose of this research was to design new symbols for improving the traditional flow mapping approach and enhancing exploratory analysis of flow maps in a GIS environment. Furthermore, a flow mapping module implementing the new design methods was developed within a commercial GIS package. The developed flow mapping module will contribute to popularizing the use of flow mapping within a contemporary GIS.

This paper is structured as follows. The next section critically reviews the associated literature on this topic. Following this, alternative design methods to the traditional flow mapping are suggested, data structure for flow mapping is discussed, and a flow mapping module developed is described. Implementation details for the module and application results are also given, and the paper ends with a summary and discussion.

FLOW MAPPING TECHNIQUES

Recently, interests in alternative flow mapping to the traditional approach are gradually increasing. The main foci of existing techniques are on reducing complexity of flow maps and extracting meaningful spatial patterns. These interests have been triggered by Tobler's (1987, 2003) seminal work. *Flow Mapper*, which was developed by Tobler, is the one of the earliest computer systems for making flow maps, and its approach is close to the traditional one. Using *Flow Mapper*, it is possible to visualize gross, net and two-way flows. The main feature of *Flow Mapper* is the attribute selection above some threshold quantity for visualizing flows. The theoretical basis of the selection is that relatively few large flows form a large percentage of the total flows. As a result, *Flow Mapper* emphasizes a few dominant flows and disregards the remaining ones. This approach can effectively reduce the complexity of flow maps while maintaining the principal patterns of geographical movement. Interaction with the flow maps, however, is very limited because *Flow Mapper* is based on graphics rather than geographic features. *Flowmap*, which was developed by Breukelman *et al.* (2009) based on Windows platform for displaying and analyzing interaction data, also employed a similar approach for selecting only specific flows.

Phan *et al.* (2005) and Cui *et al.* (2008) adopted aggregation strategies instead of selection as a way of reducing visual clutter. Phan *et al.* (2005) developed a method for generating flow maps using a hierarchical clustering given a set of nodes, positions, and flow data between the nodes. The flow map visualizes the movement of objects from one location to another. Cui *et al.* (2008) also proposed a geometry-based edge-clustering framework that can group edges into bundles to reduce the overall edge crossings. While Phan *et al.*'s (2005) method groups edges between one location and others, Cui *et al.*'s (2008) one does all flow edges between all locations by using a control mesh. The methods suggested by Phan *et al.* (2005) and Cui *et al.* (2008) are effective to reduce visual clutter on a flow map without attribute filtering. These methods, however, could not visualize directionality of flows, and aggregated flow lines make it difficult to recognize the correspondence between origin and destination (Guo, 2009). In addition, users may misinterpret the flow of intermediate regions where many lines cross, but have small flows.

Guo (2009) aggregated geographic units instead of flow lines so as to reduce the complexity of flow maps. Observed $n \times n$ spatial interaction matrix was reduced into a coarse $m \times m$ one by merging adjacent geographic units with some objectives. Then, flow maps were produced based on the reduced matrix. His basic idea was that meaningful spatial patterns may be revealed at other scales than with only a geographic unit of the observation data. Despite the meaningfulness of extracted spatial patterns from some perspective, the patterns may be unattractive to users who are familiar with the observation spatial units like administrative districts.

While origin and destination in geographical space generally are connected with a straight line in traditional flow mapping, Xiao and Chun (2009) represented geographical movement with a half-circle connecting two points in linearly projected space. They suggested a new method for visualizing geographical movement called kriskogram. In order to produce kriskograms, geographical units are projected on a straight line, and two points on the line are connected with a half-circle. Using kriskograms, bidirectional flows can be effectively visualized without arrows because half-circles at upper and lower sides indicate out- and in-flows respectively. By projection, however, kriskograms may lose some important information like the adjacency between geographical units, which is sometimes very useful to understand spatial interaction. In other words, a one-dimensional line could not preserve every spatial relationship in two-dimensional geographical space. In addition, kriskograms are not free from the occlusion problem either when the number of geographic units increases and all half-circles are expressed.

Nielsen and Hovgesen (2008), Rae (2009) and Wood *et al.* (2010) utilized completely different representation methods from lines. Nielsen and Hovgesen (2008) and Rae (2009) represented geographic movement with raster rather than vector. After generating flow line features between places from a flow table, the features and a regular grid are overlaid, and then the density of a cell is measured by the sum of total number of flows moving along flow lines through the cell. The flow line density map can effectively visualize very large spatial interaction datasets and can extract overall spatial interaction patterns. Straight lines connecting origin and destination, however, are conceptual representations rather than actual flows. This implies that if other types of lines such as a curve are used, then the calculated density of a cell and overall spatial patterns will be changed. In addition, intermediate regions where many lines cross but have small movements, may have a high line density. This means that it is not always evident whether the higher flow line densities in central regions are due to migrants at those locations or simply because they happen to be placed on the path of other origins and destinations (Wood *et al.*, 2010). High flow line density derived by the latter may be a misrepresentation of movements.

Recently, Wood *et al.* (2010) represented spatial interaction with a regular tessellation and suggested a new visualisation method, called OD maps which combine geographic space representation with OD matrix representation. Geographic space is partitioned into square polygons called the origin space, and then each polygon is partitioned again into small square polygons called the destination space. Movement from an origin to a destination is represented in each polygon. An OD map is very scalable, effective in removing occlusion problem and very useful for exploring overall spatial structure of geographical movement, particularly spatial dependency. This representation, nonetheless, causes a loss of detail in origin and destination locations. In addition, an unfamiliar and complicated OD map is difficult to interpret. A more critical problem of OD maps is that representation depends upon the size of tessellation (i.e. scale) and the method allocating underlying geographic units to tessellated polygons.

To sum up, even though the alternative visualisation techniques are obviously successful as strategies for reducing the visual cluttering problem and exploring overall spatial structure of geographical movement, they also possess some limitations. Some of these techniques could not represent two-way flows, and others have the critical risk of misrepresenting geographical movement. Moreover, sometimes it is difficult to interpret the represented results of some techniques due to the loss of important information such as detail in the location and correspondence of origin and destination or the adjacency information of geographical units by projection, abstraction or aggregation, as well as the utilisation of unusual representations like tessellation. Therefore, the alternative techniques can provide various options for visualizing or exploring geographical movement, but they are insufficient for the basic flow mapping functionality in GIS. This implies that the traditional flow mapping approach, which has the visual cluttering problem but can represent all kinds of flows and

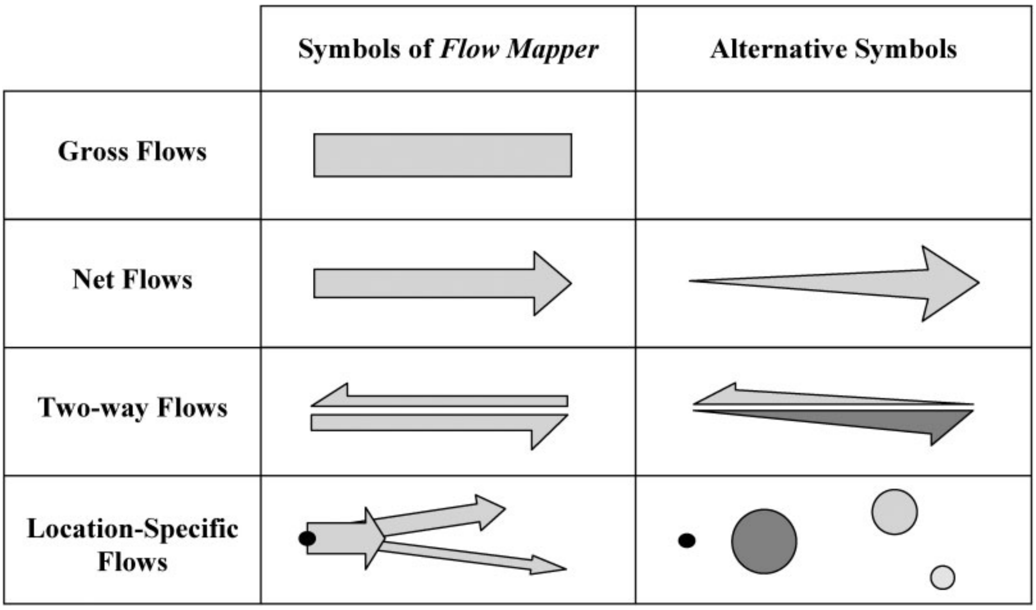

Figure 2. *Flow Mapper*'s symbols and their alternatives

is straightforward, is still valid as the starting point of understanding geographical movement in GIS environment.

FLOW MAPPING IN GIS

Alternative symbols design

Before suggesting new symbols for the traditional flow mapping, it is necessary to briefly discuss a symbol design limitation found in traditional flow mapping in general and *Flow Mapper* suggested by Tobler (2003). In the case of two-way, net and location-specific in- or out-flows, rectangular polygons may increase occlusion. Particularly, even flow maps showing only flows related to an origin or a destination suffer from the occlusion problem. This outcome primarily results from the short-distance movement tendency. Consequently, the occlusion problem of *Flow Mapper* is related to the utilisation of the rectangular polygon to a certain degree.

Alternative symbols are suggested instead of line bands with the same width or the rectangular polygons used in *Flow Mapper*. Figure 2 presents symbols utilized in *Flow Mapper* and alternative symbols devised in this research except gross flows. The best feature of the suggested symbols is that in order to represent flows, the width of a polygon gradually increases towards a destination. Gradually increasing polygons can be combined with proportional circles or arrow heads to indicate the flow direction. In the case of one-way directional movements like net flows, the gradually increasing polygons can significantly reduce the occlusion of symbols near starting points compared with rectangular polygons. This echoes *Flowmap*'s approach (Breukelman *et al.*, 2009) which utilizes a wedge symbol to visualize both the magnitude and direction in the case of net flows.

Another potential symbol for visualizing the direction and amount of flows is a circle. The circle is particularly useful to express location-specific flows as Rae (2009) has tried. Location-specific flows are not necessarily represented with flow lines connecting origin and destination because all flows are related to an origin or a destination. In other words, it is possible to visualize the flows related to a specific place without the connecting information, such as line bands, which typically cause occlusion. The circles are selected as an alternative symbol due to their compactness, users' preference and visual stability (Slocum *et al.*, 2009). In addition, the circles are relatively effective in estimating the amount of values under the overlap of symbols compared with the other symbols (Dent *et al.*, 2008).

Flow data model

In Tobler's *Flow Mapper*, input information for generating flow maps are interaction tables and location points. A flow line connecting an origin and a destination is drawn graphically, and its width is proportionally determined according to flow magnitudes. The data structure and flow mapping approach of *Flow Mapper*, however, are not compatible with existing GIS software (Glennon and Goodchild, 2004) where thematic maps are usually made based on geometries such as points, lines or polygons. Due to graphic-based flow mapping, interaction with the flow maps, such as zooming in areas where flow lines are severely overlapped or identifying specific flows, is very restrictive in the maps made by *Flow Mapper*. Flow mapping based on spatial features can retain compatibility for GIS software and improve interaction with maps like other thematic mappings. In the situation that flow mapping functionality has remained underdeveloped in modern GIS, it is a task of great significance to develop a new data model and a flow mapping module suitable for GIS environment.

Flow Data Model (FDM) suggested by Glennon and Goodchild (2004) could be a data model for flow mapping in GIS environment. FDM defines the relationship between flow tables and the reference data containing the location of flow's origin and destination in order to generate the geometry of flow and determine the magnitude of the interactions assigned to it. In other words, by using FDM, polyline features connecting origin and destination with an attribute field representing flow magnitude are created from flow tables and reference data. Rae (2009) generated ESRI shapefile, which is a source of calculating flow line density, using FDM. The original FDM, however, has a drawback when visualizing two-way flows simultaneously. When visualizing two-way flows between places, visualized symbols of two polylines with the same geometry, that is, one is from an origin to a destination and the other is one from the destination to the origin, are totally overlapped. To solve this problem, the paper modified FDM so that each polyline feature has two flow attributes, i.e. flows from origin to destination and counterpart flows. When visualizing two-way flows, each field is symbolized only on one side of the polyline without overlap. The general procedure of creating polyline features using the modified FDM is as follows:

1. Create a new polyline feature file with code of origin (OID), code of destination (DID), flows from origin to destination (FLOW), counterpart flows (CFLOW), net flows (NET) and gross flows (GROSS) fields, and allow it to be edited.
2. Read a record from a flow table and set code of origin (O), code of destination (D) and magnitude of flows (Mag). If end-of-file, then stop editing, and quit the procedure.
3. Search the magnitude of counterpart flow record from the flow table. If the record count is 0, then C_Mag is equal to 0, or otherwise set the magnitude to C_Mag.
4. Search polygon features with O and D from reference spatial data, respectively.
5. Create a polyline feature. Its FROM and TO nodes are the centroid of a feature with O and the one of a feature with D in the reference spatial data, respectively.
6. Assign attributes to the polyline feature.

Figure 3 shows the flowchart of creating a polyline feature file from a flow table and reference data. Point feature data also can be used as reference data. In this case, the location of points instead of centroids become FROM and TO nodes. The implementing code of the procedure was written by using Visual Basic for Applications within ESRI ArcGIS 9.3.

Flow mapping module

A flow mapping module that visualizes flow data with alternative symbols and is capable being integrated with contemporary GIS software, specifically ESRI ArcGIS desktop, was developed using C# in Microsoft Visual Studio.NET 2005 providing access to.NET Framework and the functions that ESRI has made for.NET. The flow maps made by the module are distributive flow maps depicting the movement between geographic regions among the five types of flow maps suggested in Slocum *et al.* (2009). Particularly, the general direction and magnitude are more important components rather than the exact route of flow. The module is an extension for ArcGIS 9.3[1]. If the module is installed, Flow Maps appears as a sub-category for thematic mapping under the 'Show' panel of the 'Symbology' tab (Figure 4).

Users can set magnitude fields, size of symbols, symbol types and colours according to flow type to be visualized. One of the most important characteristics of the module is that two magnitude values can be simultaneously visualized by setting both 'Forward' and 'Backward' fields. 'Forward' is for an attribute field indicating flows from an origin to a destination, and 'Backward' is for an attribute field representing counterpart flows. Both should be selected together when representing two-way flows. While directional marks appear at the TO node of the geometry of flow in the case of forward flow, they are displayed at the FROM node for backward flow.

The width of symbols is proportional to the magnitude of flows. If the symbol size of maximum values is determined, all other values are scaled downward from this size. In addition, to compare different flow maps, users can enter a specific maximum value by checking 'Set Max Value'. Besides the size variable, the module utilizes colour in order to show quantitative differences in map symbols as visual variables (Bertin, 1983). If 'Use Gradient Color' is checked, then the colour lightness (or value) of polygon or circle symbols is varied according to the magnitude of flows. Users can change the colour of forward and backward flows. Of course, distinction of in- and out-flows with colour may not be meaningful in understanding overall spatial patterns because they are relative concepts. That is, the out-flow of one place is the in-flow of another place. However, the distinction might be useful when in- and out-flow patterns of a specific place are understood.

If 'Sort Values by Magnitude (ascending order)' is checked, then symbols with larger flows crossover the top of ones with smaller flows. As a result, the more important flows become more noticeable, otherwise *vice versa*. Researchers have different opinions about what is the better strategy. Tobler (1987, 2003) emphasized more significant flows by showing the largest flows on top. Dent *et al.* (2008), however, argued that smaller flow lines should appear on top of larger flow lines as one of essential design decisions in creating flow maps.

Selectable symbol types and their shapes are classified according to the number of attribute field to be visualized in Figure 5. As can be seen, all alternative symbols as well as the symbols used in *Flow Mapper* suggested in Figure 2 are included as selectable options. The selection of symbols depends on the type of flows. The bold type indicates the most suitable flow type for each symbol. Rectangular polygons can be used to represent gross flows. Net flows can be effectively visualized with four different options: gradually increasing polygons, rectangular polygons with arrows, gradually increasing polygons with arrows and gradually increasing polygons with circles. In the case of gradually increasing polygons with circles, both polyline and point renderings are applied. Proportional circle symbols are generated at the TO node of a polyline. These symbols are partially appropriate to location-specific flows because arrows or circles are extremely occluded at the location when visualizing in-flows. The proper symbols for visualizing

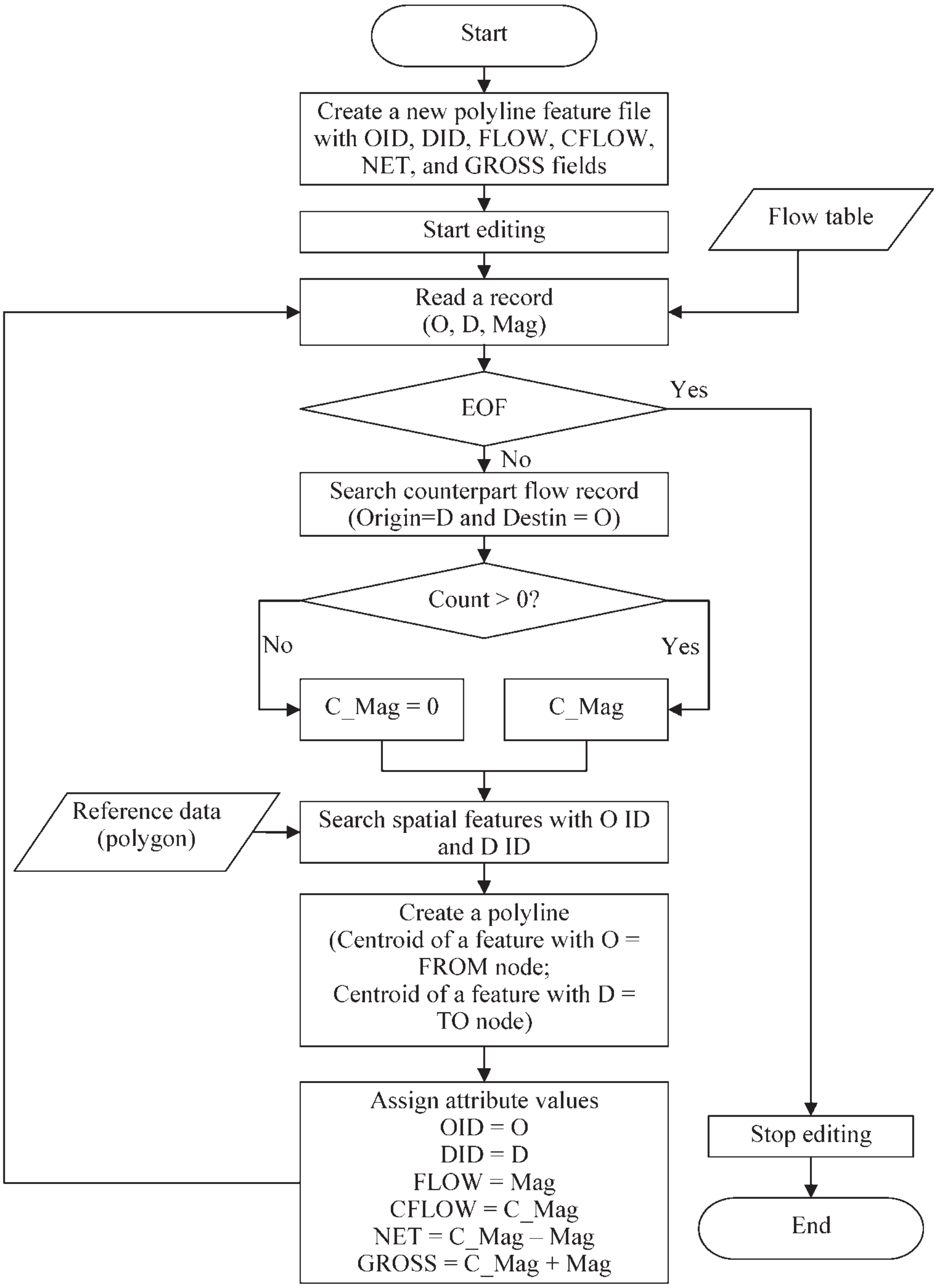

Figure 3. Flowchart of creating a polyline feature file

two-way flows are similar to ones for net flows. The most appropriate symbol for representing location-specific flows is a circle, but they are applied when visualizing only one attribute field, that is, in- or out-flows. While proportional circle symbols are created at FROM nodes (origin) when in-flows to a place are visualized, in the case of out-flows from a place, TO nodes (destination) are rendered.

In order to make flow maps with the designed symbols, selection functionality is essential. The reasons why selection is important in the developed module are, first, like Tobler's *Flow Mapper*, to reduce the complexity of the maps and to emphasize salient flows, and second, to select targeted features from the spatial data created by the modified FDM, which includes redundant information in terms of geometry, for example, two polylines connecting regions A and B, and B and A. However, a separate selection functionality does not need to be developed because GIS provides excellent query functions. Particularly, ESRI ArcGIS provides an easy dialogue box called Query Builder. It is possible to display features with certain attributes using this function.

APPLICATIONS

Our developed module was applied to visualize the interstate migration data of US Census 2000. For

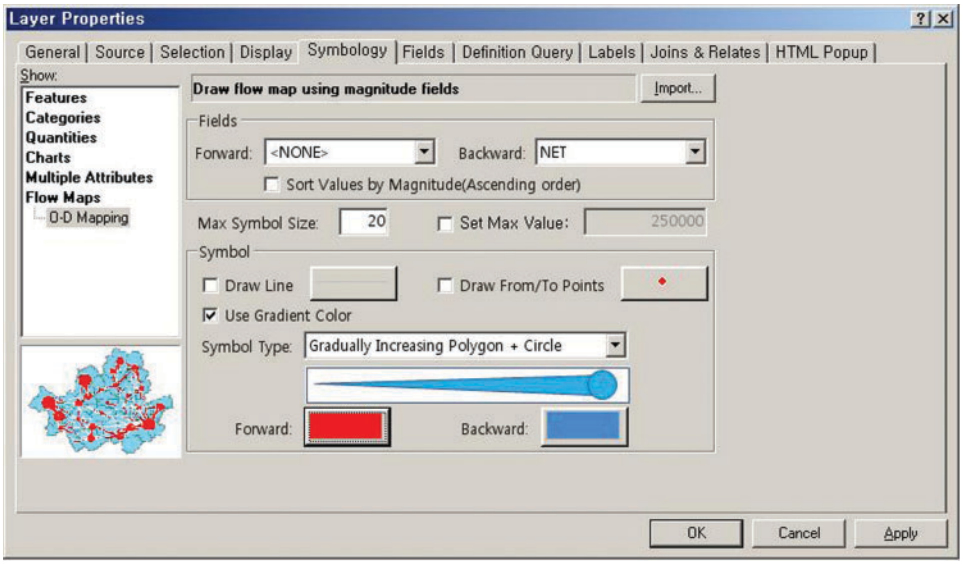

Figure 4. Interface of flow mapping module

convenience of visualisation, only migrations among the 48 US mainland states are considered.

Figure 6 shows flow maps for gross migrations from 1995 to 2000 made by setting 'Forward' with GROSS and using the rectangular polygon symbol. In these maps, only the features satisfying the following condition, that is, gross migrations are greater than or equal to 100,000 and identification of origin is greater than one of destination, have been rendered. The latter condition is included in order to select one of polylines with the same geometry for a pair of regions. While the first map displays large flows on top, the second one does small flows on top. In Figure 6a,

Symbol Types	Number of Attribute Field	
	One	Two
Rectangular polygon	**Gross**	
Gradually increasing polygon	**Net** Location-specific	Two-way Location-specific
Rectangular polygon with arrow	**Net** Location-specific	**Two-way** Location-specific
Gradually increasing polygon with arrow	**Net** Location-specific	**Two-way** Location-specific
Gradually increasing polygon with circle	**Net** Location-specific	**Two-way** Location-specific
Circle	**Location-specific**	

Figure 5. Selectable symbol types and their suitable flow types (bold type: the most suitable flow type)

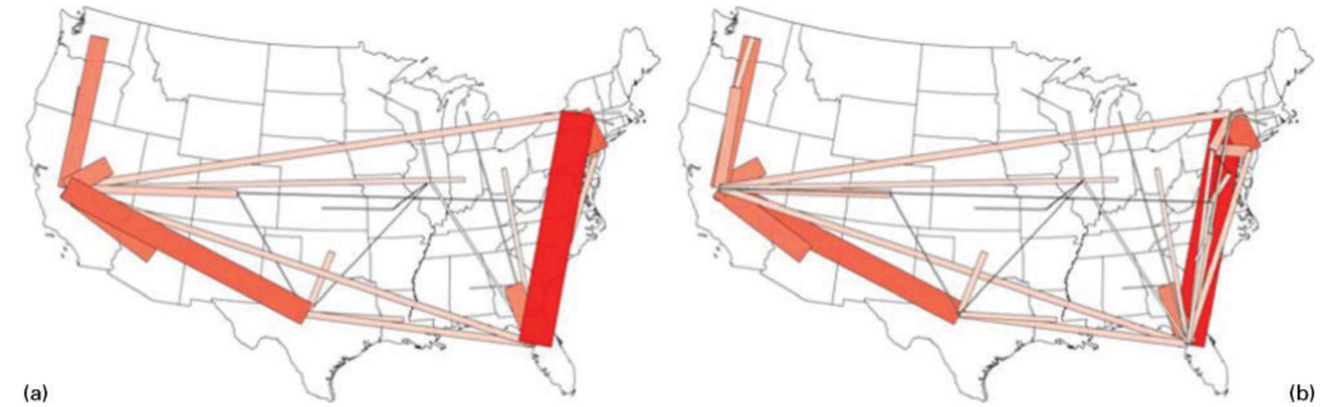

Figure 6. Flow maps for gross migrations (1995–2000): (a) large flows on top, (b) small flows on top

the salient flows are emphasized, but the small flows among intermediate states being placed between New York and Florida have been covered by the largest polygon symbol. In contrast, the blocked small symbols are shown in Figure 6b. The larger flows are less remarkable, but still identifiable. From this observation, if large spatial interactions take place between two geographically remote regions, highlighting them by placing them on the top of flow maps seems to be less suitable. Definitely, the variation in colour value of symbols facilitates extracting spatial patterns and recognizing differences in visualized values.

Figure 7 shows four different types of flow maps for net migrations from 1995 to 2000 where the values are greater than or equal to 30,000. The NET attribute field of the spatial data includes both positive and negative values, but only positive ones are the target of display. Because net flow is calculated by subtracting forward flow from backward flow according to the definition (i.e. in-flows minus out-flows), if the backward flow is greater than forward flow, then the value of the net flow is positive and the directionality of the positive value is from destination to origin. NET field, therefore, had to be set as a 'Backward' field in order to indicate the directionality. As can be seen, by comparing Figure 7b with the others, the gradually increasing polygons clearly alleviated the occlusion problem. However, in the case of using only the gradually

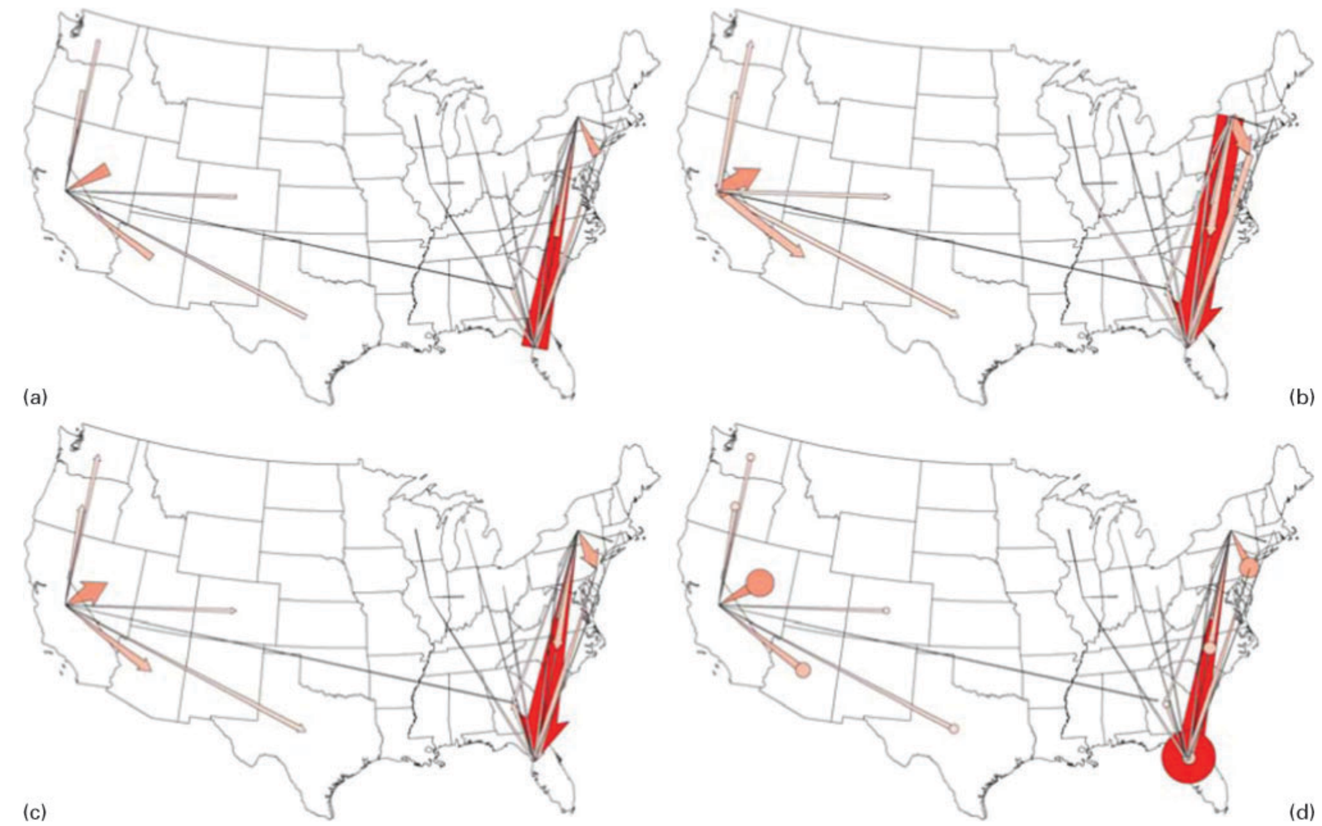

Figure 7. Flow maps for net migrations (1995–2000): (a) gradually increasing polygons, (b) rectangular polygons with arrows, (c) gradually increasing polygons with arrows, (d) gradually increasing polygons with circles

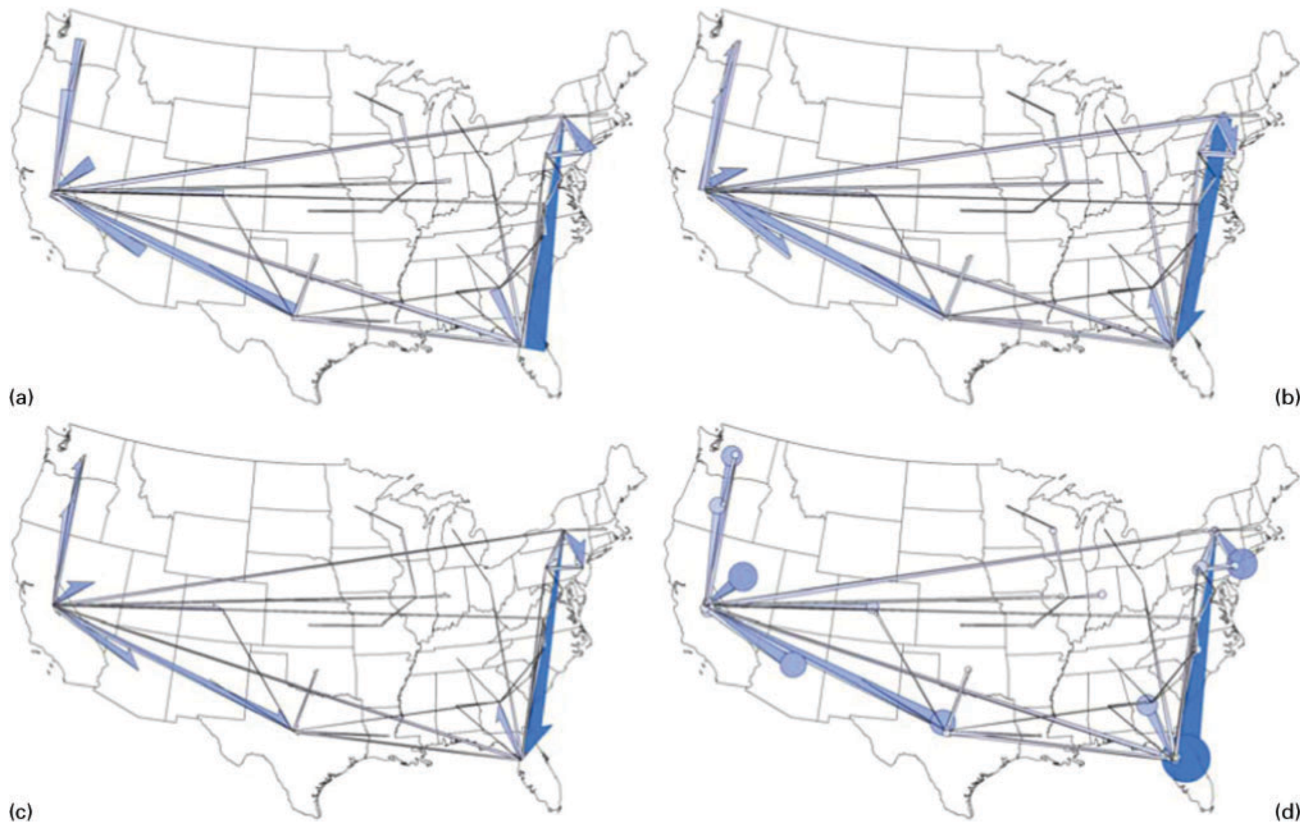

Figure 8. Flow maps for two-way migrations (1995–2000): (a) gradually increasing polygons, (b) rectangular polygons with arrows, (c) gradually increasing polygons with arrows, (d) gradually increasing polygons with circles

increasing polygons (Figure 7a), it is difficult to estimate values from the symbols due to the relatively small variation in the width of a symbol. However, using additional symbols such as arrows and circles can reduce the difficulty in value estimation resulting from the use of gradually increasing polygons (Figure 7c and d).

Flow maps visualizing two-way migrations with different symbols are suggested in Figure 8. FLOW and CFLOW are chosen as 'Forward' and 'Backward' fields, respectively, and features whose values of forward and backward flow fields are greater than or equal to 40,000 are portrayed. Similar to net migrations, the gradually increasing polygons reduced the overlap of symbols. The estimation of values from narrow polygon symbols, however, is increasingly more difficult when many symbols are related to a region like California or Florida in Figure 8. While the wide and recognisable parts of the symbols are severely overlapped with each other near the region, the part not being overlapped is too narrow to make out the value. Certainly, the circles are more effective in communicating magnitudes in overlapping situations than the arrows (Figure 8d).

The distinction of in- and out-flows by using different colours might be useful for understanding of migration pattern related to a specific location. In Figure 9, the

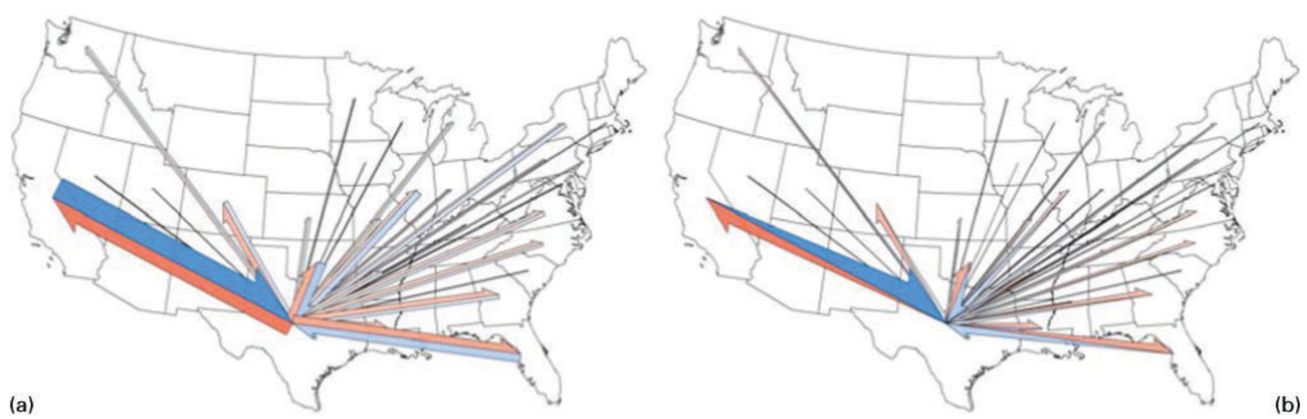

Figure 9. Two-way migrations of Texas (1995–2000): (a) rectangular polygons with arrows, (b) gradually increasing polygons with arrows

Figure 10. Flow maps for out-migrations from California (1995–2000): (a) gradually increasing polygons, (b) rectangular polygons with arrows, (c) gradually increasing polygons with arrows, (d) gradually increasing polygons with circles

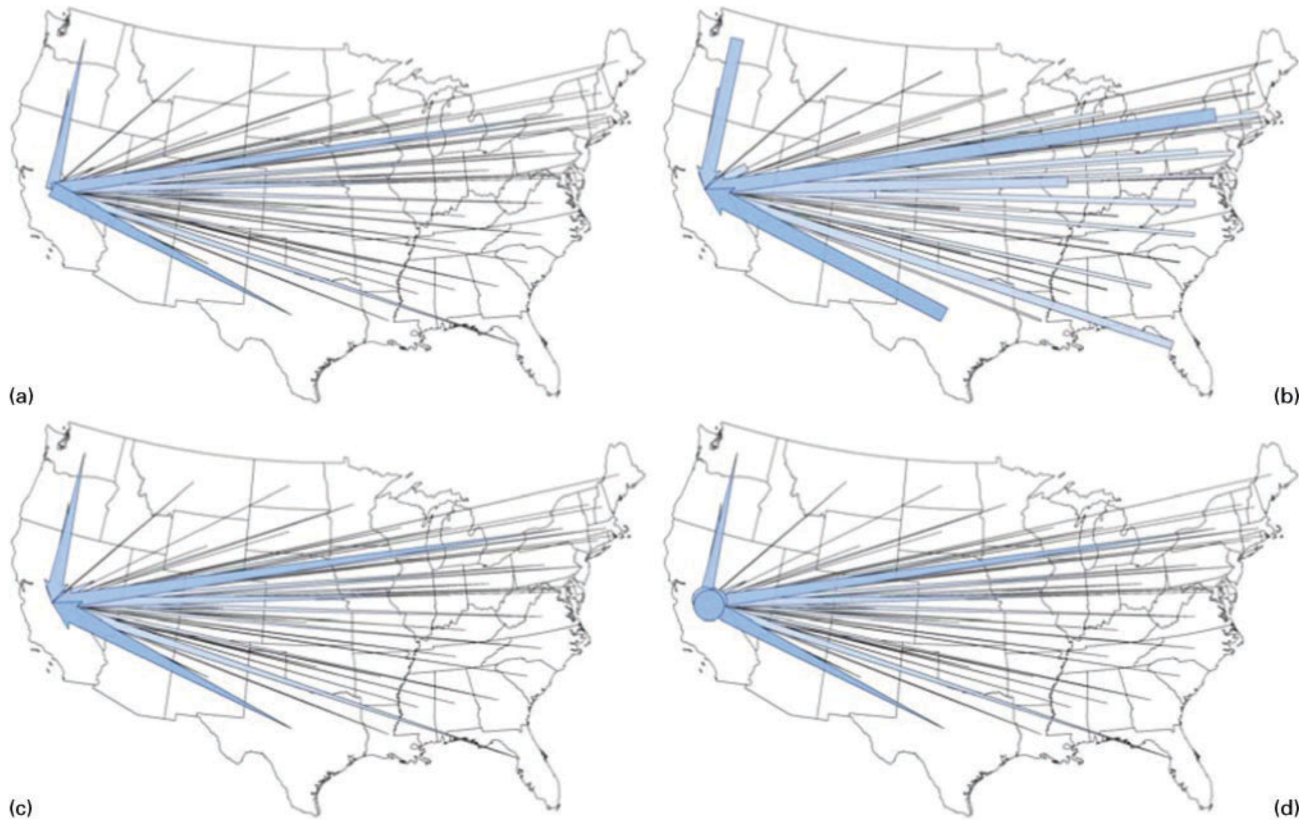

Figure 11. Flow maps for in-migrations to California (1995–2000): (a) gradually increasing polygons, (b) rectangular polygons with arrows, (c) gradually increasing polygons with arrows, (d) gradually increasing polygons with circles

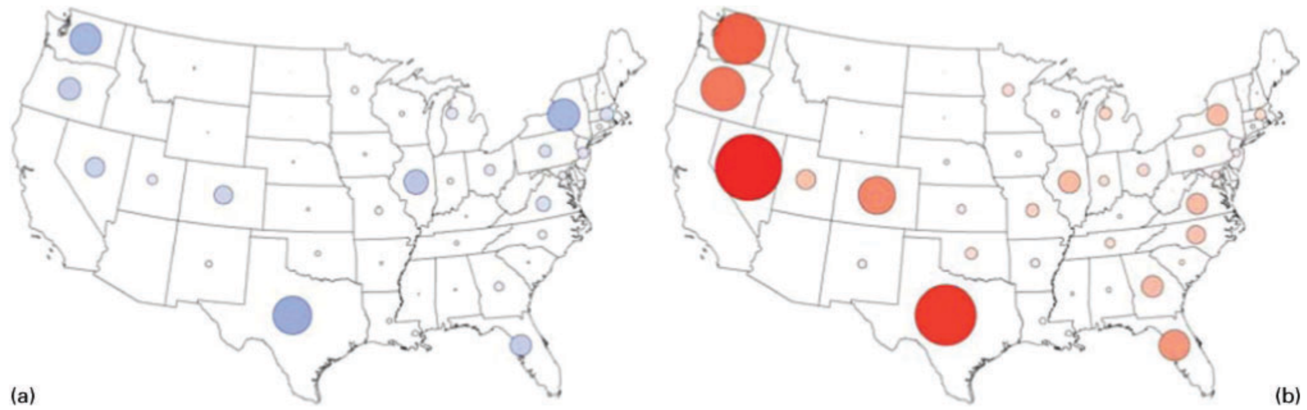

Figure 12. Proportional symbol maps for in- and out-migrations of California (1995–2000): (a) in-migrations, (b) out-migrations

conditions of the query is that the origin is Texas and gross flows are greater than or equal to 25,000, and both FLOW and CFLOW fields are chosen as 'Forward' and 'Backward' values to show the migration relationship between Texas and the other regions. Out-migrations from Texas is portrayed in blue, and in-migrations to Texas is rendered in red. As a result, flow attributes are depicted with three visual variables, i.e. size, value and hue.

Figures 10 and 11 illustrate four different flow maps visualizing out-migrations from California and in-migrations to it, respectively. Polygon symbols rendering line features are suitable for portraying out-migrations from a specific location despite their overlap because directional symbols are placed at destinations. However, in the case of in-migrations, as all directional symbols are placed at one destination, it is difficult to estimate values from the symbols. This problem is more serious when gradually increasing polygons are used.

An alternative way to solve the occlusion problem and to facilitate the estimation of values in location-specific in- or out-flow maps is to make proportional circle symbol maps (Figure 12). When visualizing location-specific in- or out-migrations, as all flow features visualized are related to a place, connection information between the place and the other ones which causes the occlusion problem does not necessarily have to be visualized. For in-flows mapping, the field indicating counterpart flows (CFLOW) is set in 'Backward' so as to generate circles at origins (Figure 12a). In contrast, the field indicating forward flows (FLOW) is chosen as 'Forward' and circle symbols are portrayed at destinations (Figure 12b). Figure 12 demonstrates that it is possible to effectively visualize location-specific flows without line symbols.

CONCLUSIONS

The traditional method of flow mapping that commonly represents geographical movement with straight lines of varying width suffers from a visual cluttering problem due to inherent characteristics of the phenomena such as visualisation of $n(n-1)$ flow lines and short-distance movement tendency. Several alternatives reducing the problem or enhancing pattern detection by selecting subsets of dominant flows, aggregating flow lines or geographical units, or utilizing new representations such as half-circle, raster or regular tessellation, have been suggested. The alternatives, however, have their shortcomings, such as partial representation or misrepresentation of geographical movement and difficulty in interpreting visualized results due to projection, abstraction or aggregation. In addition, flow mapping functionality has yet to be fully implemented in modern GIS applications compared with other thematic mappings. In this situation, development of flow mapping functionality easily integrated with contemporary GIS software as one of thematic mappings is needed.

Given this, this paper designed new symbols for improving the traditional flow mapping and developed a flow mapping module implementing them within commercial GIS software. Alternative symbols, such as gradually increasing polygons or circles, to the rectangular line band used in the traditional flow mapping, were suggested for reducing the occlusion problem of flow maps. Then, a new flow data model where each feature has forward, counterpart, gross and net flow attributes was proposed. A flow mapping module was developed as an extension of GIS software based on the flow data model and the alternative symbols. In order to portray flow attributes, three different visual variables such as size, value and hue were utilized in the module. Application results demonstrated that the developed module can represent gross, net, two-way and location-specific flows from a single dataset using the query functionality of GIS and can make flows maps with the alternative symbols as well as the symbols suggested by Tobler (1987, 2003). The alternative symbols considerably reduced the overlap of symbols and improved the estimation of values. In addition, use of the visual variables and functionality of sorting value facilitated to the detection of spatial patterns from flow maps.

In conclusion, the developed module could be a standard framework to make flow maps in GIS environments like other thematic maps. Contrary to desktop software like *Flow Mapper*, various interactions with the flow maps made by our module are possible, because it is based on spatial features similar to other thematic mappings. The interaction will enhance exploration of geographical movement. As a result, the developed flow mapping module will contribute to popularizing the use of flow mapping within a contemporary GIS.

BIOGRAPHICAL NOTES

Kamyoung Kim is an assistant professor in Department of Geography Education at Kyungpook National University, Korea, with interests in spatial modelling and geovisualisation. Sang-Il Lee is an associate professor in Department of Geography Education at Seoul National University, Korea, with interests in spatial statistics and geographic information science. Jungyeop Shin is an associate professor in Department of Geography Education at Seoul National University, Korea, with interests in spatial analysis and urban geography. Eunyoung Choi is a researcher at Statistical Research Institute, Korea, with interests in migration and spatial analysis.

NOTES

[1] Readers can get the extension by personal contact with the lead author.

ACKNOWLEDGEMENTS

This research was supported by Kyungpook National University Research Fund, 2010.

REFERENCES

Andrienko, G., Andrienko, N., Dykes, J., Fabrikant, S. I. and Wachowicz, M. (2008). 'Geovisualization of dynamics, movement and change: key issues and developing approaches in visualization research', **Information Visualization**, 7, pp. 173–180.

Bertin, J. (1983). **Semiology of Graphics: Diagrams, Networks, Maps**, University of Wisconsin Press, Madison, WI.

Breukelman, J., Brink, G., de Jong, T. and Floor, H. (2009). Flowmap 7.3 Manual, http://flowmap.geog.uu.nl/downloads/FM731_Manual.pdf.

Cui, W., Zhou, H., Qu, H., Wong, P. C. and Li, X. (2008). 'Geometry-based edge clustering for graph visualization', **IEEE Transactions on Visualization and Computer Graphics**, 14, pp. 1277–1284.

Dent, B. D. (2008). **Cartography: Thematic Map Design**, 6th ed., McGraw-Hill, Boston, MA.

Glennon, A. and Goodchild, M. (2004). A GIS Flow Data Model, http://dynamicgeography.ou.edu/flow/.

Guo, D. (2009). 'Flow mapping and multivariate visualization of large spatial interaction data', **IEEE Transactions on Visualization and Computer Graphics**, 15, pp. 1041–1048.

Holland, S. C. and Plane, D. A. (2001). 'Methods of mapping migration flow patterns', **Southeastern Geographer**, 41, pp. 89–104.

Nielsen, T. and Hovgesen, H. (2008). 'Exploratory mapping of commuter flows in England and Wales', **Journal of Transport Geography**, 16, pp. 90–99.

Phan, D., Xiao, L., Yeh, R., Hanrahan, P. and Winograd, T. (2005). 'Flow Map Layout', in **IEEE Symposium on Information Visualization (Infovis 2005)**, pp. 219–224, Minneapolis, MN, Oct 23–25.

Rae, A. (2009). 'From spatial interaction data to spatial interaction information? Geovisualisation and spatial structures of migration from the 2001 UK Census', **Computers, Environment and Urban Systems**, 33, pp. 161–178.

Ravenstein, E. (1885). 'The laws of migration', **Journal of the Statistical Society**, 48, pp. 167–235.

Slocum, T. A., McMaster, R. B., Kessler, F. C. and Howard, H. H. (2009). **Thematic Cartography and Geovisualization**, 3rd ed., Pearson Prentice Hall, Upper Saddle River, NJ.

Tobler W. (1970). 'A computer movie simulating urban growth in the Detroit region', **Economic Geography**, 46, pp. 234–240.

Tobler, W. (1987). 'An experiment in migration mapping by computer', **The American Cartographer**, 14, pp. 155–163.

Tobler, W. (2003). Flow Mapping, http://www.csiss.org/clearinghouse/FlowMapper/MovementMapping.pdf.

Wood, J., Dykes, J. and Slingsby, A. (2010). 'Visualisation of origins, destinations and flows with OD maps', **The Cartographic Journal**, 47, pp. 117–129.

Xiao, N. and Chun, Y. (2009). 'Visualizing migration flows using kriskograms', **Cartography and Geographic Information Science**, 36, pp. 183–191.