

Mapping Social Relationships across Space and Time

Caglar Koylu, Diansheng Guo, Alice Kasakoff

ABSTRACT: The wide use of social networking applications such as Facebook, Twitter, Foursquare and genealogy applications such as Family Search and Geni have made it possible to construct large and dynamic social networks that evolve over space and time. Understanding the structure and evolution of such networks is crucial for various fields such as disaster evacuation planning and provision of care to elderly. Visualization can potentially play a key role in analyzing and understanding such dynamic and complex networks embedded in space and time. With the advancement of graph drawing algorithms, current methods of network visualization are effective in discovering kinship patterns. However, such methods are inadequate in discovering how those kinship patterns change over space and time.

To address the problem, we introduce a new spatial density mapping approach for visualizing social relationships across space and time and a new kin proximity index which takes into account of space and time in quantifying social relationships. Our approach consists of two major steps. First, we partition the data into time windows and regions, and then we build a kin relationship graph in which the weight of each edge is determined by the kin proximity value between the two nodes (individuals). Second, we smooth each of the kin relationship graphs that belong to a space (region) and time (window) interval; and calculate a kin concentration value for each location using the smoothed graph. The result of our approach is a collection of kin concentration surfaces over a sequence of time windows. We demonstrate the approach using a family network in the U.S. North which includes information on migration and expands over three hundred years. The results highlight that family hubs with high concentration of kin proximity is not necessarily correlated with the high density of people and tend to relocate between regions as a result of processes such as chain migration, return migration and fluctuations in fertility and mortality.

KEYWORDS: Space-time visualization, social networks, spatial density mapping, kin proximity, migration

Introduction

The wide use of social networking applications such as Facebook, Twitter, Foursquare and genealogy applications such as Family Search and Geni have made it possible to construct large and dynamic social networks that evolve over space and time. Understanding the structure and evolution of such networks is crucial for various fields such as disaster evacuation planning and provision of care to elderly. A dynamic social network consists of nodes (e.g., individuals) and edges that metaphorically represent relations (e.g., family relations or friendship) among individuals. Both the nodes and edges of such a network evolve (change) over time as individual nodes move over space, new nodes are added or removed, and relationships change over time. Visualization can potentially play a key role in analyzing and understanding such dynamic and complex networks embedded in space and time. With the advancement of graph drawing

algorithms, current methods of network visualization (Lewis, Gonzalez, & Kaufman, 2012; Patil, 2011) are effective in discovering kinship patterns (e.g., clusters of connected members, or commonalities between friends who share interests and groups in a social networking application). However, such methods are inadequate in discovering how those kinship patterns change over space and time.

To address the problem, we introduce a new spatial density mapping approach for visualizing social relationships across space and time. Conventional spatial density mapping methods, calculate a density of features in a neighborhood (bandwidth) by fitting a smoothly curved surface over each estimation location. Instead, we smooth edges of a kin relationship graph and then derive a kin concentration surface using the smoothed graph. We also introduce a new kin proximity index that takes into account of space and time in quantifying social relationships. Our approach consists of two major steps. First, we partition the data into time windows and regions, and then we build a kin relationship graph in which the weight of each edge is determined by the kin proximity value between the two nodes (individuals). Second, we smooth each of the kin relationship graphs that belong to a space (region) and time (window) interval; and calculate a kin concentration value for each location using the smoothed graph. The result of our approach is a collection of kin concentration surfaces over a sequence of time windows. To demonstrate the approach, we use a family network derived from the published genealogies of a family from the U.S. North over a span of three hundred years. The data also include information on migration of individuals. The results highlight that family hubs with high concentration of kin proximity is not necessarily correlated with the high density of people and tend to relocate between regions as a result of processes such as chain migration, return migration and fluctuations in fertility and mortality.

Kin Proximity

We argue that the spatial proximity of family members affects their ability to interact. Despite our current ability to communicate over great distances, certain kinds of help and support can only be provided in person. For example, familial relationships provide social security during crises of various sorts and the ability to assess these is significantly affected by distance. The different forms family (kinship) networks take in space have not been well described partly due to lack of data on the whereabouts of relatives (kin), especially more distant ones. However, the wide use of social networking applications and sharing of genealogical research through web sites have made this information easily available for the first time. The spatial distribution of relatives is the result of both migration and the differential fertility and mortality of different family members. Thus, as migration patterns change, one would expect this to be reflected in changes in kin proximity.

Visualizing Social Networks

With the advancement of graph drawing algorithms, current methods of network visualization (Lewis et al., 2012; Patil, 2011) are effective in discovering kinship patterns (e.g., clusters of connected members, or commonalities between friends who share

interests and groups in a social networking application) in social networks. However, such methods are inadequate in discovering complex patterns that exist in network space, multivariate space, geographic space and time.

Dynamic network visualization methods (Federico, Aigner, Miksch, Windhager, & Zenk, 2011; Moody, McFarland, & Bender-deMoll, 2005; Shi, Wang, & Wen, 2011) allow discovery of such complex patterns in network over time using animation (network movies) (Moody et al., 2005) and “small multiple displays” (Robertson, Fernandez, Fisher, Lee, & Stasko, 2008). However, dynamic network visualization methods also disregard the geographic dimension of such networks. Alternatively, the layout of a dynamic network could be determined by the diffusion of the phenomena in geographic space rather than a graph drawing algorithm (Demoll & McFarland, 2005) that aims to enhance the perception of changes by considering additional criteria such as minimizing edge crossings and ensuring repeatability and stability. However, visual display becomes even more cluttered with edges and nodes overlaid on top of each other as a result of the tendency (spatial auto-correlation) of the phenomena to cluster around certain locations (e.g., cities attract a large number of individuals).

Interactive and integrated approaches (Guo, 2009; Luo, MacEachren, Yin, & Hardisty, 2011) which allows discovery of complex patterns by dynamically linking multiple views of network space, geographic space and multivariate space, have great potential in tackling the problem of discovering complex patterns. However, incorporating time into such integrated approaches is a significant challenge for usability and comprehension of patterns over multiple views.

Data and the Case Study

To demonstrate the approach, we use family tree data derived from the published genealogies of Chaffee family which only include males. The genealogy of the family was selected among 8 other genealogies on the basis of better temporal resolution and the information on migration. The data include 1225 individuals and approximately 2387 geo-coded moves over a three hundred year period and 856 distinct locations where the family members lived. The animation of the family flows and the relationships for Chaffee (CFE) family is available at the link (Koylu, 2012a).

Methodology

We introduce a new spatial density mapping approach for visualizing social relationships across space and time. Conventional spatial density mapping methods, calculate a density of features in a neighborhood. Instead, we smooth edges of a kin relationship graph and then derive a kin concentration surface using the smoothed graph. We also introduce a new kin proximity index that takes into account of space and time in quantifying social relationships. The new kin proximity index is a product of the kinship between individuals (e.g., the strength of the connection or the closeness of the relation between two individuals), the duration of time that the two individuals spend together and how close they were geographically within a region. Our approach consists of several steps. First, partition the spatial extent of the dataset into a fine resolution grid and set the

center of each cell as an estimation point (region center). For the case study, we used a cell size of 14km x 14km to distinguish towns from each other. Second, for each cell in the dataset, identify the spatial neighborhood based on a geographic distance threshold (fixed-bandwidth). The result of the second step is a collection of circular overlapping regions where each region is centered on a grid cell. Third, partition the data into 20 year time windows (i.e., rectangular windows) with a 50% (10 year) time-window overlap. Fourth, construct a kin relationship graph for all regions in all time intervals. A kin relationship graph is a weighted undirected multi-graph where a node represents an individual and a link represents a space-time interval of a kin relationship between two individuals at a unique pair of locations over a period of time. Fifth, we smooth each of the kin relationship graphs that belong to a space (region) and time (window) interval. Sixth, calculate a kin concentration value for each grid cell using the corresponding smoothed relationship graph. The result of the sixth step is a collection of kin concentration surfaces over the predefined time intervals. In the following subsections, we introduce each of the steps in detail.

Bandwidth selection and distance weighting

Bandwidth and kernel function are two critical parameters for spatial kernel smoothing. The bandwidth determines the radius (e.g., the extent of the neighborhood) or the number of neighbors that is considered to have an effect on the estimation point while the kernel function determines how each observation will be weighted within the kernel. A variety of adaptive bandwidth approaches are commonly used to balance the number of observations for each bandwidth and limit the bias caused by differences in size and density. However, an adaptive bandwidth might result in greater distances over areas with lower density of observations and the choice must be evaluated based on its applicability to the data and analysis. Although it usually causes biased estimations when there is a great deal of spatial heterogeneity, a fixed-bandwidth is more appropriate to observe the change of an attribute (i.e., kin proximity) over space and time. However, the decision to select a fixed-bandwidth is quite challenging and should be backed up by domain knowledge.

We argue that the spatial proximity of family members affects their ability to interact and certain kinds of help and support can only be provided in person. Therefore, a measure of kin proximity must incorporate spatial proximity and limit the bandwidth so that kin relationships over great distances are disregarded. Following the idea that the spatial distribution of relatives is greatly affected by migration, we evaluated the histogram of moves by distance travelled. The histogram highlights that a great portion of moves happened within 60 km of distance, and we observe a gradual decline in the number of moves after 60 km. Therefore, we chose 60 km as a fixed-threshold distance throughout the entire dataset.

In order to select a kernel function, we experimented on the most commonly used kernel functions which are the uniform kernel, the Gaussian kernel and triangular kernel. Previous research and our experiments indicate that the choice of models do not have a significant effect on the performance of the smoothing (Bors & Nasios, 2009; Silverman, 1986). We chose the Gaussian kernel function represented in Figure 1 below.

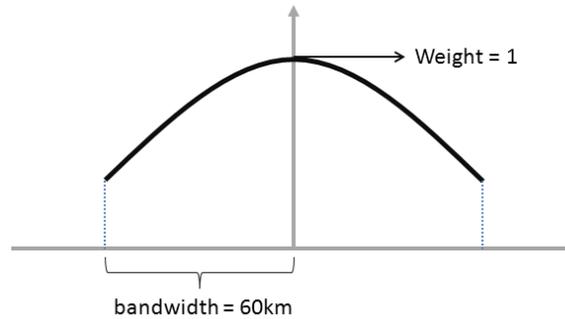


Figure 1: The Gaussian kernel function: $W_{ri} = \exp(- (d_{ri}/B_r)^2)$ if $d_{ri} \leq B_r$, else 0. B_r is the bandwidth and d_{ri} is the distance between location r (center of the kernel) and location i .

Generating time windows

A dynamic social network which expands over time is usually high dimensional due to the large number of time units. For example, the family tree dataset we use to demonstrate our approach includes many generations over 300 years. Thus, a comparison of kin relationships over time necessitates subdividing the data into smaller periods. Because some patterns may fall between two time windows and not appear, the decision to determine the size of a period (window) is critical and must be supported by domain knowledge. In our case dataset, on average a man is 35 years old when a son is born and 20 years is nearly the smallest generation i.e., the youngest a man might be when he has a son. Also, a period of 20 years kind of divides the life course into meaningful stages: age 1-20 would be before marriage, child bearing should stop by age 60 etc. So people in different 20 year windows should be in different life stages. Therefore, we partition the data into a time window (interval) length of 20 years with a rectangular window function which considers observations equally within an interval while disregarding the ones outside the interval. To mitigate the potential loss at the edges and obtain a smoother transition between the surface results of different time windows, we apply 50% (10 years) window overlap.

Constructing Kin Relationship Graphs

Given regions and time windows, we construct a kin relationship graph for each region in each time window based upon the kin proximity between individuals. Kin proximity could be measured in a variety of ways for dynamic social networks that evolve over time. Since the spatial proximity of family members affects their ability to interact and to assist each other in various sorts of crisis, we hypothesize that kin proximity is a product of the closeness of their kinship (e.g., how those individuals are connected within their family tree), the duration of time (e.g., number of years) that the two individuals spend together within a region (neighborhood) and how close they were geographically within a region. Based on our conceptualization of kin proximity, a kin relationship graph is a weighted undirected multi-graph where a node represents an individual and a link (edge) represents a space-time interval of a kin relationship between two individuals at a unique pair of locations over a period of time. For example, two individuals e.g., a father and a son live in the same location for a period of time then the son moves out to another

location 40 km away. As a result of the son's relocation, there are two space-time intervals of the father-son relationship; the first space-time interval is the period of time where they lived at the same location while the second one is the period where the son lived 40km away. A kin relationship graph is a multi-graph because multiple edges (space-time intervals of a kin relationship) exist between two nodes as a result of geographic relocation of either or both of the nodes (individuals) over time.

To construct a kin relationship graph, we first identify all space-time intervals that the two individuals lived in different locations within the same region. Then we use the equation to calculate the kin proximity measure between individuals i and j for the space-time interval s , within region r in time window t .

$$KinProximity_{ij}^{rts} = kinScore_{ij} * s_t_Weight_{ij}^s$$

where $kinScore_{ij}$ is the kin relationship score between individual i and j , $s_t_Weight_{ij}^{rt}$ is a composite weight of space and time which is calculated separately for each space-time interval (s). If the two people never move within the region then there is only one space-time interval. The components of the kin proximity index are explained in the following sections.

Calculating Kin Score

Kin score is the core component for a kin proximity measure and it is used to differentiate the intensity of kin (e.g., family, social) relationships between individuals. In our case study, we construct the kinship score based on how a pair of individuals are connected within the family tree. A variety of measures could be used to obtain a degree of a relation in a family. Consanguinity which counts the number of steps to a common ancestor is a common measure used in law and genetics. For example, consanguinity measure classifies the parent-child relation as the first degree, while relations between siblings or the grandparent-grandchildren relation as the second degree. However, within the context of family migration we argue that using consanguinity might underestimate the importance of siblings as opposed to parents. Therefore, we count them equally rather than counting the number of steps to reach a common ancestor.

According to our definition, the first degree consists of the closest relatives who are parent, child and full siblings. The second degree relatives are people who trace back to the same grandparents, e.g., grandparents, grandchildren, uncles, nephews and first cousins. Then the third degree relatives are those that can trace back to great grandparents and great grandchildren, great uncles, second cousins and etc. Following the same logic, higher degrees were obtained for more distant relatives. After determining the kinship degree between a pair of individuals we assign a score for each relationship by taking the inverse of the degree. For example, the score for a first degree relationship (e.g., siblings, father-son) is $1/1 = 1$, whereas the score for a second degree relationship (e.g., first cousins) is $1/2 = 0.5$, and a third degree relationship (e.g., great grandparent-great grandchildren) is $1/3 = 0.33$. Thus, the score is inversely proportional to the degree of kinship.

Calculating Space-Time Weight

The second component of the measure is a composite weight of space and time which is used to amplify the intensity of a relationship. Given an incidence $s \in \{window\ t\ and\ region\ r\}$, space-time weight for space-time interval s of the relationship between individuals i and j is calculated using the equation below.

$$s_t_Weight_{ij}^s = timeWeight_{ij}^s * edgeDistWeight_{ij}^s$$

where $timeWeight_{ij}^s$ is the total number of years that individual i and j in space-time interval s , and $edgeDistWeight_{ij}^s$ is the distance weight based on the geographic distance between the two individuals in space-time interval s .

We illustrate the calculation of this kin proximity measure using a simple family network in a single region over a single time window. Figure 2 illustrates the temporal dimension of the sample network. The horizontal bars in Figure 2 show the time periods that individuals (i.e., A, AA, AB, AC and AAA) spent within the region. For example, individual AC lived in the region between 1674 and 1700 whereas individual AB lived in the region for two time periods 1672-1685 and 1692-1700. Additionally, the two solid vertical lines represent the beginning (1675) and end (1695) of a time window. Notice that within the specified window, individual AB was in the region in two periods: 1675-1685 (10 years) and 1692-1695 (3 years).

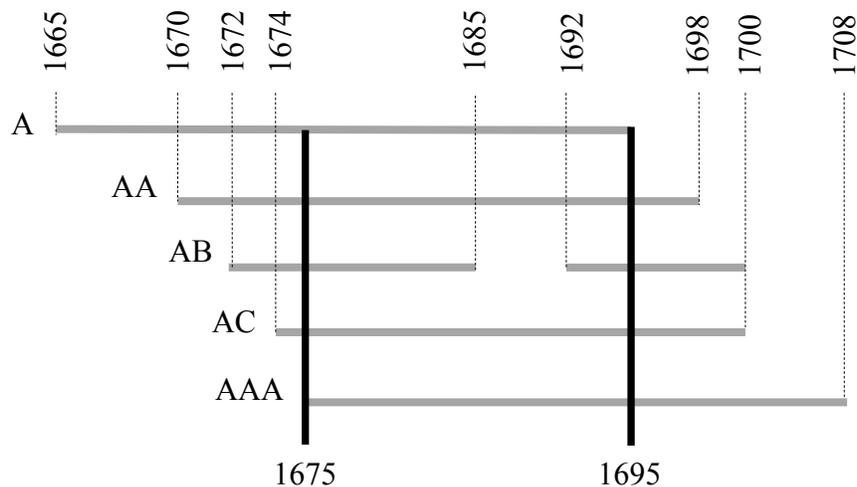


Figure 2: Temporal dimension of a sample family network in a region. The horizontal bars in the figure show the time periods that individuals (i.e., A, AA, AB, AC and AAA) spent within the region. Vertical dashed lines mark beginning and end points of each individual's space-time intervals within the region, whereas vertical solid lines mark the beginning and end of a time window.

The layout of a social (family) network is dynamic as a result of relocation of individuals, addition (birth) of new individuals and removal (death) of the existent ones. To simplify the illustration, we assume that the individuals in the sample network did not move except individual AB who left the region between 1685 and 1695 and moved back to a

different location within the same region. Given a time window 1675-1695, Figure 3 illustrates the a space-time series snapshot of a stationary period of the kin relationship graph with multiple edges partitioned into a sequence of graph layouts in which there is only one edge between individuals and nodes are placed according to their spatial coordinates.

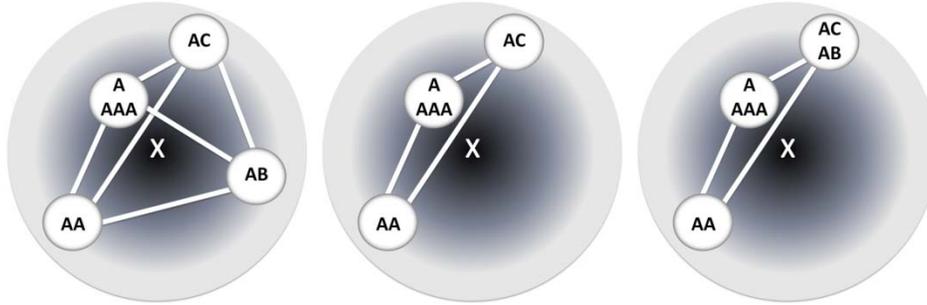


Figure 3: Illustration of a sequence of layouts within the time interval 1675-1695. Each layout is a space-time series snapshot of a stationary period of the kin relationship graph. The period for the left: 1675-1685, the period for the middle: 1685-1694 and the period for the right layout: 1692-1695.

Given the three spatial layouts in Figure 3, we demonstrate the calculation of the space-time weight and kin proximity measure for the relationship between AB and AC. The first step is to find out the space-time intervals where and when the two individuals lived within the region. There are two space-time intervals for AB-AC relationship. The first space-time interval is between 1675 and 1685 where they lived in different locations in the region and the second space-time interval is between 1692 and 1695 where they lived at the same location. Time weight for the first space-time interval (s_1) is 10 while the time weight for the second space-time interval (s_2) is 3. We calculate the edge distance weight between the two individuals by using the Gaussian kernel function and the bandwidth (the radius of the region). Notice that some of the distances between the individuals within the region such as the distance between AA and AC could be above the bandwidth. To account for such relationships, we do not apply the cut-off parameter where we assign 0 weigh for the distances farther than the bandwidth. Instead, we derive the edge distance weight using the Gaussian function which can provide weights beyond the bandwidth. Below equations illustrate the distance weights for the two space-time intervals s_1 and s_2 . While the distance between AB and AC is 60 km in s_1 (notice the distance is equal to the bandwidth), the distance between the individuals in s_2 is 0 since they lived in the same place.

$$edgeDistWeight_{AB,AC}^{s_1} = \exp(-(D_{AB,AC}^{s_1}/B_r)^2) = \exp(-(60/60)^2) = 0.36$$

$$edgeDistWeight_{AB,AC}^{s_2} = \exp(-(D_{AB,AC}^{s_2}/B_r)^2) = \exp(-(0/60)^2) = 1$$

$$s_t_Weight_{AB,AC}^{s_1} = 10 * 0.36 = 3.6$$

$$s_t_Weight_{AB,AC}^{s_2} = 3 * 1 = 3$$

After deriving the space-time weight, we calculate the kin score which forms the core component of the measure. The kin score between AB and AC is 1 since the individuals have a first degree relationship (siblings). Given the kin score and space-time weights, we now calculate the kin proximity measure for the space-time intervals s_1 and s_2 of AB-AC relationship by plugging them into the equations below:

$$KinProximity_{AB,AC}^{r(1675-1695)s_1} = kinScore_{AB,AC} * s_t_Weight_{AB,AC}^{s_1} = 1 * 3.6 = 3.6$$

$$KinProximity_{AB,AC}^{r(1675-1695)s_2} = kinScore_{AB,AC} * s_t_Weight_{AB,AC}^{s_2} = 1 * 3 = 3$$

Smoothing Kin Relationship Graphs

We smooth edges of a kin relationship graph by fitting the Gaussian kernel function over the region center and by weighting the value of each edge based on the average (geographic) distance of its nodes (individuals) to the region center. Equation below illustrates the smoothing process.

$$SmoothedKinProximity_{ij}^{rts} = KinProximity_{ij}^{rts} * distToCenterWeight_{ij}^s$$

where $KinProximity_{ij}^{rts}$ is the kin proximity measure between individuals i and j for the space-time interval s within region r in time window t , $distToCenterWeight_{ij}^s$ is the distance weight based on the average distance of individuals i and j to the region center. To illustrate, given the distances $AB_To_Center_{(s_1)} = 40\text{km}$, $AC_To_Center_{(s_1)} = 50\text{km}$, $AB_To_Center_{(s_2)} = 50\text{km}$, $AC_To_Center_{(s_2)} = 50\text{km}$, we smooth the kin proximity measure for the incidences s_1 and s_2 of AB-AC relationship below.

$$distToCenterWeight_{AB,AC}^{s_1} = \exp(-(45/60)^2) = 0.57$$

$$distToCenterWeight_{AB,AC}^{s_2} = \exp(-(50/60)^2) = 0.50$$

$$SmoothedKinProximity_{AB,AC}^{rs_1} = 3.6 * 0.57 = 2.05$$

$$SmoothedKinProximity_{AB,AC}^{rs_2} = 3 * 0.50 = 1.5$$

Calculating Kin Concentrations

Our goal is to discover the change of kin concentrations over space and time. Thus, we calculate a kin concentration value using the equation below.

$$KinConcentration_c^{rt} = \frac{\sum_{s=1}^{S^{rts}} SmoothedKinProximity_{ij}^{rts}}{N^{rt}}$$

where $KinConcentration_c^{rt}$ is a grid cell c in region r in time window t ; $SmoothedKinProximity_{ij}^{rts}$ is an edge of the smoothed relationship graph; N^{rt} is the number of individuals lived within region r and time window t . Going through each space-time interval s in region r in time window t , the smoothed kin proximity value for

each space-time interval s is summed and then normalized by the number of individuals lived within region r in time window t . Normalization decreases the effect of differences in numbers of individuals between regions and also time periods.

Results

The result of the spatial density mapping approach on the family network consists of 29 surfaces of kin concentrations over time. In order to enable comparison between surfaces over time, we select a constant sequential classification scheme in which darker colors indicate higher kin concentration. In order to identify the relationship between the population of individuals and the kin concentration over space, we also map the total number of individuals at distinct locations by using proportional hollow point symbols overlaid on top of each concentration surface. The larger the size of a circle, the more people live in that location. We provide an interactive animation of the kin concentration surfaces on a web-based application (Koylu, 2012b) where users can compare the kin concentrations between any pair of time windows. Due to the limited space, we only illustrate and report the results of 8 time windows which highlight the events in the history such as the first settlements and the American Revolution.

The Earliest Family Hub (1714-1734, 1724-1744)

Figure 4 illustrates the kin concentrations over time windows of 1714-1734 and 1724-1744. The figure highlights the first appearance of a hot spot in the data. For the first 80 years, this family did not have any hubs. Within the earlier window (1714-1734) most of the individuals stayed around of Rehoboth, while some individuals started moving out to Woodstock. As a result of the birth of new family members and the migration to Woodstock and Ashford, the family hub shifted from the east to the west while leaving some family members in the east.

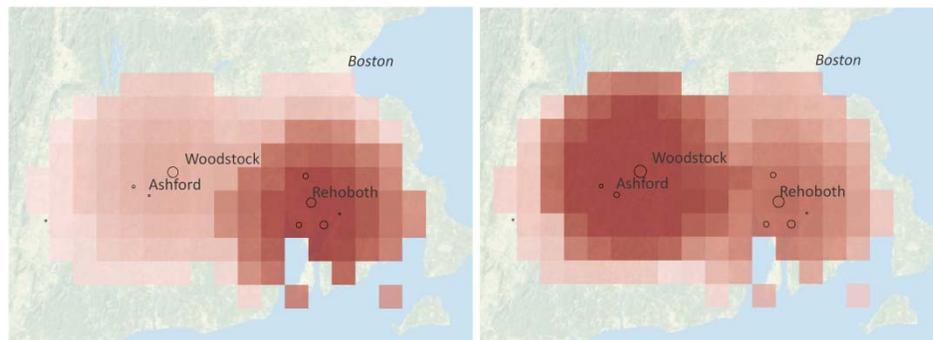


Figure 4: Kin concentration surfaces for two time intervals: Left: 1714-1734, Right: 1724-1744

Expansion after the American Revolution (1764-1784, 1794-1814)

Figure 5 illustrates the kin concentrations over time windows of 1716-1784 and 1794-1814. The comparison of the two time windows highlights the early stages of expansion after the American Revolution which opened up new lands in New York and Northern New England. But at this point in the earlier interval, as earlier, there are no areas of very high kin proximity in the new areas. It is too early to see them. On the other hand, 1794-

1814 highlights the first hubs in areas settled after the Revolution. Although individuals started moving to Vermont in the time period of the earlier interval, it took 30 to 40 years to develop hubs after people moved there. So the same process we saw when the family settled its hearth in southern Massachusetts and Connecticut was repeated as settlement moved North and West after the Revolution. The time window of 1794-1814 also highlights Berkshire as a strong family hub however, the number of individuals within the region around Berkshire is much less than the number of individuals around less stronger hubs around Chittenden and Westminster.

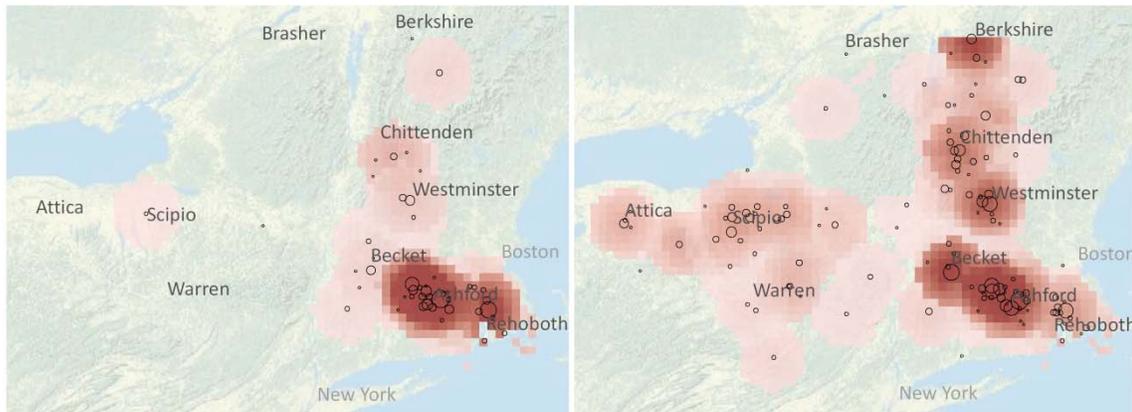


Figure 5: Kin concentration surfaces for two time intervals: Left: 1764-1784, Right: 1794-1814

Beginning settlement of the Middle West (1824-1844, 1834-1854)

Figure 6 illustrates the kin concentrations over time windows of 1824-1844 and 1834-1854. When we compare the two time windows, we observe that areas are settled in the Middle West in the earlier time window however no family hubs exist. However, in the second time window, we see the family hubs in the Middle West a generation after first settlement by this family. But what is so interesting is that the old hubs do not disappear. In fact they have become even stronger. This is due to some descendents of the earliest settlers never having left and, of course, the pattern we see here is a testament to the very high fertility and low mortality rates of this population. Only a few of the older hubs are fading: Westminster and Chittenden.

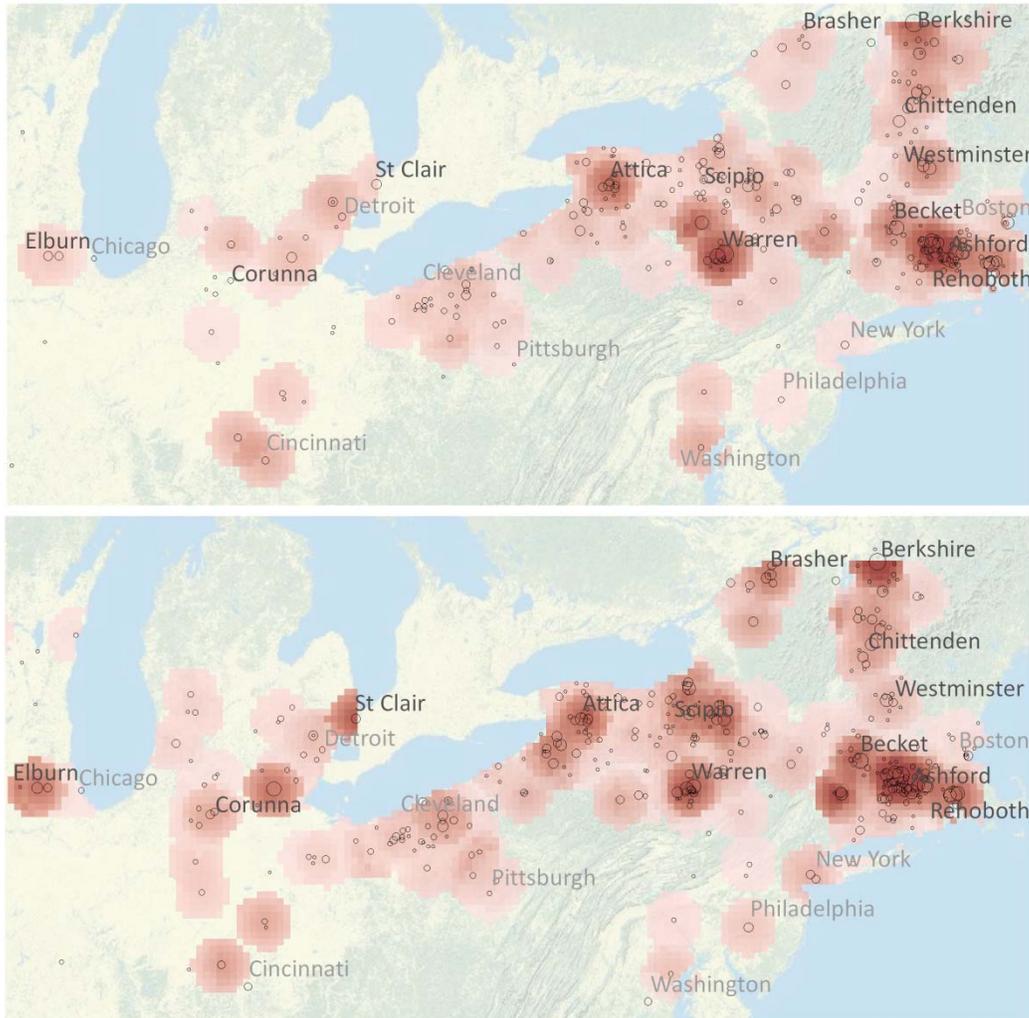


Figure 6: Kin concentration surfaces for two time intervals: Up: 1824-1844, Down: 1834-1854

Beginning Urbanization (1874-1894, 1884-1904)

Figure 7 illustrates the kin concentrations over time windows of 1874-1894 and 1884-1904. In the earlier time window, there are small outposts in many major cities in the East. But these are not yet family hubs. Hubs might not develop due to lower fertility and higher mortality rates. The original family hub was just North of Providence and perhaps it is being now fed by new migrants to that city. But even at this late date, the older rural hubs have succeeded in holding onto people. However, for the later time window, hubs disappear. The reason for the disappearance of hubs here is probably that the sample was designed to end with the cohort born in 1860 although we also included all the siblings of men born prior to 1840, so there were some later born men included. Still, by this point the effect of our cut off is seen. However, at the same time the fertility rates of this population were falling and it will be important to see how that might have contributed to the fading of hubs.

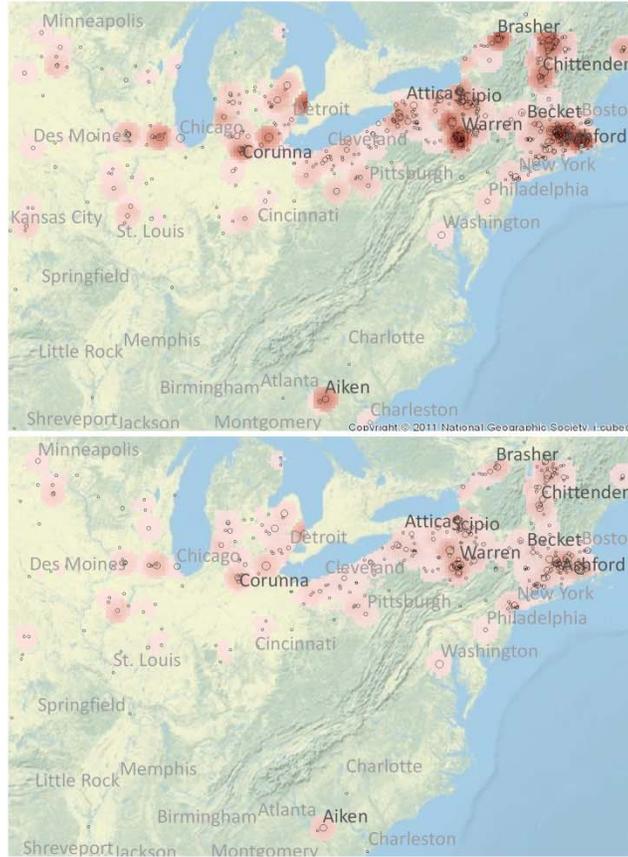


Figure 7: Kin concentration surfaces for two time intervals: Up: 1874-1894, Down: 1884-1904

Conclusions

We introduced a new spatial density mapping approach for visualizing social relationships across space and time. Conventional spatial density mapping methods, calculate a density of features in a neighborhood (bandwidth) by fitting a smoothly curved surface over each estimation location. Instead, we smooth edges of a kin relationship graph and then derive a kin concentration surface using the smoothed graph. We also introduce a new kin proximity index that takes into account of space and time in quantifying social relationships. Our approach consists of two major steps. First, we partition the data into time windows and regions, and then we build a kin relationship graph in which the weight of each edge is determined by the kin proximity value between the two nodes (individuals). Second, we smooth each of the kin relationship graphs that belong to a space (region) and time (window) interval; and calculate a kin concentration value for each location using the smoothed graph. The result of our approach is a collection of kin concentration surfaces over a sequence of time windows. To demonstrate the approach, we use a family network derived from the published genealogies of a family from the U.S. North over a span of three hundred years. The data also include information on migration of individuals. The results highlight that family hubs with high concentration of kin proximity is not necessarily correlated with the high density of people and tend to relocate between regions as a result of processes such as chain migration, return migration and fluctuations in fertility and mortality.

Reference

- Bors, Adrian, & Nasios, Nikolaos. (2009). Bayesian Estimation of Kernel Bandwidth for Nonparametric Modelling. In C. Alippi, M. Polycarpou, C. Panayiotou & G. Ellinas (Eds.), *Artificial Neural Networks – ICANN 2009* (Vol. 5769, pp. 245-254): Springer Berlin / Heidelberg.
- Demoll, Bender, & McFarland, D. (2005). The Art and Science of Dynamic Network Visualization. *JoSS: Journal of Social Structure, Volume 7*. doi: citeulike-article-id:3478454
- Federico, Paolo, Aigner, Wolfgang, Miksch, Silvia, Windhager, Florian, & Zenk, Lukas. (2011). *A visual analytics approach to dynamic social networks*. Paper presented at the Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies, Graz, Austria.
- Groh, Georg, Hanstein, Holger, & Wörndl, Wolfgang. (2009). *Interactively Visualizing Dynamic Social Networks with DySoN*. Paper presented at the Computer-Human Interaction (CHI), Boston.
- Guo, D. (2009). Flow Mapping and Multivariate Visualization of Large Spatial Interaction Data (Vol. 15, pp. 1041-1048). *IEEE Transactions on Visualization and Computer Graphics*: IEEE.
- Koylu, C. (2012a). Animation of Family Relations and Migration. Retrieved 5/5, 2012, from <http://129.252.37.169:8400/flowvis/familyflows/index.html>
- Koylu, C. (2012b). Kin Concentration Surfaces over Time. Retrieved 8/15, 2012, from <http://129.252.37.169:8400/flowvis/kinconcentration/index.html>
- Lewis, Kevin, Gonzalez, Marco, & Kaufman, Jason. (2012). Social selection and peer influence in an online social network. *Proceedings of the National Academy of Sciences, 109*(1), 68-72. doi: 10.1073/pnas.1109739109
- Luo, Wei, MacEachren, Alan M., Yin, Peifeng, & Hardisty, Frank. (2011). *Spatial-social network visualization for exploratory data analysis*. Paper presented at the Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Location-Based Social Networks, Chicago, Illinois.
- Misue, Kazuo, Eades, Peter, Lai, Wei, & Sugiyama, Kozo. (1995). Layout Adjustment and the Mental Map. *Journal of Visual Languages & Computing, 6*(2), 183-210. doi: citeulike-article-id:161303
- Moody, J., McFarland, D., & Bender-deMoll, S. (2005). Dynamic network visualization. *American Journal of Sociology, 110*(4), 1206-1241.
- Patil, D.J. (Producer). (2011, 08/19/2012 11:00 pm). Visualize your LinkedIn network with InMaps. Retrieved from http://www.youtube.com/watch?v=PC99Nw2JX8w&feature=player_embedded
- Robertson, George, Fernandez, Roland, Fisher, Danyel, Lee, Bongshin, & Stasko, John. (2008). Effectiveness of Animation in Trend Visualization. *IEEE Transactions on Visualization and Computer Graphics, 14*(6), 1325-1332. doi: 10.1109/tvcg.2008.125
- Shi, Lei, Wang, Chen, & Wen, Zhen. (2011). *Dynamic network visualization in 1.5D*. Paper presented at the Proceedings of the 2011 IEEE Pacific Visualization Symposium.
- Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. London ; New York: Chapman and Hall.

Caglar Koylu, Ph.D., Candidate, Email koylu@mailbox.sc.edu

Diansheng Guo, Associate Professor, Email guod@sc.edu

Alice Kasakoff, Ph.D., Department of Geography, Email Kasakoff@sc.edu

Department of Geography, University of South Carolina, Columbia, SC 29208.