Network boundaries version 2

Network Boundaries and System-ness

Robert A. Hanneman

Hiroko Inoue

Christopher Chase-Dunn

Department of Sociology  
University of California, Riverside

# Abstract

Systems of human settlements are conceptualized as multi-level network of population concentrations linked by economic, cultural, and political relations. Regions are defined as areas in the network where ties are denser than expected in a random graph. The “modularity” index is proposed as a criterion for identifying the number of regions and their membership – and hence for drawing boundaries. Modularity, as well as other network metrics (density, path length, and centralization) provide indexes of the degree of “system-ness” of the settlement system. Two examples (African and Middle-eastern settlements during the Roman Empire; Central European large urban places c. 1500) illustrate the approach. The examples treat each level (relation) of multi-level networks separately, and treat them jointly. The discussion of the results addresses the relevance of the current approach to “place-centric” alternative approaches. It also discusses the relevance of “multi-layer” or embedded network approaches, and the possible applicability of “small world” network topologies in understanding settlement networks.

Paper prepared for presentation at the *Institute for World-Systems Research and International Studies Association Workshop on Systemic Boundaries*. Riverside, California, March 5th, 2015.

# Social network s and the boundary/system problem

The size and spatial distributions of the global human population display continuous change. These changes are both causes and consequences in changes in the social structure of the global social system (i.e. the patterns of social relations among the members of the population). Our focus here is on how one might describe the “texture” of this system of social relations, with a particular attention to the problem of identifying the extent to which the global social system is closely or loosely coupled (cite close coupling) and identifying the regions of the system.

Social network analysis (SNA) provides one way of thinking about the “system-ness” of social structures. It has well developed tools and concepts for identifying the boundaries of sub-structures or regions in social systems, and for describing the degree of close-coupling among these regions. In the discussion below, we will outline how some key concepts from social network analysis can be applied to networks of human settlements.

We begin with a discussion of how the social structure of the global human population could be approached as a network of relations. Then we examine some network analysis metrics/algorithms that might be applied to identify “system-ness” (or cohesion), and to locate the boundaries of regions in the larger network (sub-structures or communities). We follow this conceptual discussion with two examples: trade and cultural relations among some places in the Roman Empire (c. 150CE) and trade, cultural, and political relations among some central European places (c. 1500CE).

# The global settlement system as a network

A network is formally defined as a set of nodes (vertices) and relations connecting those nodes (arcs if the relation is asymmetric or directed, edges if the relation is symmetric or un-directed). For our purposes, we take the nodes to be population concentrations or settlements. Nodes may be characterized by attributes (or “colors” or “scores on variables,” such as their size and geo-location). We consider that these settlements may be connected by multiple types of relations simultaneously (e.g. trade flows, cultural diffusion, affiliation with the same political community).

In network analysis, attention focuses on the pattern of relations among nodes, rather than the attributes of nodes and their distributions. The attributes of nodes are treated either as causes or consequences (or both) of relations among nodes. The larger the size of a settlement, for example, might be treated as a cause of an increase in the probability that it trades with larger numbers of other settlements. And, if a settlement is connected with many other settlements in a trade relation, this might be hypothesized to be a cause of its relatively large size.

A network of settlements is described by the collection of dyadic ties, or relations, between all pairs of settlements. Macro-analyses have identified many different types of relations as contributing to the structure of the global system. So, the network analyst would say that the global system of settlements is “multi-plex” or a “multi-level” network in which all dyads are connected by multiple relations. For example, two settlements may trade in luxury goods, diffuse technical knowledge, and be affiliated with the same political community.

At the highest level of abstraction, the types of ties or relations among entities might be classified as “conserved” flows (usually of material things) from one node to another; “non-conserved” sharing (usually of information things) of two nodes; and “affiliation” (or embedding) of two nodes within an qualitatively different type of entity (mode).

Conserved flows connect two settlements by the movement of some quantity from one to the other (the flows may go in both directions, but are the flow AB may differ from the flow BA). The quantity flows from one place to another, and can only be in one place at a time. The movement of material goods (including books and money) and persons are examples. Relations of this type are represented in graphs as single-headed arrows from the source to the destination node.

Non-conserved relations connect two settlements by a commonality, bonding, or sharing of some quantity. Informational or cultural quantities that connect two settlements are often treated in this way. For example, one community may diffuse the practice of Buddhism to another. What differs here from conserved quantities is that when Buddhism diffuses from one settlement to another, it does not disappear from the first (i.e. it is not “conserved”). Such relations are represented in graphs by undirected line segments (edges) connecting the members of dyads.

Ties of affiliation (or embedding, or two-mode relations) also create non-directed ties between pairs of settlement – but differ, at least conceptually, from non-conserved relations. Two settlements might be connected by being affiliated with (or members of) a higher-level entity of a different type. In the current case, we might say that two settlements have a relation if they are both under the political control of the same political community (e.g. “state”). Such relations are represented as “bi-partite” or “2-mode” graphs, with edges connecting entities of one type (e.g. settlement) to entities of another (e.g. states), but without direct ties among entities of each type.

Individual human settlements are frequently embedded in (i.e. affiliated with) “states,” “empires,” or other superordinate political units. In cases like these, the political relations defining regions may lie between the embedding units, not the individual settlements. For example, wars in the modern system are frequently actions of nation states, not individual settlements. These type of effects are readily incorporated into network approaches by treating all units embedded in one higher-level entity as having a relation with all units embedded in the other. For example, a war between Germany and France would be represented as a relationship of conflict between each German settlement and each French settlement. An example of this type of multi-layer or embedded relationship is seen in our second example – where the Holy Roman Empire partially embeds subordinate political affiliations.

The full representation of a settlement system as a network then would consist of a number of arrays of data. The largest part would be a set (or stack) of settlement-by-settlement matrices, each containing data on one conserved or non-conserved relation. Additional arrays to describe bi-partite or embedding relations would be rectangular, each displaying the affiliation of settlements with the embedding entities. In most analyses, these two-mode arrays would be converted into one-dimensional settlement-by-settlement arrays (containing the number of common embeddings for each pair of settlements. In addition, a rectangular, settlement-by-variable, matrix would contain the attribute data. For some purposes, an array of the attributes of embedding entities (e.g. states) might also be used; and, in some cases, arrays of relations between embedding entities (e.g. state-to-state relations) might also be recorded.

In network analysis, the relations to be examined may be either directed or undirected, and both types can readily be included in the same analyses. And, relations may be either binary (i.e. there is, or isn’t a tie of a given type between two settlements), or valued (i.e. the strength, not just presence of a relation is measured). Often the data on historical settlement systems is inherently binary. Even when a relation is not inherently binary, we may not have enough information available to note differences in strength. Commonly, valued relations are binarized at one (or more) cut-off points in most network algorithms (like those shown in our examples). There is, however, no difficulty in principle in utilizing both binary and valued relations.

Once a set of places and their attributes and relations are represented as a network, a number of concepts and algorithms may be applied to identify regions and their boundaries, and to index the “system-nesss” of the network.

# Network approaches to the problem of boundedness: Sub-structures in graphs

A “region” of a system of human settlements might be defined as a set of connected places, such that the strength or density of ties among the places within the region is greater than the strength or density of ties of places within to places without the region. The “boundedness” of a region is probably best thought of as a continuous quantity; the greater the ratio of the strength/density of ties within to the ties without the region, the greater the boundedness of the region.

This definition of a “region” has some important implications. The whole network of the global settlement system could be composed of one, or many regions. These regions could vary in size, and, more importantly, their degrees of boundedness. If we consider a multi-relational network, the density/strength of ties that define one region might differ qualitatively from those that define the boundedness of another. One region might, for example, be primarily defined by tight political integration while another region might be defined by strong and dense trading ties.

In networks formed by a single kind of ties among entities, the regional boundary and system-ness problems are relatively straight-forward. Human settlements, however, are not connected by a single kind of relation.

## Multiple relations

An important issue in the application of SNA ideas to the problem of boundedness in the global settlement system is how to deal with the multi-level, or multi-relational nature of ties connecting settlements. This is not, primarily, a methodological question. Rather, it is conceptual.

One general approach is to analyze each relation separately. That is, the global settlement system might be seen as composed of a set of regions based on trade relations, an alternative set of regions based on information flows or cultural influences and a third alternative set of regions based on political relations. This general conceptual approach seems broadly consistent with the thinking of scholars like Michael Mann (1986) or the Charles Tilly et al. (1975). It allows that social systems may have very different “textures” of cohesion based on different types of relations. It also allows that boundaries may be “fuzzy” and even contradictory based on the different types of relations. An interesting and important question raised by this separate-relation approach is the degree of overlap among the regions defined by alternative relations.

An alternative general approach is to consider all the forms of relations among settlements simultaneously. There are a variety of ways that one might do this, and each conceptualizes the meaning of “region” and “boundary” is somewhat different ways. One method would be to scale the multiple relations to create a single quantitative index of the strengths of dyadic ties. A second approach could be to characterize the relation between the members of each dyad as having a qualitative type or profile, according to which types of ties predominate. Equivalence analysis could also be applied. Structural equivalence methods would identify settlements as being in the same region if they had the similar patterns of ties to other specific settlements.

In the approach outlined below, we have chosen to pursue the strategy of identifying a single set of regions and boundaries based on considering all relations simultaneously.

## Identifying regions

The general notion of a “region” (a.k.a. cluster, community, dense sub-structure) in a graph is that the nodes within the region are more tightly connected among themselves than they are to nodes without the region. A graph can have a single or many regions, and the regions may contain unequal numbers of nodes. A graph is generally considered to display regions if the average density (or probability of a tie) between pairs of nodes within a “region” is significantly greater than one would expect if the ties in the graph were randomly distributed. The larger the difference between the observed clustering and that of a random graph of the same density may be taken as a measure of the degree to which the graph is clustered, or divided into clearly bordered regions.

Network analysts have developed a considerable number of definitions of what it means for a set of nodes to have relatively more dense connection, i.e. to be considered a “region.” One general approach is “bottom-up” (Hanneman and Riddle, 2005), or agglomerative. This approach follows the logic of how networks grow as their density increases from zero. A variety of algorithms and definitions have been proposed to identify the maximally sized regions in a graph based on aggregation (e.g. cliques, N-cliques, K-cores, etc.).

An alternative general approach to identifying regions in graphs is “top-down,” or divisive. In this class of methods, one begins with an existing graph, and identifies regions by removing nodes or relations that most strongly connect the graph. As one removes nodes or edges that are most “between” other nodes or edges, the graph develops local clusters or regions that have greater relative density within, than without. Top-down methods may be thought of as looking at the “robustness” of the connections in a graph.

In the examples, below, we follow the “divisive” or “top-down” logic of identifying regions in graphs. More specifically, we will identify the regions of settlement systems by locating the partitions of the graphs that maximize “modularity” (Newman, 2006; Wikipedia, 2015). The algorithm used to search graphs for the maximal modularity is that Girvan and Newman (Wikipedia, n.d. accessed, Jan. 2016). We choose the “top down” approach because, for our problem, it is reasonable to consider regions as existing at any one point in time, and to ask the question of how coherent or robust their boundaries are in the face of possible disruption.

Newman’s (2006) modularity approach seeks to find the grouping of nodes into communities (or, here, settlements into regions) such that the proportion of ties among nodes that fall within communities is as different as possible from proportion of ties that would fall within the communities if the distribution of ties were random. A summary measure of the degree of modularity (i.e. departure from randomness) for any proposed set of regions is called “Q.” Q may range from a negative .50 (indicating that ties within communities are less dense than a random distribution) to positive 1.0, indicating that ties are far more dense within communities than would be expected by random distribution. The Q statistic can be calculated for any possible number of communities, and definition of community members. The maximum value of Q across these possibilities identifies the “optimal” number, and membership of communities.

For a graph of any size, of course, there are an extremely large number of possible partitionings. The computational challenge of finding the choice of the “optimal” number of regions and their memberships is most commonly solved by using an algorithm developed by Girvan and Newman. In this method, each edge of the graph is evaluated for its “between-ness” (i.e. the proportion of pairs of other nodes where the focal edge falls on the shortest path between them). Edges that have high between-ness are edges that connect groups of nodes that have few connections other than the focal edge. The Q statistic is then calculated for a given number of communities. The next most between edge is removed, and the process repeated. At some point in this process, the Q statistic reaches a maximum value for a given number of communities. The process is then repeated across alternative hypotheses about the number of communities in the graph (i.e. two, three, up to nodes minus one). Q indexes can be directly compared across solutions with differing numbers of communities, so the optimal number of communities, as well as the membership of communities can be identified using the Girvan-Newman algorithm to maximize modularity.

### “System-ness”

The notion that connected human settlements form a “system” seems obvious, but what does “system-ness” mean? Network theory provides one approach (though hardly the only approach) to defining and indexing the idea. Network theory, here, is very closely connected with the central notions of “complex systems” or “complexity” theory. A network is likely to behave in a coordinated or synchronized way (i.e. to be “systemic”) to the extent that the network is “cohesive.”

A network is composed of nodes (in our case, settlements), and relations among the nodes. Both the nodes and the relations may have “attributes.” For example, settlements may vary in size, technology, and social institutions. Relations may have cultural, material, and political attributes. Complexity theory sees the behavior of the entire network as emerging from both the attributes of nodes and relations and, most importantly, from the ways in which these relations are structured or configured. For example, a “system” that has low density of connections may have a lower tendency toward cascading failures (e.g. epidemics) than a system that is “closely coupled.”

A network might be thought of as displaying greater “system-ness” to the extent that all nodes influence, directly or indirectly, all others. That is, a system is systemic to the degree that its parts display “close” coupling. Structures that are dense, display short path lengths (low between-ness), and little hierarchy or clustering display greater “cohesion” or “system-ness.” Network analysts have developed a considerable number of indexes of these properties of graphs, which can readily be applied to describe the extent to which the settlements in a system are closely-coupled.

Among the most useful index numbers for describing the cohesion or “system-ness” of networks are *density*, *average path length*, *centralization*, and *modularity*.

*Density* is simply the number of observed ties among settlements as a proportion of the maximum possible number of ties (i.e. a hypothetical situation in which every settlement was directly connected to every other one). As density increases, what happens in one settlement is increasingly likely to affect what happens in others.

The *average path length* of a network is the mean of the geodesic (shortest path) distances of all pairs of nodes in the graph. If every settlement is directly tied to every other one, the average path length is 1. To the extent that changes in one settlement affect changes in others only indirectly, path lengths become longer. The greater the average path length, the longer, less robust, and weaker the effect of changes in one settlement on another; the less “systemic” the network becomes.

In most settlement systems, there is a rough central-place hierarchy. That is, some places have larger numbers or stronger relations with other places; some have less. If a system has “central” or more influential nodes, there is a tendency for action to be more coordinated, synchronized, and homogeneous. This occurs because more nodes are receiving influence from the same sources. *Centralized* networks then have the capacity to display more “systemic” behavior.

Most networks display a (variable) tendency toward clustering. That is, a tie between A and C is more likely if there are also ties from A to B and from B to C. In settlement systems, clustering may occur because of geographic, political, cultural, or economic factors that make it more likely that communities will tend to form dense local connections. To the extent that a network displays a strong tendency toward clustering or *modularity*, action in one part of the system takes longer to have system wide effects. Thus, networks that display higher modularity (all else being equal) have less “system-ness.”

Treating settlement systems as multi-relational networks composed of places connected by economic, cultural, and political ties allows the application of standard algorithms to identify regions (communities, clusters, or modules), identify the members of these communities (and hence define the boundaries of regions), and index the extent to which the system is “closely coupled” or is capable of displaying coordinated and emergent macro behavior. In the next two sections we will apply these ideas to a couple data sets to illustrate how the network approach could be applied to human settlement systems.

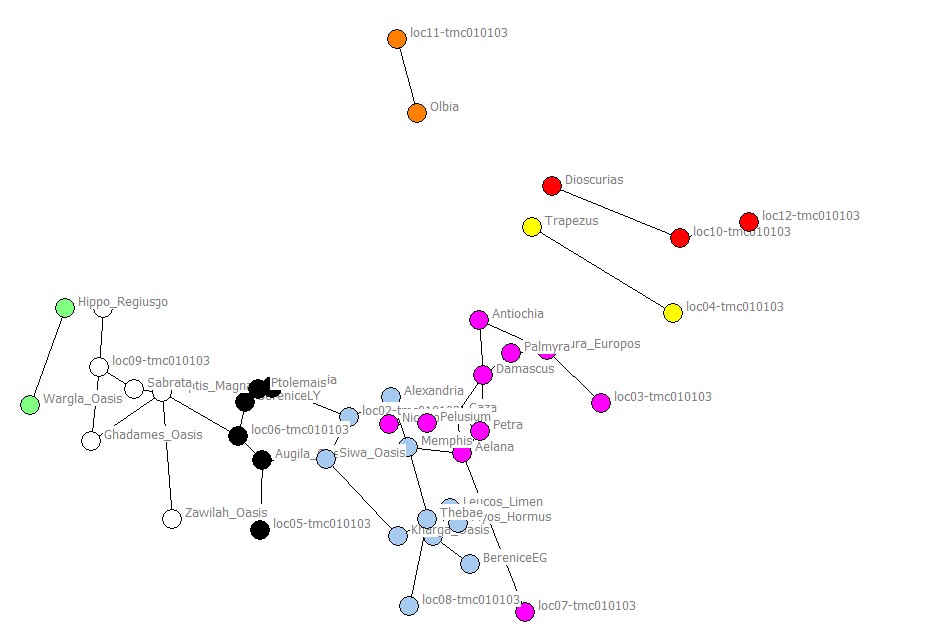
# Single and multi-relational approaches to a portion of the Roman Empire

The data here are from the early Christian era of the Roman Empire (Scullard, 1970). Information is available on trade routes in portions of the Empire (Casson, 1984; Ciolek, 2000). To identify cultural networks among the 44 places identified in the trade networks, the predominant language(s) of each was identified (Stone, 1989). If two places shared a language, a cultural link between them was inferred. The network, then, has 44 nodes and two relations. Let’s examine each relation separately, and then combine both relations using the modularity approach to identifying substructures (and, as a result, boundaries of regions).

## Trade network

Trade is the key economic relation among settlements. In figure 1, the network of known trading relationships among the 44 places is shown. In the diagram (and the others that follow) the nodes have been placed in their geo-locations, and colored according the Girvan-Newman identification of 8 communities as providing an optimal partition into regions.

Figure 1. Trade network among Roman settlements\*

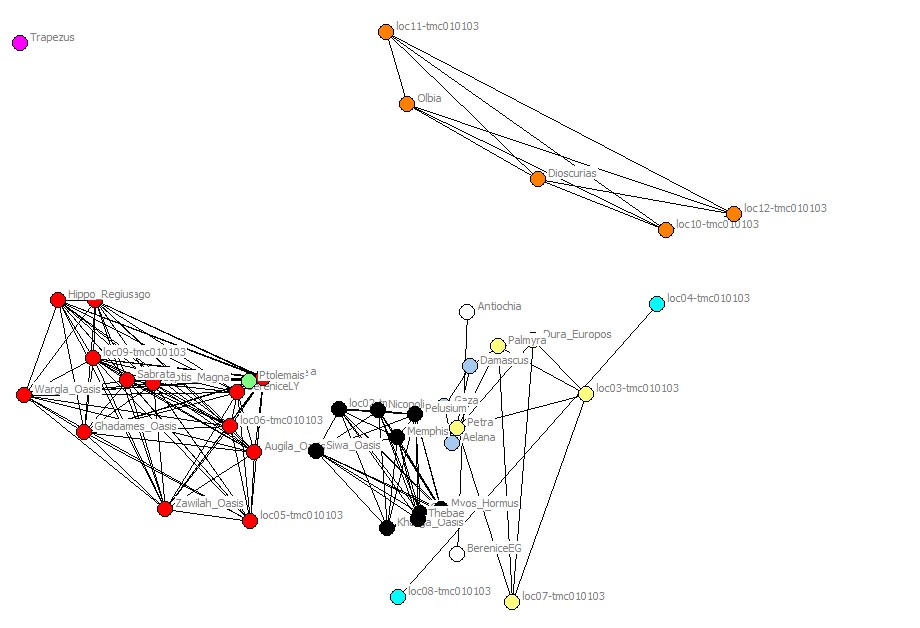
  
\* 8 Girvin-Newman communities, with modularity of .67 are shown by node colors.

The trade network among the 44 settlements has low density (.049), and displays considerable inequality of eigenvector centrality (67%). The average path length is quite great (9.2 steps, when the longest geodesic plus one is assigned to the distance between un-connected places). The Girvan-Newman algorithm has a substantial degree of maximum modularity (Q = .67), with eight communities. Four of these are separate components (dyads and a triad that are not connected to the main cluster). The main cluster is partitioned into four regions, and these regions are (largely) geographically proximate, as one would expect in a trading network.

## Cultural network

Figure 2 displays the cultural network of the communities, where a connection is shown if two communities have a shared language (Vischer, 2011). These data are far from ideal. There are many forms of cultural connection among settlements, not just shared language. The masses of two settlements may speak local dialects, but elites may share a common communication medium across settlements. Even though two settlements may share a language, this does not mean that they, necessarily, directly communicate with one another. So, these data are far from ideal. Again, the communities are shown in their geo-locations, and colored to show the Girvan-Newman communities.

Figure 2. Cultural network among Roman settlements\*

  
\* 8 Girvan-Newman communities, with modularity of .60 are shown by node colors.

Since we have assumed that all communities sharing the same major language communicate with all others directly, the density of the cultural network is higher (.183) than that of the trade network. For the same reason, the degree of Eigen-centralization is lower (31%) than in the trade network. The average path length (3.4) is also shortened by the assumption that all communities sharing the same language are directly connected. The Girvan-Newman algorithm identifies either seven or eight communities with the same (rather high) modularity (Q = .60).

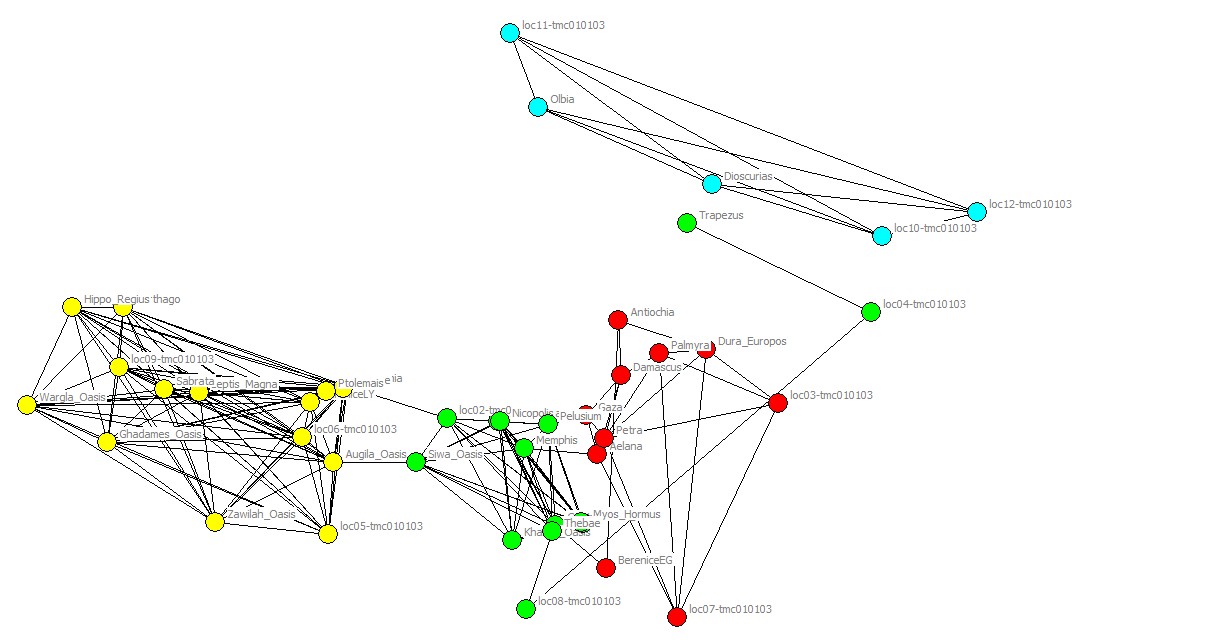
There is, as one would expect, a positive correlation between the trade network and the cultural network of the settlements. But, the correlation is modest (r = .30, p = .002 by permutation trials tests). The correlation between two networks is simply the degree to which the dyadic ties that are present/absent in one network correspond to the dyadic ties being present or absent in the other. Permutation trials are used calculate standard errors from a distribution of the same densities, in which dyadic ties are distributed randomly. While eight communities are identified in each network, the correspondence between community memberships, and hence the boundaries of the regions identified, differ.

## Trade and cultural networks combined

The identification of regions and their boundaries can produce complex solutions when multiple relations among the settlements are considered, as we see above. One may wish to accept this as the best solution – the real world is messy. Regions may, indeed, overlap to varying degrees, and some places may be parts of multiple regions, depending on which relation is considered.

It is also possible to combine the relations and seek an optimal partitioning based on multiple dimensions simultaneously. In figure 3, we show the results of applying the modularity maximization approach to both networks, combined.

Figure 3. Combined trade and language networks among Roman settlements\*



\* 4 Girvan-Newman communities, with modularity of .60 are shown by node colors.

The Girvan-Newman modularity maximization approach to the combined trade and language networks produces a four, five, or six region solution, but all have the same (rather strong) degree of modularity (Q = .60). This modularity is similar to either of the single relations, and the simplest (four community) solution is shown in figure 3. The rather high modularity of the data produces relatively un-ambiguous communities and boundaries. The relatively high modularity also indexes a fairly low “system-ness” in the sense of “close coupling.” The very modest correlation between the two relations (r = .30), also indicates relatively loose coupling or a low level of “system-ness.”

In the Roman Empire example we see that setting boundaries in cultural relations of settlements and economic relations of settlement may produce different (and only partially overlapping) maps of regions.

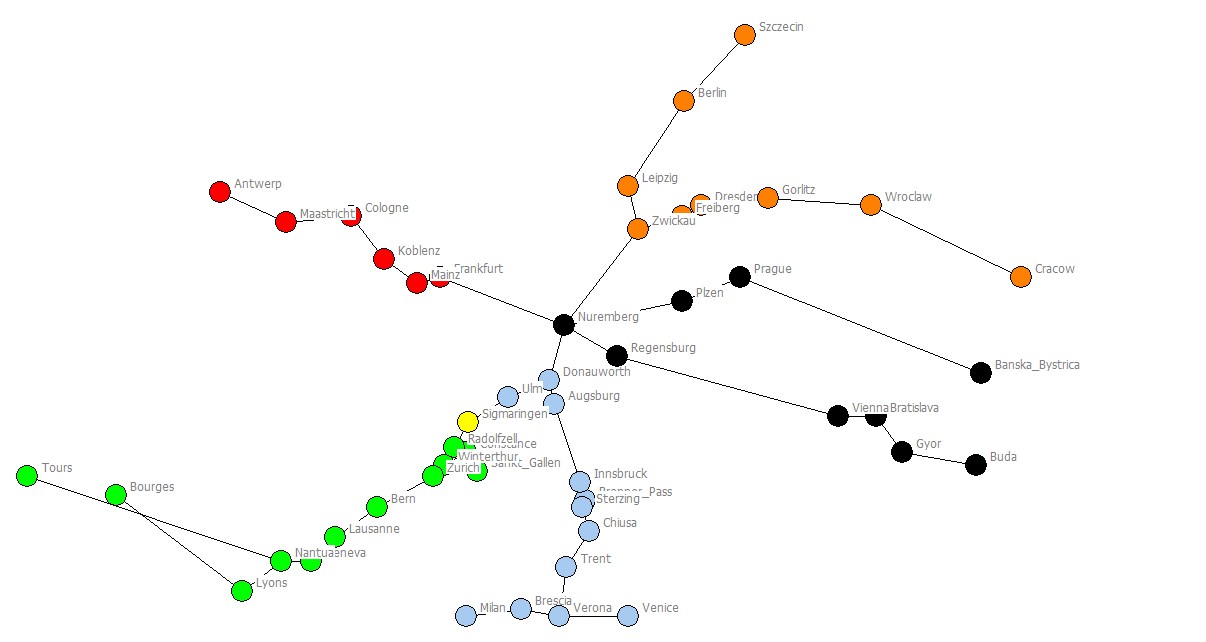
# Single and multi-relational approaches to central Europe

Political relations, as well as economic and cultural may also integrate places into regional systems. As a second example of our approach, we examine a set of data that describes trade ties, linguistic ties, and also embeddedness of settlements in political units in Europe circa 1500 (Spufford, 2002). Spufford provides a map of trading relations centered on the south German states (and hence, is not a complete network map of all the 49 settlements). We have again constructed a cultural network based on shared dominant languages (SUNY-Buffalo, accessed 2016). In addition, we consider a political dimension in this example by locating each of the 49 settlements within a political community (euratlas, accessed 2016; globalsecurity, accessed 2016).

## Trade network

Figure 4 shows the trading network centered on south Germany. The overall density of the network is quite low (.041), and the Eigen-centralization is very high (84.8%). The average path length is long (8.3). All of these values are, partially, artifacts of the way in which the network is defined (as an ego-network of sorts, focused on south German settlements, rather than a complete network (as was the case for the Roman data).

Figure 4. Trade network among central European settlements\*



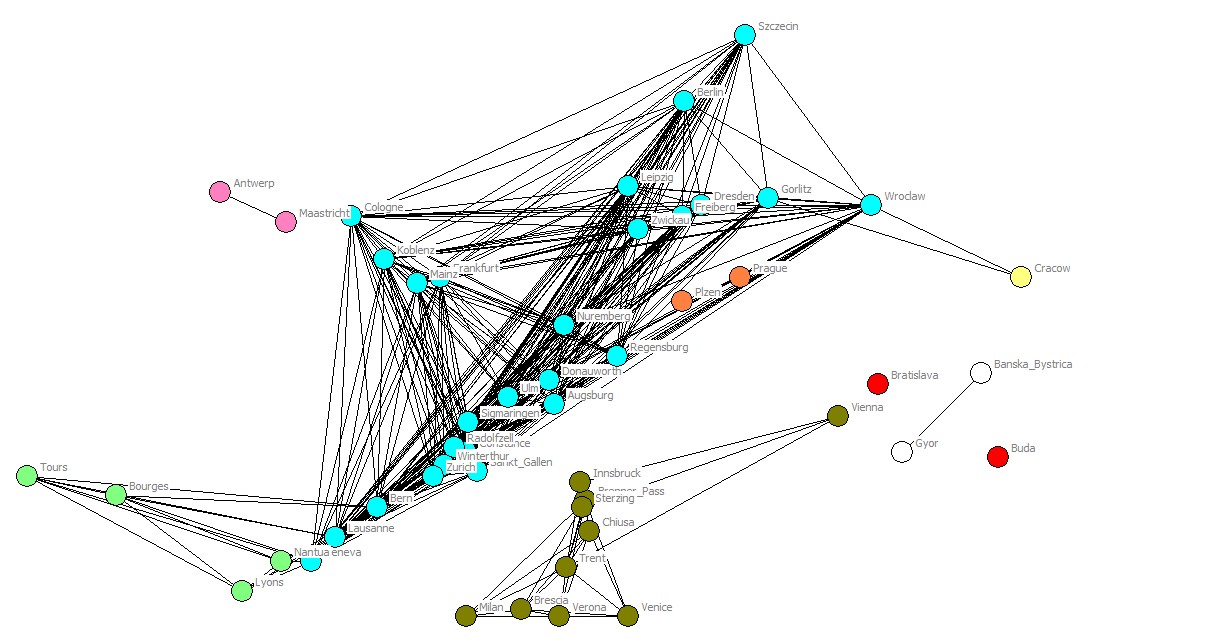
\* 6 Girvin-Newman communities, with modularity of .69 are shown by node colors.

The Girvan-Newman algorithm reaches a maximum modularity of .69 with six communities for the trading network. This relatively high figure indicates that the network can be partitioned into regions that have much more connection of the settlements within each region than between settlements in different regions. That is, in one sense, the “system-ness” of trade is relatively low.

## Cultural network

Figure 5 displays the cultural connections among the settlements. The density is quite high (.322), but this is at least partially an artifact of our assumption that all settlements sharing a language directly communicate with one another. The Eigen-centrality of the network is lower (19.2%); this, too, is partially an artifact of the assumption that all communities sharing the same language are directly connected. The average path length of the cultural network is 2.8 (assuming that un-connected components are at the maximum observed geodesic distance, plus one). All three of these indexes suggest rather high cohesion.

Figure 5. Cultural network among central European settlements\*



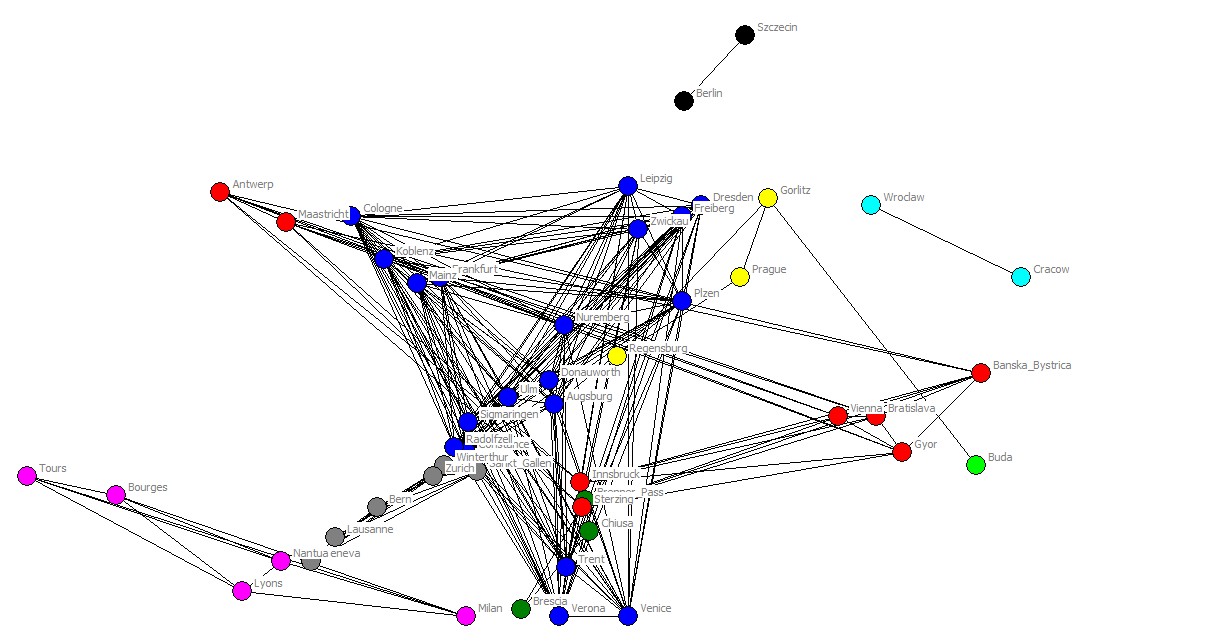
\* 7 Girvin-Newman communities, with modularity of .17 are shown by node colors.

The Girvan-Newman algorithm identifies 7 or 8 communities (both solutions have Q = .17). The relatively low modularity is a result of a large central component in which German-speaking overlaps with other language groups. This central component displays low modularity, or high clustering, or a tendency toward “system-ness.” The cultural network then displays greater cohesion and system-ness than the trade network.

## Political network

Figure 6 is a colored map of the political regions. One might expect a very high degree of modularity in a political map, because individual settlements are embedded in particular states. In the period of these data, however, the Holy Roman Empire overlaps with other political entities. This phenomenon would also be observed in the modern period, if one chose to include super-national political entities (e.g. alliances, associations such as the United Nations) as defining political connections. The density of the political network is relatively high (.215), which is also partially an artifact of the method of treating all communities affiliated with the same ruler as connected to one another. The Eigen-centrality of the network (28.8%) is rather low – partially due to the overlapping effects of the Holy Roman Empire, and partially due to the method of connecting all communities under the same ruler to one another. The average path length in the political network is quite short (2.6), reflecting, in part, the spanning influence of the Hapsburgs.

Figure 6. Political network among central European settlements\*

****  
\* 9 Girvan-Newman communities, with modularity of .44 are shown by node colors.

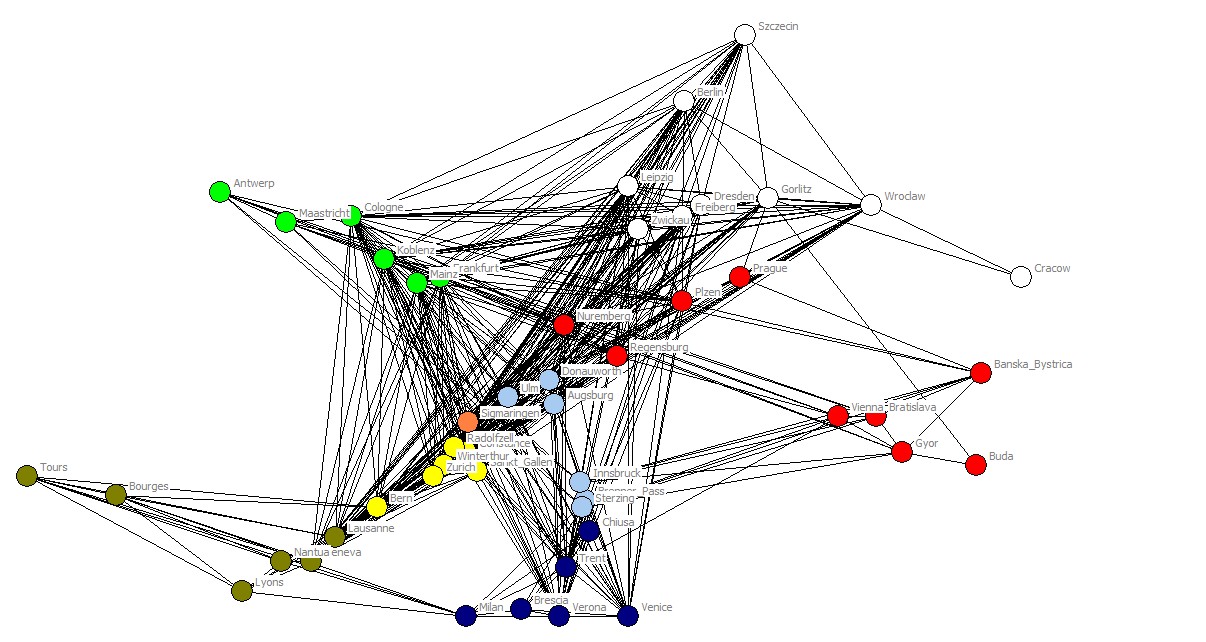
The Girvan-Newman partitioning of the political network shows only moderate modularity (Q = .44), and identifies nine communities. Overall, the political network displays a rather substantial degree of cohesion, or “system-ness.”

As the diagrams of the three networks suggest, there is some degree of correspondence of the boundaries of regions when defined by economic, cultural, and political relations. But, as in the Roman case, the association is modest. The trade and cultural networks correlate at .22; the trade and political networks correlate at .25; and, the cultural and political networks correlate at .28 (all significant a p < .05 by permutation tests).

## **Trade, cultural, and political networks combined**

An alternative to describing the data as partially overlapping regions with inter-penetrating boundaries is to combine all three relations. Figure 7 shows the results of this approach.

Figure 7. Combined networks among central European settlements\*

****  
\* 8 Girvin-Newman communities, with modularity of .71 are shown by node colors.

Combining all three relations results in a solution that identifies roughly the same number of regions as the single-relation approaches. The regions, however, do not correspond in detail to any one of three single-relation solutions. The modularity of the combined relational network is quite high (.71). This indicates that economic, political, and cultural regionalities tend to re-enforce one another (despite the modest correlation of the networks). Despite overall modestly dense connection, the “system” does not display high “system-ness.”

Discussion

A multi-level network approach to identifying regions and boundaries provides a clear conceptual definition of the notion of a “region.” A region is a set of nodes and their connecting relations that has greater density within than without the set. Regions may be identified using a single type of tie, or multiple types of ties. The ties between settlements may be due to either direct relations, or common affiliation of embedding.

The notion of a region provided by network analysis suggests that settlement systems may vary from having a single region (as in a lattice or clique structure) to having as many regions as settlements (as in a network with nodes, but no ties). Regions, in the approach we’ve taken here, may have strong or weak boundaries. The factors that make one “region” a region need not be identical to those that provide the cohesion and separateness of another. The notion of the “modularity” of a graph captures provides a continuous metric to capture this feature of settlement systems. The notion of modularity has the additional major advantage of allowing the application of computational algorithms to determine the optimal number (and node membership) of regions.

The multi-level network approach also provides an answer to the question of the “system-ness.” From the perspective advanced here, a network is more “systemic” if a signal arising at any point in the network rapidly and robustly affects all the other points. To the extent that an overall settlement network is “cohesive,” it is likely to be highly responsive, and may display synchronization and/or emergent phase-shift behavior. That is, the more cohesive the system, the less it is true that the total is equal to the sum of the parts.

Our perspective also directly connects the regionality-boundary issue to the “system-ness” issue. We have argued that several properties of networks affect the extent to which they display emergent systemic behavior (i.e. short path lengths, high density, low centralization, and modularity). Of these, modularity may be particularly important. A network that displays high modularity is one that has clearly bounded regions. Such a network has un-even, and weak connections among its parts, and is less likely to display “system-ness.”

But, we must add a very important caveat to the conclusion that high modularity implies low system-ness. This caveat is that many systems may display “small world” properties in which high clustering (high modularity) may combine with short-path lengths. With the proper configuration of a small number of highly-between paths, a graph may display both strong regionality and a quite high degree of cohesion (though robust-ness may be low). Many natural systems have evolved toward “small-world” structures because they display the survival advantages of both robust and dense local connection along with fairly rapid access to full-system resources. One might hypothesize that such adaptive advantages of network structure would also be selected for in human settlement systems.

The network approach we’ve discussed here differs in some important ways from other approaches to identifying regional boundaries and indexing the “system-ness” of settlement systems.

In identifying “regions,” many approaches may focus exclusively on one relational dimension; or may give primacy to one type of relation. It is common, for example, to focus first on political affiliations (e.g. states or empires). Economic and/or cultural ties may then be seen as “softening” or “re-enforcing” the political boundaries (and hence the “system-ness” of the entire system). Typically, anthropologists gravitate to the cultural/informational regionality as primary; the Marxist gravitates to the commodity-production and trade relations as primary; the political scientist (or conflict theorist) to the political. The network analysis approach discussed here is agnostic. Different relations may be more important in defining optimal regionality in different networks; and even within the same network different regions may be defined, primarily, by different relations (i.e. one region may be primarily “cultural” while another, in the same settlement system may be primarily defined by economic ties).

This is to say that the network approach is a descriptive, data-driven, bottom-up way of approaching the questions of region and system. Many other common approaches are much more driven by prior theory. The network approach then is both more closely connected to data, and more dependent upon it. If one is data-driven, it matters very much what dimensions and relations are included, the reliability and validity of measurement, and sampling of time and place.

Many approaches to regionality are “place-centric.” That is, there is a prior notion of “key nodes.” These nodes are often (explicitly or implicitly) chosen because of their nodal attributes rather than their relational positions. The network approach that we have put forward here does not pre-suppose that the settlement of Rome is any more important, or a starting point for analysis than any other settlement in 1500 CE. Nor does the network approach “care” that Rome had a large population, complex and advanced technology, unique political institutions or cultural features.

In network analysis, the connection between the attributes of a node (e.g. Rome is big, complex, rich) and it’s relational position (e.g. Rome militarily dominates other populations, is a central hub in trade, and a source of religion and culture) is treated as problematic. Certainly, the attributes of places matter. Large population centers have the potential to create large armies that allow them to dominate others. But a place becomes large precisely because it is in favorable position in networks of trade, culture, and power. Network analysis sees the attributes of places and the relations among places as a co-evolving system. The “central place hierarchy” and the “exponential distribution of settlement sizes” are both caused by, and causes of centralization in relational networks.

Conclusion

Conceptualizing structure as a multi-relational web or network of geo-located human populations seems intuitive from an empiricist and human-ecological perspective (though not necessarily from a functional “systems” perspective – eg. Core-periphery). If one takes the empiricist/structuralist approach, formal network analysis has a useful tool-kit for identifying boundaries and boundedness; as well as for describing the complexity of the system as a set of weakly connected dense sub-structures of interactions of populations. Readily measurable network properties allow the identification of regions and their boundaries within settlement systems, as well as comparisons of regionality and system-ness across settlement systems.

But, there are major issues and questions in the actual application of the approach discussed here:

What are the sets of relations that define the interactions of human settlements? How can these relations be measured? What are the effects of invalidity and un-reliability of data on the robustness of the network approach?

What kind of an answer do we want to the question of sub-structures – different approaches yield solutions that are more qualitative or quantitative: each might have advantages. Our examples illustrate that the separate analysis of each relation in a multi-relational network does not lead to the same result and insights as combining multiple relations. Analyzing multiple networks separately seems more nuanced, and gives a richer picture of regionality and the reasons why some regions are more cohesive than others. But, separate network analysis does not directly address the question of “system-ness.”

Making choices about some of these issues will need to be informed by thinking more deeply about why we care about “regionality” and “system-ness.” Are these properties of human settlement systems the thing to be explained? Are these properties that determine other outcomes of interest? Or it both?

# References

Casson, Lionel. 1984. *Ancient Trade and Society*. Detroit: Wayne State University Press.

Ciolek, T. Matthew. 2000. *Georeferenced data set (Series 1 - Routes): Maritime trade routes in the Mediterranean c. 150-200 CE.* OWTRAD Dromographic Digital Data Archives (ODDDA). Old World Trade Routes (OWTRAD) Project.

Euratlas.org. (Retrieved 2016). [http://www.euratlas.net/history/europe/1500/index.html](https://post.ucr.edu/owa/redir.aspx?SURL=kjwikfEDP0BPuKQhI3OPue5bXyox1gUhRac4_DU9wVMNUxarIyDTCGgAdAB0AHAAOgAvAC8AdwB3AHcALgBlAHUAcgBhAHQAbABhAHMALgBuAGUAdAAvAGgAaQBzAHQAbwByAHkALwBlAHUAcgBvAHAAZQAvADEANQAwADAALwBpAG4AZABlAHgALgBoAHQAbQBsAA..&URL=http%3a%2f%2fwww.euratlas.net%2fhistory%2feurope%2f1500%2findex.html)

Globalsecurity.org. (Retrieved 2016). [http://www.globalsecurity.org/jhtml/jframe.html#http://www.globalsecurity.org/military/world/europe/images/map-europe-1500-2.jpg|||European%20History%20Map%20-%201500%20AD](https://post.ucr.edu/owa/redir.aspx?SURL=k9BWFqNEI0BItc-5Jbp1NOfoooVp5PLB_LCNSnhi38kNUxarIyDTCGgAdAB0AHAAOgAvAC8AdwB3AHcALgBnAGwAbwBiAGEAbABzAGUAYwB1AHIAaQB0AHkALgBvAHIAZwAvAGoAaAB0AG0AbAAvAGoAZgByAGEAbQBlAC4AaAB0AG0AbAAjAGgAdAB0AHAAOgAvAC8AdwB3AHcALgBnAGwAbwBiAGEAbABzAGUAYwB1AHIAaQB0AHkALgBvAHIAZwAvAG0AaQBsAGkAdABhAHIAeQAvAHcAbwByAGwAZAAvAGUAdQByAG8AcABlAC8AaQBtAGEAZwBlAHMALwBtAGEAcAAtAGUAdQByAG8AcABlAC0AMQA1ADAAMAAtADIALgBqAHAAZwAlADcAQwAlADcAQwAlADcAQwBFAHUAcgBvAHAAZQBhAG4AJQAyADAASABpAHMAdABvAHIAeQAlADIAMABNAGEAcAAlADIAMAAtACUAMgAwADEANQAwADAAJQAyADAAQQBEAA..&URL=http%3a%2f%2fwww.globalsecurity.org%2fjhtml%2fjframe.html%23http%3a%2f%2fwww.globalsecurity.org%2fmilitary%2fworld%2feurope%2fimages%2fmap-europe-1500-2.jpg%257C%257C%257CEuropean%2520History%2520Map%2520-%25201500%2520AD)

Hanneman, Robert and Mark Riddle. 2005. *Introduction to Social Network Methods*. <http://faculty.ucr.edu/~hanneman/nettext>

Mann, M. 1986. *The Sources of Social Power: A History of Power from the Beginning to AD* 1760. Cambridge: Cambridge University Press.

Newman, M.E.J. 2006. “Modularity and Community Structure in Networks*.” Proceedings of the National Academy of Sciences of the United States*. 103, 23: 8577-8696.

Scullard, H.H. 1970. *From the Gracchi to Nero: a History of Rome from 133 B.C. to A.D. 68.* London: Methuen & Co. Ltd.

Spufford, Peter. 2002. *Power and Profit: The Merchant in Medieval Europe*. London: Thames & Hudson.

State University of New York, Buffalo. (Retrieved 2016). *Demographic Ethnicity and Linguistic Maps*. [http://info-poland.buffalo.edu/classroom/maps/task5.html](https://post.ucr.edu/owa/redir.aspx?SURL=VJvivJV5HntKXzbnfNvhfABHRyw7xOLszTA3OKr7Ey0NUxarIyDTCGgAdAB0AHAAOgAvAC8AaQBuAGYAbwAtAHAAbwBsAGEAbgBkAC4AYgB1AGYAZgBhAGwAbwAuAGUAZAB1AC8AYwBsAGEAcwBzAHIAbwBvAG0ALwBtAGEAcABzAC8AdABhAHMAawA1AC4AaAB0AG0AbAA.&URL=http%3a%2f%2finfo-poland.buffalo.edu%2fclassroom%2fmaps%2ftask5.html)  
[http://info-poland.buffalo.edu/classroom/maps/europe1815\_1914lang.jpg](https://post.ucr.edu/owa/redir.aspx?SURL=DrAlblVz0c1O89V7xMxCwiXV5E5rnbw9aZZC5Dj_OOUNUxarIyDTCGgAdAB0AHAAOgAvAC8AaQBuAGYAbwAtAHAAbwBsAGEAbgBkAC4AYgB1AGYAZgBhAGwAbwAuAGUAZAB1AC8AYwBsAGEAcwBzAHIAbwBvAG0ALwBtAGEAcABzAC8AZQB1AHIAbwBwAGUAMQA4ADEANQBfADEAOQAxADQAbABhAG4AZwAuAGoAcABnAA..&URL=http%3a%2f%2finfo-poland.buffalo.edu%2fclassroom%2fmaps%2feurope1815_1914lang.jpg)

Stone, Norman (ed.). 1989. *"The Times" Atlas of World History*. Third edition. London: Times Books Ltd. pp.91.

Tilly, C., L.A. Tilly, and R.H. Tilly. 1975. *The Rebellious Century*. Harvard: Harvard University Press.

Vischer, M.S. 2011. Landscape of Languages The position of provincial languages in the Roman Empire. (Master’s Thesis, University of Leiden). https://openaccess.leidenuniv.nl/bitstream/handle/1887/18922/ScriptieMSVisscher.pdf?sequence=1

Wikipedia. (Retrieved 2015). “Modularity (Networks).”

Wikipedia. (Retrieved 2016). “Girvan-Newman algorithm.”