Social Conductance: An Adaptation of Ohmic Circuit Theory To Measure Tie Strength and Importance in Relational Data

Andrew C. Thomas

June 27, 2009

Graph-theoretic interpretations of relational data sets have proven popular due in part to their clarity of explanation. The geodesic path length between individuals can be considered to be a measure of closeness between individuals; the degree of an individual is a proxy for both gregariousness and popularity. Additionally, statistical measures for characteristics of the whole ensemble, such as global transitivity and reciprocity, may model the tendencies of individuals to form particular relationships given others present in the system.

There are several shortcomings in this method of analysis that cannot easily be rectified. Geodesic path lengths may not give an accurate measure of social distance if multiple shortest paths exist, such as the case where two unconnected people have one mutual friend, or ten. In order to measure the importance of a tie rather than a node, a measure of betweenness can be constructed using geodesic paths [Freeman, 1979], though the issue of multiple geodesic paths, possibly sharing edges between them, requires solutions that may prove difficult to interpret.

As an alternative, I consider another well-studied model where multiple paths between nodes have a straightforward interpretation. When considering social ties as conduits for information transfer, there is an immediate analogue to electrical circuitry, in which the connection points for circuit components are seen to be the nodes of a network, and a signal is sent from one node to another in the form of an applied voltage difference. The strength of this signal, in terms of the current flow that results, depends both on the properties of each edge in the system as well as their topological arrangement.

I briefly outline the concept of electrical conductance in terms of a fixed potential difference applied across a pair of nodes. I then demonstrate how this measure allows us to determine a refined measure of social distance between two individuals in a network, as well as the relative importance of a tie for information transfer using an analogue of electrical power.

0.1 Previous Work

Approaches for considering electrical circuits as networks are at the root of Kirchhoff's Circuit Laws, discussed in detail in Section 0.3, so it is not surprising that the extension to social networks has been considered repeatedly in the past 150 years. Among these considerations are those by Freeman [1979], in which degree-based and betweenness centrality are discussed. These were natural precursors to Freeman et al. [1991], which dealt with the notion of network "flow" as a method of connectivity between two points, in which the value of a network tie is taken to be a flow capacity – essentially, that all ties in a network have equal length but varying diameter.

Circuit-law based derivations are found by Stephenson and Zelen [1989], and rediscovered in the form of random walks on a graph by Newman [2005]. The underlying methods for calculating current flow are essentially identical to those presented in this paper, as they are based on the Kirchhoff methodology, but with differing interpretations which I elaborate in this work.

0.2 Social Pathways, Milgram's Experiment and the Flow of Information

As an oft-mentioned study of cross-country sociology and community, Stanley Milgram's small-world experiment [Milgram, 1967] was aimed to determine the minimum number of connections between individuals separated by geography. The experiment had a number of people from Nebraska try to send a package to a person in Boston, with the constraint that the package could only be passed to an intermediary known personally by each respective sender in the chain. [1]

One crucial piece of this experiment is that each person in the experiment, with the exception of the final recipient, was able to communicate only through one more person. As a result, only single paths were traced between the originators and the

[1] The observed average of the number of senders was roughly six, yielding the popular aphorism "Six Degrees of Separation", a point made in detail throughout the networks literature.

recipient. While this model may be useful for demonstrating a short distance in social space along a single path, it does not accurately represent the mechanism highlighted by the experiment – the transfer of information along social ties. Replication, division and transmission of information are qualities that cannot be duplicated in an experiment with this physical limitation.

As a thought experiment, consider what could happen if the small world experiment were repeated, but where each source of information were given, say, one million small packages to send to the target in Boston. The previous goal, that they attempt to have the transmission take as few paths as possible, remains in effect, though there are also practical problems with sending all the packages to one local recipient for fear of a backlog problem. Over time, each recipient can report back to the original sender with their ultimate capacity and how much traffic they can handle; this is then taken into account by the original sender as they plan for future deliveries.

This is a crude but accurate representation of the flow of current in an electrical circuit. Resistance to current represents the energy required for a signal to be transmitted at every step, and nodes have a potential energy level that governs the degree to which current flows. As seen in Figure 1, and as I explain throughout this note, this model illustrates the importance of each node and edge in the flow of information, with a natural extension to edges that have greater transmission power than others.

0.3 Ohmic Circuits and Kirchhoff's Laws

This method of exploring networks requires three quick definitions from the circuit theory literature.

Electric current I is the result of an electric field being applied to a conductor, in which particles of one charge are free to move and others are fixed in place, and is defined as the rate at which charge moves past a reference point per unit of time.

> Kirchhoff's Current Law states that charge cannot accumulate at any point in a completed circuit. This implies that with respect to any fixed point, the net current is zero; that is, any current going in must also be balanced by an equal amount of current going out.

The energy carried by an electric current is most often expressed in terms of an "electric potential difference" or "voltage" V, which is equal to the amount of energy transferred per unit of charge.

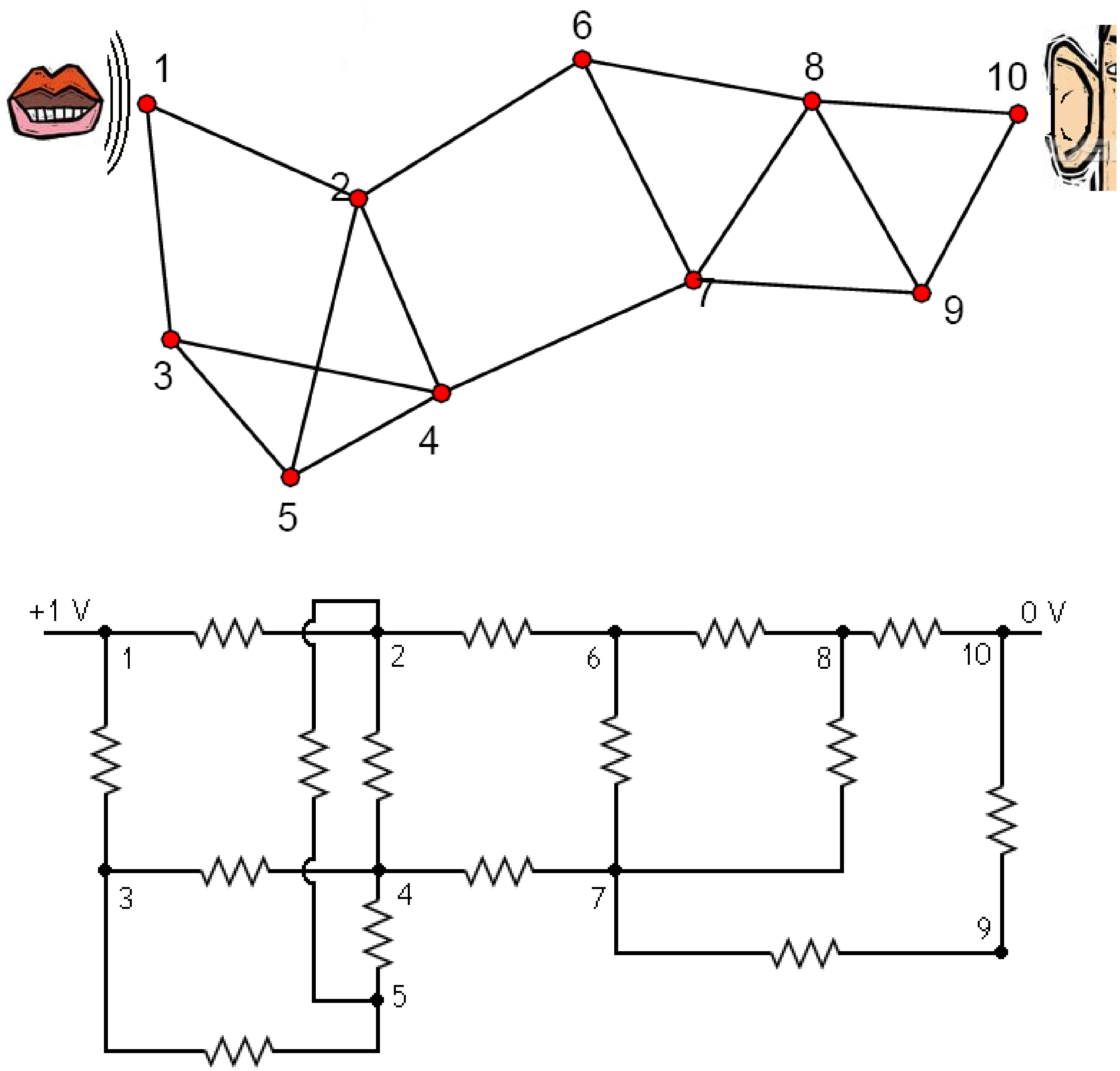

Figure 1: Representing the analogy between information transmission through a social network and current flow through a compound electrical resistor.

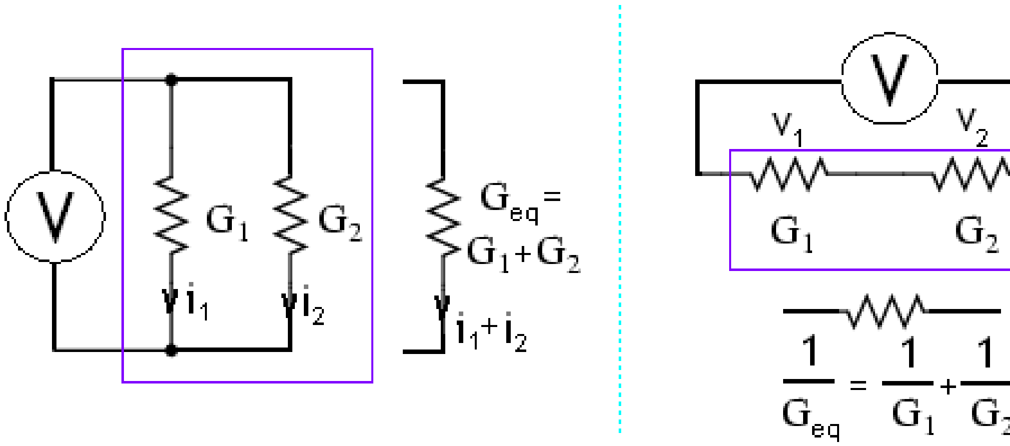

Figure 2: Simple circuits demonstrate compound conductances. In the parallel case on the left, the total conductance is the sum of the respective conductances because the induced current is higher; in the serial case on the right, the total conductance is the reciprocal sum due to the drop in induced current.

> Kirchhoff's Voltage Law states that in any closed loop, the total electric potential difference must be zero. As a corollary, any two paths taken from one point in a circuit to another must have identical potential differences.

Ohm's Law relates the behaviour of current to that of potential difference along a circuit element:

> The ratio of current to potential difference, I/V, is called the conductance, with symbol G. In so-called Ohmic circuits, a conductor is a device for which this ratio of current to potential difference is maintained for all feasible potential differences across its length. [2]

From these three pieces of information, we can define the equivalent conductance of an arrangement of conductors in terms of the current produced by a specified applied potential difference (as in, say, a common household battery). By determining the total current I_{total}, we calculate the equivalent conductance of the circuit as $G_{eq} = I_{total}/V_{total}$.

[2] This is usually represented canonically as the "resistance", or R, defined as the reciprocal of conductance. I use conductance in this paper for its isomorphism to the strength of a connection, which in the sense of information takes the form of current.

In the case of two parallel conductors attached to a common potential difference, the total current is found to be

$$I_{total} = VG_1 + VG_2 = V(G_1 + G_2),$$

yielding the equivalent conductance of two parallel conductors as their sum, $G_{eq} = G_1 + G_2$.

In the case of two serial conductors, we note that the total potential difference across the conductors must equal that of the source:

$$V = \frac{I}{G_1} + \frac{I}{G_2} = I\left(\frac{1}{G_1} + \frac{1}{G_2}\right);$$

this yields the equivalent conductance, given by $1/G_{eq} = V/I = (1/G_1 + 1/G_2)$.

Similar expansions can be calculated for more complicated circuit arrangements, but the equivalent conductance for any pair of nodes, even those without direct connections, can be calculated using the following algorithm:

1. Create a vector $(V_1, ...V_n)$ specifying each intersection node in the system.

2. Choose the nodes (a, b) across which the equivalent conductance is to be found. Set $V_a = 1$ and $V_b = 0$.

3. Set up a system of equations using Kirchhoff's Current Law and Ohm's Law for all nodes (except the source and sink nodes a and b):

$$0 = I_k = \sum_{j \neq k} \frac{(V_j - V_k)}{G_{jk}}$$

   Note that the currents through the source and sink nodes, I_a and I_b, are by definition non-zero, reflecting the induced current due to the applied potential difference across the circuit.

4. Solve this system of equations for the remaining V_k, noting that all values must be in the interval $[0, 1]$.

5. Determine the total induced current

$$I_a = \sum_{j \neq a} (1 - V_j) G_{aj}$$

and set $G_{eq} = \frac{I_a}{V_a - V_b} = I_a$.

As this yields a set of electric potentials, we can also identify the current across any conductor as being $I_{jk} = (V_j - V_k) C_{jk}$.

0.3.1 A Note of Comparison

The method suggested here is identical in solution to that proposed by Stephenson and Zelen [1989] and Newman [2005], except that these works suggest a fixed current approach rather than a fixed voltage. In fact, as the voltage-based method requires a matrix inversion for every node pair considered, it is less than practical for a full analysis of a system with more than 200 nodes on currently available computing hardware, and the implementations in this work have been computed using fixed current.

For the sake of the analyses that follow, both the fixed-voltage and fixed-current approaches will be used to investigate networks. Since the relationship between voltage and current is linear, the results only differ by a factor of the equivalent conductance and are immediately deriveable from each other.

1 Social Conductance

In a social network setting, when individuals are considered to be nodes in a graph, the absence or presence of an edge determines the ability to conduct information directly between the individuals in question. Previous explorations have considered "degree", or the minimum path length between two individuals, to represent the effective distance between individuals; this representation assumes, however, that information can not be replicated and sent separately along multiple channels.

We hence define **social conductance** as being the rate at which two individuals can share information directly between them, and **equivalent social conductance** as the total information flow rate between individuals when accounting for all possible paths of transmission and conductance. In the case of a binary network, the social conductance of a tie is set as equal to one, and for the lack of a tie is equal to zero.

It must also be noted that there is an immediate extension to univariate non-binary relations, since we can represent any nonnegative tie strength as a social conductance.

Once we have solved for the potentials at each node in the system, given an applied potential difference across two nodes, we can measure the current flow across any edge. This current then represents a fraction of information that travels from one person to another within the network, and are internally comparable for determining which edges are the most influential in conducting information. See Figure 4 for a comparison of several paths.

1.1 A Measure of the Effective "In-Network" Strength of a Tie

For each pair of nodes in the network, we determine their equivalent social conductance using the aforementioned procedure. As a measure of social connectivity, this now represents not only the strength of a direct connection but of all social pathways connecting two individuals. This interpretation immediately lends credence to Simmel's hypothesis [Simmel, 1955] that influence through a single tie is insufficient to capture sociological phenomena without considering common connections. Because this measure considers the entire network when calculating the degree of connection between two individuals, this effectively extends the measurement of social distance to the consideration of multiple middlemen alongside a relationship.

To demonstrate, consider the effective tie strength between two members of an n-clique. As demonstrated in Figure 3, when inducing a potential difference across any edge, the remaining points each take an electric potential halfway between the source and sink nodes. Because no current flows between points of equal potential, the only remaining edges that carry current are the direct path and the remaining $n-2$ two-step paths, all in parallel. As the conductance of two serial conductors is half that of each component, the total conductance of the remaining assembly is $G + \frac{n-2}{2}G = \frac{nG}{2}$.

It follows immediately that the observation of a binary tie within a highly interconnected component may in fact be illusory; that this is in fact a long-distance tie in true magnitude that is only observed due to the influence of their intermediaries. This "local amplification" effect is purely endogenous, making it difficult to disentangle in the case of binary tie observation.

An additional visual comparison between the geodesic and electro-social models is shown in Figure 3, in which five individuals form a clique but each maintain one outside

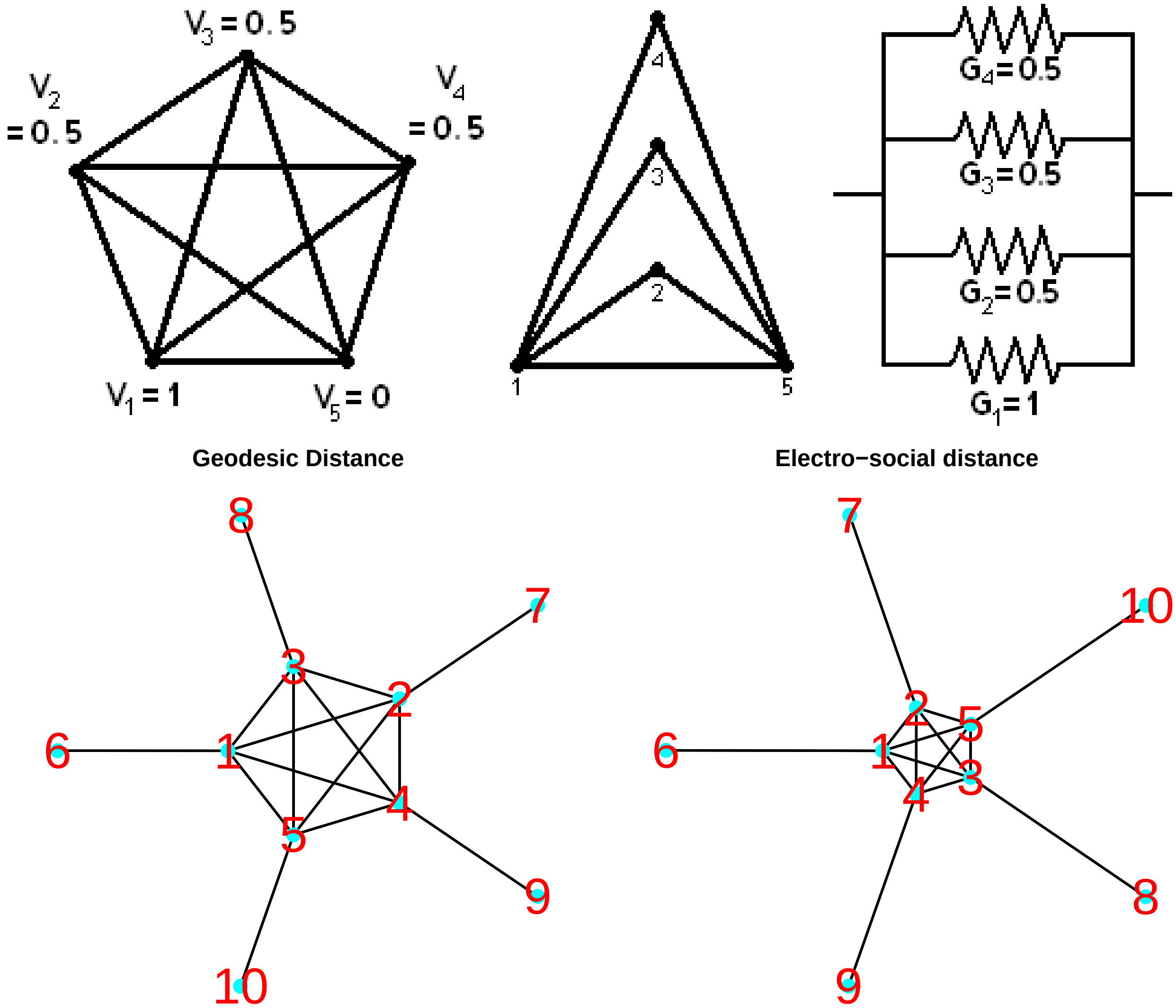

Figure 3: Above, the equivalent social conductance between members of a five-clique. As additional middlemen are added, the conductance increases, reinforcing the theory of Simmelian ties. Below, a visual comparison of social distance under the geodesic and electro-social models.

friend. If the underying tie strengths are all identical, the geodesic method does not consider the outside friendships to be any closer than those within the clique, whereas the social conductance method allows for this possibility.

1.2 Tie and Node Importance As Average Observed Current

Consider that with a fixed potential difference across nodes a and b, there is an induced potential at each node in the graph V_i between 0 and 1, and an induced current between the nodes specified as

$$I_{ij}^{ab} = (V_i^{ab} - V_j^{ab})G_{ij}.$$

If the total induced current in the system is I^{ab}, define the **embedded importance** of an edge in this configuration as the fraction

$$M_{ij}^{ab} = \frac{I_{ij}^{ab}}{I^{ab}},$$

and the average embedded importance of an edge as

$$M_{ij} = \frac{1}{\binom{n}{2}} \sum_{a} \sum_{b \neq a} M_{ij}^{ab}.$$

This is defined with respect to directed edges, so that one arc in a dyad may be more important than another for whatever reason of topology, as discussed ahead.

This measure is *not necessarily* equal to the drop in current that would be observed if the tie were removed from the system, defined as the **tie-cut importance**,

$$T_{ij}^{ab} = \frac{1 - (I_{ab}|G_{ij} = 0)}{I_{ab}|G_{ij} = G_{ij}(true)};$$

in fact, as part of this current may be directed to other less-used nodes, it can be shown that the embedded importance is an upper bound on the true tie-cut importance, $M_{ij}^{ab} \geq T_{ij}^{ab}$.

As a comparative measure, the embedded importance has many desirable properties. It can be calculated simultaneously for all ties in the system given a source-sink pair, whereas tie-cut importance requires at most $2\binom{n}{2}$ calculations, one for each tie to be removed. Additionally, the relative order of tie importance will be extremely close in most circumstances.

To demonstrate, Figure 4 shows the sample network from Figure 1 with potential differences applied across pairs of nodes, where line and node thicknesses represent current flows and therefore importances to information transfer.

2 Betweenness Measures For Nodes Based On Current Flow

The concepts in the previous section to evaluate edge importance can be extended immediately to nodes by taking the sum of all currents entering or leaving a node during an electro-social path test, as seen in Figure 5.

The measure of betweenness centrality for a node, as defined in Freeman [1979], is an average of the fraction of shortest paths between two other nodes that contain the third node. Since non-geodesic paths may also play an important role in node-to-node communication, an approach that considers multiple paths appropriately weighted can be valuable. Current-based measures for centrality are the prime focus of Newman [2005], and occupy a full chapter of Bollobas [1998].

The nature of how the averaging is conducted, however, is paramount to the calculation of a centrality statistic. The assumptions of Newman [2005]; Bollobas [1998] are that each node pair has equal weight when composing betweenness centrality; that is, that a pair of nodes of large geodesic distance (say, 12 intermediate nodes) carry as much weight for calculating betweenness as two adjacent points.

By considering equivalent distances based on a voltage measure, the current produced effectively decreases with distance, though it also multiplies with the number of paths between targets. Therefore, care must be taken when applying a measure of betweenness based on current flow. I present three electro-betweenness measures based on what quantity should be constant over averaging for each node-pair trial: current flow (I=1), potential difference(V=1), or electrical power (VI=1).

Having defined the current through edge (i, j) as

$$I_{ij}^{ab} = (V_i^{ab} - V_j^{ab})\mathbb{I}(V_i^{ab} > V_j^{ab})C_{ij}$$

with respect to terminal potentials ($V_a = 1, V_b = 0$) and a total induced current I_a, the centrality measures are defined with respect to these terms.

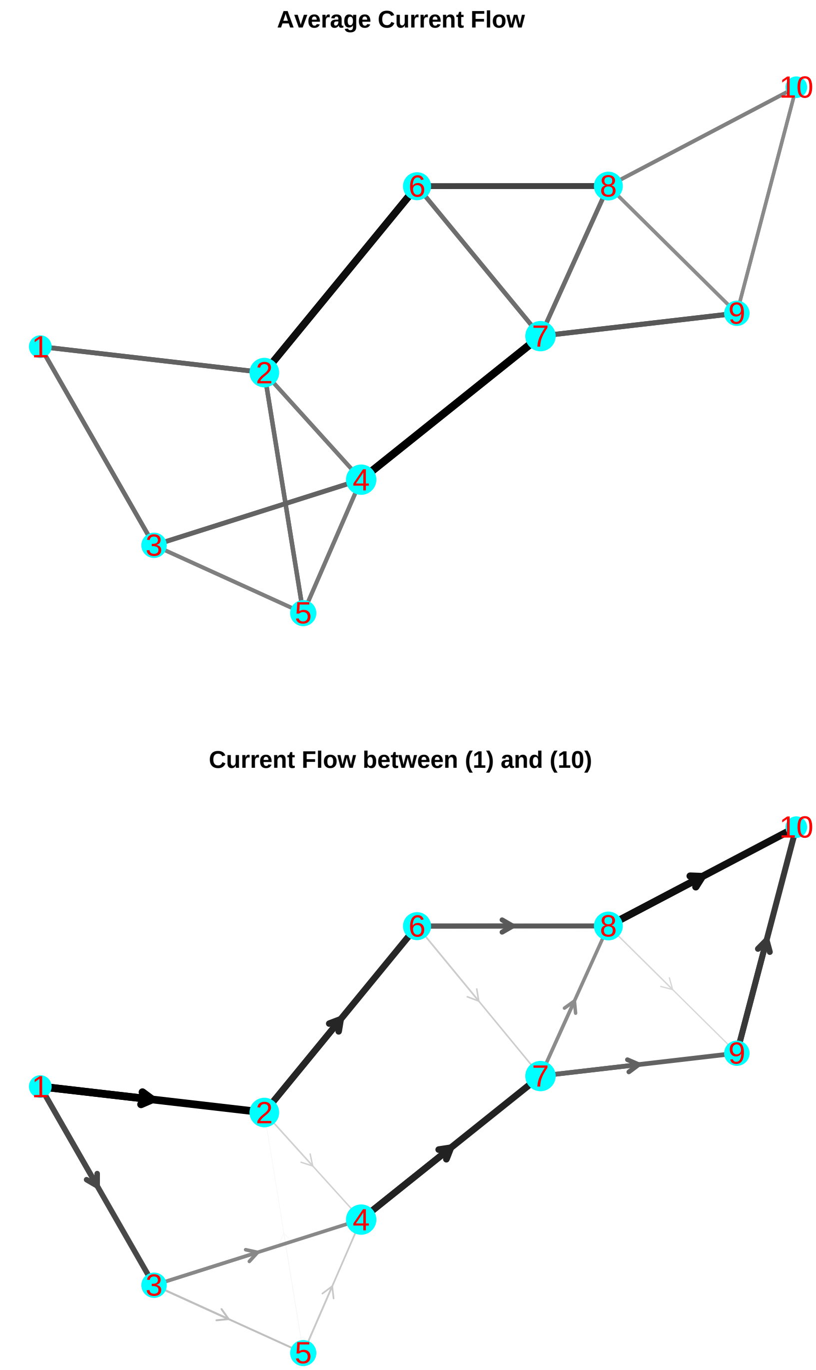

Figure 4: Current through nodes and edges of a sample network. Line thickness represents the fraction of current going through an edge; node size represents the total current flowing into (and out of) the node. Top: The current along each edge of the network, averaging together every pair of source and sink. Bottom: The current flow for the source at node 1 and sink at node 10.

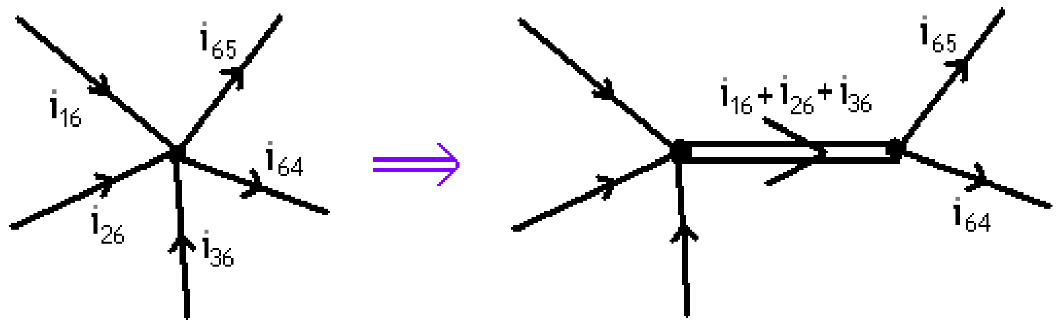

Figure 5: Current through a node as importance of the node in communication.

The notation of Freeman [1979] gives $C_D(i)$, $C_C(i)$ and $C_B(i)$ as the degree, closeness and (geodesic/shortest path) betweenness centralities respectively for a node labelled i.

2.0.1 Fixed-Voltage Electro-Betweenness Centrality

Define $C_V(i)$ as the centrality of node i as determined by the sum of the current flowing from it, averaged with respect to an applied voltage of 1 across all pairs of nodes:

$$C_V(i) = \sum_a \sum_{b \neq a} \sum_{j \neq i} I_{ij}^{ab}.$$

2.0.2 Fixed-Current Electro-Betweenness Centrality

Define $C_I(i)$ as the centrality of node i as determined by the sum of the current flowing from it, averaged with respect to an applied current of 1 across all pairs of nodes:

$$C_V(i) = \sum_a \sum_{b \neq a} \frac{1}{G_{eq}} \sum_{j \neq i} I_{ij}^{ab}.$$

2.0.3 Fixed-Power Electro-Betweenness Centrality

The power through an electrical circuit is equal to the current passing through it multiplied by the potential difference, or $P = VI$; factoring in Ohm's Law, we have $P = V^2 G$. A unit power is achieved by setting the potential difference to $V_a = \frac{1}{\sqrt{G_{eq}}}$. This yields

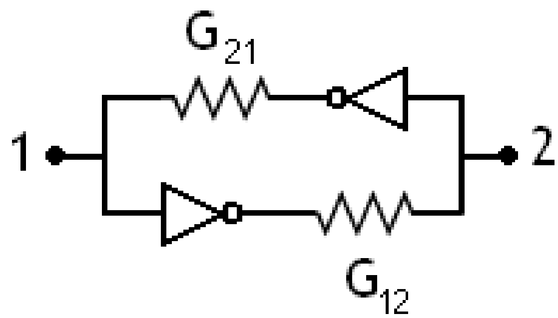

Figure 6: A differential resistor, with diodes set up to enforce a one-way flow of current; if current flows from node 1 to 2, the conductance equals G_{12} and current only goes through the bottom path; likewise, the conductance is G_{21} if current flows in that direction.

$$C_P(i) = \sum_a \sum_{b \neq a} \frac{1}{\sqrt{G_{eq}}} \sum_{j \neq i} I_{ij}^{ab}.$$

The choice of class of electro-betweenness centrality depends on the application at hand, though each has a natural explanation: in order, the multiplicative effect of parallel paths, the effect of larger distances requiring more connections, and a constant amount of energy for any individual connection.

3 Extension to Directed Graphs

The above methods have derived from standard direct current circuit theory, in which the conductance in each direction is equivalent. To apply this method to directed graphs, including those where conductance in each direction is non-zero but not necessarily equal, I introduce the notion of a differential resistor, in which the resistance/conductance of an edge in the system depends on the direction of current; see Figure 6 for a sample electrical diagram.

Because this introduces a non-linearily into the system, a single algebra operation will be unable to solve for the state of the system. However, because current is continuous with respect to voltage in this element (though non-differentiable at zero), the following iterative procedure can solve for the equilibrium voltages and currents:

1. For each asymmetric edge, create a vector of indicators whether current should flow from the lower-numbered node to the higher one, or vice versa. (As a default, set all to be "ascending" – however, this method can be sped up by cleverly guessing which elements will have current flow in each direction in advance.)

2. Create a symmetric sociomatrix by replacing all asymmetric elements with those for the indicated flow direction, and solve this system to obtain the equilibrium voltages.

3. Determine whether any of the true current flows violate their assigned path by examining the differences in voltages in the asymmetric edges.

4. If there are no conflicts, the procedure is finished; if there are, reverse the incorrect flows in the indicator vector, and repeat the previous two steps (and this one) until equilibrium is reached.

All previous measures considered for nodes and edges can still apply, since with each potential difference applied, the system behaves as if it were a traditional Ohmic circuit.

4 Extension to Stochastic Relational Data

The purpose of assembling this toolkit has been to evaluate practical tie-level statistics under conditions when a tie strength is stochastic in nature. Given an (exogenous) generative model for tie strengths, a set of networks can be drawn from the underlying parameters. This produces a joint distribution of ties that can be examined for their relevant statistical observations.

For example, the most "important" tie in one instance of a network may prove to be that way due entirely to the underlying uncertainty, and that other ties may prove to be just as important in other instances. Rank-based measures, such as centrality and edge importance, can immediately be compared across separate instances of a generative graph process.

5 Implementation: R Package ElectroGraph

The routines and methods described within have been implemented in an R package titled ElectroGraph. The manual is included as Appendix **??**, and full instructions and demonstrations are included within. The package is available on CRAN for download.

References

BOLLOBAS, B. (1998). *Modern Graph Theory*. Springer.

FREEMAN, L. C. (1979). Centrality in social networks: Conceptual clarification. *Social Networks*, **1** 215–239.

FREEMAN, L. C., BORGATTI, S. P. and WHITE, D. R. (1991). Centrality in valued graphs: A measure of betweenness based on network flow. *Social Networks*, **13** 141–154.

MILGRAM, S. (1967). The small world problem. *Psychology Today*, **2** 60–67.

NEWMAN, M. (2005). A measure of betweenness centrality based on random walks. *Social Networks*, **27** 3954.

SIMMEL, G. (1955). *Conflict and the Web of Group Affiliations*. Free Press.

STEPHENSON, K. and ZELEN, M. (1989). Rethinking centrality: Methods and examples. *Social Networks*, **11** 1–37.