Connections 17(1): 45-46

How To Test the Strength of Weak Ties Theory

Stephen P. Borgatti
Sociology Dept., University of South Carolina

Scott L. Feld
Sociology Dept., Louisiana State University

Techniques *is a regular column edited by Stephen P. Borgatti. Contributions to the column are appreciated. Contributions should be short (normally 2 pages or less), and should consist of tips on how to do things — ways of collecting data, methods of analysis, data management tricks, and so on.*

One component of Granovetter's (1973) Strength of Weak Ties (SWT) theory is the proposition that the stronger the tie between two people, the more their social neighborhoods overlap. In this article, we show how to test this proposition using QAP correlation (Hubert & Schultz 1976).

The basic strategy is as follows. First, obtain the strength of ties among all pairs of actors, and arrange these data as an actor-by-actor matrix called STRENGTH. Second, obtain a measure of the extent to which the neighborhoods of each pair of actors overlap. In other words, for each pair, count the number of people that both actors are connected to. Arrange these data as an actor-by-actor matrix as well, taking care that the order of actors is the same as in the STRENGTH matrix. Call this matrix OVERLAP. Third, compute the pearson correlation between the two matrices, using an appropriate randomization test (such as QAP) to assess the significance of the observed correlation.

Data. Typically, the STRENGTH matrix is obtained directly from respondents. An example is given by Zachary (1977) who collected strength of ties among members of a Karate club. These data are available on INSNALIB (described elsewhere this issue), and as part of the UCINET IV (Borgatti, Everett and Freeman 1992) software package.

The OVERLAP matrix requires some processing. Let us assume that in the Zachary data, a tie strength of zero indicates the lack of a tie. We could then create a dichotomized version (call it A) of the STRENGTH matrix such that A(i,j) = 1 if STRENGTH(i,j) > 0 and A(i,j) = 0 otherwise. The matrix A, then, is a simple adjacency matrix. Now, to compute OVERLAP, postmultiply A by its transpose as follows:

$$\text{OVERLAP} = AA'$$

The effect of this matrix multiplication is to count the number of times that each pair of rows in A has a 1 in the same column, indicating that the two actors are connected to the same third party. We're assuming here that the matrix A is symmetric: otherwise, AA' only gives the overlap in the out-neighborhoods of each pair of actors.

Analysis. Now we use the QAP technique to (a) compute the pearson correlation between the elements of STRENGTH and the corresponding elements of OVERLAP, and (b) assess the significance of the correlation. If we were only interested in (a), we could compute the correlation using any statistical package by stringing out the elements of each matrix into a long vector with $N*(N-1)/2$ elements (again assuming symmetric data). In other words, we create a dataset with two variables (STRENGTH and OVERLAP) in which the cases are dyads. Then we correlate the variables.

If, however, we are also interested in assessing the significance of the correlation, we cannot use the significance test computed by standard statistical packages, since these assume that the observations (which in this case would appear to be dyads) are independently sampled from a population, which is clearly not the case here. Our dyads comprise a census of all dyads given the actors sampled, and furthermore, dyads located in the same row or column of the data matrices are not independent since they share the same actors as endpoints.

So we use the QAP procedure which, in essence, compares the observed correlation with a distribution of random correlations generated according to the null hypothesis of no relationship between the matrices. The *p*-value is given by the proportion of random correlations that are as large or larger than the observed correlation.

The QAP procedure works by permuting the rows and columns (together) of one of the input matrices, and then correlating the permuted matrix with the other data matrix. This process is repeated hundreds of times to build up a distribution of correlations under the null hypothesis.

In the Zachary data, the correlation between strength of tie and neighborhood overlap is 0.416, $p < 0.001$, indicating strong support for the theory.

Extension. Another interpretation of the sense of the SWT theory is that the stronger the tie between two people, the less their tie serves as a connection between their non_overlapping social connections to others. The SWT hypothesis as applied to non_overlap is only equivalent to its application to overlap under the unlikely condition that overlap is perfectly negatively correlated with non_overlap (i.e. only when there is no variation in acquaintance volume). In general, it may be useful to separately explore the implications of strength of ties for non_overlap as well as for overlap.

To extend the previous technique for application to non_overlap, we need to construct a NONOVERLAP matrix that can be correlated with STRENGTH. Recall the dichotomized matrix, A, and its transpose, A'. We can construct a new matrix, B, which is the complement of A, so that B has 1's where A has 0's and has 0's where A has 1's. One way to do this is to subtract A from a matrix full of 1's. Alternatively, we can go back to the dichotomization step and set B(i,j) = 1 if STRENGTH = 0 and B(i,j) = 0 otherwise. In either case, though, we must be sure to put zeros down the main diagonal. Then compute C = AB' and, NONOVERLAP = C + C'.

Once the NONOVERLAP matrix is obtained, we can correlate it with the STRENGTH matrix in the same way that we did with the OVERLAP matrix, except this time we expect a negative correlation, and the *p*-value is given by the proportion of random correlations that are as <u>small</u> or smaller than the observed.

Applying this technique to the Zachary data, we observe a <u>positive</u> correlation of 0.2 between strength of tie and <u>non</u>overlap. These results do not support the theory, and instead suggest that strong ties tend to occur between actors who have many ties in general, so that their neighborhoods both overlap and do not overlap a great deal. ‘

References

Borgatti, S.P, M.G. Everett, and L.C. Freeman. 1992. *Ucinet IV Version 1.35.* Columbia, SC: Analytic Technologies.

Granovetter, M. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78:1360-1380.

Hubert, L. and J. Schultz. 1976. "Quadratic Assignment as a General Data Analysis Strategy." *British Journal of Mathematical and Statistical Psychology* 29:190-241.

Zachary, W. 1977. "An Information Flow Model for Conflict and Fission in Small Groups." *Journal of Anthropological Research* 33:452-473.