An Information Flow Model for Conflict and Fission in Small Groups

By: **Wayne W. Zachary**

Presenter: **Saverio Giallorenzo**

Context • General Context

Anthropology

The **scientific study of humanity**, concerned with human:

- biology
- behaviour, societies, and culture
- linguistics

in both the present and past (archaeology).

Context • Specific Context

Social Anthropology

Social anthropology is the study of patterns of behaviour in human societies and cultures.

Social anthropology is different from the neighbouring fields of economics and sociology because of its holistic range and methods, based on long-term participant observation.

The field is characterised by a commitment to the relevance of micro studies and many social anthropologists use quantitative methods to objectively measure data collected through polls, questionnaires, and surveys, or by manipulating pre-existing statistical data using computational techniques.

Context • Specific Application

Zachary studies the problem of

Characterising (how) and explaining (why) group scission/fission takes place in small (bounded) groups

To do that, he presents data from a university-based karate-club group, in which a concrete political discussion led to an ideological fracture and eventually to a formal separation of the club into two organisations.

The political organisation of the club was informal and most decisions were made by consensus at club meetings. The two factors formed around the political rivalry between the club instructor and the manager.

Problem and Motivation

Problem: explaining how and why fission takes place in small bounded groups

Importance: a (back then) long central issue in social anthropology

Contributions of the paper:

- Present a new model to explain and characterise group fission, based on a social network approach;
- Present a measure, applied to the model, shown to be a good predictor of group membership and able to characterise the phenomenon (who goes where) — second part: the measure works also as split predictor;
- Present (network) data on a small group in which a factional division led to a formal separation into two organisations.

Data · Collection

Collected from a university-based karate club, in a period of three years.

During the collection, the club maintained between 50 and 100 members.

The data collected considered activities in which the club members attended both lessons and other social events (tournaments, parties, dances, banquets, etc.).

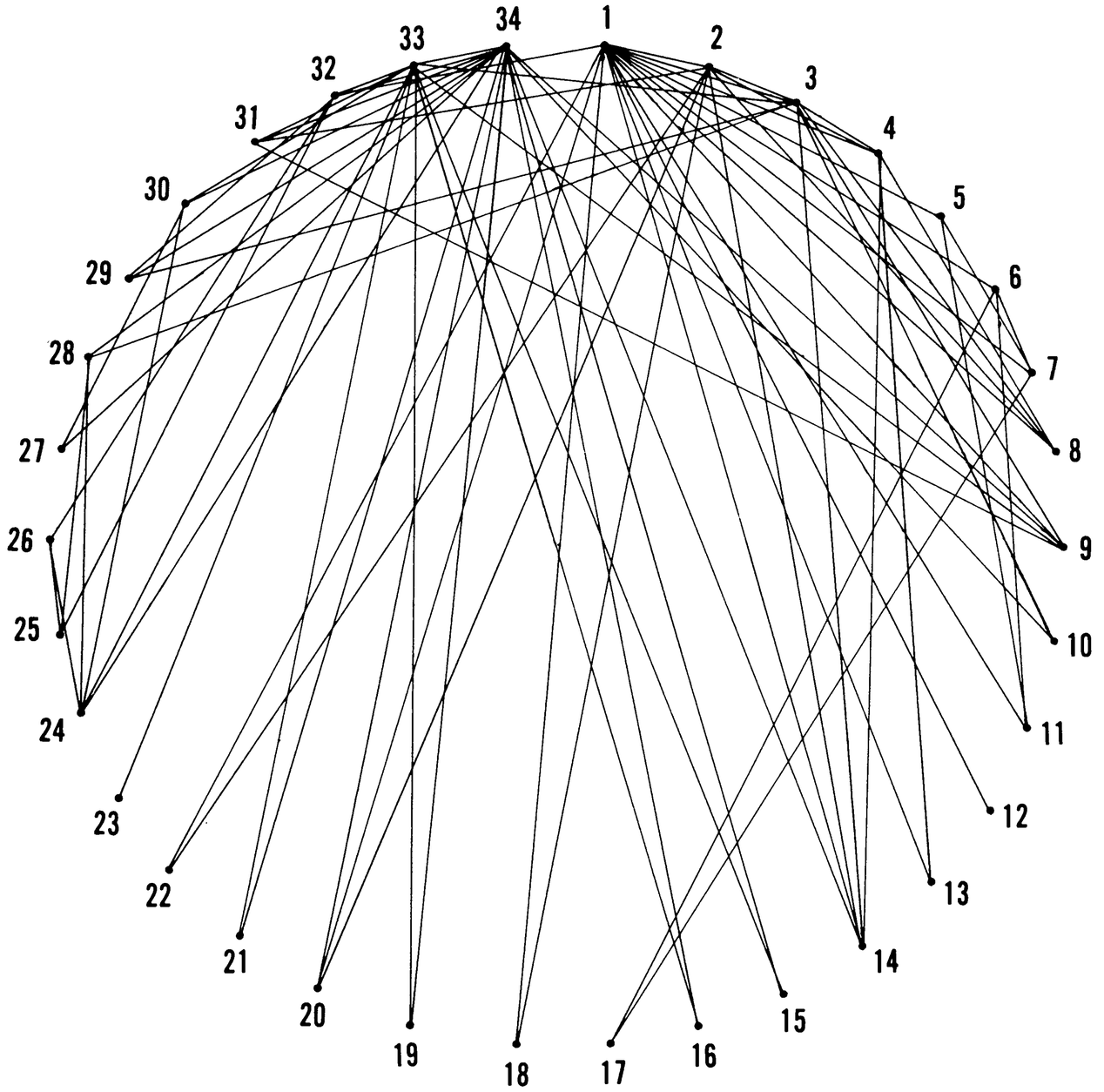
The data collected represent an information flow network proxied through co-attedance among the members of the club.

The network is a scalar one, where links between nodes are weighted and the weight is quantified by the number of events both nodes attended.

Data

Of the (fluctuating) total number of club members who joined and departed the club, only **34 individuals** are considered in the study (principled network boundary definition).

The reason is that the remaining members did not interact with other club members outside the context of meetings and classes.



Data • Measures

As a proxy of group division, Zachary employed the NETFLOW algorithm, which uses the maximum information flow between two given nodes to separate a given network between two groups, either closer to a *source* or a *sink* nodes.

The premise to use NETFLOW is that Zachary knows that the group could be torn apart by the political tension between two important nodes in the network: on one side the **manager** of the club and on the other the club **instructor**.

The hypothesis (we omit to present the second hypothesis on group-split determination) of Zachary is that the affiliation of a node to either faction can be determined by the NETFLOW algorithm, which implements the **maximum flow-minimum cut labelling procedure**.

Data • Measures, NETFLOW

NETFLOW uses the Ford-Fulkerson procedure to determine the maximum information flow between two nodes in the network.

Let $G = (V, E, C)$ be a graph with V vertices, E edges and pairwise flow-capacity C, the maximum flow between two nodes i (called source) and j (called sink) corresponds to the result of the algorithm `maxflow`(G, i, j), described by the pseudocode:

maxflow(G, i, j)

$G' \leftarrow G$

$flow_{ij} \leftarrow 0$

$p \leftarrow$ findAugmentingPath(G', i, j)

While $\exists\ p$

$flow_{ij} \leftarrow flow_{ij} + \min($ residual_capacity(G', p))

$G' \leftarrow$ computeResidualGraph(G', p)

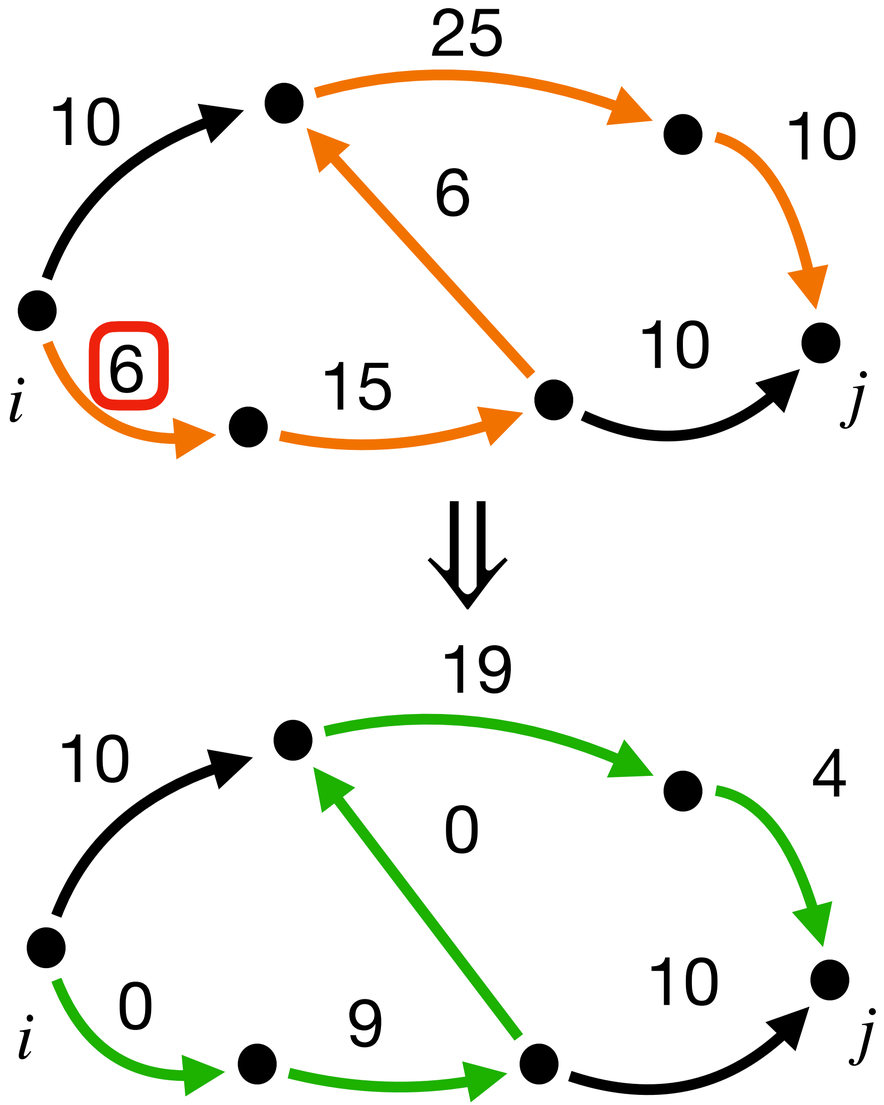
$p \leftarrow$ findAugmentingPath(G', i, j)

return $flow_{ij}$

e.g., via a depth-first search

the residual capacity of all the pairs of edges in p

Data • Measures, NETFLOW



maxflow(G, i, j)

$G' \leftarrow G$

$flow_{ij} \leftarrow 0$

$p \leftarrow$ findAugmentingPath(G', i, j)

e.g., via a depth-first search

While $\exists\ p$

$flow_{ij} \leftarrow flow_{ij} +$ min(residual_capacity(G', p))

the residual capacity of all the pairs of edges in p

$G' \leftarrow$ computeResidualGraph(G', p)

$p \leftarrow$ findAugmentingPath(G', i, j)

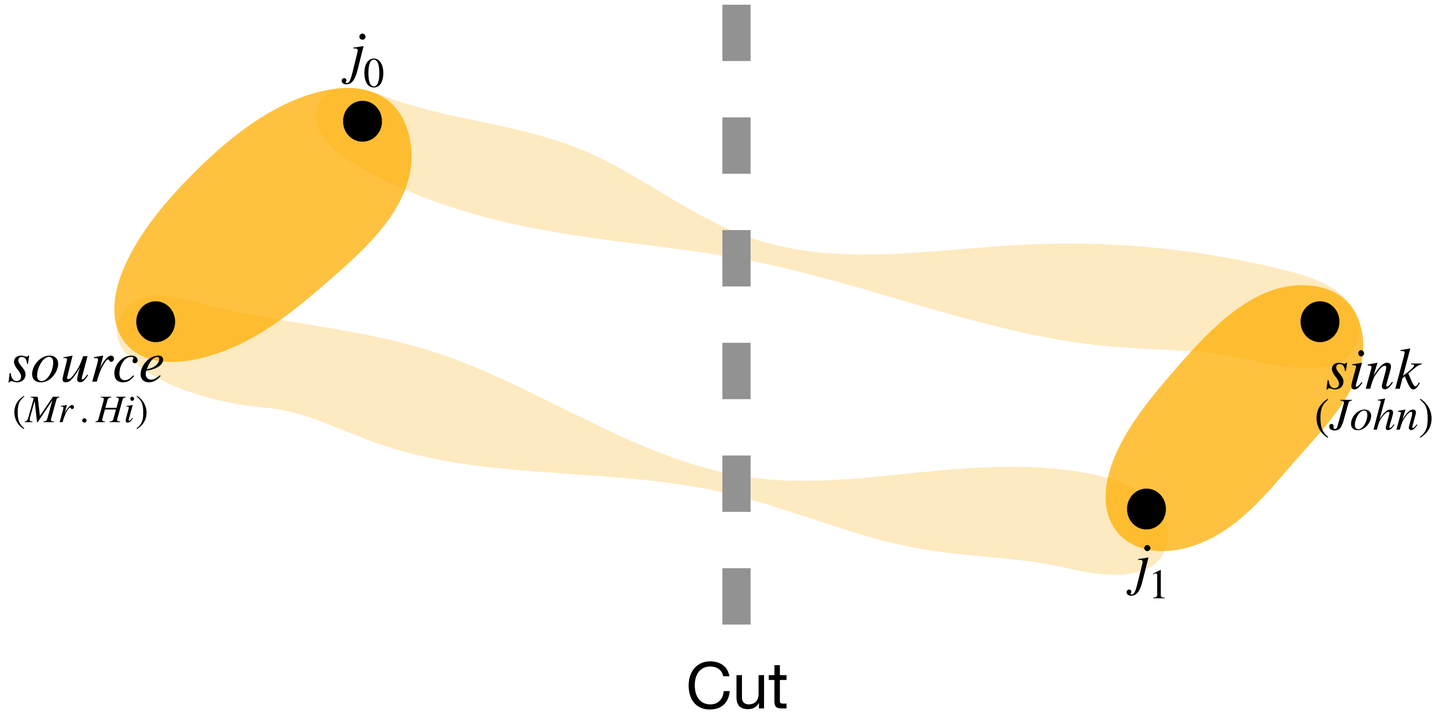
return $flow_{ij}$

Data • Measures, NETFLOW

NETFLOW uses maxflow to determine the maximum-flow-minimum-cut labelling between a *source* and a *sink* nodes. The labelling intuitively corresponds to labelling a given node i as having either a greater maxflow value from the *source* or towards the *sink* (here the distinction in moot, since the network is symmetrical)

$$maxflow_{source,\ j_0} > maxflow_{sink,\ j_0}$$

$$maxflow_{sink,\ j_1} > maxflox_{source,\ j_1}$$



Results

INDIVIDUAL NUMBER	SIDE OF CUT	FACTION	CLUB AFTER FISSION
1	Source	Mr. Hi - Strong	Mr. Hi's
2	Source	Mr. Hi - Strong	Mr. Hi's
3	Source	Mr. Hi - Strong	Mr. Hi's
4	Source	Mr. Hi - Strong	Mr. Hi's
5	Source	Mr. Hi - Strong	Mr. Hi's
6	Source	Mr. Hi - Strong	Mr. Hi's
7	Source	Mr. Hi - Strong	Mr. Hi 's
8	Source	Mr. Hi - Strong	Mr. Hi's
9	Sink	John - Weak	Mr. Hi's
10	Sink	None	Officers'
11	Source	Mr. Hi - Strong	Mr. Hi's
12	Source	Mr. Hi - Strong	Mr. Hi's
13	Source	Mr. Hi - Weak	Mr. Hi's

Results

From the application of the measure, the model (the data representing the analysed system) and the measure (NETFLOW) were 100% accurate in predicting faction membership (with respect to the membership data gathered from the surveyed individuals) — and, second hypothesis, 97% accurate in predicting club membership after the split.

INDIVIDUAL NUMBER IN MATRIX *C*	FACTION MEMBERSHIP FROM DATA	FACTION MEMBERSHIP AS MODELED	HIT/ MISS	CLUB AFTER SPLIT FROM DATA	CLUB AFTER SPLIT AS MODELED	HIT/ MISS
1	Mr. Hi	Mr. Hi	Hit	Mr. Hi's	Mr. Hi's	Hit
2	Mr. Hi	Mr. Hi	Hit	Mr. Hi's	Mr. Hi's	Hit
3	Mr. Hi	Mr. Hi	Hit	Mr. Hi's	Mr. Hi's	Hit
4	Mr. Hi	Mr. Hi	Hit	Mr. Hi's	Mr. Hi's	Hit
5	Mr. Hi	Mr. Hi	Hit	Mr. Hi's	Mr. Hi's	Hit
6	Mr. Hi	Mr. Hi	Hit	Mr. Hi's	Mr. Hi's	Hit
7	Mr. Hi	Mr. Hi	Hit	Mr. Hi's	Mr. Hi's	Hit
8	Mr. Hi	Mr. Hi	Hit	Mr. Hi's	Mr. Hi's	Hit
9	John	John	Hit	Mr. Hi's	Officers'	Miss
10	John	John	Hit	Officers'	Officers'	Hit
11	Mr. Hi	Mr. Hi	Hit	Mr. Hi's	Mr. Hi's	Hit
12	Mr. Hi	Mr. Hi	Hit	Mr. Hi's	Mr. Hi's	Hit
13	Mr. Hi	Mr. Hi	Hit	Mr. Hi's	Mr. Hi's	Hit
14	Mr. Hi	Mr. Hi	Hit	Mr. Hi's	Mr. Hi's	Hit
15	John	John	Hit	Officers'	Officers'	Hit
16	John	John	Hit	Officers'	Officers'	Hit
17	Mr. Hi	Mr. Hi	Hit	Mr. Hi's	Mr. Hi's	Hit
18	Mr. Hi	Mr. Hi	Hit	Mr. Hi's	Mr. Hi's	Hit
19	John	John	Hit	Officers'	Officers'	Hit
20	Mr. Hi	Mr. Hi	Hit	Mr. Hi's	Mr. Hi's	Hit
21	John	John	Hit	Officers'	Officers'	Hit
22	Mr. Hi	Mr. Hi	Hit	Mr. Hi's	Mr. Hi's	Hit
23	John	John	Hit	Officers'	Officers'	Hit

Critique

The solution is only partial to characterising (how) and explaining (why) group scission/fission takes place in small (bounded) groups.

Just one case, almost anecdotical, it does not provide a large-enough body of evidence to assess whether NETFLOW is a good predictor or not for small-group fission.

There is no permutation study to make sure the results are non-accidental.

The hypothesis needs more cases to strengthen its reliability.