

Multi-Data Mining for Understanding Leadership Behavior

CODATA2006, Oct. 23-25, 2006

Naohiro Matsumura and Yoshihiro Sasaki
Graduate School of Economics, Osaka University



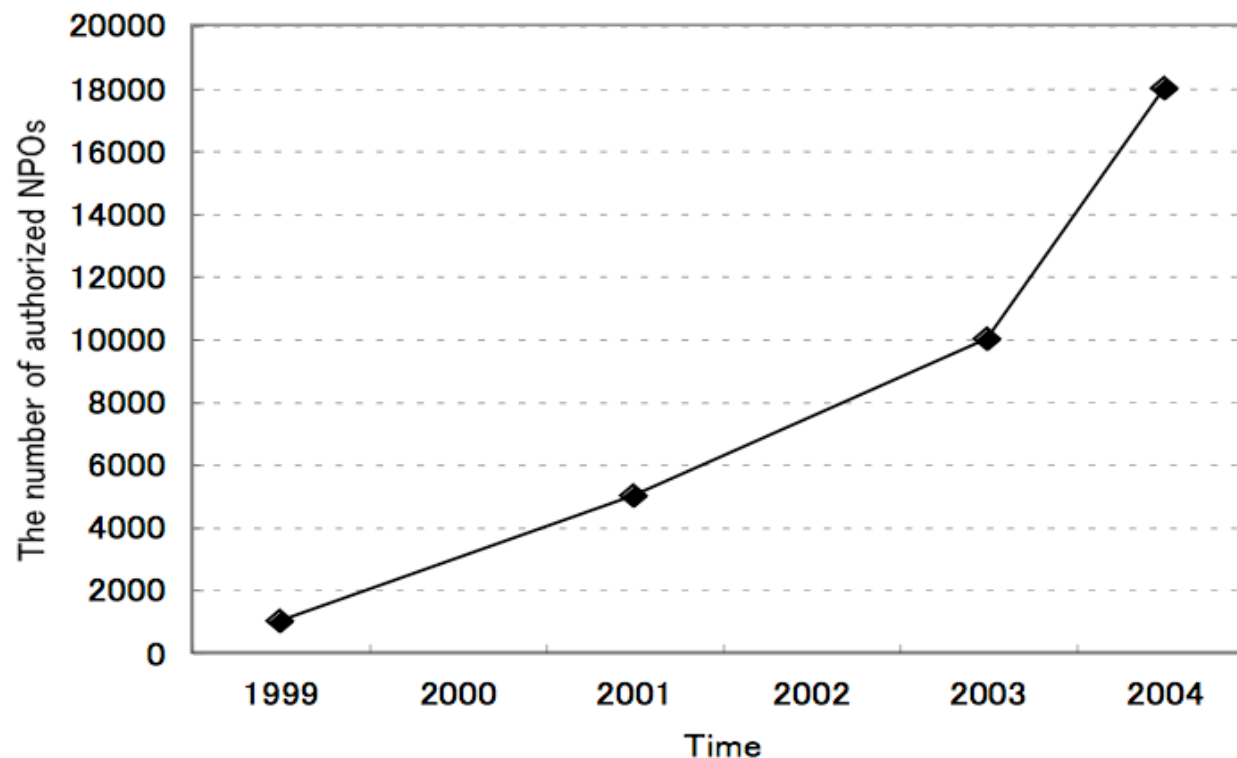
Outline

- Background
- Our approach: Multi-Data Mining
 - Questionnaire survey
 - E-mail mining
 - Integration of two approaches
- Results and Discussions
- Conclusions



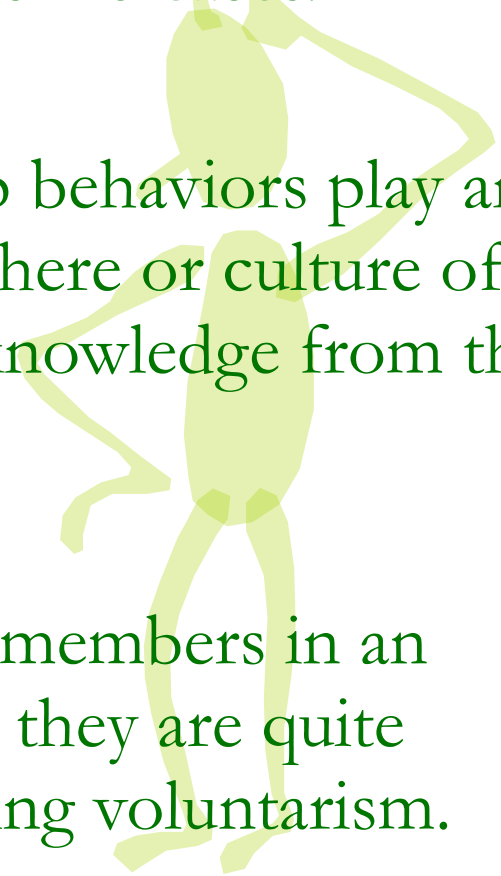
Background

- After NPO (non-profit organization) law was established in 1998 in Japan, more NPOs have been established than before.



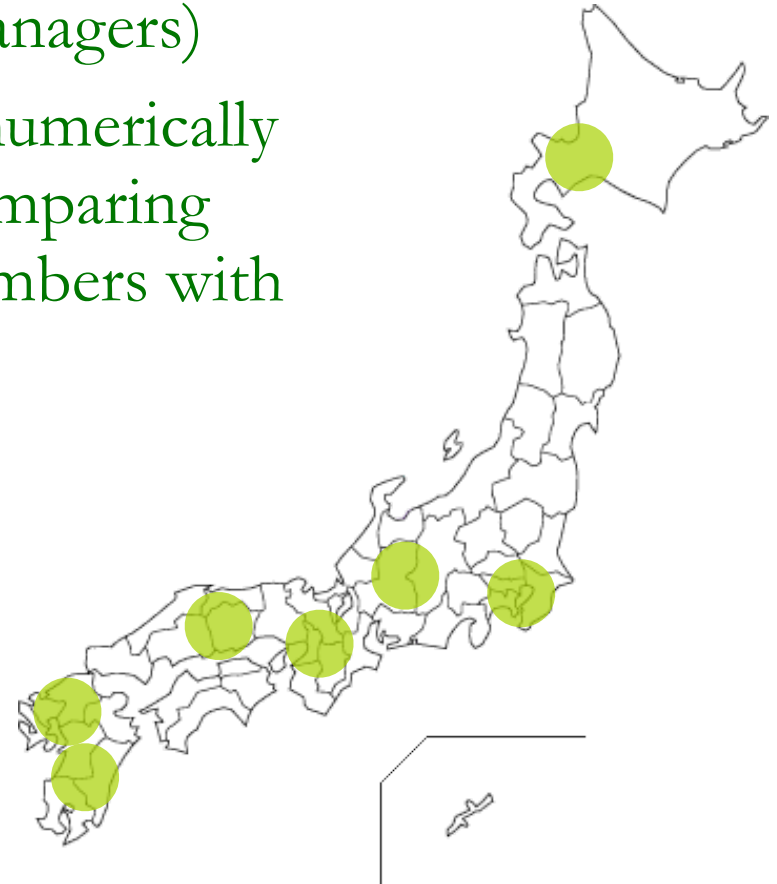
Difficulties in Managing NPO

- The commitment of staff members to an NPO depends on their **voluntarism**, not on obligation. However, keeping the voluntarism is difficult as the number of staff members increases.
- Much research has showed that leadership behaviors play an important role in determining the atmosphere or culture of an organization, and the ability to create knowledge from the experiences of staff members.
- However, the relationships between staff members in an NPO are not being investigated, although they are quite relevant for sharing missions and motivating voluntarism.



Overview of dot-jp

- The mission of “dot-jp” is to hold seminars and internship programs.
- Seven branch offices are distributed all over Japan.
- 104 staff members (including 7 area managers)
- The activity of a branch office can be numerically evaluated by **Achievement rate** (by comparing actual number of students and diet members with the desired number.)



Multi-Data Mining

- Conduct different data mining approach for different data, then integrate the results to create knowledge.
 - **Questionnaire** for understanding qualitative relations between staff members
 - **E-mail mining** for obtaining quantitative relations between staff members
- Reliable leadership behavior is derived from the cross-point of the results.

Questionnaires

- We sent questionnaires to all the staff members (104 in total) in dot-jp, and the results were received from 97 staff members (completion rate was 93%)

- Questions:
 - *Q-1.* Please rate your degree of satisfaction with your branch office.
 - *Q-2.* Please rate your degree of satisfaction with the area manager in your branch office.
 - *Q-3.* Please list up to three substantive leaders in your branch office.

(All answers are selected from five alternatives (1: Very dissatisfied, 2: Dissatisfied, 3: Neutral, 4: Satisfied, 5: Very Satisfied)).

Questionnaire Results and Achievement Rate

	Averaged satisfaction	Achievement rate (%)
Area Manager in Branch A	5.00	135
Area Manager in Branch B	4.37	109
Area Manager in Branch C	3.00	106
Area Manager in Branch D	4.50	97
Area Manager in Branch E	2.75	99
Area Manager in Branch F	3.69	79
Area Manager in Branch G	4.50	69

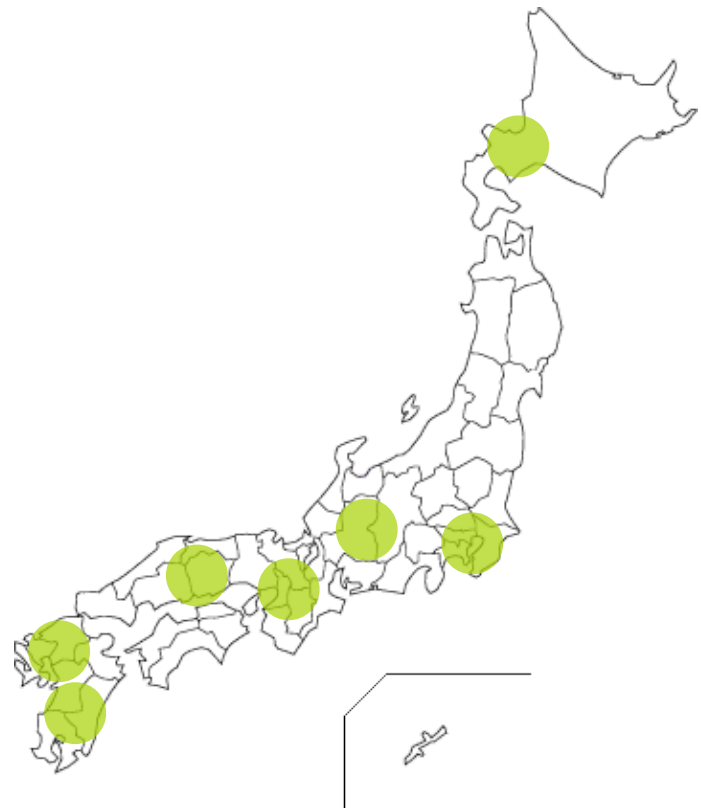
1:Very dissatisfied, 2:Dissatisfied, 3:Neutral, 4:Satisfied, 5:Very satisfied

Mining E-mail Archives

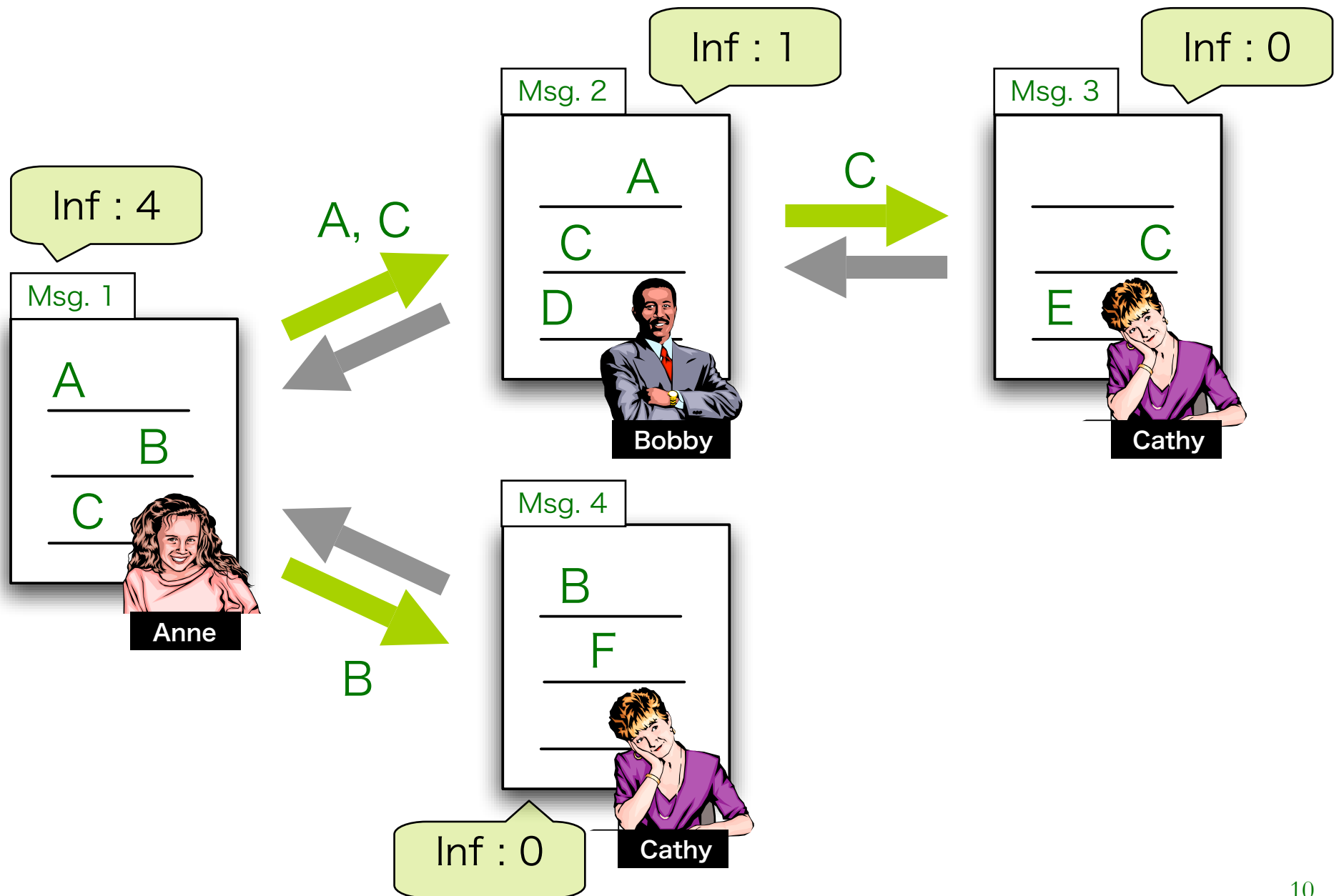
- To investigate further, we apply IDM for mining e-mail archives to extract human influence network.

Branch	# of members	# of e-mails
Branch-A	21	2297
Branch-B	9	1198
Branch-C	16	2465
Branch-D	14	2076
Branch-E	16	3258
Branch-F	16	1309
Branch-G	12	1717

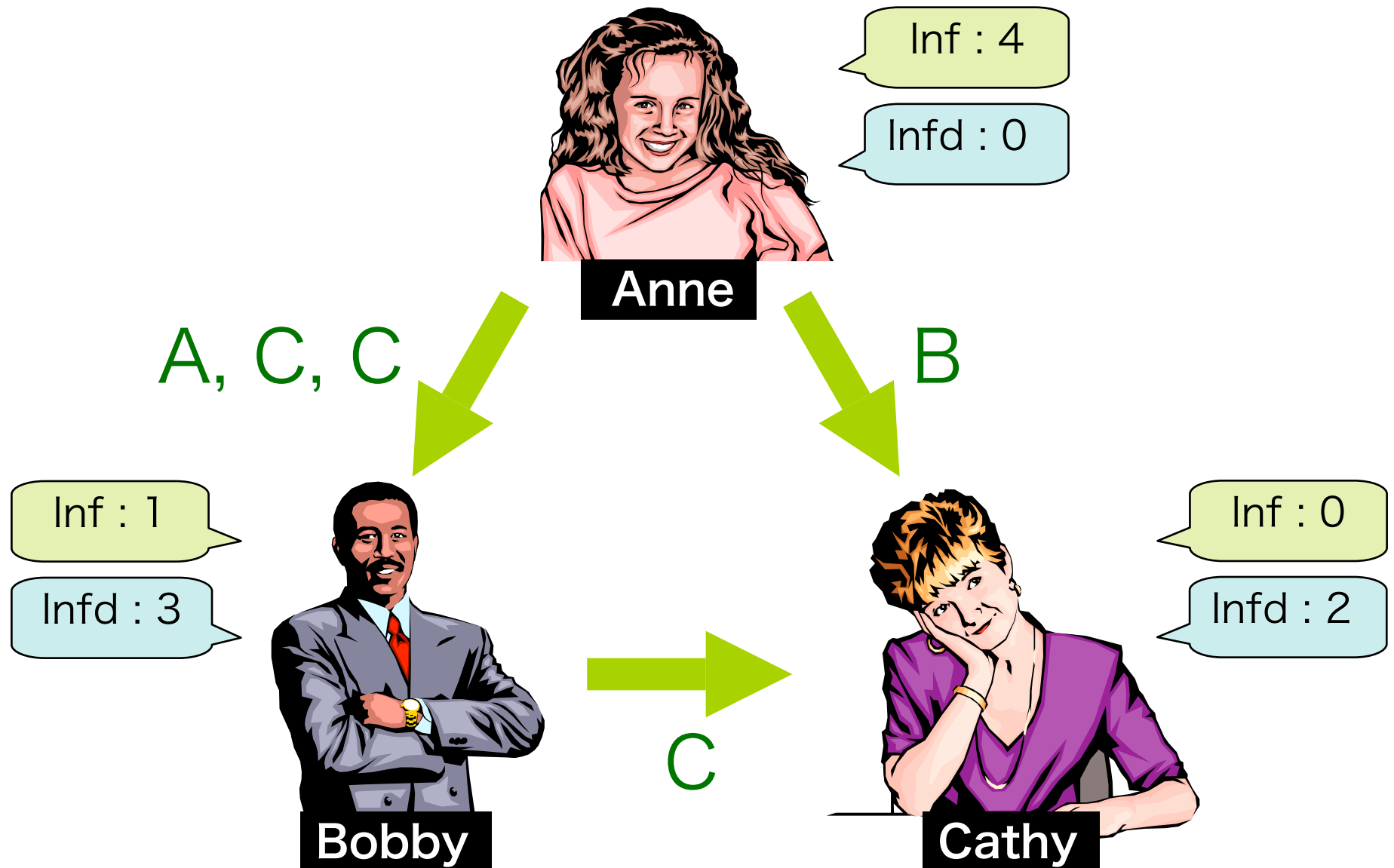
(Oct. 2004 – Mar. 2005)



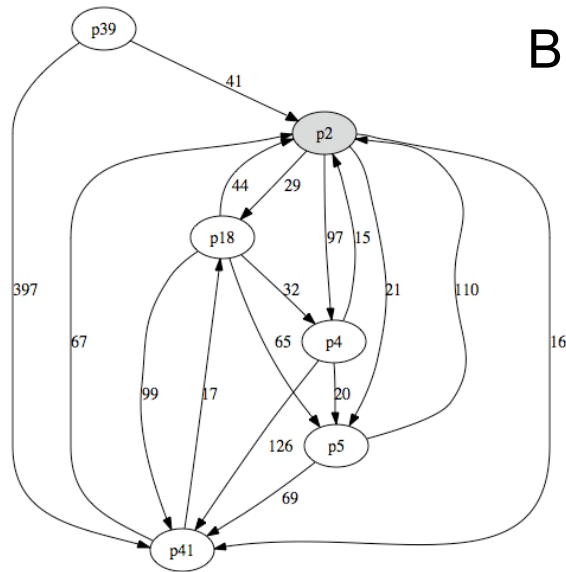
IDM Algorithm



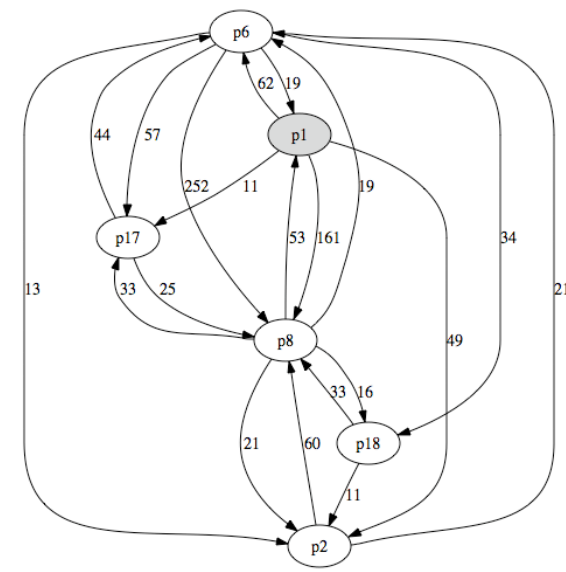
Human Influence Network



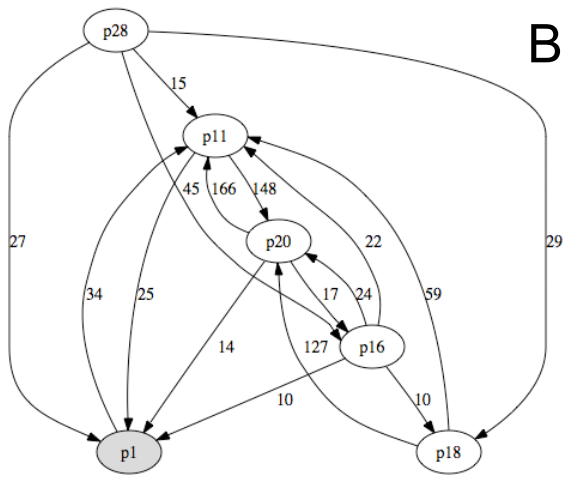
Examples of Influence Human Networks



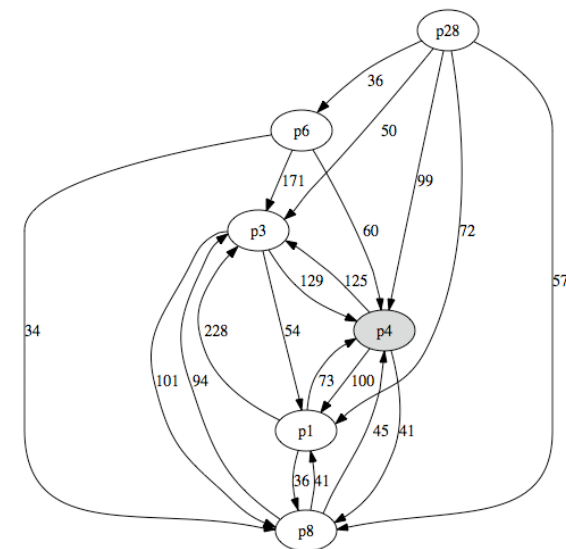
Branch A



Branch C



Branch F



Branch-G

Inf/InfD Values for Each Branch Offices

Staff in Branch A	$I_p^{<out>}$	$I_p^{<in>}$	Staff in Branch B	$I_p^{<out>}$	$I_p^{<in>}$	Staff in Branch C	$I_p^{<out>}$	$I_p^{<in>}$
p2*	541	827	p1*	841	385	p6	1113	403
p39	518	0	p9	772	1051	p1*	836	142
p18	506	177	p5	542	317	p8	515	1257
p4	470	244	p2	484	666	p18	490	95
p41	454	910	p11	304	198	p2	470	163
p5	417	235	p17	251	537	p17	405	261

Staff in Branch D	$I_p^{<out>}$	$I_p^{<in>}$	Staff in Branch E	$I_p^{<out>}$	$I_p^{<in>}$	Staff in Branch F	$I_p^{<out>}$	$I_p^{<in>}$	Staff in Branch G	$I_p^{<out>}$	$I_p^{<in>}$
p1*	1495	717	p4	1022	8	p11	578	524	p3	779	1270
p4	923	68	p3	1019	1032	p20	506	871	p4*	559	623
p13	893	1880	p11	754	563	p18	432	52	p6	537	36
p3	794	927	p7	737	664	p1*	404	263	p8	521	466
p5	790	386	p1*	686	849	p16	355	98	p1	513	480
p14	655	205	p2	560	491	p28	325	0	p28	445	0

Ranking of Area Managers

	Vote*	Inf	Infd
Area Manager in Branch A	1st	1st	3rd
Area Manager in Branch B	2nd	1st	4th
Area Manager in Branch C	N/A	2nd	14th
Area Manager in Branch D	2nd	1st	6th
Area Manager in Branch E	4th	5th	4th
Area Manager in Branch F	N/A	4th	8th
Area Manager in Branch G	1st	2nd	3rd

* Q-3. Please list up to three substantive leaders in your branch office.

Quick Reference Matrix for Reliable Leadership Behavior

	High Inf	Middle Inf	Low Inf
High Inf	Branch A (1st by Vote)	Branch G (1st by Vote)	Branch F (N/A by Vote)
Middle Inf	Branch B, D (both 2nd by Vote)	Branch E (4th by Vote)	
Low Inf	Branch C (N/A by Vote)		

High: 1st, 2nd, 3rd Middle: 4th, 5th, 6th Low: 7th, 8th, 9th, ...

Conclusions

- By integrating qualitative analysis with quantitative analysis, we obtained a quick reference matrix for reliable leadership behavior.
- The discoveries might be applicable only in an NPO. Further investigation is necessary for other types of organization.

