

## Introduction

Collaboration is playing an increasingly important role in scientific research because an individual scholar may not possess all the expertise and knowledge to address complex research issues. This work proposes to investigate the question of whether conference will breed new scientific collaboration based on the focal closure theory. Focal closure theory means that new connections may emerge between people who have joined in the same community. The basic idea of triadic closure and focal closure can be seen from Figure 1. Through the analysis of conference closure on individual and community level, we show that academic conference can promote new scientific collaborations, and conferences with more attendees and higher field ratings bring more new scientific collaborations.

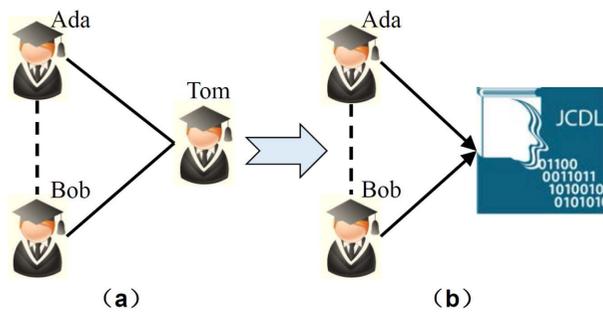


Fig.1 (a) Triadic closure;  
(b) Focal closure based on conference.

## Methods

Our study utilizes 12 influential conferences in Data Mining field from the computer science bibliography DBLP. We first extracted all the authors who have published papers in these conferences in 2010. The basic features of selected conferences are presented in Table 1.

Conference	Founded time	Field rating <sup>1</sup>	No. of authors	No. of papers
FSKD	2002	11	1703	635
CIKM	1992	67	1103	374
ICDM	2001	56	960	353
ICDE	1984	104	660	217
KDD	1995	122	414	125
ICETET	2008	7	403	166
ADMA	2005	10	336	118
PAKDD	1998	33	292	100
SDM	2001	45	243	82
InCDM	2001	9	207	73
ICWSM	2007	19	197	72
DMIN	2005	7	165	62

<sup>1</sup> Field rating is got from Microsoft Academic Search.

Table 1: Statistics of selected conferences in 2010.

Conference closure means that two academic strangers who have not coauthored before are possibly to collaborate after attending a same conference. The conference closure can be studied at two different levels: (i) individual level and (ii) community level.

### ◆ Individual Level

$$CC_{\text{indiv}} = \sum_{i=1}^n \frac{N_i^c}{nN_i}$$

where,  $N_i$  is the total number of new collaborators of scholar  $i$  after attending conference  $C$ ,  $N_i^c$  is the number of scholar  $i$ 's new collaborators who simultaneously attend  $C$ , and  $n$  is the total number of attendees of  $C$ .

### ◆ Community Level

$$CC_{\text{com}} = \frac{N^c}{n}$$

where  $N^c$  is the number of scholars who cooperate with other unfamiliar attendees of  $C$ .

## Simulation and Results

Conference	2011	2012	2013	2014	2015
FSKD	0.061	0.052	0.039	0.041	0.032
CIKM	0.077	0.063	0.053	0.029	0.032
ICDM	0.061	0.061	0.045	0.023	0.023
ICDE	0.059	0.052	0.052	0.035	0.021
KDD	0.048	0.055	0.044	0.021	0.017
ICETET	0.029	0.006	0.010	0.013	0.000
ADMA	0.038	0.019	0.009	0.005	0.009
PAKDD	0.032	0.022	0.012	0.020	0.007
SDM	0.035	0.023	0.020	0.010	0.012
InCDM	0.028	0.019	0.023	0.021	0.009
ICWSM	0.071	0.049	0.025	0.031	0.013
DMIN	0.018	0.018	0.009	0.013	0.004

Table 2: Shown the ratios of conference closure at individual level in next 5 years.

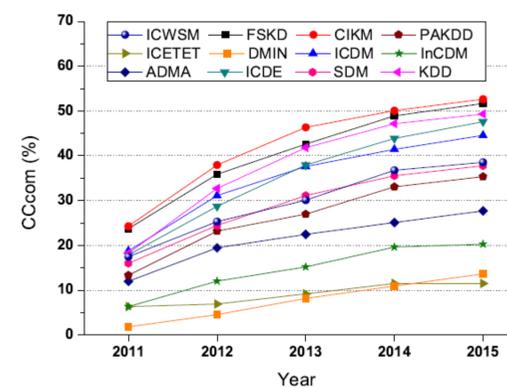


Fig.2 Shown the ratios of conference closure at community level in next 5 years.

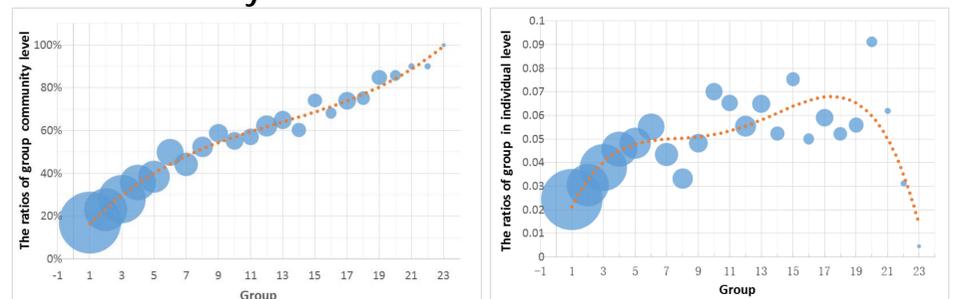


Fig.3 Shown the average ratio of conference closure at community and individual level in different group. Scholars are grouped according to the papers they published.

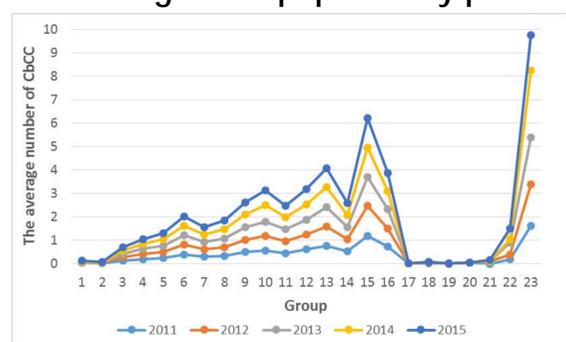


Fig.4 Shown the average number of collaborator born of conference closure of different group in next 5 years.

## Conclusion

The theory behind our analysis is that new links may emerge between people who have joined in the same community. We verified this phenomenon based on the behavior of co-attending a conference. In future work, we will propose new mechanisms to promote scientific collaboration based on this verified theory.