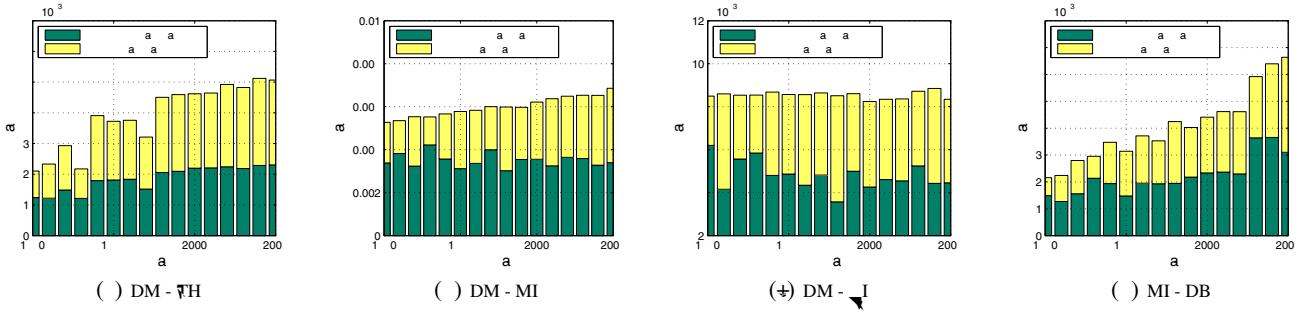


# Cross-domain Collaboration Recommendation

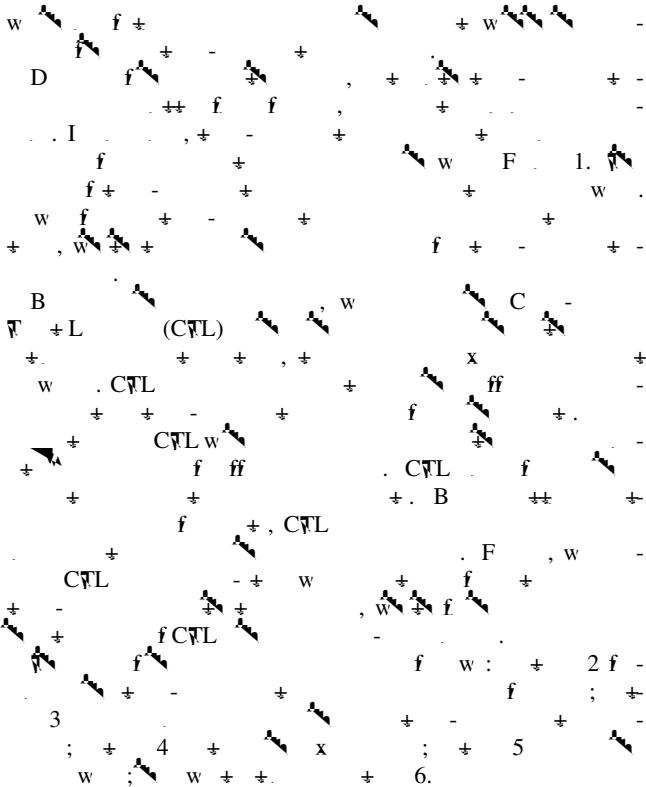
a<sup>†</sup>, a<sup>†</sup>, u<sup>†</sup>, u<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, u<sup>†</sup>, u<sup>†</sup>  
a<sup>†</sup>, a<sup>†</sup>, u<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, u<sup>†</sup>, u<sup>†</sup>, u<sup>†</sup>  
a<sup>†</sup>, a<sup>†</sup>, u<sup>†</sup>, u<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, u<sup>†</sup>, u<sup>†</sup>, u<sup>†</sup>  
a<sup>†</sup>, a<sup>†</sup>, u<sup>†</sup>, u<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, u<sup>†</sup>, u<sup>†</sup>, u<sup>†</sup>  
a<sup>†</sup>, a<sup>†</sup>, u<sup>†</sup>, u<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, u<sup>†</sup>, u<sup>†</sup>, u<sup>†</sup>  
a<sup>†</sup>, a<sup>†</sup>, u<sup>†</sup>, u<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, a<sup>†</sup>, u<sup>†</sup>, u<sup>†</sup>, u<sup>†</sup>

## ABSTRACT

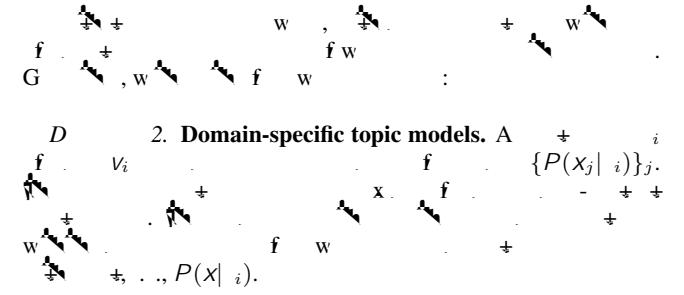
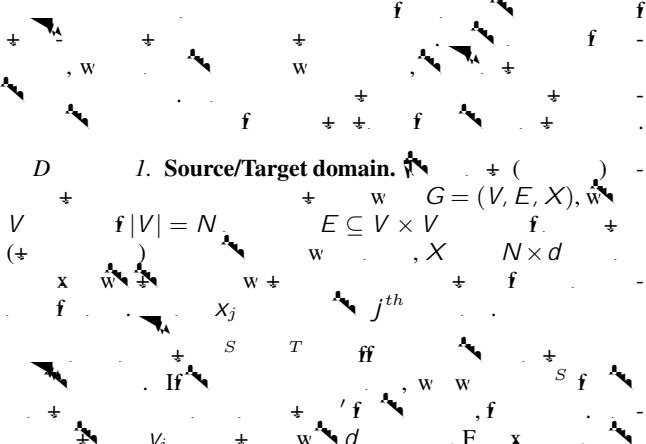
I  
+ . H w , f  
+ - + ? H w f ? C w +  
C - + + ff : 1) sparse  
connection: + - + f ; 2) complementary expertise: + - + ff x-  
; 3) topic skewness: + - +  
f I , w f + . A  
f C T + L (CTL) , CTL +  
F F , w , CTL , w  
f + x , , CTL , w + -



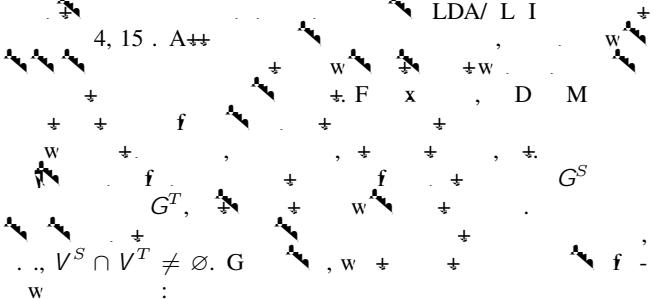
**Figure 1:** The comparison of existing collaboration and new collaboration trends over years. DM - Data Mining domain; MI - Medical Informatics domain; TH - Theory domain; VIS - Visualization domain; DB - Database domain. The trends of cross-domain collaborations in all but one case are growing (The exception between DM and VIS remain roughly constant over time). Newly formed cross-domain collaborations are significantly in all cases.



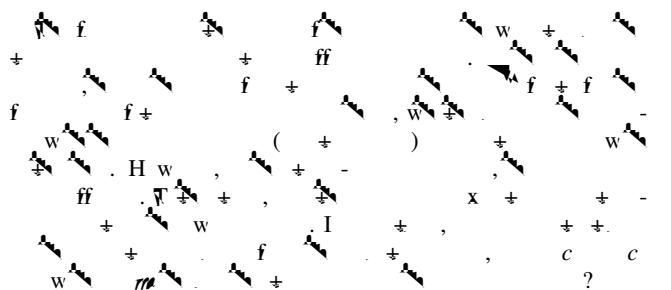
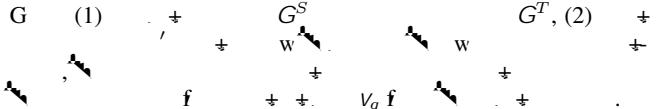
## 2. PROBLEM DEFINITION



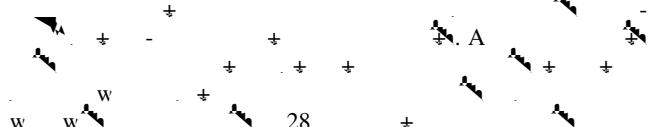
### 2. Domain-specific topic models.

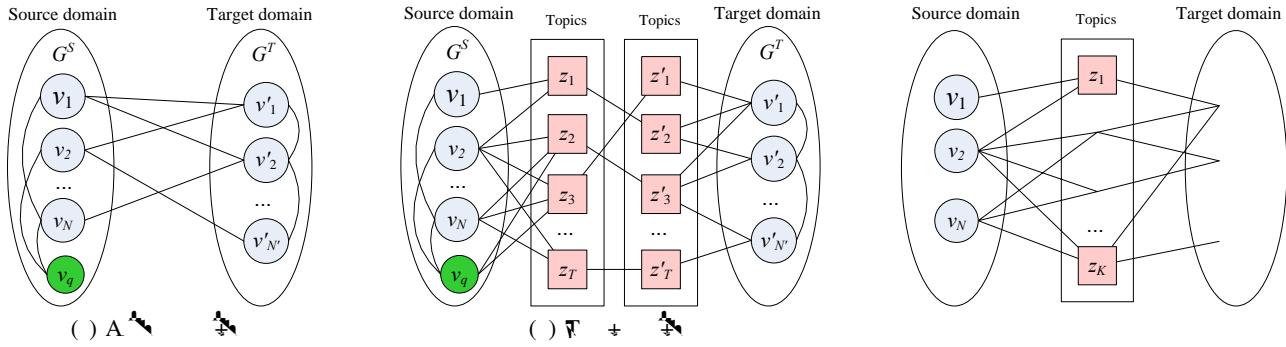


### 3. Cross-domain collaboration recommendation.



## 3. CROSS-DOMAIN TOPIC LEARNING





**Input:**  $G^S$ ,  $G^T$ ,  $\theta, \theta', \phi, \vartheta, \lambda$

**Output:**  $f$ ,  $w$

```

I   ACT $_{G^S}$   $f$   $+ w$ 
, ACT $_{G^S}$   $f$   $+ G^T;$ 
foreach  $c$   $ab$   $a d d c m d$  do
  foreach  $d$   $di \in d$  do
    if  $di = 0$  then
       $D_w \vartheta_{vv'} \sim l_i(\vartheta_{vv'}) f$ 
    end
    if  $di = 1$  then
       $D_w \vartheta_{vv'} \sim l_i(\theta_v) f$ 
    end
  end
   $D_w \vartheta_{vv'} \sim l_i(\phi_{z_{di}}) f$ 
end

```

Algorithm 1: CTL.

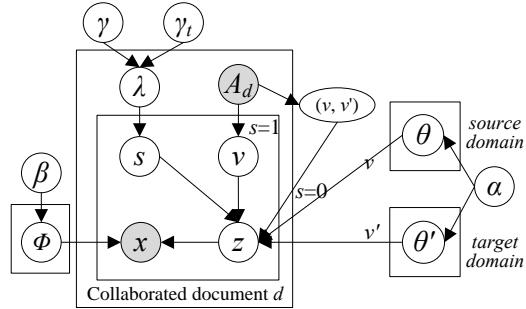
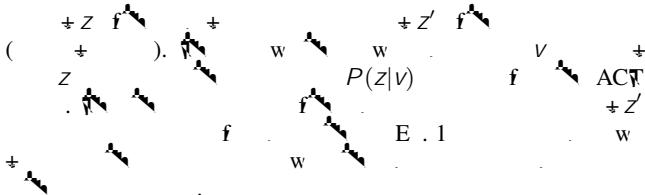
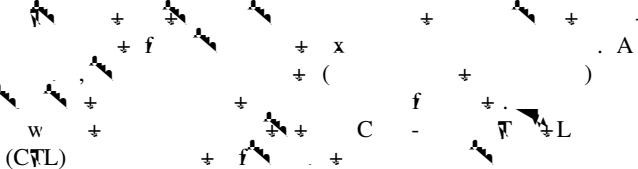


Figure 3: Graphical representation of CTL model.



### 3.3 Cross-domain Topic Learning (CTL)



#### Model description.

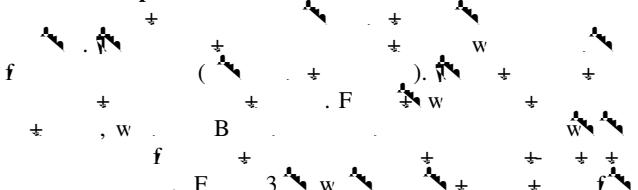
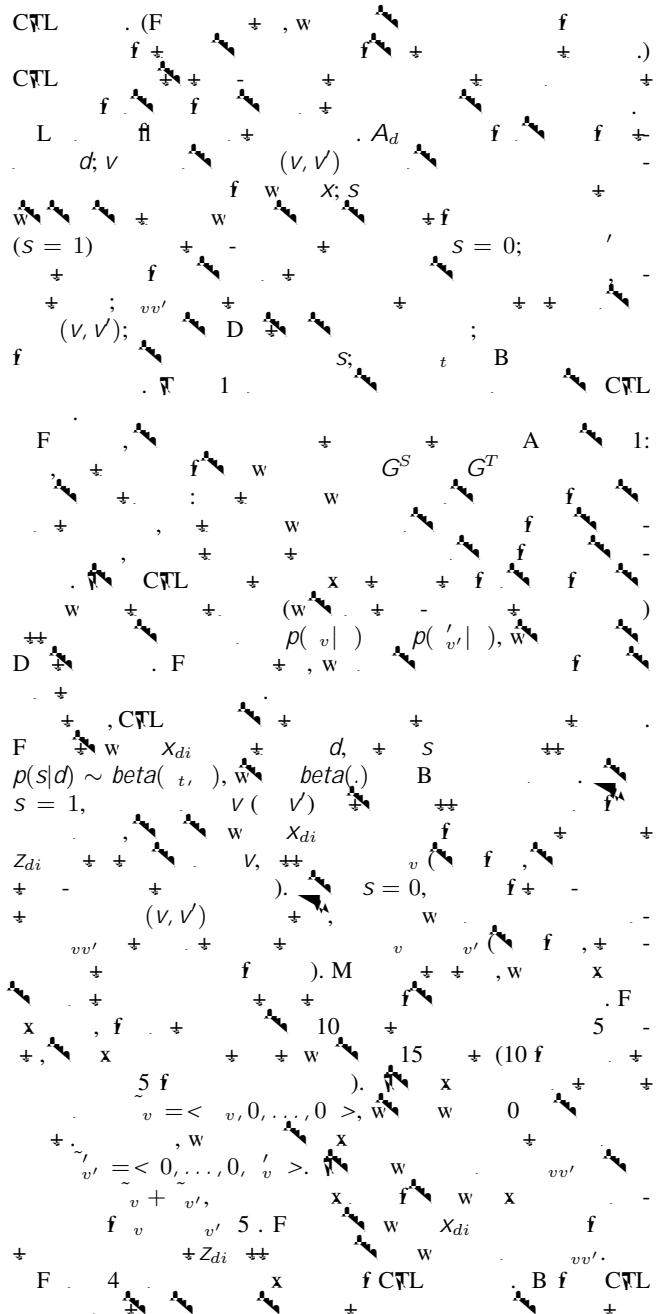
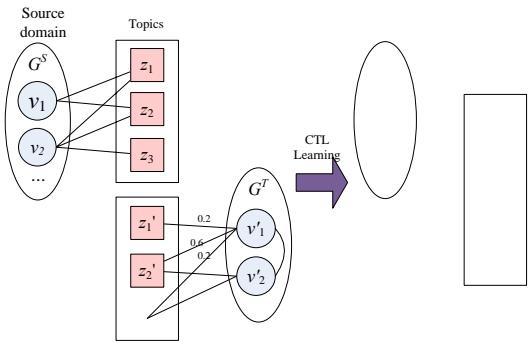


Table 1: Notations in the CTL model.

MB	L	DE	C	I	VI
$T$		$f$	$+$		
$d$	$+$		$+$		
$A_d$		$f$	$f$	$+$	$d$
$di$		$i$	$(w)$	$+$	$d$
$di$		$+$			$di$
$di$		$f$	$di$	$w$	$f$
$\theta_v$				$+$	$+$
$\vartheta_{vv'}$			$(,')$		$+$
$\phi_z$				$w$	$+$
$\alpha, \beta$		$D$			$\theta, \theta' \phi$
$\lambda$			$f$		
$\gamma, \gamma_t$		$B$			$\lambda$





1,932,442 + -  
f 1990 2005. w f w x - :  
+ f + : KDD, DM, ICDM, DM KDD

22,862 + -

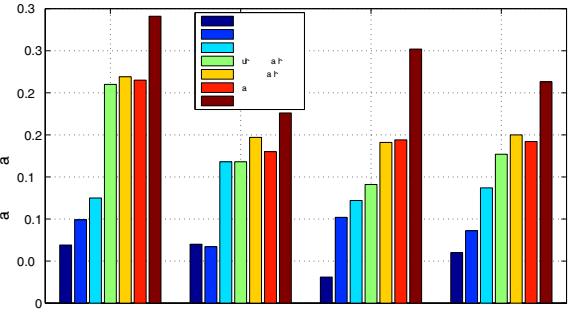
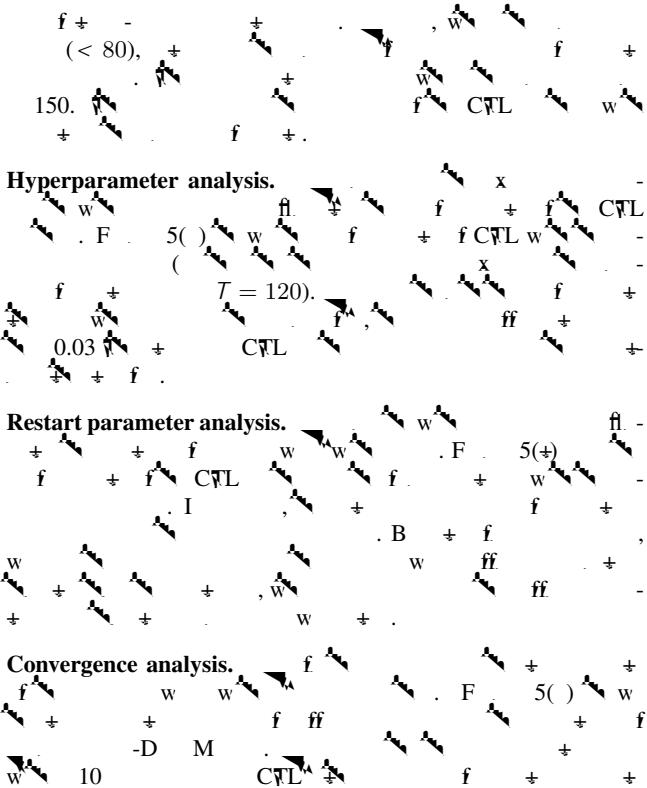
• Data Mining:  
J f A + M + I f + A + ,  
J f B + I f + , A + I +  
M + IEEE f . M I IEEE f +  
I f f 9,150 B + , f  
w w 31,851 + -

• Medical Informatics:  
F C DA, f w w + f + , .., f C,  
27,712 + -

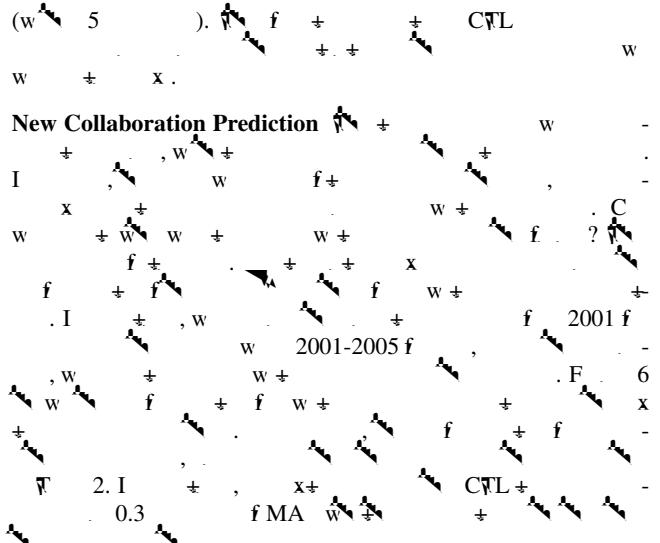
• Theory:  
+ f w + f +

**Table 2: Recommendation performance by different methods on the four cross-domain test cases (%). Content – Content Similarity; CF – Collaborative Filtering; Author – Author Matching; Topic – Topic Matching.**

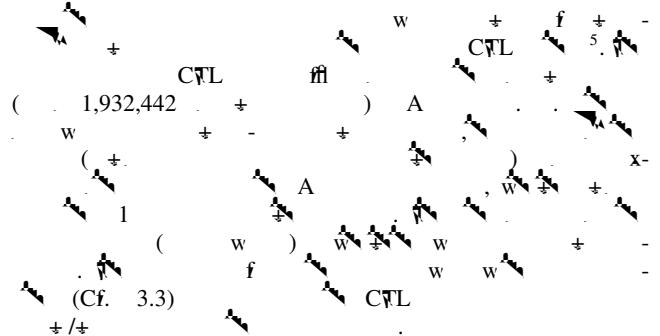
Cross domain	ALG	P@10	P@20	MAP	R@100	ARHR -10	ARHR -20
D ( )	C	10.3	10.2	10.9	31.4	4.9	2.1
	CF	15.6	13.3	23.1	26.2	4.9	2.8
	H	17.4	19.1	20.0	29.5	5.0	2.4
	A	27.2	22.3	25.7	32.4	10.1	6.4
	F +	28.0	26.0	32.4	33.5	13.4	7.1
	K	30.4	29.8	31.6	27.4	11.2	5.9
	CTL	<b>37.7</b>	<b>36.4</b>	<b>40.6</b>	<b>35.6</b>	<b>14.3</b>	<b>7.5</b>
M + I f . ( )	C	10.1	10.9	12.5	45.9	3.6	2.1
	CF	18.3	20.2	21.4	47.6	5.3	3.9
	H	25.0	26.5	28.4	59.1	6.4	4.2
	A	26.2	29.6	32.2	54.8	10.5	<b>5.4</b>
	F +	29.4	26.3	34.7	59.3	<b>11.5</b>	5.2
	K	27.5	28.3	30.7	57.2	10.5	5.0
	CTL	<b>32.5</b>	<b>30.0</b>	<b>36.9</b>	<b>59.8</b>	11.4	<b>5.4</b>
M + I f . ( )	C	5.8	5.7	9.5	19.8	1.9	0.9
	CF	13.7	17.8	18.9	34.3	2.7	1.3
	H	18.0	19.0	19.8	36.7	3.4	1.3
	A	20.1	23.8	29.3	<b>64.4</b>	5.3	2.1
	F +	26.0	<b>25.0</b>	33.9	48.1	10.7	5.6
	M	21.2	23.8	32.4	48.1	10.2	4.8
	CTL	<b>30.0</b>	24.0	<b>35.6</b>	49.6	<b>12.2</b>	<b>6.0</b>
M (F)	C	9.6	11.8	13.2	18.9	3.1	1.8
	CF	14.0	20.8	26.4	29.4	6.9	4.3
	H	16.0	20.0	27.6	30.1	6.3	4.4
	A	22.0	25.2	27.7	31.1	11.9	6.7
	F +	26.3	25.0	32.3	31.4	13.2	8.8
	K	23.0	25.1	29.3	30.2	10.4	5.4
	CTL	<b>28.3</b>	<b>26.0</b>	<b>32.8</b>	<b>36.3</b>	<b>14.0</b>	<b>9.1</b>



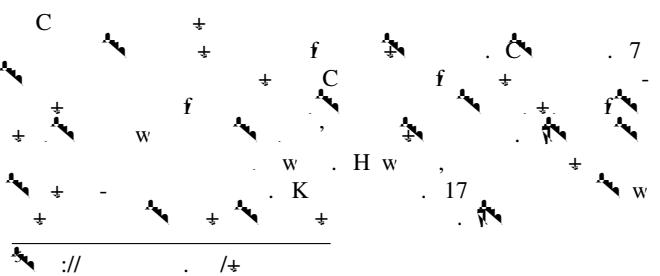
**Figure 6: Performance on new collaboration prediction of all algorithms.**



### 4.3 Prototype System



## 5. RELATED WORK





- 22 L. L. . w : A . . C m b a c ,  
2(1):1?6, 1993.
- 23 D. M. A. M+C . Ex f .  
w w . I KDD'07, 500 509, 2007.
- 24 J. + . M+ . N  
E a d J a M d c , 354:2463 2472, J 2006.
- 25 D. + . C . I KDD'10,  
979 988, 2010.
- 26 . , D. A. G . , . A + .  
f + . w . I SIGIR'11, 1337 1338, 2011.
- 27 M. f f . I KDD'04, + - +  
f f . I KDD'04, 306 315, 2004.
- 28 J. , H. , D. C . , C. F . . I ICDM'05,  
f + . I ICDM'05, 418 425, 2005.
- 29 J. T , L. , D. C . M . - + .  
I SDM'09, 1147 1158, 2009.
- 30 J. T , J. , J. , , K. C , L. , . T +  
x w . Mac L a J a , 82(2):211 237, 2011.
- 31 J. T , J. , L. , J. L. , , . A :  
Ex + f + + + w . I KDD'08,  
990 998, 2008.
- 32 L. T H. L . + .  
I KDD'09, 817 826, 2009.
- 33 T , J. T T. L. C. T , B. G , T. L .  
x w . N c m ,  
76(1):71 83, 2012.
- 34 C. D. M. B. C f .  
+ + + + . I KDD'11, 448 456, 2011.
- 35 . , L. C . , . F + . . :  
f + . I R cS '11, 245 252, 2011.
- 36 J. , J. T , J. L. Ex + w . I  
DASFAA'07, 1066 1069, 2007.

## 8. APPENDIX

A++ (B ) + , w + .