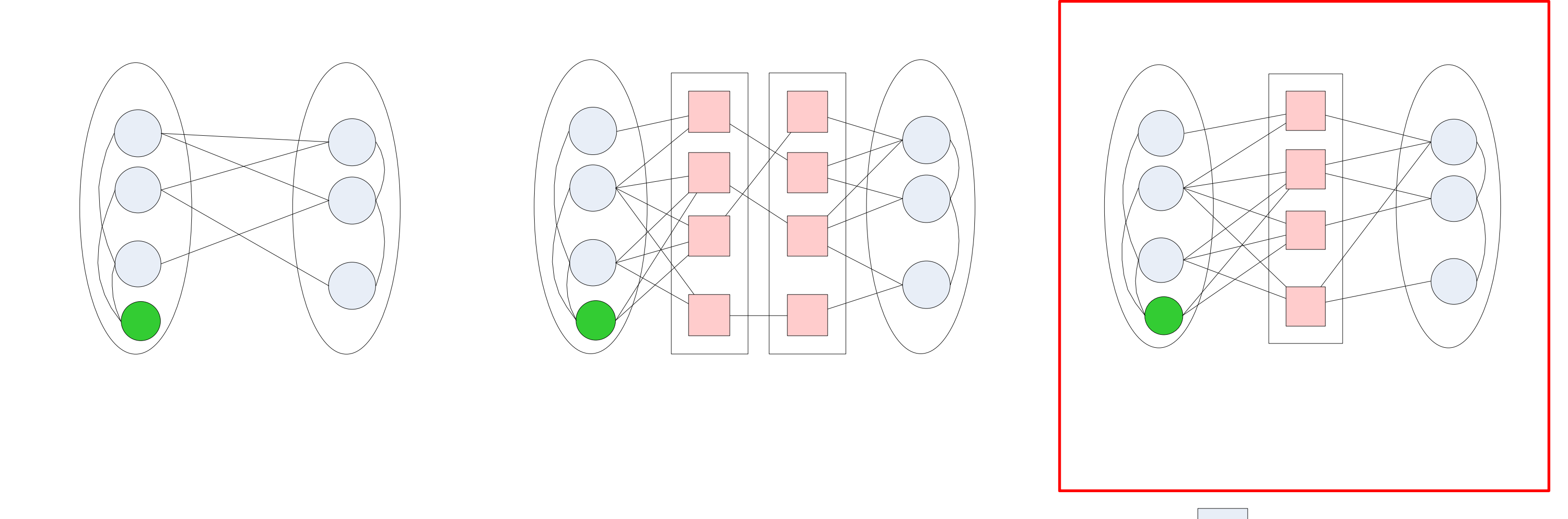


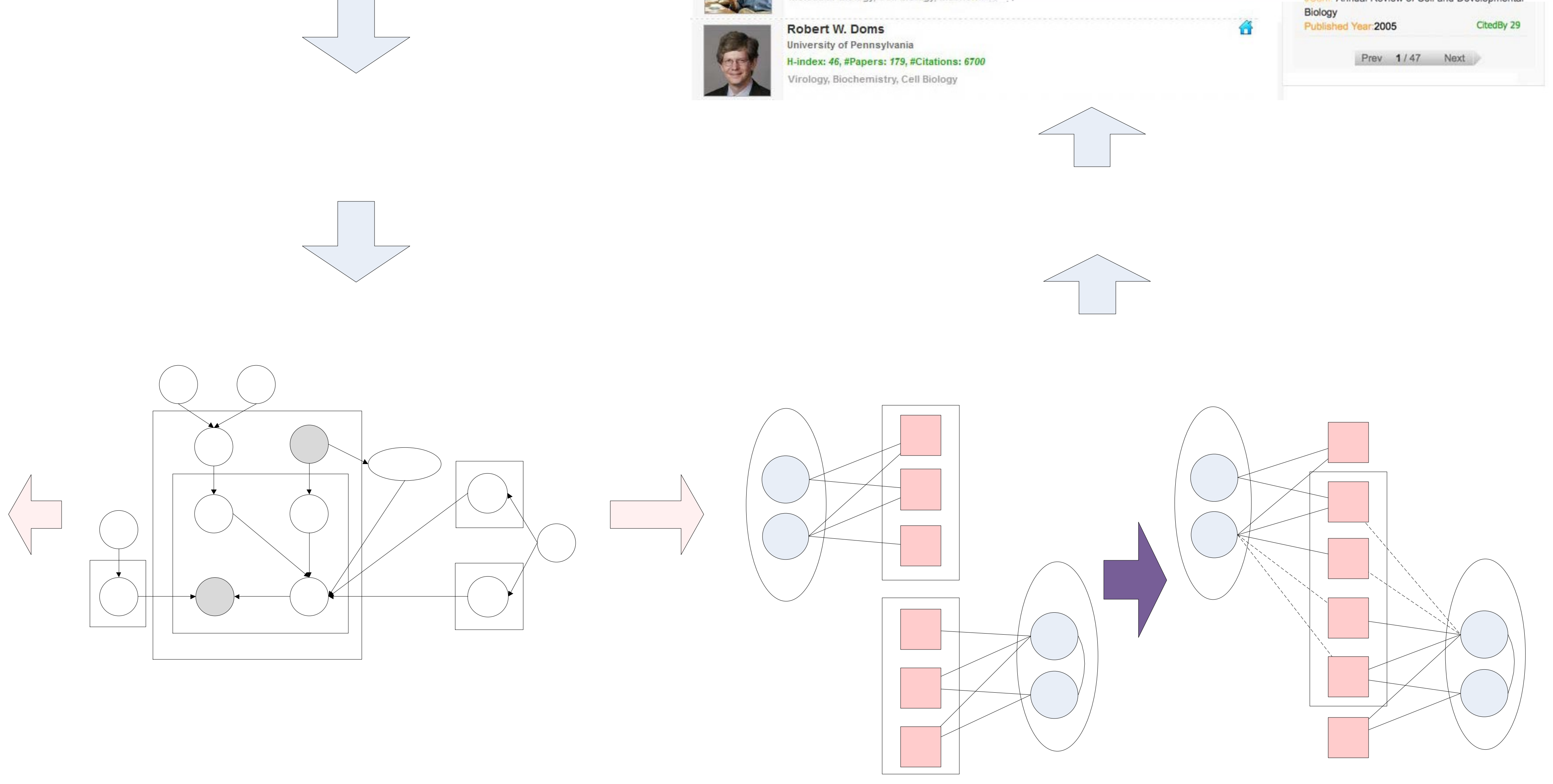
Cross-domain Topic Learning (CTL)



SYMBOL	DESCRIPTION
T	number of topics
A_d	set of authors of domain d
$\phi_{z,d}$	multinomial distribution over words specific to topic z and domain d
$\theta_{v,d}$	multinomial distribution over topics specific to author v and domain d
ϕ_z	multinomial distribution over words specific to topic z
α, β	Dirichlet priors to multinomial distributions θ, ϕ and ϕ
λ	parameter for sampling the binary variable s
γ, η	Beta parameters to generate λ

```

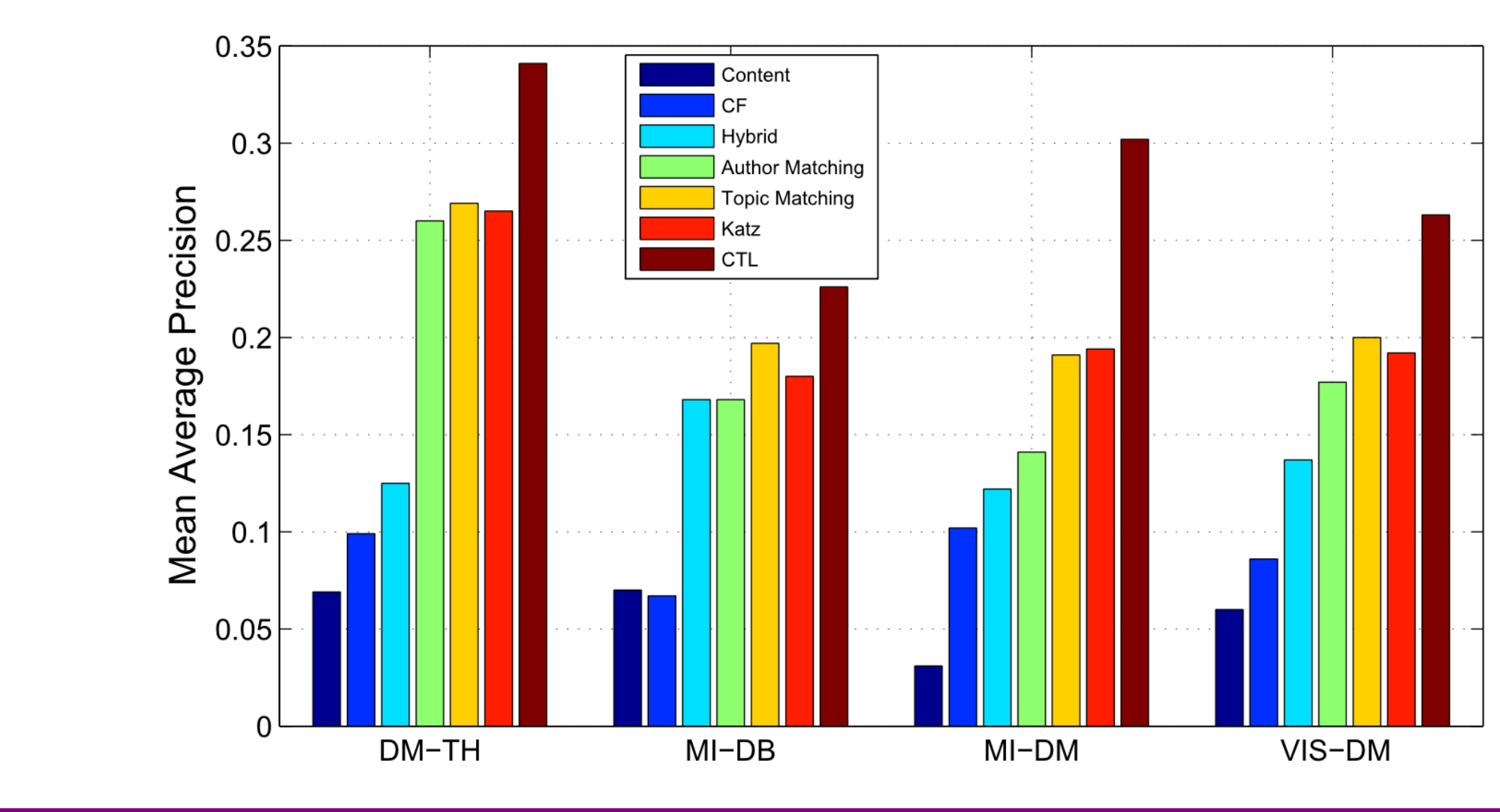
foreach word  $w_i$  in  $D$  do
    where  $\beta_{z,d}$  is a beta distribution, and  $\gamma$  and  $\eta$  are parameters:
        randomly select a pair  $(v, d)$  from  $A \times D$ , where  $v$  is an author from  $G^S$  and  $d$  from  $G^T$ ;
        Draw a topic  $z_{di} \sim \text{multi}(\theta_{v,d})$  from the topic mixture  $\theta_{v,d}$  specific to  $(v, d)$ ;
    end
    if  $s_{di} = 1$  then
        Randomly select a user  $v$ ;
        Draw a topic  $z_{di} \sim \text{multi}(\theta_v)$  from the topic model of user  $v$ ;
    end
    Draw a word  $w_{di} \sim \text{multi}(\phi_{z_{di},d})$  from  $z_{di}$ -specific word distribution;
end
    
```



Empirical Analysis

Data Mining (DM)—
Medical Informatics (MI)—
Theory (TH)
Visualization (VIS)
Database (DB)

Baselines:



domain	ALG	P@10	P@20	MAP	R@100	-10	-20
Data Mining (S) to Theory (T)	Content	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
	Hybrid	17.4	19.1	20.0	29.5	5.0	2.4
	Author	27.2	22.3	25.7	32.4	10.1	6.4
	Topic	28.0	26.0	32.4	33.5	13.4	7.1
Medical Info. (S) to Medical Info. (T)	Katz	30.4	29.8	31.6	27.4	11.2	5.9
	CTL	37.7	36.4	40.6	35.6	14.3	7.5
	Content	27.5	28.3	30.7	57.2	10.5	5.0
	CTL	32.5	30.0	36.9	59.8	11.4	5.4
	Content	5.8	5.7	9.5	19.8	1.9	0.9
Medical Info. (S) to Theory (T)	CF	13.7	17.8	18.9	34.3	2.7	1.3
	Hybrid	18.0	19.0	19.8	36.7	3.2	1.3
	Author	20.2	22.2	23.2	64.4	5.2	2.1
	Topic	28.0	26.0	32.4	48.1	10.2	5.6
	Katz	21.2	23.8	32.4	78.1	10.2	4.8
Database (S) to Database (T)	CTL	30.0	24.0	35.6	49.6	12.3	6.0
	Content	9.6	11.8	13.2	18.9	3.1	1.8
Data Mining (S) to Data Mining (T)	CTL	23.0	25.1	29.3	30.2	10.4	5.4
	Katz	28.3	26.0	32.8	36.3	14.0	9.1

