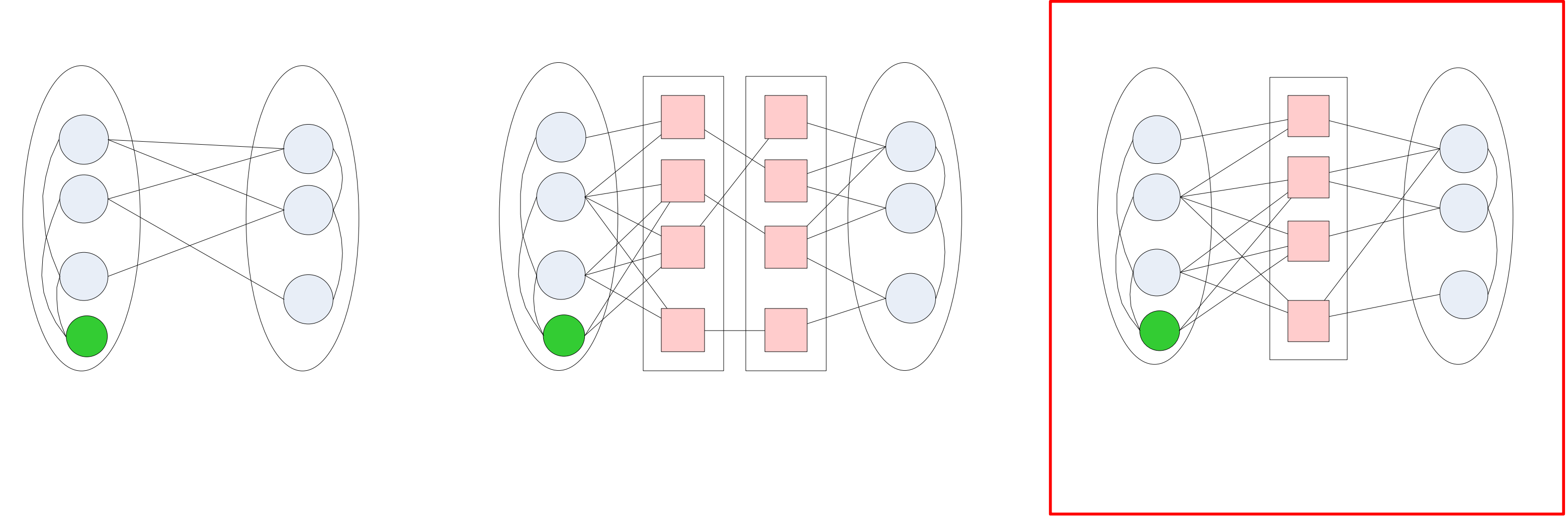


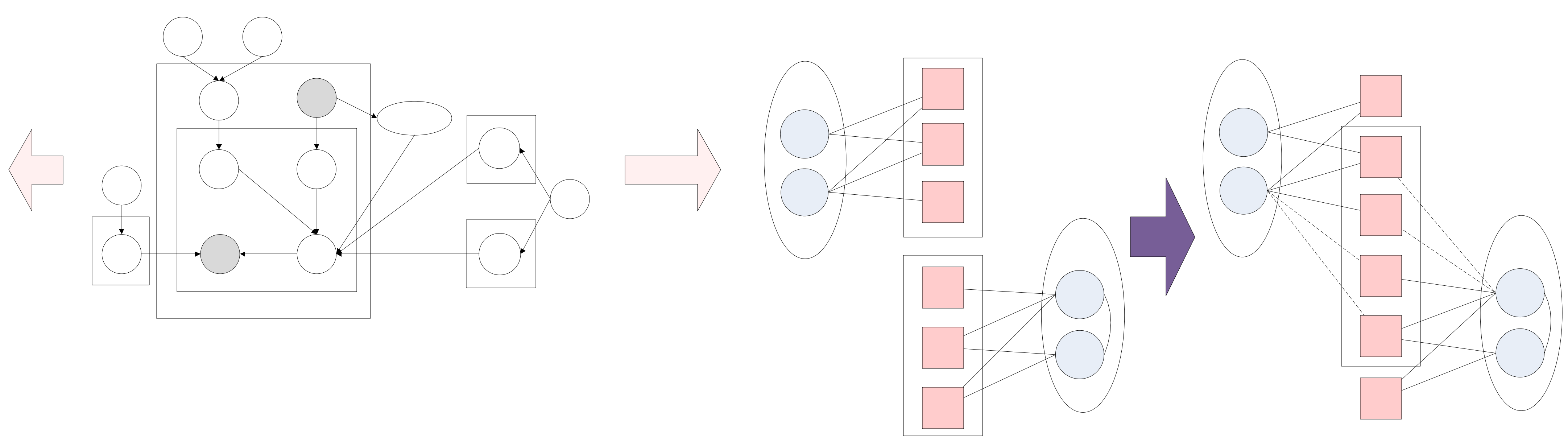
Cross-domain Topic Learning (CTL)



SYMBOL	DESCRIPTION
T	number of topics
d	a collaborated document
A_d	a set of authors of document d
$x_{d,i}$	the i th attribute (word) in document d
$w_{z,d}$	the topic z assigned to document d
$w_{z,d}$	$w_{z,d}$ is a word from a single domain or a cross domain
$\phi_{z,w}$	multinomial distribution over topics specific to word w
$\theta_{z,v}$	multinomial distribution over topics specific to author v
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

```

function G by learning from documents written by
S.
in AC model for target domain G'.
d document d do
    initialize an AC model for target domain G'.
    Similarly, initialize
    foreach collaborate
        ...
    end
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
Data Mining (S)	Content	10.3	10.2	10.9	31.4
	CF	15.6	13.3	23.1	26.2
	Hybrid	17.4	19.1	20.0	29.5
	Author	27.2	22.3	25.7	32.4
	Katz	30.4	29.8	31.6	27.4
Theory (I)	Topic	28.0	26.0	32.4	33.5
	Katz	30.4	29.8	31.6	27.4
Content	CTL	37.7	36.4	40.6	35.6
Content	CTL	10.1	10.9	12.5	45.9
Medical Info. (S)	Content	18.3	20.9	21.7	41.5
Hybrid	25.0	26.5	28.4	59.1	
Author	26.2	29.6	32.2	48.5	
Topic	27.5	28.3	30.7	57.2	
CTL	32.5	30.0	36.9	59.8	
Medical Info. (S)	Content	5.8	5.7	9.5	19.8
CF	13.7	17.8	18.9	34.3	
Hybrid	18.0	19.0	19.8	36.7	
to Author	20.1	23.8	29.3	64.4	
Topic	26.0	25.0	33.9	48.1	
Katz	21.2	23.8	32.4	78.1	
Medical Info. (I)	CTL	30.0	24.0	35.6	49.6
Content	CTL	9.6	11.8	13.2	18.9
Content	CTL	23.3	26.7	29.4	65.7
Visualization (S)	Topic	23.0	23.0	27.5	50.1
CF	23.0	25.1	29.3	30.2	
CTL	28.3	26.0	32.8	36.3	

