



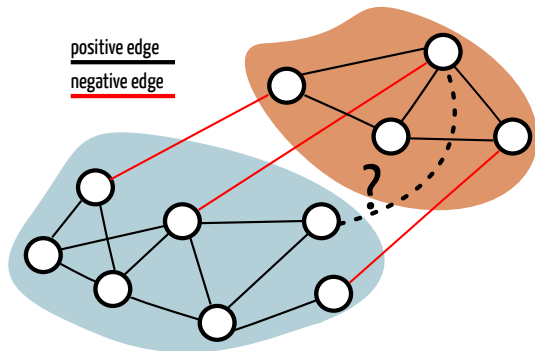
# A Fast Active Learning Algorithm for Link Classification

N.Cesa-Bianchi<sup>◦</sup>, C.Gentile<sup>\*</sup>, F.Vitale<sup>◦</sup> and **G.Zappella<sup>◦</sup>**

<sup>◦</sup> **Università degli Studi di Milano**

<sup>\*</sup> **Università dell'Insubria**

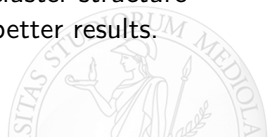
# Model



**Noise:** each label may be inconsistent with the 2-cluster structure with probability (at most)  $p$ . Lower value of  $p \rightarrow$  better results.

**Transductive setting:** graph topology is given

**Task:** predict binary labels on edges



## Algorithms

---

We present active learning algorithms with different trade-offs between number of queried labels and number of mistakes.

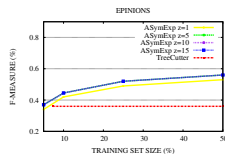
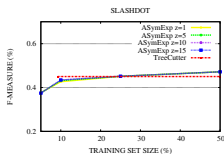
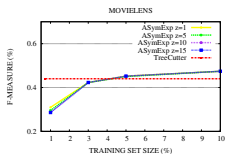
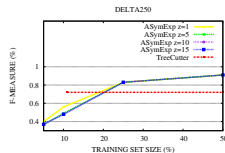
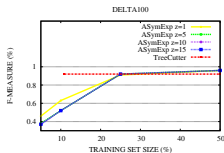
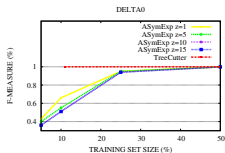
- Lower bound:  $\mathbb{E}[m] \geq \rho E_{test}$
- Dense graph:  $\mathbb{E}[m] \leq 5 \rho E_{test}$  with  $E_{train} = O(|V|^{3/2})$
- Small diameter:  $\mathbb{E}[m] \leq 2 \text{Diam} \rho E_{test}$  with  $E_{train} = |V| - 1$

Complete results available at the poster.

All these algorithms have a running time of  $O(|E|)$ .



# Experimental results



THANK YOU!

