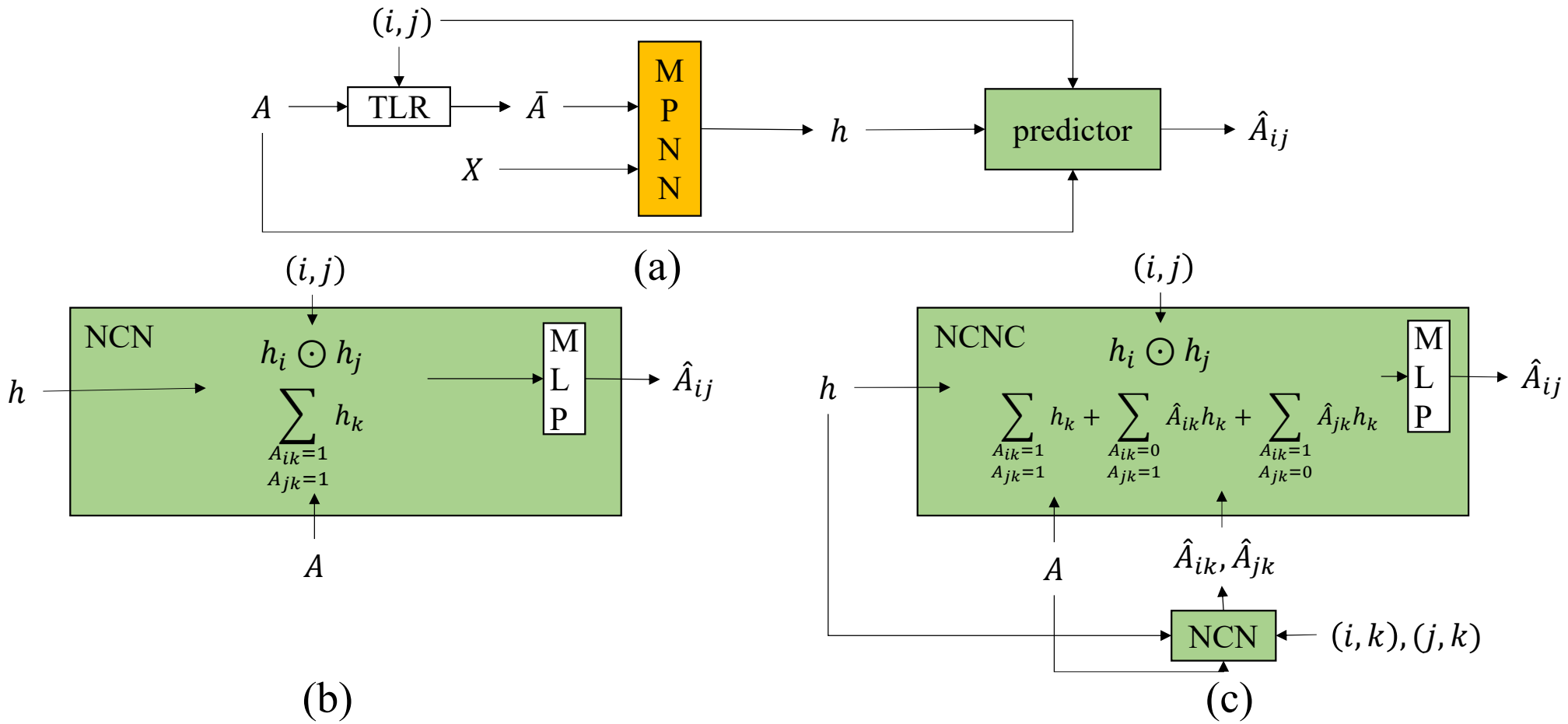

Figure 4: Architecture of our models. (a) The overall architecture. Given node feature X, adjacency matrix A and target link (i,j), models first set $A_{ij} = 0$ (Target link removal, TLR). Then, $\bar{A}, X$ are feed to a vanilla MPNN for node representations h. With (i,j), h and A as input, the predictor produces $\hat{A}_{ij}$, the probability that edge (i,j) exists. (b) The NCN predictor. It use node representation of target nodes i,j and their common neighbors to produce edge representations. Then, it feeds the edge representation to a MLP to produce final prediction. (c) The NCNC predictor. It first uses NCN to predict unobserved links $\hat{A}_{ik}, \hat{A}_{jk}$, which is then used to complete unobserved common neighbors.