

Appendix: Spatial-Temporal Adaptive Graph Convolutional Network for Skeleton-based Action Recognition

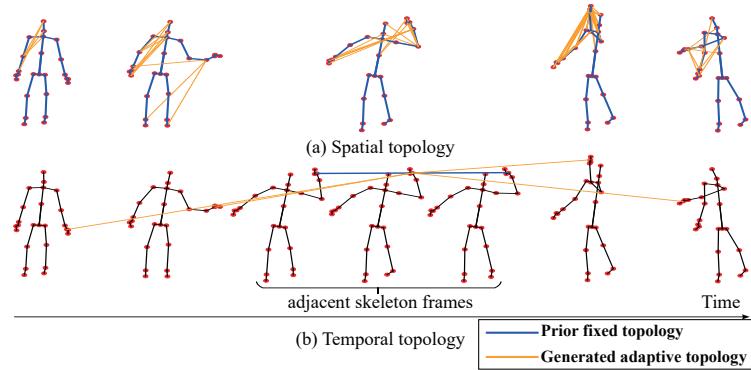


Fig. 1. Visualization of (a) spatial topology and (b) temporal topology. The blue line indicates the prior fixed topology and the orange line indicates the generated adaptive topology by our model. For clarity, we only show the right hand's temporal topology on one skeleton frame (the middle one in the bottom row).

Visualization. As shown in Fig. 1, we visualize the adaptive spatial and temporal topology of the "throw" action by saving the adjacency matrix generated from our model. For a better view, a threshold is set to filter the weak connections in the adjacency matrix. Compared to the fixed topology, the spatial adaptive topology can connect any two joints within a human skeleton, while the temporal adaptive topology can connect any two joints in a long-range trajectory rather than adjacent at time interval. This characteristic enables the model effectively aggregate the spatial and temporal features for skeleton-based action recognition.