Our learning problem: Node classification of weighted and undirected graphs based only on the network topology. This is useful in several domains, including document networks, social networks, biological networks.

Protocol: online learning. Vertices are issued one by one in an arbitrary (possibly adversarial) order \( v_1, v_2, \ldots, v_N \). At each time step, the learner predicts the binary label of \( v_t \) and then observes its true value.

Transductive classification: the entire unlabeled graph is known in advance.

Goals: few prediction mistakes and scalability

The Algorithm

Shazoo operates on weighted trees with a combination of mincut and nearest neighbor strategies: mincut is used to assign the label \( y_{\text{min}} \) to forks; nearest neighbor with resistance distance is used to label all nodes of a hinge tree.

Multiclass: in this paper we focus on binary classification, but it is possible to extend Shazoo to multiclass just by adding a quasilinear factor to its time complexity.

Pseudocode

```
for t=1\ldots|V|
    if \( v(t) \) is next node to be predicted:
        predict with \( y_{\text{min}}(v(t)) \);
    else
        \( \tau_{v(t)} \leftarrow \text{connection node closest to } v(t) \);
        predict with \( y_{\text{max}}(\tau_{v(t)}) \);
```

Experiments

We ran our experiments in \textit{batch mode}, using different sizes of randomly selected training sets (5%, 10%, 25%).

The results are averaged over \textit{10 runs} for each combination of train set size and algorithm.

We tested WTA and SHAZOO on \textit{different trees}:
- MST: the minimum spanning tree, generated in time \( O(|E| \log |V|) \).
- NWRST: random spanning tree, generated in time \( O(|V|) \) for most weighted graphs.
- NWRST: random spanning tree of the unweighted graph, generated in time \( O(|V|) \) for most graphs.

Batch case: SHAZOO+NWRST takes constant time per prediction on most graphs.

Shazoo is suitable to large scale networks

Results

1. SHAZOO outperforms WTA irrespective of the type of spanning tree being used
2. The predictive performance of SHAZOO+MST is comparable to, and sometimes better than, that of LABPROP (which is slower)
3. Committees of SHAZOO are effective: \textit{they outperform} LABPROP, when the training set is small
4. NWRST is extremely fast to generate and in our experiments is only slightly inferior to RST

Essential Bibliography