Elastic Graph Neural Networks

Algorithm
Motivated by the idea of (graph) trend filter, we propose new smoothing priors which favor adaptive piecewise constant signal over the graph:

\[
\lim_{\gamma \to 0} \frac{1}{2} \left( |\langle \gamma \rangle_{x_n} |^2 - |\langle \gamma \rangle_{x_n} |^2 \right) + \frac{1}{2} \left( |\langle \gamma \rangle_{x_n} |^2 - |\langle \gamma \rangle_{x_n} |^2 \right)
\]

Option I
\[
\min_{x \in \mathbb{R}^n} \frac{1}{2} \left( |\langle \gamma \rangle_{x_n} |^2 - |\langle \gamma \rangle_{x_n} |^2 \right) + \frac{1}{2} \left( |\langle \gamma \rangle_{x_n} |^2 - |\langle \gamma \rangle_{x_n} |^2 \right)
\]

[Normalized incident matrix]

Option II
\[
\min_{x \in \mathbb{R}^n} \frac{1}{2} \left( |\langle \gamma \rangle_{x_n} |^2 - |\langle \gamma \rangle_{x_n} |^2 \right) + \frac{1}{2} \left( |\langle \gamma \rangle_{x_n} |^2 - |\langle \gamma \rangle_{x_n} |^2 \right)
\]

[Coupling multi-dimensionality]

Elastic Message Passing
\[
\begin{align*}
Y^{k+1} &= Y^k + (1 - \gamma) \Delta P^k \\
Z^{k+1} &= Z^k + \beta \Delta F^k \\
\end{align*}
\]

Theorem (Convergence)
Under the stepsize setting \( \gamma \leq \frac{2}{12 + \lambda} \) and \( \beta \leq \frac{4}{\|\Delta F\|_2} \), the elastic message passing scheme (EMP) converges to the optimal solution of the elastic graph signal estimator. It is sufficient to choose any \( \gamma \leq \frac{2}{12 + \lambda} \) and \( \beta \leq \frac{4}{\|\Delta F\|_2} \) since \( \|\Delta u\|_2 \leq \|\Delta u\|_2 \leq \|\Delta u\|_2 \leq 2 \).

Interpretation
The first step in EMP is the standard message passing in existing GNNs, and the following steps accumulate \( \Delta F \) to promote the sparsity in \( \Delta F \) and to preserve jump signals over edges that might be noise or adversarial. The sparsity patterns in \( \Delta F \) indicates the piecewise constant behavior of the graph (e.g., clusters in the graph), and the preserved large node differences prevent the inappropriate diffusion of features over edges.

Elastic GNNs
\[
Y_{pre} = EMP(h_0(X_{test}), K, \lambda_1, \lambda_2)
\]

Experiment
Classification accuracy (%) on benchmark datasets with 10 times random data splits.

<table>
<thead>
<tr>
<th>Model</th>
<th>Data</th>
<th>Class</th>
<th>Leave-out</th>
<th>Training</th>
<th>Test</th>
<th>accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCN</td>
<td>Cora</td>
<td>88.4</td>
<td>88.4</td>
<td>88.4</td>
<td>88.4</td>
<td></td>
</tr>
<tr>
<td>APPNP</td>
<td>Cora</td>
<td>88.4</td>
<td>88.4</td>
<td>88.4</td>
<td>88.4</td>
<td></td>
</tr>
</tbody>
</table>

Conclusion
- Introduce \( L_1 \)-based graph smoothing in GNNs to enhance local smoothness adaptively.
- Derive a novel, general and efficient message passing scheme (EMP) that promote piecewise behavior in information diffusion.
- Develop a family of GNNs (Elastic GNNs) that are intrinsically more robust to noise and adversarial graphs.