Graph Embedding using the Edge-based Wave Kernel

Hewayda El Ghawalby and Edwin Hancock
Department of Computer Science
University of York

Supported by a Royal Society Wolfson Research Merit Award
Motivation
Learning with graph data

• Problems based on graphs arise in areas such as language processing, proteomics/chemoinformatics, data mining, computer vision and complex systems.

• Relatively little methodology available, and vectorial methods from statistical machine learning not easily applied since there is no canonical ordering of the nodes in a graph.

• Can make considerable progress if we develop permutation invariant characterisations of variations in graph structure.
Learning with graph data

- Problems based on graphs arise in areas such as language processing, proteomics/chemoinformatics, data mining, computer vision and complex systems.

- Relatively little methodology available, and vectorial methods from statistical machine learning not easily applied since there is no canonical ordering of the nodes in a graph.

- Can make considerable progress if we develop permutation invariant characterisations of variations in graph structure.
Learning with graph data

• Problems based on graphs arise in areas such as language processing, proteomics/chemoinformatics, data mining, computer vision and complex systems.

• Relatively little methodology available, and vectorial methods from statistical machine learning not easily applied since there is no canonical ordering of the nodes in a graph.

• Can make considerable progress if we develop permutation invariant characterisations of variations in graph structure.
Structural Variations
Problem

In computer vision graph-structures are used to abstract image structure. However, the algorithms used to segment the image primitives are not reliable. As a result there are both additional and missing nodes (due to segmentation error) and variations in edge-structure. Hence image matching and recognition can not be reduced to a graph isomorphism or even a subgraph isomorphism problem. Instead inexact graph matching methods are needed.
Problem

In computer vision graph-structures are used to abstract image structure. However, the algorithms used to segment the image primitives are not reliable.

As a result there are both additional and missing nodes (due to segmentation error) and variations in edge-structure.

Hence image matching and recognition can not be reduced to a graph isomorphism or even a subgraph isomorphism problem. Instead inexact graph matching methods are needed.
Problem

In computer vision graph-structures are used to abstract image structure. However, the algorithms used to segment the image primitives are not reliable. As a result there are both additional and missing nodes (due to segmentation error) and variations in edge-structure. Hence image matching and recognition can not be reduced to a graph isomorphism or even a subgraph isomorphism problem. Instead inexact graph matching methods are needed.
Learning is difficult because

- **Graphs are not vectors:** There is no natural ordering of nodes and edges. Correspondences must be used to establish order.

- **Structural variations:** Numbers of nodes and edges are not fixed. They can vary due to segmentation error.

- **Not easily summarised:** Since they do not reside in a vector space, mean and covariance hard to characterise.
Learning with graphs

Work with (dis) similarities: Can perform pairwise clustering or embed sets of graphs in a vector space using multidimensional scaling on similarities. Non metricity of similarities may pose problems.

Embed individual graphs in a low dimensional space: Characterise structural variations in terms of statistical variation in a point-pattern.

Learn modes of structural variation: Understand how edge (connectivity) structure varies for graphs belonging to the same class. Requires correspondences of raw structure or alignment of an embedded one. Can also be effected using permutation invariant characteristics (path length, commute-times, cycle frequencies).

Construct generative model: Borrow ideas from graphical model to construct model for raw structures or point distribution model to for embedded graphs.
Problem studied

- How can we find efficient means of characterising graph structure which does not involve exhaustive search? Enumerate properties of graph structure without explicit search, e.g. count cycles, path length frequencies, etc..

- Can we analyse the structure of sets of graphs without solving the graph-matching problem? Inexact graph matching is computational bottleneck for most problems involving graphs.

- Past: Explored how diffusion processes based on heat equation can be used for this purpose.
Alternative paradigms

• Recently notion of performing calculus on graphs (Friedman) has meant that richer family of physical operators can be extended from continuous domain to graphs.

• Wave equation can be extended to graphs through concept of edge-based Laplacian and used to analyse vibrational modes of graphs.

• Can this lead to interesting characterisations and new embeddings?
Aims
Friedman’s edge-based wave equation

- Edges of graph are taut strings joined at vertices (nodes) of graph.

- Construct edge-based Laplacian

\[
L_V = 1 - \cos \sqrt{L_E}
\]

- Solve wave equation

\[
\ddot{W} = -L_E W
\]
- Describe a new approach for embedding graphs on pseudo-Riemannian manifolds based on the wave kernel.

- Wave-kernel eigensystem determined by the eigenvalues and the eigenfunctions of the normalized adjacency matrix; can be used to solve the edge-based wave equation.

- Factorising wave kernel Gram-matrix for gives the embedding co-ordinates.

- Investigate the utility of this new embedding as a means of gauging the similarity for graphs by representing the proximity of image features in different views of different objects.
Method

• Treat normalised (dis) similarity matrix as wave kernel.

• Compute spectrum of edge-based Laplacian using analysis given by Friedman.

• Perform Krein-space decomposition on kernel.
The vertex Laplacian and the heat equation
Laplacian Matrix

• Weighted adjacency matrix

\[ A(u, v) = \begin{cases} w(u, v) & (u, v) \in E \\ 0 & \text{otherwise} \end{cases} \]

• Degree matrix

\[ D(u, u) = \sum_{v \in V} A(u, v) \]

• Laplacian matrix

\[ L = D - A \]
Laplacian spectrum

• Spectral Decomposition of Laplacian

\[ L = \Phi \Lambda \Phi^T = \sum_{k} \lambda_k \phi_k \phi_k^T \]
\[ 0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{|V|} \]
\[ \Lambda = \text{diag}(\lambda_1, \ldots, \lambda_{|V|}) \]
\[ \Phi = (\phi_1 | \ldots | \phi_{|V|}) \]

• Element-wise

\[ L(u, v) = \sum_{k} \lambda_k \phi_k (u) \phi_k (v) \]
Properties of the Laplacian

• Eigenvalues are positive and smallest eigenvalue is zero

\[ 0 = \lambda_1 < \lambda_2 < \ldots < \lambda_{|V|} \]

• Multiplicity of zero eigenvalue is number connected components of graph.

• Zero eigenvalue is associated with all-ones vector.

• Eigenvector associated with the second smallest eigenvector is Fiedler vector.
Heat Kernels

• Solution of heat equation and measures information flow across edges of graph with time:

\[
\frac{\partial h_t}{\partial t} = -L h_t
\]

\[
L = D - W = \Phi \Lambda \Phi^T
\]

• Solution found by exponentiating Laplacian eigensystem

\[
h_t = \sum_k \exp[-\lambda_k t] \phi_k \phi_k^T = \Phi \exp[-\Lambda t] \Phi^T
\]
Heat kernel and random walk

- State vector of continuous time random walk satisfies the differential equation

\[
\frac{\partial p_t}{\partial t} = -Lp_t
\]

- Solution

\[
p_t = \exp[-Lt] p_0 = h_t p_0
\]
The wave equation and the edge Laplacian
Heat Equation and Wave Equation Compared

• Heat equation models diffusion over the graph with time:

\[ \dot{H} = -L_v H \]

• Wave equation models vibrational modes of graph, eigenvalues are frequency modes:

\[ \ddot{W} = -L_E W \]
Solution of the wave equation

Equation plus boundary conditions

\[ W_{tt} = -L_E W \]
\[ W|_{t=0} = f \]
\[ W_t|_{t=0} = g \]

General solution

\[ W = \frac{\sin \left( \sqrt{L_E} t \right)}{\sqrt{L_E}} g + \cos \left( \sqrt{L_E} t \right) f \]
Simplification

- Work with boundary conditions

\[ W |_{t=0} = 0 \]

\[ \dot{W} |_{t=0} = 1 \]

- Simplified solution is

\[ W = \frac{\sin t \sqrt{L_E}}{\sqrt{L_E}} \]

- Since Laplacian is positive semidefinite, have MacLaurin series approximation

\[ W = t(I - \frac{1}{6}L_E t^2 + ...) \]
Constructing the Edge-Based Laplacian

• Commence by constructing the normalised adjacency matrix

\[ W = \tilde{A} = D^{-1/2} A D^{-1/2} \]

• Construction is implicit since it can be computed from spectrum of vertex Laplacian.
Spectral relations

• Normalised edge and vertex Laplacians

\[ \tilde{L}_V = 1 - \cos \sqrt{\tilde{L}_E} \]
\[ \cos \sqrt{\tilde{L}_E} = 1 - \tilde{L}_V = \tilde{A} \]

• Spectra

\[ \cos \sqrt{\lambda} \in \text{Spec}[\tilde{A}] \iff \lambda \in \text{Spec}[L_E] \]

• Eigenfunctions

\[ \tilde{A}f = f \cos \lambda \iff L_E f = \lambda f \wedge L_V f = 0 \]
Spectrum of Edge-Based Laplacian

• Spectrum of harmonics

\[ \text{Spec}[L_E] = \{ (\cos^{-1} \sqrt{\lambda} + 2\pi Z_{>0}), 2\pi - \cos^{-1} \sqrt{\lambda} + 2\pi Z_{>0}, \]
\[ (\pi + 2\pi Z_{>0}, \pi + 2\pi Z_{>0})^{[E]-[V]} \} \]
Physical Interpretation

Describes the vibrational modes associated with a taut strings on each edge that are joined together at the vertices. If we excite or ”pluck” the system, it would produce tones with frequencies whose spectrum ranging over the edge-based eigenvalues.
What does it mean

• Project nodes of graph into eigenspace that corresponds to harmonic frequencies of graph edges.

• Nodes are stationary under vibration.

• Negative eigenspace corresponds to imaginary non-physical vibration modes.
Experiments
Datasets

1) The standard CMU, MOVI and chalet house image sequences, with thirty different views of three model houses (10 for each) from equally spaced viewing directions.

2) Five objects selected from the COIL image sequences, with 72 different views of each object from equally spaced viewing directions.

From each image, corner features are extracted, and Delaunay graphs representing the arrangement of feature points are constructed.
Processing Steps

- Compute the eigensystem of the adjacency matrix.

- Construct the edge-based Laplacian matrix.

- Compute the edge-based wave kernel using $t = 10.0$, $t = 1.0$, $t = 0.1$, and $t = 0.01$.

- Construct the embedding coordinates matrix, whose columns are the coordinates of the embedded nodes in a Pseudo-Euclidean space.

- Project the coordinate vectors onto a Pseudo-Euclidean space with low dimension using an orthonormal basis.

- Construct the distance matrices between the different graphs using the modified Hausdorff distance.
MDS embedding obtained when using the Wave Kernel for the houses data
MDS embedding obtained when using the Wave Kernel for the COIL data
Rand Index

- Compute the mean for each cluster;
- Compute the distance from each point to each mean;
- If the distance from correct mean is smaller than those to remaining means, then classification is correct, if not then classification is incorrect;
- Compute the Rand-index

\[ R = \frac{\# \text{ incorrect}}{\# \text{ incorrect} + \# \text{ correct}} \]

The following table shows the results:

<table>
<thead>
<tr>
<th></th>
<th>t=10</th>
<th>t=1.0</th>
<th>t=0.1</th>
<th>t=0.01</th>
</tr>
</thead>
<tbody>
<tr>
<td>Houses data</td>
<td>0.2333</td>
<td>0.0000</td>
<td>0.0333</td>
<td>0.1000</td>
</tr>
<tr>
<td>COIL data</td>
<td>0.3333</td>
<td>0.3333</td>
<td>0.3333</td>
<td>0.7000</td>
</tr>
</tbody>
</table>
Conclusions
• A new approach for embedding graphs on pseudo-Riemannian manifold based on the wave kernel was presented.

• Embedding the Pseudo Riemannian Manifold in a Pseudo Euclidean space gives us the advantage to suggest an evaluation for the similarity of graphs using a measure of distance based on that space.

• Based on experiments on objects from two datasets (the York Model House and COIL datasets), we are confident that an edge-based wave kernel embedding can be used for the purpose of graph characterization.