LOCALIZED ITERATIVE METHODS FOR INTERPOLATION IN GRAPH STRUCTURED DATA

Sunil K Narang, Akshay Gadde, Eduard Sanou and Antonio Ortega

Ming Hsieh Department of Electrical Engineering
University of Southern California
kumarsun@usc.edu, agadde@usc.edu, eduardsanou@gmail.com, ortega@sipi.usc.edu

ABSTRACT

In this paper, we present two localized graph filtering based methods for interpolating graph signals defined on the vertices of arbitrary graphs from only a partial set of samples. The first method is an extension of previous work on reconstructing bandlimited graph signals from partially observed samples. The iterative graph filtering approach very closely approximates the solution proposed in the that work, while being computationally more efficient. As an alternative, we propose a regularization based framework in which we define the cost of reconstruction to be a combination of smoothness of the graph signal and the reconstruction error with respect to the known samples, and find solutions that minimize this cost. We provide both a closed form solution and a computationally efficient iterative solution of the optimization problem. The experimental results on the recommendation system datasets demonstrate effectiveness of the proposed methods.

1. INTRODUCTION

The field of *graph signal processing* extends signal processing tools designed for regularly sampled signals to graph datasets [1]. In the graph representation, the data points are represented as nodes connected to each other via links. The weights of the links usually represent similarity between the data points. Each node stores a sample, and the collection of these samples is referred to as a *graph signal*. In this paper we discuss an important problem, namely that of interpolation of missing values from known samples, which appears in various applications, such as matrix/vector completion, sampling of high-dimensional data, semi-supervised learning etc. Inspired by standard signal processing approaches, we formulate the data interpolation problem as a signal reconstruction problem on a graph. This is an extension of our previous work in [2], where we used sampling results in graphs to find classes of bandlimited (BL) graph signals that can be reconstructed from their partially observed samples. A class of BL graph signals is specified by the cut-off graph frequency ω (denoted as ω-BL), and the interpolated signal is obtained by projecting the input signal onto the appropriate ω-BL subspace using a *least square* method. The value of ω is estimated using the topology of the underlying graph and location of known samples in the graph. The method proposed in [2] provides exact reconstruction of ω-BL graph signals and the best approximation (in the least square sense) of arbitrary signals as ω-BL graph signals.

However, this method of reconstruction is computationally expensive for large graphs as it involves eigenvalue decomposition of Laplacian matrix, followed by inverse of a square matrix of the size of the graph. Therefore, in this paper we formulate the interpolation problem on graph as an iterative graph filtering problem, where the graph filter is designed as an ideal low-pass graph filter with cut-off frequency ω as computed in [2]. The proposed iterative algorithm is faster and converges to the least square reconstruction method in [2]. Also, to avoid eigenvalue decomposition of Laplacian matrix, the ideal low pass filter is approximated with a polynomial of the Laplacian matrix, which can be computed efficiently as matrix-vector product without the need of eigenvalue decomposition.

Further, the estimated cut-off frequency ω is only an estimate, and the actual signal may not be ω-BL. Therefore, we set up a regularized cost that exploits the trade-off between signal smoothness, and the reconstruction errors at the known samples. The proposed cost function is based on the data fitting error at the known samples and the energy of the reconstructed signal outside the ω-BL subspace. The solution of the regularization is computed first as an exact solution, followed by an approximate solution based on the iterative graph filtering approach. The rest of the paper is organized as follows: in Section 2, we briefly explain the interpolation method proposed in [2], in Section 3, provide an iterative graph filtering based solution of this interpolation method. In Section 4, we describe a second method for graph signal interpolation, based on a regularization framework, in Section 5 we discuss application of proposed method to item-recommendation systems, and compare results with respect to existing methods in Section 6.

2. SAMPLING THEOREM FOR BAND-LIMITED GRAPH SIGNALS

A graph $G = (\mathcal{V}, E)$ is a collection of nodes $\mathcal{V} = \{1, 2, ...N\}$ connected together by set of links $E = \{(i, j, w_{ij})\}$, $i, j \in \mathcal{V}$. (i, j, w_{ij}) denotes the link between nodes i and j having weight w_{ij}. The adjacency matrix $\mathbf{W}$ of the graph is an $N \times N$ matrix such that $W(i, j) = w_{ij}$. The degree d_i of node i is the sum of link-weights connected to node i. The degree matrix $\mathbf{D} = diag\{d_1, d_2, ..., d_N\}$ is a diagonal matrix. The combinatorial Laplacian matrix is defined as $\mathbf{L} = \mathbf{D} - \mathbf{W}$. The corresponding symmetric normalized Laplacian matrix is $\boldsymbol{\mathcal{L}} = \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2}$. We use the normalized Laplacian matrix because it is closely related to the random walk matrix and is shown to produce superior classification results [3]. We consider only undirected graphs without self loops for which $\boldsymbol{\mathcal{L}}$ is a symmetric positive semi-definite matrix. Therefore, it has the eigenvalue decomposition:

$$\boldsymbol{\mathcal{L}} = \mathbf{U}\Lambda\mathbf{U}^t = \sum_{i=1}^{N} \lambda_i \mathbf{u}_i \mathbf{u}_i^t, \tag{1}$$

This work was supported in part by NSF under grant CCF-1018977.

with a diagonal eigenvalue matrix $\mathbf{\Lambda}$ containing non-negative eigenvalues $\{\lambda_1, \lambda_2 \dots \lambda_N\}$ arranged in a non-decreasing order at the diagonal, and a unitary matrix U containing corresponding eigenvectors $\mathbf{u}_i$. A graph signal is a function $f : \mathcal{V} \rightarrow \mathbb{R}$ defined on the vertices of the graph. It can be represented as a vector $\mathbf{f} \in \mathbb{R}^N$ where the ith component represent the function value on the ith vertex. Eigenvectors and eigenvalues of $\boldsymbol{\mathcal{L}}$ are used to define Fourier transform for graph signals [1,4,5]. Eigenvalues λ_i are the graph frequencies which are always in the range $[0, 2]$, and eigenvectors serve as the corresponding basis vectors. Every graph signal can be represented with basis $\mathbf{U}$ as $\mathbf{f} = \sum_i \tilde{f}(\lambda_i)\mathbf{u}_i$, where $\tilde{f}(\lambda_i) = \langle \mathbf{f}, \mathbf{u}_i \rangle$ is the *graph Fourier transform* (GFT) of $\mathbf{f}$.

In classical signal processing, the signal being bandlimited implies that the energy of the signal is zero above a certain frequency. The spectral analysis of graph signals offers a similar interpretation. Following definitions and results were used in [2] to design reconstruction algorithm for graph signals.

Definition 1 (Band-limited graph signal [5])**.** *A signal on a graph G is said to be band-limited to the graph frequency band $[0, \omega)$, if its GFT has support only at frequencies $[0, \omega)$.*

The space of ω-bandlimited signals is called Paley-Wiener space and is given by

$$PW_\omega(G) = \{\mathbf{f} : \ \tilde{f}(\lambda) = 0 \quad \text{if} \quad \lambda \geq \omega\} \tag{2}$$

Definition 2 (Λ-set)**.** *A set $\mathcal{Q} \in \mathcal{V}$ is a Λ-set if all graph signals $\boldsymbol{\phi}$ with support on $\mathcal{Q}$ (i.e. $\boldsymbol{\phi}(v) = 0$ if $v \notin \mathcal{Q}$) satisfy*

$$\|\boldsymbol{\phi}\| \leq \Lambda \|\boldsymbol{\mathcal{L}}\boldsymbol{\phi}\| \quad \dots (\Lambda > 0) \tag{3}$$

Theorem 2.1 (Sampling theorem [5])**.** *All graph signals $\mathbf{f} \in PW_\omega(G)$ can be uniquely recovered from a subset of its samples on $\mathcal{S}$ if $\mathcal{S}^c = \mathcal{V} - \mathcal{S}$ is a Λ-set such that $0 < \omega < 1/\Lambda$.*

The following result [2] computes the maximum ω such that any signal in $PW_\omega(G)$ can be reconstructed given a subset of known samples $\mathcal{S}$ on any graph G.

Proposition 1 (Cut-off frequency)**.** *[2] Let $(\boldsymbol{\mathcal{L}}^2)_{\mathcal{S}^c}$ be the submatrix of $\boldsymbol{\mathcal{L}}^2$ containing only the rows and columns corresponding to unknown set $\mathcal{S}^c$. Let σ^2_{min} to be the smallest eigenvalue of $(\boldsymbol{\mathcal{L}}^2)_{\mathcal{S}^c}$. Any $\mathbf{f} \in PW_\omega(G)$ with $\omega = \sigma_{min}$ can be uniquely recovered from its samples on $\mathcal{S}$.*

2.1. Least Squares Reconstruction

Proposition 1 gives a condition on the GFT of a graph signal such that unique reconstruction is possible from its given known subset of samples. A simple way to do this reconstruction is to solve a least-squares problem in the spectral domain as explained below.

Let λ_k be the largest eigenvalue of $\boldsymbol{\mathcal{L}}$ less than ω. An ω-bandlimited signal can be written (under appropriate permutation) as

$$\begin{bmatrix} \mathbf{f}(\mathcal{S}) \\ \mathbf{f}(\mathcal{S}^c) \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1(\mathcal{S}) & \mathbf{u}_2(\mathcal{S}) & \cdots & \mathbf{u}_k(\mathcal{S}) \\ \mathbf{u}_1(\mathcal{S}^c) & \mathbf{u}_2(\mathcal{S}^c) & \cdots & \mathbf{u}_k(\mathcal{S}^c) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix} \tag{4}$$

Let $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_k]^t$ and

$$\begin{bmatrix} \mathbf{u}_1(\mathcal{S}) & \mathbf{u}_2(\mathcal{S}) & \cdots & \mathbf{u}_k(\mathcal{S}) \\ \mathbf{u}_1(\mathcal{S}^c) & \mathbf{u}_2(\mathcal{S}^c) & \cdots & \mathbf{u}_k(\mathcal{S}^c) \end{bmatrix} = \begin{bmatrix} (\mathbf{U}_k)_{\mathcal{S}} \\ (\mathbf{U}_k)_{\mathcal{S}^c} \end{bmatrix}$$

$\boldsymbol{\alpha}$ can be obtained by calculating a least squares solution to $\mathbf{f}(\mathcal{S}) = (\mathbf{U}_k)_{\mathcal{S}}\boldsymbol{\alpha}$. Then, the unknown signal values are given by

$$\mathbf{f}(\mathcal{S}^c) = (\mathbf{U}_k)_{\mathcal{S}^c} \left((\mathbf{U}_k)^t_{\mathcal{S}}(\mathbf{U}_k)_{\mathcal{S}}\right)^{-1} (\mathbf{U}_k)^t_{\mathcal{S}}\mathbf{f}(\mathcal{S}) \tag{5}$$

The sampling theorem guarantees that the there exist a unique solution to the above least squares problem, which is equal to the original signal $\mathbf{f}$ if $\mathbf{f} \in PW_\omega(G)$. On the other hand, if $\mathbf{f} \notin PW_\omega(G)$, we still get a unique least square approximation of $\mathbf{f}$ in $PW_\omega(G)$ space. The choice of the cut-off frequency ω (estimated from Theorem 2.1) is still crucial, even though the reconstructed signal in this case may not be the best solution in terms of reconstruction errors. This is because, for a frequency ω' higher than ω, there exists a LS solution but the sampling theorem guarantee fails. This means that there may be infinitely many LS solutions in the $PW_{\omega'}(G) \supset PW_\omega(G)$ space, each giving a different interpolation result at the unknown samples. Therefore, in [2] we used ω as the cut-off frequency for all reconstructed graph signals. The proposed method in [2] provides good interpolation results when applied to item-recommendation problem. However, the algorithm is computationally expensive as it requires computation of eigenvalues of the Laplacian matrix. In the next section, we provide an iterative method for solving the above reconstruction problem.

3. ITERATIVE LEAST SQUARE RECONSTRUCTION

Our proposed method is similar to the Papoulis-Gerchberg algorithm [6–8] in classical signal processing which is used to reconstruct a band-limited signal from irregular samples. It is a special case of projection onto convex sets (POCS) [9], where the convex sets of interest in this case are:

$$C_1 = \{\mathbf{x} : \mathbf{J}\mathbf{x} = \mathbf{J}\mathbf{f}\} \tag{6}$$

$$C_2 = PW_\omega(G) \tag{7}$$

Here $\mathbf{J} : \mathbb{R}^N \rightarrow \mathbb{R}^M$ denotes the downsampling operator where M is the size of the known subset $\mathcal{S}$ of samples. At k^{th} iteration, the solution $\mathbf{f}_k$ is obtained from $\mathbf{f}_{k-1}$, and satisfies the following two constraints: (1) the signal equals the known values on the sampling set (i.e., $\mathbf{f}_k \in C_1$). (2) the signal is ω-bandlimited, where ω is computed using Proposition 1 (i.e., $\mathbf{f}_k \in C_2$). We define $\mathbf{P} : \mathbb{R}^N \rightarrow PW_\omega(G)$ to be the low-pass graph filter such that

$$\mathbf{y} = \mathbf{P}\mathbf{x} \Rightarrow \mathbf{y} \in PW_\omega(G) \tag{8}$$

$\mathbf{P}$ can be written in graph spectral domain as $\mathbf{P} = \mathbf{H}(\boldsymbol{\mathcal{L}}) = \sum_{i=1}^N h(\lambda_i)\mathbf{u}_i\mathbf{u}_i^t$ where

$$h(\lambda) = \begin{cases} 1 & \text{if } \lambda < \omega \\ 0 & \text{if } \lambda \geq \omega \end{cases} \tag{9}$$

We define the downsample then upsample (DU) operation as

$$\mathbf{f}_{du} = \mathbf{J}^t\mathbf{J}\mathbf{f} \Rightarrow \mathbf{f}_{du}(\mathcal{S}) = \mathbf{f}(\mathcal{S}) \text{ and } \mathbf{f}_{du}(\mathcal{S}^c) = \mathbf{0}. \tag{10}$$

With this notation the proposed iterative algorithm can be written as

$$\begin{aligned} \mathbf{f}_0 &= \mathbf{P}\mathbf{f}_{du} \\ \mathbf{f}_{k+1} &= \mathbf{P}(\mathbf{f}_k + \mathbf{J}^t\mathbf{J}(\mathbf{f}_{du} - \mathbf{f}_k)) \end{aligned} \tag{11}$$

At each iteration the algorithm resets the signal samples on $\mathcal{S}$ to the actual given samples and then projects the signal onto the low-pass space $PW_\omega(G)$.

3.1. Convergence

We define the operators $\mathbf{B} : \mathbb{R}^N \rightarrow \mathbb{R}^N$ and $\mathbf{T} : C_2 \rightarrow C_2$ corresponding to iteration in (11) as

$$\mathbf{Bx} = \mathbf{x} + \mathbf{J}^t\mathbf{J}(\mathbf{f}_{du} - \mathbf{x}) \tag{12}$$

$$\mathbf{Tx} = \mathbf{P}(\mathbf{x} + \mathbf{J}^t\mathbf{J}(\mathbf{f}_{du} - \mathbf{x})) = \mathbf{PBx} \tag{13}$$

It has been shown [9] that an iterative algorithm of the form $\mathbf{x}_{k+1} = \mathbf{Tx}_k$ converges to a fixed point of $\mathbf{T}$ if

1. $\mathbf{T}$ is non-expansive, i.e., $\|\mathbf{Tx} - \mathbf{Ty}\| \leq \|\mathbf{x} - \mathbf{y}\|$
2. $\mathbf{T}$ is asymptotically regular, i.e., $\|\mathbf{Tx}_{k+1} - \mathbf{Tx}_k\| \rightarrow 0$ as $k \rightarrow \infty$.

$\mathbf{P}$ is a bandlimiting operator and hence is non-expansive. $\mathbf{B}$ is non expansive because $\|\mathbf{Bx} - \mathbf{By}\| = \|(\mathbf{I} - \mathbf{J}^t\mathbf{J})(\mathbf{x} - \mathbf{y})\| \leq \|\mathbf{x} - \mathbf{y}\|$. Since both $\mathbf{P}$ and $\mathbf{B}$ are non-expansive, $\mathbf{T}$ is also non-expansive. Asymptotic regularity of $\mathbf{T}$ can also be proved as shown in [8]. Note that if $\mathbf{f}$ is a fixed point of $\mathbf{T}$ then $\mathbf{f} \in C_1 \cap C_2$. From the sampling theorem $\mathbf{f} \in C_1 \cap C_2$ is unique. So the asymptotic solution of the proposed algorithm converges to the solution of the least square projection method described in previous section.

3.2. Iterative reconstruction with polynomial low pass filter

The low pass filter $\mathbf{P}$ above is a spectral graph filter with an ideal brick wall type spectral response. Thus, the exact computation of $\mathbf{P}$ requires eigenvalue decomposition of the Laplacian matrix, which is computationally very expensive for large matrices. However, it is possible to approximate the ideal filtering operation as a matrix polynomial in terms of $\boldsymbol{\mathcal{L}}$, that can be implemented efficiently using only matrix vector products. Thus we replace $\mathbf{P}$ in (11) with an approximate low pass filter $\mathbf{P}^{poly}$ given as:

$$\mathbf{P}^{\text{poly}} = \sum_{i=1}^{N} \left(\sum_{j=0}^{k} a_j \lambda_i^j \right) \mathbf{u}_i \mathbf{u}_i^t = \sum_{j=0}^{k} a_j \boldsymbol{\mathcal{L}}^j \tag{14}$$

We specifically use the truncated Chebychev polynomial expansion of any spectral kernel $h(\lambda)$, as proposed in [4], in our experiments. The proposed iterative least square method with polynomial low pass filter, is termed as iterative least square (ILSR) in this paper.

4. INTERPOLATION BASED ON REGULARIZATION

The method presented above does not allow solutions from outside the $PW_\omega(G)$ space. This is advantageous if the input signal belongs to or is close to the subspace spanned by ω-BL signals. In general, for real world datasets such as recommendation systems, the graph signals tends to be smooth but not exactly band-limited. Therefore, we use a graph regularization framework in which we set up the cost of reconstruction as:

$$\mathbf{f}^* = \underset{\mathbf{x}}{\operatorname{argmin}} \underbrace{\|\mathbf{J}(\mathbf{f}_{du} - \mathbf{x})\|^2}_{A} + \alpha \underbrace{\|\mathbf{Hx}\|^2}_{B} \tag{15}$$

where A is the *data-fitting* term which computes the error between reconstructed signal and the original signal at the known samples and B is the Euclidean norm of the output of a highpass graph filter $\mathbf{H}$. Thus, the term A in the cost function penalizes the signals that are different from original signal at the known nodes, and the term B penalizes signals that have significant high frequency components. **Note that the optimal solution of (15) converges to the least square solution computed in (11), if $\mathbf{H} = \mathbf{I} - \mathbf{P}$ and $\alpha \rightarrow \infty$.** In our experiments, $\mathbf{H}$ is a chosen as a spectral graph transform with spectral kernel $h(\lambda) = \exp(-1/\lambda)$. The problem in (15) has a well known closed form solution given as:

$$\mathbf{f}^* = (\mathbf{J}^t\mathbf{J} + \alpha\mathbf{H}^t\mathbf{H})^{-1}\mathbf{J}^t\mathbf{J}\mathbf{f}_{du} = (\mathbf{J}^t\mathbf{J} + \alpha\mathbf{H}^t\mathbf{H})^{-1}\mathbf{f}_{du} \tag{16}$$

However, a direct implementation is computationally expensive, as it involves both the eigenvalue decomposition of the Laplacian matrix (to compute highpass filter $\mathbf{H}$) and inversion of a graph size matrix. Therefore, we propose an approximate iterative solution of the optimization problem in (15), similar to the method based on POCS [10] described in Section 3.

$$\begin{aligned} \mathbf{f}_0 &= \mathbf{f}_{du} \\ \mathbf{f}_{k+1} &= (\mathbf{I} - \beta\alpha\mathbf{H}^t\mathbf{H})\mathbf{f}_k + \beta\mathbf{J}^t\mathbf{J}(\mathbf{f}_{du} - \mathbf{f}_k) \end{aligned} \tag{17}$$

The parameter β is chosen to ensure convergence and maximize the rate of convergence. Replacing the spectral transform $\mathbf{H}^t\mathbf{H}$ by its polynomial approximation, we get a local iterative method for regularized graph signal recovery. Since we use a continuous function of λ to construct the regularization term, even a low degree polynomial approximation does not greatly affect the solution.

5. APPLICATION: RECOMMENDATION SYSTEMS

We apply the proposed interpolation method for collaborative filtering in recommendation systems. The input in this problem is a partially observed user-item rating matrix $\mathbf{R}$, such that $\mathbf{R}(u, m)$ is the rating given by user u to the item m. Based on this information, the system predicts new user-movie ratings. Following the setup in [2], an item-item graph G_0 is computed using partially observed rating matrix $\mathbf{R}$. The weight of the link between each pair of items i and j is computed as the cosine similarity [11] between i and j based on the training samples. For each test user u, we define $\mathcal{S}$ to be the set of items with known ratings, and and $\mathcal{U}$ to be the set of test items. We compute the subgraph $G_u = (\mathcal{S} \cup \mathcal{U}, E_u)$ of G_0, corresponding to the subset $\mathcal{S} \cup \mathcal{U}$ of nodes. We define DU signal $\mathbf{f}_u$ for u to be of size $|\mathcal{U} \cup \mathcal{S}|$, with $\mathbf{f}_u(\mathcal{U}) = 0$ and $\mathbf{f}_u(\mathcal{S})$ equal to known ratings. Subsequently, we compute interpolated signal $\hat{\mathbf{f}}_u$ by using graph based interpolation.

5.1. Graph Simplification

The item-item graphs computed using cosine similarity (as above), usually end up being highly connected if the rating matrix $\mathbf{R}$ is not sparse. The graph frequencies of very dense graphs are not uniformly distributed and hence not very informative in describing the smoothness of the signal. We observe that simplification of the item-item graph as a K *nearest neighbor* (KNN) leads to more uniform and informative distribution of graph frequencies. Therefore, we sparsify the subgraph G_u obtained for user u by connecting each item i in G_u by at most top K of its known neighbors, (ordered according to the decreasing link weights with item i). The best value of K is determined empirically to be around 30 in this paper.

5.2. Bilateral Link-Weight Adjustment

In addition to the sparsification step, the weights of the links between known samples $\mathcal{S}$ in the subgraph G_u are adjusted to reflect the user u's preferences, as is done in [2]. This adjustment step makes sense since subgraph G_u is the result of observing average correlation over a set of training users (multiple instances), and the

signal $\mathbf{f}_u$ corresponds to a single test user u. Specifically, we use bilateral-like weights for the links between known set of nodes, the exact implementation of which can be found in [2].

6. EXPERIMENTS

In our experiments, we use three different recommendation system datasets to evaluate the performance of proposed algorithms. Each dataset contains a reduced set of $100k$ randomly selected entries of user-item-ratings. The properties of the datasets are given in Table 1. In each case, we perform a 5 fold cross-validation, in which

Dataset	# users	# items	rating range	$mod[u]$	$mod[i]$
Movielens [12]	943	1682	1–5	215	57
Jester [13]	1412	100	0–20	80	1104
BX-Books [14]	6299	7046	1–10	80	5

Table 1: Datasets used in the experiments. $mod[u]$ and $mod[i]$: mode number of the ratings per user and per item, respectively. The ratings of Jexter datasets are originally fractional values in the range -10 to 9, which are rescaled to the range 0 to 20 and rounded to integer value.

we split the rating entries into 5 sets of approximately the same size. Then we evaluate the dataset 5 times, always using one set for testing and all other sets for training. In each iteration, an item-graph is formed from the training samples, as described in Section 5. The accuracy of the proposed methods depends to a large extent on the accuracy of computing link weights between items. Comparing the three databases in Table 1, the Jester database contains ratings of only 100 items (jokes). This is also the dataset with the highest votes per item. Therefore, we expect the link weights in the item graph, as computed from the training data to be highly accurate. On the contrary, the books database has the smallest number of ratings per item, which means that the weight of the links in the item-graph may be noisy and not very accurate. The movielens dataset seems to have enough ratings per item to properly compute the weights. Note that, insufficient training ratings is common problem in all collaborative filtering methods. Further, the accuracy of the proposed methods also depends on the number of movies rated by each user. In all the above databases, each user ranked enough items to give us a good prediction. Table 2 shows the RMSE of the proposed methods with some of the existing methods. To fairly compare the performance, the actual RMSE obtained for each dataset is normalized to be between 0 and 1, by dividing it with the maximum possible error (i.e., maximum rating - minimum rating). The best RMSE obtained in each dataset is represented with bold letters. It can be seen that the proposed regularized based kernel method (RBM) performs the best in the MovieLens and books dataset, and very close to the best method (PMF) in the Jester dataset. The iterative approximation of RBM (i.e., IRBM) also performs very close to the RBM method. However, in case of least square methods the iterative algorithm (ILSR) performs better than the exact method(LSR), on the movie and jokes datasets. This is because the ILSR method uses approximate low-pass filters which allow some energy to be in the frequencies bands higher than the cutoff ω, and is therefore closer to the RBM method.

7. CONCLUSIONS

In this paper, we presented two localized iterative graph filtering based methods for interpolation of graph signals from partially observed samples. The methods are implemented on recommendation system datasets, and provide reasonably good results when compared with the existing methods.

Dataset	KNN	PMF	RBM	IRBM	LSR	ILSR
Movielens	0.2482	0.2513	**0.2415**	0.2450	0.2514	0.2466
Jester	0.2348	**0.2299**	0.2304	0.2341	0.2344	0.2315
BX-Books	0.2677	0.2093	**0.1966**	0.2138	0.2651	0.2828

Table 2: Normalized RMSE results of the algorithms applied to the different datasets. KNN: K nearest neighbor method, PMF: probabilistic factorization method, RBM: Regularization based method, IRBM: Iterative regularization based method, LSR: Least Square Reconstruction, ILSR: iterative least square reconstruction.

8. REFERENCES

[1] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE sig. proc. mag.*, vol. 30, no. 3, pp. 83–98, May 2013.

[2] S.K. Narang, A. Gadde, and A. Ortega, "Signal processing techniques for interpolation of graph structured data," *ICASSP*, May 2013.

[3] D. Zhou and B. Schölkopf, "Learning from labeled and unlabeled data using random walks," in *Pattern Recognition*, pp. 237–244. Springer, 2004.

[4] D. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129 – 150, 2011.

[5] I. Pesenson, "Sampling in Paley-Wiener spaces on combinatorial graphs," *Transactions of the American Mathematical Society*, vol. 360, no. 10, pp. 5603–5627, 2008.

[6] A. Papoulis, "A new algorithm in spectral analysis and band-limited extrapolation," *Circuits and Systems, IEEE Transactions on*, vol. 22, no. 9, pp. 735–742, 1975.

[7] R.W. Gerchberg, "Super-resolution through error energy reduction," *Journal of Modern Optics*, vol. 21, no. 9, pp. 709–720, 1974.

[8] K. Sauer and J. Allebach, "Iterative reconstruction of bandlimited images from nonuniformly spaced samples," *Circuits and Systems, IEEE Transactions on*, vol. 34, no. 12, pp. 1497–1506, 1987.

[9] D.C. Youla and H. Webb, "Image restoration by the method of convex projections : Part I - Theory," *Medical Imaging, IEEE Transactions on*, vol. 1, no. 2, pp. 81–94, 1982.

[10] A.K. Katsaggelos, J. Biemond, R.W. Schafer, and R.M. Mersereau, "A regularized iterative image restoration algorithm," *Signal Processing, IEEE Transactions on*, vol. 39, no. 4, pp. 914–929, 1991.

[11] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," in *WWW '01*, 2001, pp. 285–295.

[12] "MovieLens dataset," 2003, `http://www.grouplens.org/`.

[13] K. Goldberg, T. Roeder, D. Gupta, and C. Perkins, "Eigentaste: A constant time collaborative filtering algorithm," *Information Retrieval*, vol. 4, no. 2, pp. 133–151, July 2001.

[14] C. Ziegler, S. M. McNee, J. A. Konstan, and G. Lausen, "Improving recommendation lists through topic diversification," *WWW'05*, pp. 22–32, July 2005.