

Т.М. Попова, М.Д. Жидков

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ НА ГРАФАХ

Аннотация. Рассмотрены понятия цифровой обработки сигналов, которые представлены в виде взвешенных графов, применены различные фильтры для очистки зашумления выборки сигналов в виде сенсорного и кластерного графов. Используются фильтры на основе метода теплового ядра, алгоритма Таубина и свертки этих алгоритмов. Проведен сравнительный анализ применения этих методов для сигналов в виде сенсорного и кластерного графов.

Ключевые слова: цифровые сигналы, цифровая обработка сигнала, фильтр, взвешенные графы, методы теплового ядра, алгоритм Таубина, произведение фильтров, моделирование.

Т.М. Popova, M.D. Jidkov

DIGITAL SIGNAL PROCESSING ON GRAPHS

Abstract. The concepts of digital signal processing, which are presented in the form of weighted graphs, are considered, various filters are applied to clean up the noise of a signal sample in the form of sensory and cluster graphs. Filters based on the thermal kernel method, the Taubin algorithm and the convolution of these algorithms are used. A comparative analysis of the application of these methods for signals in the form of sensor and cluster graphs is carried out.

Keyword: digital signals, digital signal processing, filter, weighted graphs, thermal core methods, Taubin algorithm, product of filters, modeling.

Введение

В современном обществе практически вся информация находится в цифровом пространстве. Цифровые данные окружают нас повсюду, и каждый аспект человеческой жизни может быть записан: внутриклеточные процессы, действия и запросы в интернете, финансовые транзакции, путешествия, приложения для отслеживания здоровья и др. Сложность таких сетей означает, что данные теперь хранятся на нестандартных и сложных конструкциях, не всегда поддающихся классическим статистическим методам обработки данных.

Представление сигналов в виде графа позволяет моделировать подобные данные со сложными взаимодействиями между ними. Например, пользователи приложений могут быть смоделированы как вершины, их запросы и обращения к тем или иным услугам, вкладкам, помощи, внутренним или внешним ресурсам – как ребра графа. Обработка сигналов на графах позволяет добавить определенные атрибуты к таким вершинам и моделировать их как сигналы на графах. При обработке цифровых сигналов их можно фильтровать, семплировать, то есть управлять начальной выборкой при известной цели, и это одна из наиболее сложных задач; их можно очищать от шума, изучать их структуру, моделировать процессы, происходящие внутри данных или при выходе из структуры графа. Обработка сигналов на графах плотно связана со многими теоретическими и практическими областями исследований; с ее помощью можно разработать новые инструменты и подходы к решению существующих проблем. Основой цифровой обработки сигналов стали работы Оппенгейма [6] и других авторов.

Попова Татьяна Михайловна

кандидат физико-математических наук, доцент, доцент кафедры прикладной математики, Тихоокеанский государственный университет, город Хабаровск. Сфера научных интересов: математическое моделирование физических, гидрологических процессов, обработка и анализ данных. Автор более 60 опубликованных научных работ.

Электронный адрес: 000511@pnu.edu.ru

Жидков Максим Дмитриевич

студент, Тихоокеанский государственный университет, город Хабаровск. Сфера научных интересов: компьютерное моделирование, обработка и анализ данных.

Электронный адрес: 2017100606@pnu.edu.ru

Сигналом на графе назовем множество значений, принадлежащее множеству вершин, которые связаны взвешенными ребрами. Сами сигналы могут брать свое начало в различных областях, но в отличие от классической обработки сигналов, лежащий в основе граф несет в себе определенную информацию о них. Разные типы графов позволяют моделировать разные типы связей, которые эти узлы представляют. В основном цифровая обработка сигналов изучает: сигналы и их представление; системы, обрабатывающие сигналы (фильтры); преобразование сигналов; семплирование сигналов. Фильтр – это система, избирательно меняющая форму сигнала (амплитудно-частотную или фазово-частотную характеристики). Цифровой фильтр – это определенная аппаратная или программная реализация алгоритма фильтрации. В цифровых фильтрах используются оцифрованные отсчеты аналоговых сигналов или хранящиеся в памяти числа.

Обработка дискретных сигналов данных со сложной структурой

Опираясь на концепцию теории алгебраических и спектральных графов, ряд авторов предложили новую структуру обработки дискретных сигналов для представления и анализа наборов данных со сложной структурой [7–9]. Они расширили традиционную теорию дискретной обработки сигналов до структурированных наборов данных, рассматривая их как сигналы, представленные графами так, что выборки сигналов указываются вершинами графа, а отношения между ними представлены взвешенными ребрами. Были определены понятия сигналов и фильтров на графах, а также понятия спектра и преобразования Фурье для сигналов на графах. Вопросы расширения идей и концептов стандартной обработки сигналов на графы могут быть разделены на категории стационарности и локализации.

Стационарность обычного сигнала во времени проверяется либо влиянием сдвига на статистические характеристики сигнала, либо путем наблюдения за самим сигналом в разные промежутки времени. Также можно определять стационарность, основываясь на спектральных свойствах оператора сдвига на вершинах [4]. Для решения проблем с существующими операторами сдвига было предложено использовать другой оператор [3]. Важная проблема семплирования – выборка значимых значений, по которой можно восстановить всю структуру с минимальными погрешностями – следует из той же проблемы классической обработки сигналов. Но на графах не проходит обычная модель семплирования, необходимо определить методы отбора наиболее информативных вершин.

Рассмотрим сигнал $s = (s_0, s_1, \dots, s_{N-1})$. Обозначим $\mathbf{s} = [s_0 s_1 \dots s_{N-1}]^T \in \mathbb{C}^N$. Пусть есть N выборок сигнала $s_n, n=0, 1, \dots, N-1$. Z -преобразование $s(z)$ временного сигнала s распределяет выборки в упорядоченное множество временных выборок как

$$s(z) = \sum_{n=0}^{N-1} s_n z^{-n}. \quad (1)$$

Фильтром назовем систему, преобразующую исходный сигнал, который можно записать аналогично (1):

$$h(z) = \sum_{n=0}^{N-1} h_n z^{-n}. \quad (2)$$

Так, выходной сигнал $s_{\text{вых}}$ фильтра h , примененного к входящему $s_{\text{вход}}$,

$$s_{\text{вых}}(z) = h(z) s_{\text{вход}}(z) \quad (3)$$

с фильтром $h_{\text{сдвиг}}(z) = z^{-1}$, который называется сдвигом. Фильтр h представим как матрицу \mathbf{H} . Это позволит переписать (3) в матричном виде:

$$\mathbf{s}_{\text{вых}} = \mathbf{H} \cdot \mathbf{s}_{\text{вход}}. \quad (4)$$

Сдвиг – это основной элемент цифровой обработки сигналов, позволяющий составлять более сложные фильтры. Различные типы применения некоторых фильтров рассмотрены в работах [1; 2].

Рассмотрим граф $G = (V, E)$, где $V = \{0, 1, \dots, N-1\}$ – множество вершин, определяющие выборки сигналов $\mathbf{s} = [s_0 s_1 \dots s_{N-1}]^T \in \mathbb{C}^N$, а E – множество ребер между смежными вершинами (то есть соединенными ребром) с матрицей смежности, определяющей матрицей сдвига

$$\mathbf{A}_c = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 1 & 0 & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 \end{bmatrix},$$

а операцию сдвига – произведением

$$[s_{N-1} s_0 \dots s_{N-2}]^T = \mathbf{A}_c \cdot [s_0 s_1 \dots s_{N-1}]^T.$$

Применение фильтра является умножением каждой выборки сигнала на коэффициент $h_n = \{h_n : n=0, \dots, N-1\}$. В системе на графе это можно представить выражением

$$\mathbf{s}_{\text{вых}} = \sum_{n=0}^{N-1} h_n \mathbf{L}^n \mathbf{s}_{\text{вход}}, \quad (5)$$

где $\mathbf{L} = \mathbf{D} - \mathbf{A}$ – лапласиан графа; \mathbf{D} – диагональная матрица степеней вершин графа (количество связей с остальными вершинами графа); \mathbf{L}^n – спектральное представление Лапласиана,

Цифровая обработка сигналов на графах

$$L^n = U \Lambda U^{-1} = \sum_{n=0}^{N-1} \lambda_n u_n u_n^{-1},$$

где $U = [u_0 u_1 \dots u_{N-1}]$ – ортонормированная матрица собственных векторов;

$\Lambda = \text{diag}[\lambda_0 \lambda_1 \dots \lambda_{N-1}]$ – диагональная матрица собственных значений лапласиана.

Подставляя в соотношение (5), получим

$$s_{\text{вых}} = \sum_{n=0}^{N-1} h_n U \Lambda^n U^{-1} s_{\text{вход}} = U h(\Lambda) U^{-1} s_{\text{вход}},$$

где

$$h(\Lambda) = \sum_{n=0}^{N-1} h_n \Lambda^n,$$

$h(\Lambda)$ – передаточная функция системы на графе.

В работе [12] даны теоретические объяснения, когда и почему спектральное представление оператора Лапласа можно рассматривать как значимое преобразование Фурье для достаточно гладких сигналов.

Использование фильтров в спектральной области достаточно прямолинейно и выполняется в три этапа:

1. Вычисление преобразования Фурье на графе входящего сигнала $Fs = U^{-1}s$.
2. Умножение результата на передаточную функцию $\tilde{F}s = h(\Lambda)Fs$.
3. Получение выходного сигнала путем вычисления обратного преобразования Фурье $\tilde{s} = U\tilde{F}s$.

Проектирование фильтров это один из основных этапов моделирования сигналов на графах. Могут быть использованы уже известные фильтры, адаптированные или специально созданные для графов [5].

Рассмотрим применение различных фильтров на сенсорном и кластерном графах. На Рисунках 1, 2 смоделированы начальный и искаженный сигналы. Используются следующие фильтры: теплового ядра, алгоритм Таубина и произведение (свертка) этих фильтров. Применение таких фильтров для спирального графа рассматривалось в работе [1].

Представим граф, в котором каждая вершина имеет температуру, а тепло может переходить только по ребрам между вершинами. Распространение тепла на таком графе будет следовать закону Ньютона – Рихмана и должно быть пропорционально весу ребра и разнице температур между вершинами. Пусть $s_n^{(t)}$ – температура вершины n во время t , тогда теплопередачу на графе можно выразить уравнением

$$\frac{ds_n^{(t)}}{dt} = -k \sum_{m=0}^{N-1} A_{nm} (s_n^{(t)} - s_m^{(t)}) = -k \left[D_{nn} s_n^{(t)} - \sum_{m=0}^{N-1} A_{nm} s_m^{(t)} \right],$$

которое в матричной форме уравнение примет вид

$$\frac{ds^{(t)}}{dt} = -kLs^{(t)}.$$

При замене параметров t и k в один $t' = kt$, это уравнение можно переписать как

$$\frac{ds^{(t')}}{dt'} = -Ls^{(t')}.$$

Тепловое ядро \mathbf{H}_t – это фундаментальное решение уравнения теплообмена. Оно определяется квадратной матрицей $\mathbf{H}_t = e^{-\mathbf{L}t}$. Тепловое ядро можно рассматривать как многочлен лапласиана, значит, его можно использовать как фильтр низких частот [10; 11].

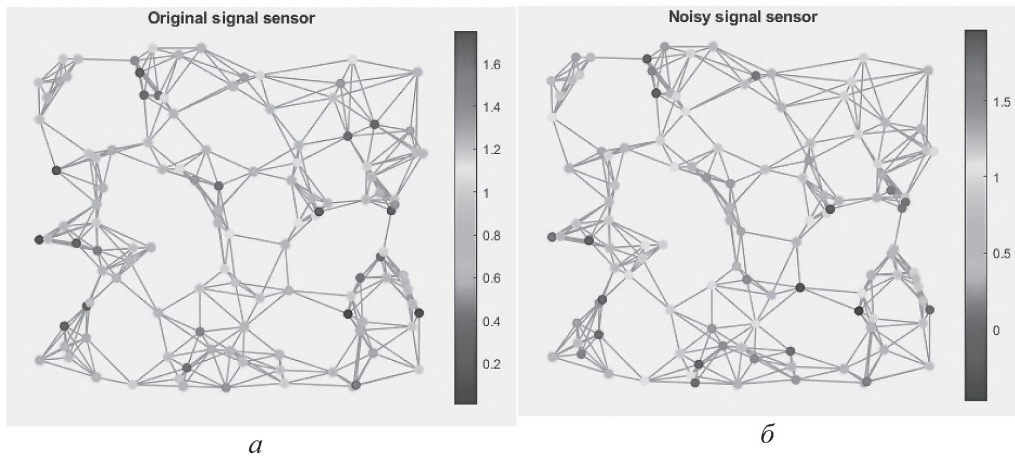


Рисунок 1. Сигнал s и шум y на сенсорном графе:
a – изначальный сигнал на графе; *b* – искаженный сигнал

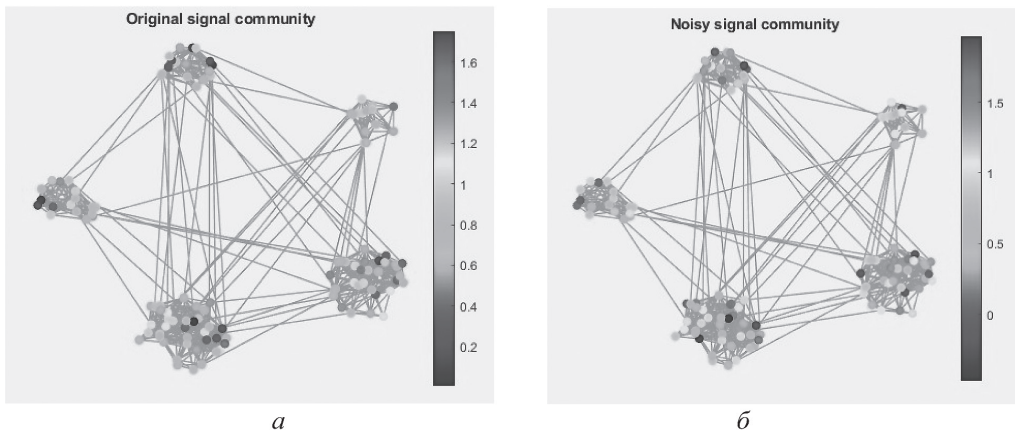


Рисунок 2. Сигнал s и шум y на кластерном графе: *a* – изначальный сигнал на графе;
b – искаженный сигнал

Алгоритм Таубина представляет собой итерационный алгоритм вида

$$\mathbf{s}_{p+1} = \mathbf{s}_p - \alpha \mathbf{L} \mathbf{s}_p = (\mathbf{E} - \alpha \mathbf{L}) \mathbf{s}_p,$$

который можно использовать в качестве простого, но эффективного фильтра сигналов на графах [14]. Так как минимумом квадратичной формы является константа, чтобы избежать получения только постоянного сигнала в результате стоит использовать его поочередно с $\mathbf{s}_{p+2} = (\mathbf{E} + \beta \mathbf{L}) \mathbf{s}_{p+1}$.

Компактная форма этого алгоритма называется алгоритмом Таубина:

$$s_{p+2} = (E - \alpha L)(E + \beta L)s_p.$$

Передаточная функция данного фильтра имеет вид

$$h(\lambda_n) = \left(1 + (\sigma_n - \beta_n)\lambda_n - \beta_n \sigma_n \lambda_n^2\right)^p,$$

где p – количество итераций.

Применение описанных фильтров на графах разной структуры

Рассмотрим применение описанных фильтров на графах разной структуры: стандартном сенсорном графе и кластерном. Количество вершин – 120. Введем определенный сигнал s и шум y , искажающий данный сигнал. Значения выборок на всех графах будут одинаковы, как и их искажение, различия на данном этапе будут только в связях, так как шум изменил значения некоторых выборок сигнала, и теперь некоторые вершины графов имеют другие цвета (см. Рисунок 1).

Рассмотрим каждый сигнал на графе в отдельности, применив каждый фильтр, чтобы очистить их от шумов, и проанализируем их эффект. Выберем значения коэффициентов таким образом: $\alpha = 0,4$, $\beta = 0,2$, $t = 0,2$; количество итераций $p = 25$. Получим следующий результат (см. Рисунки 3, 4). Для критерия эффективности применения того или иного фильтра была использована метрика расхождения между исходным и очищенным сигналом типа евклидова расстояния между соответствующими вершинами (разности в цвете).

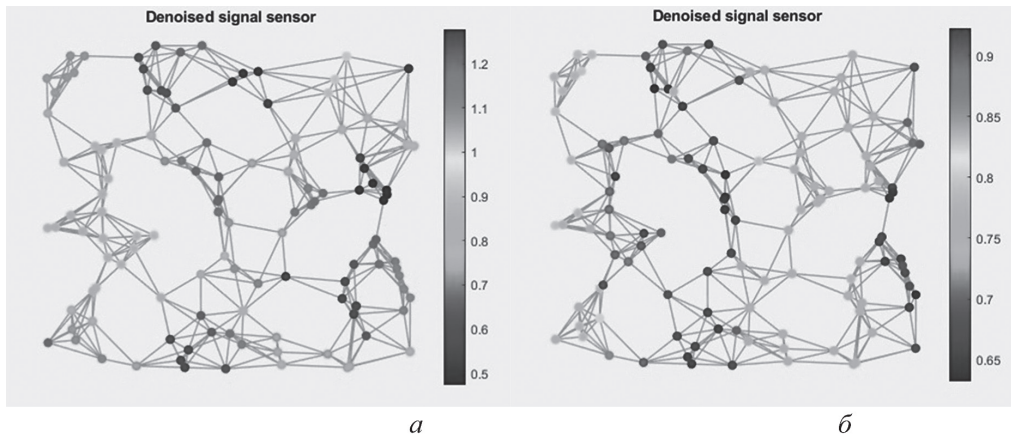


Рисунок 3. Результат применения фильтра теплового ядра (а), фильтра с алгоритмом Таубина (б)

Фильтр теплового ядра очистил сигнал недостаточно хорошо – большинство выборок приблизилось к своим изначальным значениям совсем немного, а некоторые даже отдалились. Можно заметить, что в целом сигнал стал более однородным – преобладают низкие частоты. Алгоритм Таубина дал заметно более хороший результат по сравнению с тепловым ядром. Из явных ошибок данного фильтра можно выделить только увеличение частоты некоторых выборок, предположительно, из-за большого значения β . Евклидово расстояние $D(s_{\text{вход}}, s_{\text{вых}}) = 0,089$.

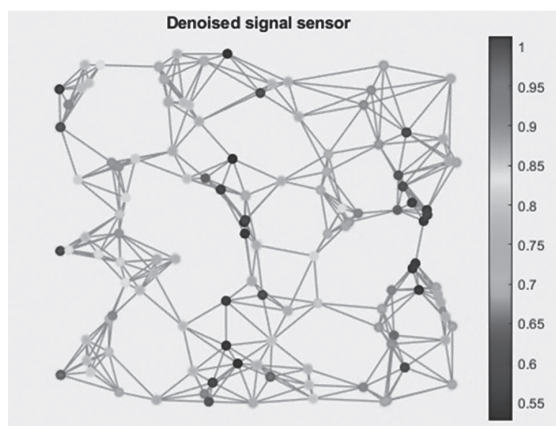


Рисунок 4. Результат применения свертки фильтров

Произведение фильтров дало схожий результат с тепловым ядром. В местах, где преобладали высокие частоты, фильтр сильно увеличил значения выборок, но в основном значения стали меньше. При данных параметрах использование произведения фильтров не оправдано. Евклидово расстояние $D(s_{\text{вход}}, s_{\text{вых}}) = 0,141$.

Далее рассмотрим результаты работ фильтров на кластерном графе (см. Рисунки 5, 6).

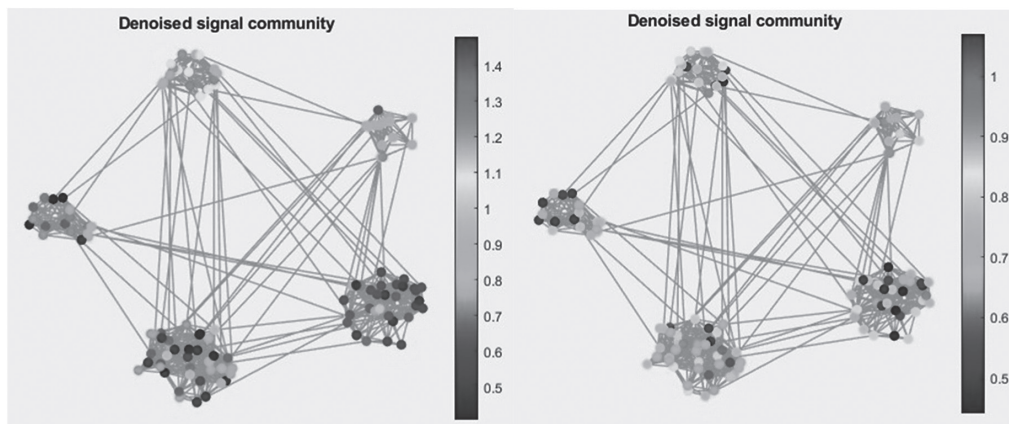


Рисунок 5. Результат применения фильтра теплового ядра (а),
фильтра с алгоритмом Таубина (б)

Фильтр Таубина на кластерном графе показал очень хороший результат, так как значения большинства выборок сильно приблизились к изначальным. Евклидово расстояние $D(s_{\text{вход}}, s_{\text{вых}}) = 0,037$, но произведение фильтров работает неудовлетворительно – практически все значения выборок исказились еще больше по сравнению с результатом шума. Но стоит обратить внимание на то, каким образом фильтр распределил значения: в рамках каждого кластера все значения схожи между собой, что позволяет выделить свойство сортировки данного фильтра на данном графе, и это может привести к различным применениям данного фильтра, несмотря на то, что шум не был очищен.

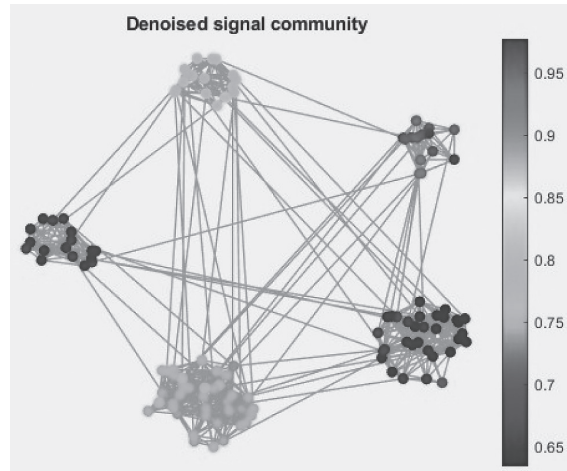


Рисунок 6. Результат применения произведения фильтров

В работе [1] показано, что на спиральном графе тепловое ядро показывает относительно низкую эффективность, а фильтр алгоритма Таубина справился достаточно хорошо, несмотря на то, что в некоторых местах с высокими частотами он сильно снизил значения выборок. Произведение фильтров выявило наименьшую из всех метрику расхождений, следовательно, дало наилучший результат из всех рассматриваемых моделей.

Заключение

Таким образом, можно сделать вывод, что применение того или иного фильтра зависит от вида графа, описывающего процесс и дает широкие возможности по компоновке разных типов фильтров. В итоге наилучшим фильтром при отсеивании шумов оказался фильтр, основанный на алгоритме Таубина: он показал отличный или хороший результат для всех рассматриваемых типов графов. Фильтр теплового ядра в данных экспериментах, напротив, дал наихудший результат, поэтому стоит провести более подробные исследования по нахождению подходящего применения этого типа фильтра.

Новый тип фильтра, представленный в данной работе и работе [1], показал смешанные результаты. Так, на сенсорном графе результат был подобен результату теплового ядра, на спиральном новый фильтр дал наилучший результат из всех рассматриваемых, а на кластерном, несмотря на наихудший результат очистки, было выявлено свойство сортировки, которое может быть полезно в определенных приложениях.

Литература

1. Жидков М.Д., Попова Т.М. Обработка цифровых сигналов методами теории графов // Far East Math – 2021: материалы национальной научной конференции. Хабаровск, 2021. С. 3–9.
2. Рыжаков В.В., Прохорова Е.В. Цифровая обработка сигналов электрокардиограммы в распределенной вычислительной системе // Системы синхронизации, формирования и обработки сигналов. 2020. Т. 11, № 5. С. 35–43.
3. Girault B., Gonçalves P., and Fleury É. (2015) Translation on graphs: An isometric shift operator. IEEE Signal Process Lett, 2015, Vol. 22, No. 12, Pp. 2416–2420.

4. Marques A.G., Segarra S., Leus G. (2017) Stationary graph processes and spectral estimation. *IEEE Trans. Signal Process*, 2017, Vol. 65, No. 22, Pp. 5911–5926.
5. Narang S. K. and Ortega A. (2012) Perfect reconstruction two-channel wavelet filter banks for graph structured data. *IEEE Trans. Signal Process*, 2012, Vol. 60, No. 6, Pp. 2786–2799.
6. Oppenheim A.V., Schaffer, R.W. (1989) *DiscreteTime Signal Processing*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1989, 893 p.
7. Püschel M., Moura J.M.F. (2003) The algebraic approach to the discrete cosine and sine transforms and their fast algorithms. *SIAM J. Comput*, 2003, Vol. 32, No. 5, Pp. 1280–1316.
8. Sandryhaila A., Moura J.M.F. (2013) Discrete signal processing on graphs. *IEEE Trans. Signal Process*, 2013, Vol. 61, No. 7, Pp. 1644–1656.
9. Shuman D.I., Narang S.K., Frossard P. (2013) The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Process. Mag*, 2013, Vol. 30, No. 3, Pp. 83–98.
10. Wagner R., Choi H., Baraniuk R. (2005) Distributed wavelet transform for irregular sensor network grids: *Proc. IEEE SSP Workshop*, 2005, Pp. 1196–1201.
11. Wagner R., Cohen A. (2006) An architecture for distributed wavelet analysis and processing in sensor networks: *Proc. IPSN*, 2006, Pp. 243–250.
12. Zhu X., Rabbat M. (2012) Approximating signals supported on graphs: *Proc. ICASSP*, 2012, Pp. 3921–3924.

Literature

1. Jidkov M.D., Popova T.M. (2021) *Obrabotka cifrovyyh signalov metodami teorii grafov* [Digital signals processing by graph theory method]. *Far East Math – 2021: Materialy nacional'noj nauchnoj konferencii*. Khabarovsk, 2021, Pp. 3–9 (in Russian).
2. Ryzhakov V.V., Prohorova E.V. (2020) [Digital processing of electrocardiogram signals in a distributed computing system]. *Sistemy sinhronizacii, formirovaniya i obrabotki signalov*, 2020, Vol. 11, No. 5, Pp. 35–43 (in Russian).
3. Girault B., Gonçalves P., and Fleury É. (2015) Translation on graphs: An isometric shift operator. *IEEE Signal Process Lett*, 2015, Vol. 22, No. 12, Pp. 2416–2420.
4. Marques A.G., Segarra S., Leus G. (2017) Stationary graph processes and spectral estimation. *IEEE Trans. Signal Process*, 2017, Vol. 65, No. 22, Pp. 5911–5926.
5. Narang S. K. and Ortega A. (2012) Perfect reconstruction two-channel wavelet filter banks for graph structured data. *IEEE Trans. Signal Process*, 2012, Vol. 60, No. 6, Pp. 2786–2799.
6. Oppenheim A.V., Schaffer, R.W. (1989) *DiscreteTime Signal Processing*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1989, 893 p.
7. Püschel M., Moura J.M.F. (2003) The algebraic approach to the discrete cosine and sine transforms and their fast algorithms. *SIAM J. Comput*, 2003, Vol. 32, No. 5, Pp. 1280–1316.
8. Sandryhaila A., Moura J.M.F. (2013) Discrete signal processing on graphs. *IEEE Trans. Signal Process*, 2013, Vol. 61, No. 7, Pp. 1644–1656.
9. Shuman D.I., Narang S.K., Frossard P. (2013) The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Process. Mag*, 2013, Vol. 30, No. 3, Pp. 83–98.
10. Wagner R., Choi H., Baraniuk R. (2005) Distributed wavelet transform for irregular sensor network grids: *Proc. IEEE SSP Workshop*, 2005, Pp. 1196–1201.

11. Wagner R., Cohen A. (2006) An architecture for distributed wavelet analysis and processing in sensor networks: Proc. IPSN, 2006, Pp. 243–250.
12. Zhu X., Rabbat M. (2012) Approximating signals supported on graphs: Proc. ICASSP, 2012, Pp. 3921–3924.