

تشخیص و مکان‌یابی ناهنجاری شبکه حسگر با استفاده از پردازش سیگنال روی گراف

حمید سعیدی سورک^۲

^۲دانشکده مهندسی برق، دانشگاه یزد

یزد، ایران

Saeedi@yazd.ac.ir

شماره تلفن: ۰۳۵۳۱۲۳۳۰۲۰

سپیده السادات شمس فیروزآبادی^۱

^۱دانشکده مهندسی برق، دانشگاه یزد

یزد، ایران

Sepideh.shams.id@stu.yazd.ac.ir

شماره تلفن: ۰۹۱۳۴۳۱۹۱۱۷

روی گراف، می‌توان مفاهیم موجود در پردازش سیگنال کلاسیک مانند فیلتر، تبدیل فوری و همواری سیگنال را به این حوزه تعمیم داد و برای کاربردهایی نظیر جایگذاری مناسب حسگرها و تشخیص ناهمواری استفاده کرد. شایان ذکر است گراف ساختار فیزیکی شبکه را محقق می‌کند و سیگنال روی گراف مقادیر اندازه‌گیری شده به وسیله هر حسگر را نشان می‌دهد. به عبارت بهتر در پردازش سیگنال روی گراف، دو دسته اطلاعات شامل گراف پس زمینه و مقدار سیگنال روی رئوس گراف در دسترس است [۲].

به یافتن الگوهایی در داده‌ها که با رفتار قابل انتظار، مطابقت ندارد، تشخیص ناهنجاری گفته می‌شود [۳]. ناهنجاری با عناوینی دیگری مانند داده‌های دورافتاده^۵، اختلالات، مشاهدات غیرواقعی و خرابی^۶ نیز نامیده می‌شود. از روش‌های تشخیص ناهنجاری در کاربردهای مختلفی مانند تشخیص نفوذ و مسائل سایبری، تشخیص پزشکی، پردازش تصویر و صوت، پردازش متنی و شبکه حسگر استفاده می‌شود [۳]. به طور کلی تشخیص ناهنجاری را می‌توان به پنج روش تقسیم‌بندی کرد: روش مبتنی بر آمار، روش مبتنی بر نزدیکترین همسایه، روش مبتنی بر خوشه‌بندی^۷، روش مبتنی بر طبقه‌بندی^۸ و روش مبتنی بر تجزیه طیفی. با به کارگیری هر کدام از روش‌ها یا ترکیب آن‌ها، تکنیک‌های متفاوتی برای تشخیص ناهنجاری ارائه شده است [۴].

در مدل‌سازی شبکه‌ها، گراف می‌تواند فواصل و شباهت بین داده‌ها را تصویر کند. حال از آنجا که روش‌های تشخیص خطا مبتنی بر نزدیکترین همسایه، بر اساس مقایسه داده با سایر داده‌ها و به خصوص داده‌هایی که در همسایگی آن‌ها قرار دارند انجام می‌شود لذا تشخیص ناهنجاری با پردازش سیگنال روی گراف را می‌توان در این گروه

چکیده — امروزه با استفاده از حسگرها می‌توان بر پدیده‌های محیطی نظیر دما، نور و فشار، کنترل یا نظارت داشت. گاهی بخشی از اطلاعات حسگرها دچار تغییرات ناخواسته‌ای می‌شوند که تشخیص مکان و زمان این ناهنجاری یکی از مباحث پرکاربرد در حوزه‌های امنیت، سلامت و نظارت است. اطلاعات ثبت شده توسط حسگرها، علاوه بر وابستگی زمانی، وابستگی مکانی نیز دارند، لذا برای پردازش آن‌ها و تشخیص تغییرات می‌توان از پردازش سیگنال روی گراف و ابزارهای آن استفاده نمود. در این مقاله داده‌های واقعی شبکه حسگر با استفاده از مفاهیم گراف، مدل شده است و سپس با پردازش سیگنال روی گراف، الگوریتمی پیشنهاد شده است که به تشخیص و مکان‌یابی ناهنجاری می‌پردازد. نتایج شبیه‌سازی، عملکرد مناسب و حجم محاسباتی کم الگوریتم را در مقایسه با الگوریتم‌های موجود تایید می‌کند.

کلید واژه — پردازش سیگنال روی گراف، تبدیل فوری روی گراف، تشخیص ناهنجاری، تغییرات محلی، شبکه حسگر

۱. مقدمه

امروزه برخی از سیگنال‌ها بر روی شبکه‌های مختلف مانند شبکه‌های اجتماعی^۱، شبکه‌های حسگر^۲ و شبکه‌های بیولوژیکی^۳ قرار دارند. این سیگنال‌ها دارای ساختار نامنظم و پیچیده‌ای هستند. در واقع برای پردازش چنین سیگنال‌هایی، مکانی که داده روی آن قرار دارد و ارتباط آن با سایر نقاط اندازه‌گیری دارای اهمیت است [۱]. استخراج اطلاعات ارزشمند از این سیگنال‌ها نیاز به روش‌های نوآورانه‌ای دارد. گراف‌ها توانایی مدل‌سازی ارتباط پیچیده بین داده‌ها را دارند و پردازش سیگنال روی گراف^۴، به منظور پردازش این داده‌ها مورد استفاده قرار می‌گیرد [۱]. با توجه به ساختار گراف و ویژگی‌های پردازش سیگنال

⁵ Outlier

⁶ Malfunction

⁷ Clustering

⁸ Classification

¹ Social Networks

² Sensor Networks

³ Biological Networks

⁴ Graph Signal Processing (GSP)

دسته‌بندی کرد. از طرفی با استفاده از روش مبتنی بر تجزیه طیفی در حوزه گراف می‌توان رفتارهای ناهنجاری را در زیرفضای جدید با دقت بیشتری مورد بررسی قرار داد [۲]. همچنین مفاهیم همواری و مؤلفه‌های فرکانس پایین و بالای سیگنال که در تشخیص ناهنجاری کاربرد دارند، را می‌توان با استفاده از پردازش سیگنال روی گراف استخراج کرد [۱]. اخیراً استفاده از پردازش سیگنال روی گراف در تشخیص ناهنجاری، مورد توجه قرار گرفته است. هدف از [۵] و [۶]، تشخیص ناهنجاری داده‌ها روی شبکه حسگر است. برای انجام این کار علاوه بر استفاده از داده‌های حسگر، می‌توان با استفاده از اطلاعاتی مانند فاصله بین حسگرها و هرگونه دانش قبلی در مورد محیط، رفتار غیرعادی موجود در شبکه را تشخیص داد. از فیلترهای مبتنی بر گراف و تعریف حد آستانه، می‌توان برای طراحی زیرفضاهای هنجار و ناهنجار از سیگنال بهره برد و از نتایج حاصل، وجود یا عدم وجود ناهنجاری را در داده‌های شبکه تخمین زد. در [۶] برای ساخت گراف علاوه بر فواصل اقلیدسی از اطلاعات دمایی گذشته نیز استفاده شده است. در [۷] پیکسل‌های تصویر و مکان قرارگیری آن‌ها به عنوان گراف وزن‌دار و سطوح خاکستری به صورت سیگنال روی گراف مدل می‌شود. با استفاده از ماتریس لاپلاسیان گراف، اطلاعات آماری پس‌زمینه تخمین زده می‌شود و با پردازش سیگنال روی گراف، تشخیص انواع ناهنجاری تصویر و یا حتی لبه‌ها^۹ و تخمین گرادیان‌ها^{۱۰} نیز امکان‌پذیر است. در [۸] نیز تشخیص داده‌های خراب ناشی از حمله به شبکه‌های هوشمند انرژی، مورد بررسی قرار گرفته است. این حمله‌ها می‌توانند داده‌های شبکه را به خطر بیندازند. در این مقاله با استفاده از مشخصه‌های الکتریکی، گراف بدون جهت تشکیل می‌شود و سپس با استفاده از تبدیل فوری روی گراف^{۱۱} و حد آستانه، حمله‌ها تشخیص داده می‌شوند. در [۹] برای مدل‌سازی گراف از فواصل اقلیدسی و ارتفاع استفاده شده است و برای تشخیص ناهنجاری از پنجره‌های تبدیل فوری و تبدیل موجک^{۱۲} و یادگیری ماشین^{۱۳} استفاده شده است. در [۱۰] روشی مبتنی بر فیلتر روی گراف برای تشخیص و مکان‌یابی ناهنجاری ارائه شده است. با استفاده از فیلتر بالاگذر، مؤلفه‌های فرکانس پایین داده‌ها حذف شده و سپس با اطلاعات گذشته دما و تعریف معیاری از اختلاف دما با میانگین سه روز گذشته و تعریف حد آستانه، ناهنجاری تشخیص و مکان‌یابی می‌شود. روش

ساخت گراف در [۱۱] بر اساس همبستگی سیگنال‌های رأس‌هاست و از معیار تغییرات کلی^{۱۴} و با استفاده از امید ریاضی و واریانس این معیار در مشاهدات گذشته، ناهنجاری را در زمان و مکان تشخیص می‌دهد. همان‌طور که اشاره شد گراف نمایش بهتری از داده‌های شبکه‌ای ارائه می‌کند، از طرفی روش‌های موجود در گراف، معمولاً به علت فیلترهای گراف و تجزیه گراف، پیچیدگی محاسباتی بالایی دارند. در این مقاله یک روش با استفاده از پردازش سیگنال روی گراف ارائه شده است که نسبت به روش‌های موجود عملکرد قابل قبول و پیچیدگی محاسباتی کمتری دارد. ابتدا یک گراف بدون جهت و وزن‌دار برای حسگر دما با استفاده از فواصل هندسی و با در نظر گرفتن تأثیر ارتفاع و نیز همبستگی دمایی هر حسگر با سایر حسگرها ساخته می‌شود. برای تشخیص ناهنجاری از معیار تغییرات محلی^{۱۵} و مقایسه آن با حد آستانه، وجود یا عدم وجود ناهنجاری تشخیص داده می‌شود و برای تشخیص مکان وقوع ناهنجاری از معیار تغییرات محلی نرمالیزه استفاده می‌شود. روش پیشنهادی، عملکرد قابل قبولی در تشخیص و مکان‌یابی ناهنجاری دارد. در ادامه مقاله، ابتدا مقدمات و مفاهیم پردازش سیگنال روی گراف در بخش دوم معرفی می‌شود. بخش سوم به توضیح ناهنجاری و وقوع آن در شبکه‌های حسگر و چگونگی ساخت گراف می‌پردازد. در بخش چهارم روشی برای تشخیص مکان ناهنجاری در شبکه حسگر پیشنهاد می‌شود و سپس نتایج شبیه‌سازی ارائه شده و مورد بحث قرار می‌گیرد و در بخش آخر نیز نتیجه‌گیری بیان خواهد شد.

۲. مبانی نظری

گراف غیرجهت‌دار $G = (V, W)$ مفروض است که در آن $V = \{v_0, v_1, \dots, v_{N-1}\}$ یک مجموعه محدود از N رأس گراف و ماتریس وزن $W \in \mathbb{C}^{N \times N}$ یک ماتریس مربعی است که هر المان ماتریس $w_{i,j} \in \mathbb{C}$ نشان‌گر یک یال وزن‌دار از گره v_i به v_j است [۲]. در این مقاله وزن‌ها مثبت فرض می‌شود و داریم $w_{i,j} \in \mathbb{R}^+$. مجموعه همسایگی v_j به مجموعه گره‌هایی مانند i اطلاق می‌شود به طوری که $w_{i,j} \neq 0$ باشد. ماتریس وزن برای گراف بدون جهت متقارن است [۱]. درجه یک گره از گراف برابر مجموع وزن یال‌هایی است که با گره مورد نظر تلاقی دارند. لذا ماتریس درجه D برای یک گراف به شکل یک ماتریس قطری

¹³ Machine Learning

¹⁴ Total Variation

¹⁵ Local Variation (LV)

⁹ Edges

¹⁰ Gradient

¹¹ Graph Fourier Transform

¹² Wavelet Transform

تعریف می‌شود که هر درایه قطر اصلی آن نمایان گر درجه گره متناظر با آن درایه است. ماتریس لاپلاسین گراف از $L = D - W$ محاسبه می‌شود. این ماتریس را می‌توان با استفاده از رابطه $L = U\Lambda U^T$ قطری‌سازی کرد که Λ ماتریس قطری از مقادیر ویژه و ستون‌های ماتریس U ، بردارهای ویژه ماتریس لاپلاسین هستند. از آنجایی که ماتریس L متقارن و نیمه معین مثبت است، مجموعه‌ی کاملی از بردارهای ویژه متعامد و یک‌ه را شامل می‌شود [۲].

یک ماتریس $X \in \mathbb{R}^{N \times M} = [x_0, x_1, \dots, x_{N-1}]^T$ را در نظر بگیرید که در آن هر سطر $x_i \in \mathbb{R}^M$ یک سیگنال زمانی روی گره i و هر یک از M ستون آن، سیگنال روی گراف در زمان $t \in \{0, 1, \dots, M-1\}$ است. تبدیل فوریه کلاسیک، سیگنال را بر پایه‌های متعامد تصویر می‌کند و محتوای فرکانسی سیگنال را نشان می‌دهد. این مفهوم در پردازش سیگنال روی گراف نیز وجود دارد که به آن تبدیل فوریه روی گراف می‌گویند. تبدیل فوریه گراف و عکس آن ابزاری برای نمایش معادل یک سیگنال در دو حوزه رأس و طیف را فراهم می‌کند. یکی از تعاریف تبدیل فوریه از سیگنال گراف با استفاده از ماتریس لاپلاسین محاسبه می‌شود. در تبدیل فوریه سیگنال روی گراف مقادیر ویژه همانند فرکانس و بردارهای ویژه همانند پایه‌های متعامد هستند [۲].

اگر سیگنال در یک لحظه، به فرم $x = [x(0), x(1), \dots, x(N-1)]^T$ بر روی رئوس گراف G تعریف شود، تبدیل فوریه و عکس تبدیل فوریه سیگنال روی گراف به ترتیب به فرم $\hat{x} = U^T x$ و $x = U\hat{x}$ بیان می‌شود که $\hat{x}$ نشان‌دهنده تبدیل فوریه سیگنال روی گراف است.

یک سیگنال روی گراف دارای پهنای باند محدود K است، اگر و تنها اگر تبدیل فوریه سیگنال روی گراف آن، K مؤلفه مخالف صفر داشته باشد. اگر این K ضریب تبدیل فوریه در ابتدای طیف قرار گیرند، سیگنال را هموار^{۱۶} گویند. در صورتی که اندازه ضرایب تبدیل فوریه گراف برای مقادیر ویژه بزرگ، مقدار ناچیزی داشته باشند، سیگنال را تقریباً هموار^{۱۷} می‌نامند [۱۲]. همواری برای سیگنال روی گراف باعث می‌شود که مقدار سیگنال یک رأس به مقادیر همسایگان آن نزدیک باشد [۲]. یکی از روابط ساده برای تشخیص همواری سیگنال نسبت به رئوس همسایه استفاده از:

$$(Lx)(i) = \sum_{j \in N_i} W_{i,j} [x(i) - x(j)], \quad (1)$$

است که اختلاف یک رأس با سایر رئوس همسایه اش و با توجه به یال وزن دار میان آن‌ها محاسبه می‌کند [۲]. طبق رابطه (۱) اگر اختلاف بین

دو سیگنال و وزن یال‌های بین آن دو گره، زیاد باشد، حاصل آن مقدار بزرگی خواهد داشت. اختلاف دو سیگنال، می‌تواند به صورت منفی نیز باشد و بنابراین امکان دارد مقادیری که در عمل جمع در رابطه (۱) شرکت دارند اثر یکدیگر را خنثی کنند و مقدار عبارت کلی، کاهش یابد. رابطه (۲) که تغییرات محلی در رأس i است و به صورت [۲]

$$\| \nabla_i x \|_2 := \left[\sum_{j \in N_i} W_{i,j} [x(i) - x(j)]^2 \right]^{\frac{1}{2}}. \quad (2)$$

تعریف می‌شود، این نقص را برطرف کرده است. به علت غیرمنفی بودن ماتریس وزن، مقدار حاصل همواره مثبت است. با استفاده از این تعریف، رابطه p -دیریکله^{۱۸} از سیگنال x به صورت زیر تعریف می‌شود [۲]:

$$S_p(x) = \frac{1}{p} \sum_{i \in V} \| \nabla_i x \|_2^p = \frac{1}{p} \sum_{i \in V} \left[\sum_{j \in N_i} W_{i,j} [x(i) - x(j)]^2 \right]^{\frac{p}{2}}. \quad (3)$$

در صورتی $p = 2$ باشد، $S_2(x)$ لاپلاسین گراف از مرتبه دوم نامیده می‌شود، یعنی

$$S_2(x) = \frac{1}{2} \sum_{i \in V} \sum_{j \in N_i} W_{i,j} [x(i) - x(j)]^2 = \sum_{(i,j) \in E} W_{i,j} [x(i) - x(j)]^2. \quad (4)$$

رابطه را می‌توان به صورت

$$S_2(x) = x^T L x, \quad (5)$$

نیز نوشت که بیان گر معیاری از همواری کل سیگنال در گره‌های مجاور است. هرچه مقدار این عبارت برای یک سیگنال روی گراف بیشتر باشد تنوع بیشتری در سیگنال‌های مربوط به گره‌های مجاور وجود دارد و سیگنال ناهموارتری بر روی گراف قرار دارد [۲].

۳. ساخت گراف

اطمینان سنجی و صحت‌سنجی داده‌های حساس قبل از استفاده آن در تصمیم‌گیری، بسیار مهم است. یکی از رویکردهای آسان برای تشخیص ناهنجاری، شناسایی مناطق عادی و سپس دسته‌بندی هرآنچه در آن منطقه قرار ندارد، به عنوان داده‌های غیرعادی است. اما تعریف و تشخیص مناطق عادی سخت است. علاوه بر این مرزبندی بین مناطق عادی و غیرعادی نیز معمولاً مشکل خواهد بود. از طرف دیگر رفتار حوزه عادی ممکن است تغییر داشته باشد و الگوی ثابتی موجود نباشد [۴].

ناهنجاری در شبکه‌های حسگر بی‌سیم ممکن است به علت خرابی، شکستگی یا تمام شدن باتری‌های حسگر اتفاق بیفتد و یا عملیات

¹⁸ p-Dirichlet

¹⁶ Smooth

¹⁷ Approximately Smooth

۴. روش پیشنهادی برای تشخیص ناهنجاری

وقوع یک ناهنجاری در یک گراف، باعث تغییراتی در همواری سیگنال می‌شود. ناهنجاری باید به صورت محلی و با استفاده از روابط همسایگی رؤس تشخیص داده شود. روش پیشنهادی در فاز اول وجود ناهنجاری و در فاز دوم مکان آن را تشخیص می‌دهد. سیگنال مورد نظر در روز t و e روز گذشته به ترتیب به صورت $x^{(t)}$ و $x^{(t-e)}$ است. با دماهای k روز گذشته، k بردار تغییرات محلی از رابطه (۲) استخراج می‌شود. به ازای هر رأس، ضریب تغییرات محلی سیگنال $x^{(t)}$ با در نظر گرفتن حد آستانه λ ، با بزرگترین ضریب تغییرات محلی در k روز گذشته $(LV_{max}(n))$ مقایسه می‌شود، حال اگر به ازای حداقل یک رأس، رابطه $LV(n) > \lambda \cdot LV_{max}(n)$ (۱۱)

برقرار باشد، به معنای وجود ناهنجاری در سیگنال روی گراف است. در فاز دوم، مکان ناهنجاری تشخیص داده می‌شود. براساس رابطه (۲)، تغییرات محلی رؤس با درجه بزرگتر، مقادیر بیشتری خواهد داشت و مقدار آن، با مجذور مجموع وزن‌ها رابطه مستقیم دارد. بنابراین لازم است ضریب تغییرات محلی هر رأس با تقسیم بر مجذور درجه آن رأس نرمالیزه شود (LV_n) . با تعریف آستانه‌های مناسب، مجموعه‌ای از رؤس که در رابطه (۱۱) صدق می‌کنند، با هم مقایسه و بزرگترین مقدار به عنوان ناهنجاری تشخیص داده می‌شود. الگوریتم ۱ مراحل تشخیص و مکان‌یابی سیگنال ناهنجار را شرح می‌دهد.

الگوریتم ۱: تشخیص و مکان‌یابی وقوع یک ناهنجاری

ورودی: گراف $G = (V, W)$ ، $x^{(t)} = [x(0), x(1), \dots, x(N-1)]^T$ ، $x^{(t-e)}$ ، $e = 1, \dots, k$

خروجی: تشخیص و تعیین مکان گره دارای ناهنجاری

گام اول: تشکیل بردار تغییرات محلی به ازای بردار $x^{(t)}$ و k بردار تغییرات محلی به ازای $x^{(t-e)}$ با استفاده از رابطه (۲)

گام دوم: انتخاب بزرگترین ضریب تغییرات محلی از k بردار به ازای هر رأس به عنوان بردار LV_{max}

گام سوم: تعیین آستانه λ به صورت تجربی

گام چهارم: اگر حداقل در یک رأس $LV(n) > \lambda \cdot LV_{max}(n)$ برقرار باشد، $x^{(t)}$ دارای ناهنجاری بوده و گام بعدی اجرا می‌شود، در غیر این صورت سیگنال بدون ناهنجاری است.

گام پنجم: تقسیم ضریب $LV(n)$ هر رأس بر مجذور درجه آن رأس به عنوان تغییرات محلی نرمالیزه $(LV_n(n))$

گام ششم: انتخاب رأس m^* به عنوان رأس ناهنجار

$m^* = \arg \max_{m \in V} \{LV_n(n) | \exists LV(n) > \lambda \times LV_{max}(n)\}$

تخریبی دشمن باشد که به منظور گمراهی، مقداری اطلاعات اشتباه را به صورت عمدی به شبکه تزریق می‌کند. نوع دیگر از ناهنجاری، رویداد خاصی نظیر آتش‌سوزی است که در اصل هدف طراح شبکه، تشخیص این رویداد است [۱۳]. در یک شبکه حسگر، یک ناهنجاری، ممکن است در یک زمان خاص و یا در یک مکان خاص اتفاق بیفتد. به عبارت بهتر در حالت زمانی، هدف تشخیص یک زمان خاص از طول بازه‌ی اندازه‌گیری است که در آن ناهنجاری رخ داده است و در حالت مکانی، هدف تشخیص یک نقطه ناهنجار در زمان مورد نظر است [۱۳].

در این مقاله تشخیص ناهنجاری در شبکه حسگر در زمان و مکان بررسی می‌شود. وقوع یک ناهنجاری با تابع ضربه در رأس مدل می‌شود و داریم

$$\delta_m(n) = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases} \quad (6)$$

سیگنال $\tilde{x}^{(t)}$ دارای ناهنجاری به مقدار a در رأس m در لحظه t با

$$\tilde{x}^{(t)} = x^{(t)} + a \cdot \delta_m, \quad (7)$$

بیان می‌شود: در تشخیص ناهنجاری با پردازش سیگنال روی گراف، به علت تاثیر رؤس همسایه، ساخت گراف و مقداردهی یال‌ها مهم است. در بسیاری از مقالات، وزن‌دهی یال‌ها بر اساس فاصله اقلیدسی رؤس است [۵-۶] و [۹ و ۱۰ و ۱۴]. مثلاً در [۱۴] ماتریس وزن از

$$W_{ij} = e^{-\frac{d_{ij}^2}{2\alpha^2}}, \quad (8)$$

محاسبه می‌شود که d فاصله اقلیدسی بین رؤس و α پارامتر ثابت است. در [۹] علاوه بر فاصله اقلیدسی در دو بُعد طول و عرض جغرافیایی، از ارتفاع به عنوان بُعد سوم استفاده شده است. همچنین می‌توان مقدار یال بین رؤس i و j را از همبستگی آماری سیگنال آن‌ها، به صورت:

$$\rho_{ij} = \frac{\mathbb{E}[(Z_i - \mu_i)(Z_j - \mu_j)]}{\sqrt{\mathbb{E}[(Z_i - \mu_i)^2] \mathbb{E}[(Z_j - \mu_j)^2]}}, \quad (9)$$

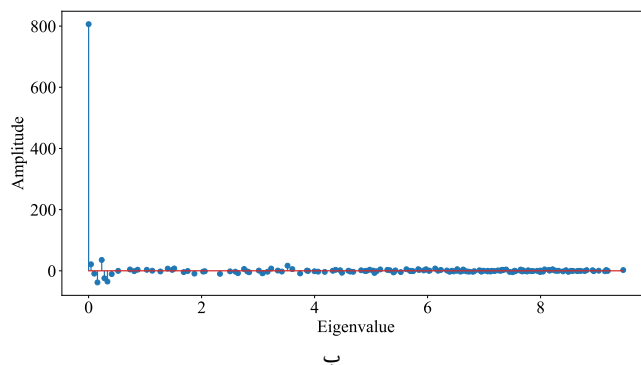
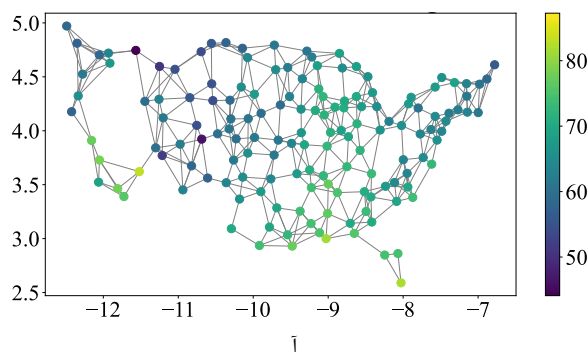
مقداردهی کرد که Z_i بیانگر متغیر تصادفی در رأس i ، $\mathbb{E}$ امید ریاضی و μ میانگین متغیر تصادفی Z_i است. مقدار همبستگی آماری دو رأس i و j برای مقادیر محدود M مشاهده از سیگنال x و تقریب میانگین توزیع با میانگین عملی، با رابطه ضرایب نیم‌پیرسون و به صورت [۱۱]

$$r_{ij|\sigma_M} = \frac{\sum_{p=1}^M (x_i^{(p)} - \hat{\mu}_i)(x_j^{(p)} - \hat{\mu}_j)}{\sqrt{\sum_{p=1}^M (x_i^{(p)} - \hat{\mu}_i)^2 \sum_{p=1}^M (x_j^{(p)} - \hat{\mu}_j)^2}} \quad (10)$$

معادل می‌شود که $x_i^{(p)}$ مقدار رأس i در p مین مشاهده، $\hat{\mu}_i$ میانگین عملی از کل مشاهدات رأس i و σ_M فضای مشاهده از سیگنال x است.

۵. شبیه‌سازی

در این قسمت، شبیه‌سازی برای ساخت گراف و تشخیص و مکان‌یابی ناهنجاری در محیط نرم‌افزار پایتون^{۱۹} اجرا شده است. ابتدا برای ساخت گراف، موقعیت ۱۵۰ ایستگاه هواشناسی در ایالات متحده آمریکا از [۱۵] استخراج شد و با استفاده از رابطه (۸) ماتریس اولیه وزن ساخته شد. برای بهبود گراف، یال‌های بین دو رأس مجاور با اختلاف ارتفاع بیشتر از ۱۵۰۰ متر، صفر در نظر گرفته شد^{۲۰}. همچنین با استفاده از همبستگی اطلاعات پیشین هر حسگر در رابطه (۱۰) در حالت استفاده از داده‌های بدون ناهنجاری، ماتریس وزن دیگری ساخته شد و در انتها حاصل ضرب هردو ماتریس به عنوان ماتریس وزن برای ساخت گراف استفاده شد. مشابه شرط موجود در [۵] مبنی بر حداکثر l نزدیک‌ترین همسایه، شرط کوچکتر بودن درجه هر رأس از عدد ۷ نیز به گراف ساخته شده اضافه شد. شکل ۱ آ. گراف ساخته شده با دمای روز ۲۰۰۳/۰۹/۱۲ و شکل ۱ ب. اندازه طیف سیگنال بر حسب مقادیر ویژه گراف در همان روز را نشان می‌دهد. همان‌طور که در شکل ۱ ب. مشاهده می‌شود، سیگنال‌های واقعی دما، داده‌های تقریباً هموار هستند.



شکل ۱ آ. گراف از موقعیت ایستگاه‌های هواشناسی با دمای روز ۲۰۰۳/۰۹/۱۲ بر روی گراف بر حسب فارنهایت ب. طیف سیگنال در همان روز.

برای آستانه مناسب، از مجموعه‌ای از دماهای سه روز قبل، بزرگترین اندازه از ضرایب تغییرات محلی به ازای هر رأس در نظر گرفته می‌شود. برای روز مورد نظر نیز ضریب تغییرات محلی محاسبه می‌شود. طبق الگوریتم ۱، در صورتی که یکی از ضرایب، بزرگتر از حاصلضرب حدود آستانه و تغییرات محلی روزهای گذشته باشد، آن روز به عنوان ناهنجاری تشخیص داده می‌شود. سپس ضرایب نرمالیزه می‌شوند و بزرگترین ضریب از بین ضرایب LV بزرگتر از حدود آستانه به عنوان رأس ناهنجار انتخاب می‌شود.

اطلاعات دمایی ۳۶۰ روز تمام سال انتخاب شده است. میانگین دمایی در تمام روزها تقریباً ۵۳ درجه فارنهایت و کمینه و بیشینه دما به ترتیب ۱۹/۹- و ۱۰۵/۸ درجه فارنهایت است [۱۶]. برنامه برای این اطلاعات دمایی و برای حالات انتخاب تصادفی با توزیع یکنواخت گسسته، مکان ناهنجاری و روزها، (۱۵۰×۳۶۰ حالت مختلف) ۱۰۰,۰۰۰ مرتبه برای ۴ مقدار مختلف ناهنجاری ۵، ۱۰، ۱۵، ۲۰ اجرا شد. برای مقایسه الگوریتم تشخیص، از معیار منحنی مشخصه عملکرد استفاده می‌شود. این منحنی نسبت بین نرخ مثبت صحیح^{۲۱} و نرخ مثبت کاذب^{۲۲} در سطوح آستانه‌های مختلف است. به عبارت بهتر نرخ مثبت صحیح، میزان تشخیص درست ناهنجاری به عنوان ناهنجاری است، در حالی که نرخ مثبت کاذب، نرخ تشخیص روزهای عادی به عنوان دارای ناهنجاری است [۱۷]. الگوریتم پیشنهادی برای تشخیص با روش [۵] و [۱۰] مقایسه شده است. نتایج به صورت شکل ۲ است. روش پیشنهادی عملکرد بهتری نسبت به روش موجود در [۵] و عملکرد ضعیف‌تری نسبت به [۱۰] دارد. حسن این روش در مقایسه با [۱۰] سادگی محاسبات و عدم نیاز به پایه‌های متعامد و یکه حاصل از تجزیه ماتریس لاپلاسی است که در گراف‌های حجیم، یکی از قسمت‌های اساسی محاسبات محسوب می‌شود. جدول ۱ نتایج عملکرد مکان‌یابی قرارگیری ناهنجاری پس از تشخیص را ارائه می‌دهد. ردیف اول در حالت وجود یک ناهنجاری و ردیف دوم برای حالت وجود دو ناهنجاری با دامنه یکسان است.

جدول ۱: درصد مکان‌یابی صحیح روش پیشنهادی در مقادیر مختلف یک ناهنجاری:

مقدار ناهنجاری	۵	۱۰	۱۵	۲۰
درصد تشخیص یک ناهنجاری	۰/۰۱	۰/۱	۰/۴	۰/۸۵
درصد تشخیص دو ناهنجاری	۰	۰/۰۳	۰/۲۰	۰/۶۸

²¹ True Positive Rate (TPR)

²² False Positive Rate (FPR)

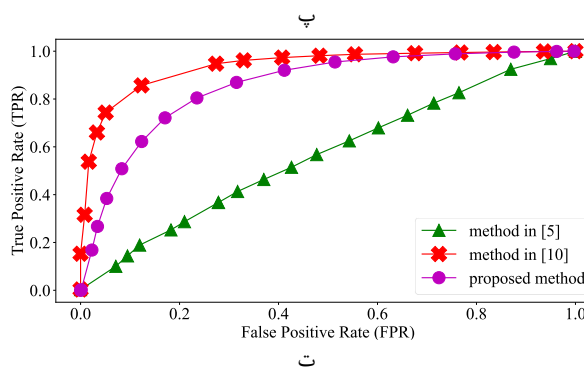
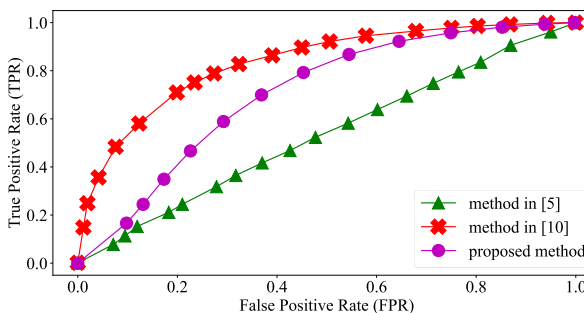
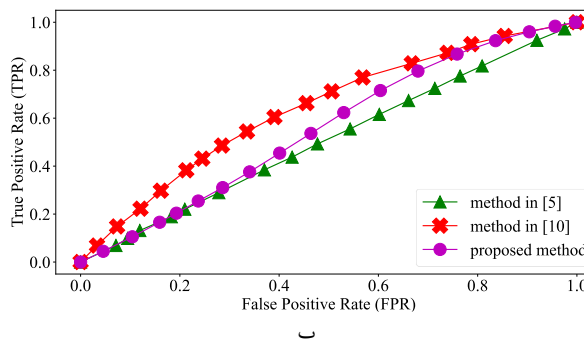
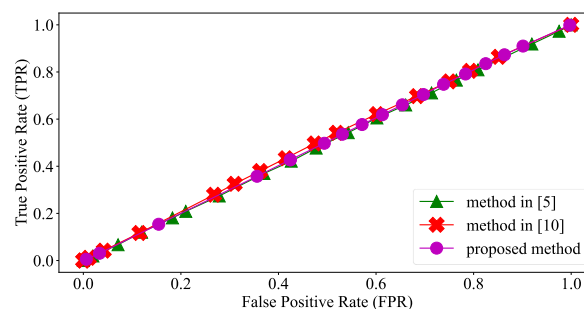
¹⁹ Python

^{۲۰} به ازای هر ۱۰۰۰ متر ارتفاع، دمای هوا ۶ درجه سانتی گراد کاهش می‌یابد.

ساخته شد، سپس با پردازش سیگنال روی گراف، معیار تغییرات محلی تعریف شد که با آن تشخیص و مکان‌یابی ناهنجاری انجام شد. روش پیشنهادی بدون نیاز به تجزیه گراف و با داشتن حجم محاسباتی کمتر نسبت به روش‌های مشابه، دارای عملکرد قابل قبولی در تشخیص و مکان‌یابی ناهنجاری در یک سیگنال نسبتاً هموار دارد.

مراجع

- [1] A. Ortega, P. Frossard, J. Kovačević, J. M. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808–828, May. 2018.
- [2] 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 signal process Mag.*, vol. 30, no. 3, pp. 83–98, May. 2013.
- [3] Y. Zhang, N. Meratnia, and P. Havinga, "Outlier detection techniques for wireless sensor networks: A survey," *IEEE commun. surveys & tuts*, vol. 12, no. 2, pp. 159–170, Feb 2009.
- [4] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM computing surveys (CSUR)*, vol. 41, no. 3, pp. 15, Jul. 2009.
- [5] A. Sandryhaila, J. M. Moura, "Discrete Signal Processing on Graphs: Frequency Analysis," *IEEE Trans. on Signal Process*, vol. 62, no. 12, pp. 3043–3054, Apr. 2014.
- [6] H. E. Egilmez and A. Ortega, "Spectral anomaly detection using graph-based filtering for wireless sensor networks," in *IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, pp. 1085–1089, May. 2014.
- [7] F. Verdoja and M. Grangetto, "Graph laplacian for image anomaly detection," *Machine Vision and Applications*, vol. 31, no. 1, pp. 11, Jan. 2020.
- [8] E. Drayer and T. Routtenberg, "Detection of false data injection attacks in smart grids based on graph signal processing," *IEEE Systems Journal*, Aug. 2019.
- [9] G. Lewenfus, W. A. Martins, S. Chatzinotas, and B. Ottersten, "On the use of vertex-frequency analysis for anomaly detection in graph signals," *Anais do XXXVII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT)*, pp. 1–5, 2019.
- [10] M. Ma, F. Zhou, J. Jiang, and S. Ouyang, "Malfunction Detection and Localization Algorithm for Wireless Sensor Network," *Circuits, Systems, and Signal Process*, pp. 1–9, Jul. 2020.
- [11] K. Gopalakrishnan, M. Z. Li, and H. Balakrishnan, "Identification of outliers in graph signals," in *IEEE 58th Conf. Decision and Control (CDC)*, pp. 4769–4776, Dec. 2019.
- [12] A. Sakiyama, Y. Tanaka, T. Tanaka, and A. Ortega, "Efficient sensor position selection using graph signal sampling theory," in *IEEE Int. Conf. Acoust. Speech and Signal Process. (ICASSP)*, pp. 6225–6229, Mar. 2016.
- [13] P. Oluwasanya, "Anomaly detection in wireless sensor networks," Univ. Edinburgh, Scotland, U.K., Tech. Rep. PGEE11110, 2017.
- [14] L. Stanković, D. Mandić, M. Daković, B. Scalzo, M. Brajović, E. Sejdić, and A. G. Constantinides, "Vertex-frequency graph signal processing: A comprehensive review," *Digital Signal Process*, pp. 102802, Aug. 2020.
- [15] National Climatic Data Center, 2020 [Online]. Available: <https://www1.ncdc.noaa.gov/pub/data/noaa>
- [16] National Climatic Data Center, 2020 [Online]. Available: <https://www1.ncdc.noaa.gov/pub/data/gsod>
- [17] A. Moumena, "Anomalies Detection Based on the ROC Analysis using Classifiers in Tactical Cognitive Radio Systems: A survey," *IAES Int. Journal of Artificial Intelligence (IJ-AI)*, vol. 5, no. 3, pp. 105–116, 2016.



شکل ۲: منحنی مشخصه عملکرد با مقادیر مختلف $a=5$ تا $a=20$

۶. نتیجه‌گیری

در این مقاله به بررسی روش پیشنهادی با استفاده از تغییرات محلی برای تشخیص ناهنجاری در شبکه‌های حسگر پرداخته شد. ابتدا از فواصل اقلیدسی حسگرها و روابط همبستگی داده‌های پیشین، گراف

