سیزدهمین کنفرانس شبکه های توزیع نیروی برق
۱۱و ۱۲ اردیبهشت ۱۳۸۷- گیلان

کاربرد شبکه‌های عصبی خود انجمنی در شناسایی و تصحیح داده‌های نامناسب بار مورد استفاده در پیش‌بینی بار کوتاه‌مدت بازار برق

عبدالحسین وهابی[۱و۲] سعیده برقی‌نیا[۱] بابک نجار اعرابی[۲] کارولوکس[۲]

۱- پژوهشگاه نیرو
۲- قطب علمی کنترل و پردازش هوشمند- دانشکده مهندسی برق و کامپیوتر- دانشگاه تهران

کلمات کلیدی: بازار برق، پیش‌بینی بار کوتاه‌مدت، شناسایی و تصحیح داده‌های نامناسب بار، مانده هنجار شده، شبکه‌های عصبی خودانجمنی، آنالیز طیف تکین

چکیده

مدیریت تولید و توزیع انرژی الکتریکی باید براساس تطبیق عرضه بر تقاضای انرژی برق، اقدام به برنامه‌ریزی، بهره‌برداری و سرمایه‌گذاری بهینه نماید. لذا در برنامه‌ریزی آینده یک سیستم قدرت، پیش‌بینی بار از اهمیت ویژه‌ای برخوردار بوده و باید میزان خطای آن تا حد امکان کاهش یابد. دقت نتایج این پیش‌بینی بر هزینه تولید و همچنین میزان خاموشی در سیستم قدرت تأثیرگذار می‌باشد.

با راه‌اندازی بازار برق در شبکه ایران، شرکت‌های برق منطقه‌ای و در ادامه با توسعه بازار برق، شرکت‌های توزیع که به منزله خریدار محسوب می‌شوند می‌بایست نیاز مصرف ساعت به ساعت کل شبکه تحت پوشش خود را در روزهای آتی پیش‌بینی و ارائه نمایند. ضمن آنکه شرکت‌های مدیریت تولید و یا بهره‌برداری نیروگاه‌ها نیز ممکن است برای پیش‌بینی و پیشنهاد قیمت به پیش‌بینی بار نیاز داشته باشند. بدین‌ترتیب دقت پیش‌بینی ضمن بهبود بهره‌برداری از شبکه تحت پوشش از تخصیص جرایم مربوطه نیز جلوگیری می‌نماید.

پیش‌بینی دقیق بار نیازمند استفاده از داده‌های مناسب است. بدین‌منظور الگوریتمی با استفاده از شبکه‌های عصبی خود انجمنی(AANN) تهیه شده است که داده‌های نامناسب را در مجموعه داده‌های بار تشخیص داده و تصحیح می‌نماید. این الگوریتم در واقع مانند یک فیلتر ورودی قبل از شروع پیش‌بینی بار عمل می‌کند.

در این مقاله پس از مروری بر روش‌های مانده هنجار شده و آنالیز طیف تکین، روش جدیدی مبتنی بر AANN معرفی شده است. عملکرد شناسایی و تصحیح داده‌های بار برق تهران با استفاده از روش پیشنهادی و الگوریتم‌های فوق مقایسه شده و روش پیشنهادی عملکرد بهتری از خود نشان داده است.

۱- مقدمه

پیش‌بینی بار کوتاه مدت در بهره‌برداری سیستم‌های قدرت، نقش اساسی ایفا می‌نماید. برای این پیش‌بینی می‌بایست از نگاشتی استفاده شود که بتواند از ورودی‌های از پیش تعیین‌شده استفاده کرده و بار را در فواصل ساعتی، حداکثر تا چند روز بعد پیش‌بینی نماید.

بار یک سیستم به پارامترهای آب و هوایی و پارامترهای گذشته بار وابستگی شدیدی دارد. داده‌های آب و هوایی نظیر دما توسط سازمان هواشناسی تهیه می‌گردند. داده‌های بار به‌ویژه داده‌های بار مناطق و شرکت‌های توزیع در شبکه سراسری ایران، علاوه بر خطاهای سیستم‌های اندازه‌گیری دارای خطاهای قرائت و دریافت اطلاعات توسط اپراتورها نیز می‌باشند. از سوی دیگر اتفاقات ناخواسته و یا خاموشی‌ها در شبکه نیز منحنی‌های بار را غیرعادی می‌کنند که برای کاربرد در پیش‌بینی بار می‌بایست تصحیح شوند. چون پیش‌بینی بار برای شرایط عادی شبکه انجام می‌گیرد. بدین‌ترتیب با توجه به تاثیر مستقیم و قابل توجه کیفیت داده‌های بار بر خطای پیش‌بینی، اهمیت شناسایی و تصحیح داده‌های نامناسب در مجموعه داده‌های بار آشکار می‌شود.

در ادامه این مقاله، در بخش‌های (۲) و (۳) به‌ترتیب به معرفی روش پیشین شناسایی داده‌های نامناسب مبتنی بر مانده هنجار شده و روش جدیدتر مبتنی بر آنالیز طیف تکین می‌پردازد و بخش (۴) به چگونگی تصحیح داده‌های نامناسب در روش‌های پیشین اختصاص دارد. بخش (۵) به معرفی روش جدید مبتنی بر شبکه‌های عصبی خود انجمنی[1](AANN) می‌پردازد. در بخش (۶) نتایج پیاده‌سازی روش بر روی داده‌های بار برق تهران ارائه گردیده است. بخش (۷) به نتیجه‌گیری و بالاخره بخش (۸) به مراجع اختصاص دارد.

۲- روش شناسایی داده‌های نامناسب مبتنی بر مانده هنجار شده

جهت شناسایی داده‌های نامناسب از میان روش‌های بکار گرفته‌شده در مراجع [۱-۳]، آزمون قدرمطلق مانده هنجار شده[2] - که در مرجع [۲] به تفصیل شرح داده شده - مورد استفاده قرار گرفته است. مطابق این آزمون، محتمل‌ترین اندازه‌گیری نامناسب طبق رابطه زیر بیان می‌شود:

محتمل‌ترین اندازه‌گیری نامناسب $= Max_i \left| \frac{Z_{M_i} - \overline{Z_i}}{\sigma_i} \right|$

که در این رابطه :

Z_{M_i} : i کمیت اندازه‌گیری شماره

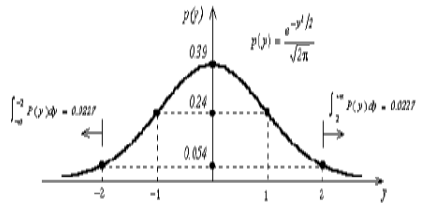
$\overline{Z_i}$: i میانگین اندازه‌گیری

σ_i : i انحراف معیار اندازه‌گیری

برای تشخیص داده‌های نامناسب، کافی است اختلاف مقادیر داده‌ها از میانگین آنها را بر انحراف معیارشان تقسیم کنیم. سپس با مقایسه قدرمطلق مقادیر حاصل، محتمل‌ترین اندازه‌گیری (داده) نامناسب مشخص می‌شود. با توجه به آنچه گفته شد، مقدار متناظر با هر اندازه‌گیری به‌صورت رابطه زیر می‌باشد:

$$Y_i = \frac{Z_{M_i} - \overline{Z_i}}{\sigma_i} \qquad (۱)$$

که این در واقع، هنجارشده اختلاف بین هر اندازه‌گیری و میانگین آن است که به آن مانده هنجارشده داده هم می‌گویند. حال با تعیین یک آستانه[3]، تنها اندازه‌گیری‌هایی حذف خواهند شد که مقدار Y_i مربوط به آنها بیشتر از آستانه باشد. مقادیر هنجارشده مانده‌ها، متغیرهای تصادفی با توزیع نرمال می‌باشند که دارای متوسط صفر و واریانس یک هستند. تابع چگالی احتمال این متغیرهای تصادفی در شکل (۱) رسم شده است:



شکل(۱): تابع چگالی احتمال برای متغیرهای تصادفی هنجارشده

مجموع خطاهای ممکن، همان مساحت زیر منحنی تابع چگالی احتمال و برابر یک می‌باشد:

$$\int_{-\infty}^{+\infty} P(y).dy = 1.0 \qquad (۲)$$

با توجه به شکل (۱) مشخص می‌شود که با داشتن مانده اندازه‌گیری و انحراف معیارش، می‌توان احتمال درست بودن آن را محاسبه نمود. بطور مثال، اگر مقدار NR اندازه‌گیری i، بزرگتر از ۲ باشد (Y_i>2)، کمتر از ۰/۰۴۵۴ احتمال درست بودن و

[1] Auto Associative Neural Network (AANN)
[2] Normalized Residual (NR)
[3] Threshold

بیش از ۹۵۴۶ /۰ احتمال نادرست بودن آن وجود دارد. این مقادیر به‌وسیله انتگرال‌گیری از منحنی (۱) دربازه $-2<Y_i<2$ بدست آمده‌اند. بنابراین، با انتخاب مقدار Y_i به‌عنوان آستانه داده قابل قبول، یک اندازه‌گیری در صورتی فاقد ارزش یا نادرست خواهد بود که مقدار NR آن Y_i، بزرگتر از این آستانه شود. جدول (۱) نشان دهنده احتمال خوب یا بد بودن یک داده با توجه به تابع چگالی احتمال گوسی با واریانس یک می‌باشد.

جدول (۱): احتمال خوب یا بد بودن یک داده براساس تابع چگالی احتمال با واریانس ۱

احتمال بد بودن اندازه‌گیری	احتمال خوب بودن اندازه‌گیری	$\lvert\frac{Z_{M_i}-\overline{Z_i}}{\sigma_i}\rvert=Y_i$
۰/۰۷۹۶	۰/۹۲۰۴	۰/۱
۰/۱۹۷۴	۰/۸۰۲۶	۰/۲۵
۰/۳۸۳۰	۰/۶۱۷۰	۰/۵
۰/۶۸۲۶	۰/۳۱۷۴	۱/۰
۰/۹۵۴۶	۰/۰۴۵۴	۲/۰
۰/۹۹	۰/۰۱	۲/۵۷
۰/۹۹۷۴	۰/۰۰۲۶	۳/۰

لازم است که برای آستانه خطای قابل قبول اندازه‌گیری، مقدار مناسبی انتخاب شود. به اندازه‌گیری که مانده هنجارشده آن بیشتر از این آستانه باشد، اندازه‌گیری نامناسب گفته می‌شود. از آنجایی که امکان دارد داده‌های خوب نیز به‌عنوان داده‌های بد جدا گردند، لذا این داده‌های نامناسب انتخاب شده با درصدی از میانگین داده‌های مناسب در هر دسته مقایسه می‌شوند که در صورت بزرگ بودن از این مقدار به‌عنوان داده‌های نامناسب نهایی شناسایی می‌گردند (مقدار درصدی از میانگین داده‌های مناسب براساس تجربه و نظر فرد خبره انتخاب می‌شود).

در کاربرد روش NR، تغییرات روند منحنی بار در طولانی‌مدت با اعمال ضریب افزایش سالانه لحاظ می‌شود. این امر باعث می‌گردد که تغییرات محلی، کمتر تاثیرگذار باشند و به‌عنوان خطا اعلام شوند. در ضمن پس از تشخیص داده‌های نامناسب، نیاز به نظر فرد خبره برای حذف یا تایید برخی از داده‌هایی که نامناسب تشخیص داده شده‌اند، دارد. از جمله دیگر مشکلات این روش می‌توان به تشخیص تعداد زیادی از داده‌ها در پیک بار به‌عنوان خطا اشاره کرد. لذا نیاز به روشی که از روند تغییرات بار استفاده کند، احساس می‌شد که منجر به ارائه روش مبتنی بر آنالیز طیف تکین[1] (SSA) گردید.

۳- روش مبتنی بر آنالیز طیف تکین

برای درنظر گرفتن روند تغییرات منحنی بار در شناسایی داده‌های نامناسب، روش دیگری بر اساس آنالیز طیف تکین در مرجع [۴] ارائه شده است. آنالیز طیف تکین ابزار قدرتمندی است که برای استخراج اطلاعات از سری‌های زمانی استفاده می‌کند [۵و۶]. این روش براساس بسط کارهونن-لوو است که روی تقریب ماتریس کوواریانس بدست آمده از تاخیرهای متوالی Mتایی سری زمانی عمل می‌کند و اجزای اصلی سری را از روی بردارهای ویژه با مقادیر ویژه بزرگتر ماتریس کوواریانس، بازسازی می‌کند.

در روش پیشنهادی برای شناسایی داده‌های نامناسب بار، از دو اجرای SSA روی سری داده‌های بار و سری خطای بازسازی سری داده‌های بار و نیز یک فیلتر گوسی مشابه روش NR استفاده می‌نماید.

ابتدا یک SSA به سری داده‌های بار یک سال اعمال می‌شود و سری بازسازی شده با استفاده از اجزای اصلی که مقادیر ویژه بزرگتری دارند و حاوی اطلاعات روند غالب هستند، تشکیل می‌گردد. خطای سری بازسازی‌شده نسبت به سری اصلی، سری خطا را بدست می‌دهد. تغییرات سریع در سری داده‌های بار، خود را در اجزای اصلی با مقادیر ویژه کوچک‌تر نشان می‌دهند که ممکن است در بازسازی حذف گردند. به‌عنوان مثال تغییرات محدوده پیک بار از این نوع هستند. در سری خطا، دامنه در این نقاط به‌صورت شبه پریودیک زیاد می‌شود. برای اینکه بتوان داده‌های نامناسب در محدوده‌های تغییرات سریع را نیز به درستی شناسایی کرد، این بار SSA به سری خطا اعمال می‌شود و سری بازسازی‌شده جدید با استفاده از اجزای اصلی با مقادیر بزرگتر بدست می‌آید. خطای این بازسازی، سری زمانی نهایی خواهد بود که برای شناسایی داده‌های نامناسب بار مورد استفاده قرار می‌گیرد.

برای شناسایی داده‌های نامناسب - مشابه روش NR- میانگین و واریانس داده‌های سری زمانی نهایی محاسبه شده و در فیلتر گوسی استفاده می‌گردد. به عبارت دیگر برای هر داده در سری، رابطه (۱) محاسبه می‌شود و اگر قدرمطلق آن از آستانه‌ای بزرگتر باشد، داده نامناسب محسوب می‌گردد. این آستانه با توجه به اعتمادی که به داده‌های اصلی وجود دارد و نظر فرد خبره تعیین می‌گردد.

برای روزهای خاص و ساعات پیک باید آستانه بزرگتری نسبت به روزها و ساعات عادی انتخاب گردد تا بتواند با افزایش

[1] Singular Spectrum Analysis (SSA)

خطایی که به‌طور طبیعی برای روزهای تعطیل و ساعات پیک وجود دارد تطبیق یابد و پس از بررسی نحوه تغییر علامت سری خطا در نقاط مذکور و مقایسه با تعدادی آستانه که در مرجع [۴] توضیح داده شده است، داده‌های نامناسب بار تعیین می‌شوند.

۴- تصحیح داده‌های نامناسب در روش‌های پیشین

پس از شناسایی داده‌های نامناسب می‌بایست آنها را با مقادیر مناسب جایگزین نمود. در روش NR، داده‌های نامناسب ساعات پیک و ۱ و۲۴ و نیز داده‌های نامناسب متوالی با میانگین‌های متناظر جایگزین شده و برای داده‌های نامناسب سایر ساعات از برازش اسپیلاین مکعبی[1] داده‌های همان روز استفاده می‌شود. از مشکلات این روش می‌توان به ایجاد دندانه‌های ناخواسته در منحنی بار اشاره کرد که آنها نیز باید تصحیح گردند.

یکی از روش‌های مرسوم برای تصحیح داده‌های خطا، درنظر گرفتن پنجره‌ای در حوالی آنها و جستجو در بین داده‌های درست موجود برای یافتن داده‌ای با پنجره مشابه و سپس جایگزینی است [۷]. این روش را می‌توان برای داده‌های بار نیز بکار برد، اما جایگزینی مشابه‌ترین داده به دلیل تغییرات روند بار و اختلاف دامنه‌ها ممکن است باعث ایجاد دندانه‌های ناخواسته در منحنی بار شود. لذا روش یاد شده برای جایگزینی داده‌های نامناسب نیاز به تغییراتی دارد.

بدین‌منظور ابتدا پنجره‌ای به طول مناسب در اطراف داده نامناسب بار در نظرگرفته شده و داده‌های موجود بر اساس شباهت از نظر نرم اقلیدسی مرتب می‌گردند. سپس تعداد مناسبی از داده‌ها که نزدیکترین به داده نامناسب هستند در نظر گرفته می‌شوند. برای این داده‌ها، جایگزینی داده نامناسب به‌صورت ترکیب خطی بقیه داده‌های همان پنجره فرض می‌شود. ضرایب این ترکیب خطی با تقریب کمترین مربعات[2] از پنجره‌های متناظر بدست می‌آید. با استفاده از این ضرایب و داده‌های پنجره مربوط به داده نامناسب، مقدار موردنظر برای جایگزینی بدست می‌آید. ترکیب خطی جایگزینی عبارت است از:

$$L(i)= a_0 + a_{-k} L(i-k) +\ldots+ a_{-1} L(i-1) + a_1 L(i+1) + \ldots+ a_k L (i+k) \quad (۳)$$

نکته قابل توجه این است که اگر داده نامناسب دیگری نیز در پنجره یک داده نامناسب باشد، داده‌های متناظر داده نامناسب در تمام پنجره‌های مشابه حذف می‌شوند تا در تصحیح داده نامناسب اصلی تاثیر نداشته باشند.

۵- روش جدید مبتنی بر شبکه‌های عصبی خودانجمنی

برای اینکه از الگوهای ساعتی موجود در سری زمانی بار بتوان برای تشخیص داده‌های نامناسب استفاده کرد، در این مقاله شبکه‌های عصبی خودانجمنی که ابزار قدرتمندی برای تشخیص خطاها و تغییرات غیرطبیعی در کاربردهای مختلف نظیر رخداد خطا در سنسورها می‌باشد، پیشنهاد می‌گردند. در این بخش، ابتدا به معرفی این روش پرداخته و سپس چگونگی استفاده از آن برای تشخیص داده‌های نامناسب بار شرح داده می‌شود.

۵-۱- شبکه‌های عصبی خودانجمنی (AANN)

شبکه‌های AANN شبکه‌هایی هستند که در آنها خروجی‌ها به‌گونه‌ای آموزش داده می‌شوند که ورودی‌ها را در محدوده دینامیکی خاصی بازسازی کنند [۸] و در مرحله آموزش روابط داخلی بین متغیرهای ورودی در ضرایب شبکه عصبی مدل شده و در واقع، خروجی برای هر ورودی، خود ورودی درنظر گرفته می‌شود.

شبکه‌های AANN همانگونه که در شکل (۲) نمایش داده شده است به فرم شبکه‌های عصبی پیشرو هستند و شامل لایه‌های ورودی، خروجی و تعدادی لایه پنهان هستند. از لحاظ نظری تعداد ۳ لایه پنهان برای این شبکه‌ها کفایت کرده و تعداد بیشتر لایه‌ها به بهبود عملکرد کمک می‌کند و باعث می‌شود تا روابط درونی بین متغیرهای ورودی دقیق‌تر مدل شود.

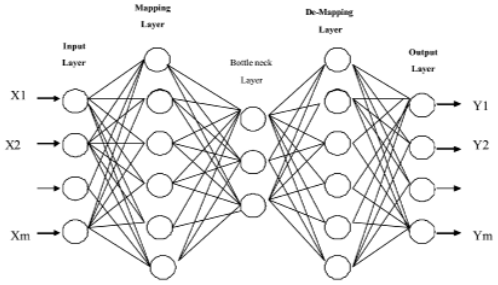
لایه پنهان اول، لایه نگارنده[1] خوانده می‌شود که تابع تبدیل این لایه معمولا سیگموئید و یا توابع مشابه آن می‌باشد. لایه بعدی گلوگاه[2] نامیده می‌شود و تابع آن می‌تواند خطی و یا غیرخطی باشد. ابعاد لایه دوم باید کمتر از تعداد ورودی‌ها باشد. لایه سوم، لایه باز نگارنده[3] نامیده می‌شود که توابع آنها نیز غیرخطی می‌باشند.

شبکه‌های AANN، نگاشت همانی از ورودی به خود را بدست می‌دهند. این نگاشت متغیرهای ورودی را بر روی خودشان می‌نگارد، بدون آنکه از نگاشت ساده یک به یک همانی استفاده کند و در این حالت هر خروجی تابعی از تمام ورودی‌ها

[1] Spline
[2] Least Square

[1] Mapping Layer
[2] Bottleneck
[3] De-Mapping Layer

می‌باشد. روش‌های گوناگونی برای این نگاشت وجود دارد ولی AANN از یک الزام درونی برای جلوگیری از همانی یک به یک شدن نگاشت استفاده می‌کند. لایه دوم این نقش را ایفا می‌کند. این لایه، AANN را در حین آموزش وادار می‌کند تا ورودی‌ها را فشرده کرده و سپس در لایه بعدی بازیابی نماید. آموزش شبکه عصبی در اینجا نیز همانند شبکه‌های پیشرو دیگر، بر اساس الگوریتم پس‌انتشار خطا[4] است و در حین یادگیری ضرایب را به نحوی تنظیم می‌کند تا خطای بازسازی ورودی از دید کمترین مربعات، حداقل شود.



شکل(۲): شمای کلی شبکهAANN [۹]

۵-۲- شبکه‌های عصبی خودانجمنی بهبود یافته (E-AANN)

برای تشخیص و تصحیح ورودی معیوب یا داده نامناسبی که به شبکه AANN اعمال شده است، روشی تحت عنوان شبکه‌های عصبی خودانجمنی بهبود یافته[5] در مرجع [۹] پیشنهاد شده است که به روش زیر عمل می‌کند:

برای تک تک ورودی‌ها بازه بین ماکزیمم و مینیمم آن را به قسمت‌های مساوی با دقت دلخواه تقسیم کرده و سپس همان ورودی را به شبکه اعمال می‌کند و تمامی بازه مربوطه را با داده بدست آمده جایگزین می‌کند و هرجا که مربعات خطا مینیمم شود به‌عنوان تخمین داده واقعی داده می‌شود. الگوریتم این روش در زیر آمده است:

اگر ورودی و خروجی شبکه عصبی به‌صورت زیر درنظر گرفته شوند، داریم:

$$X=[X_1,X_2,\dots X_m]^T \qquad (۴)$$

$$Y=[Y_1,Y_2,\dots Y_m]^T \qquad (۵)$$

که در آن X_i ورودی i ام شبکه عصبی و Y_i خروجی متناظر آن می‌باشد و m برابر با تعداد ابعاد ورودی شبکه عصبی است. هدف در این مرحله پیدا کردن X^* است که تابع هزینه J را کمینه کند:

$$J=(X-Y)^T(X-Y)=\sum_{i=1}^{m}(X_i-Y_i)^2 \qquad (۶)$$

و در واقع X^* برابر است با:

$$X^*=\arg\min(J) \qquad (۷)$$

در این الگوریتم به طریقه زیر عمل می‌شود:

برای $(X_1,X_2,X_3,\dots,X_m)$ داریم:

$$J_{\min}=\infty \qquad (۸)$$

For i=1,2,…m

{

$$\Delta X_i=\frac{\max(X_i)-\min(X_i)}{N} \qquad (۹)$$

N بسته به دقت و رزولوشنی که مدنظر طراح است، تعیین می‌شود.

$$J_{i,\min}=\infty \qquad (۱۰)$$

For k=1,2,3,…N

{

$$\hat{X}_k=[X_1,X_2,\dots X_i^{'},\dots X_m] \qquad (۱۱)$$

که $X_i^{'}$ برابر است با:

$$(X_i^{'}=\min(X_i)+\Delta X_i*k) \qquad (۱۲)$$

$$\hat{Y}_k=Net(\hat{X}_k) \qquad (۱۳)$$

$$J_k=(\hat{X}_k-\hat{Y}_k)^T(\hat{X}_k-\hat{Y}_k) \qquad (۱۴)$$

$$\text{If} \quad J_k<J_{i,\min} \text{ then} \quad J_k=J_{i,\min} \qquad (۱۵)$$

}

$$\text{If} \quad J_{i,\min}<J_{\min} \text{ then} \quad J_{i,\min}=J_{\min} \qquad (۱۶)$$

$$\text{and} \quad X^*=\arg\min(J_{\min}) \qquad (۱۷)$$

}

در کاربردهایی نظیر تشخیص خطای سنسورها و امثال آن از X^*-X برای تشخیص خطا استفاده می‌شود.

۵-۳- استفاده از E-AANN برای شناسایی و تصحیح داده‌های نامناسب بار

برای اینکه بتوان از روش فوق در شناسایی و تصحیح داده‌های نامناسب در سری‌های زمانی به‌ویژه سری زمانی بار استفاده کرد، سعی بر آن است تا الگوهای بار به شبکه عصبی AANN آموزش داده شوند و با استفاده از آن داده‌های

---

[4] Back Propagation (BP)

[5] Enhanced Auto-Associative Neural Network (E-AANN)

نامناسب تعیین گردند. الگوهای بار به‌صورت الگوهای Mساعتی به شبکه اعمال می‌گردند که M توسط فرد خبره و به‌صورت تجربی درنظر گرفته می‌شود.

برای ساختن داده‌های آموزش برای شبکه AANN، گام اول، تشکیل بردارهای Mبعدی متوالی از سری و قرارگیری در دنباله $\{\widetilde{X}(t)\}$ است.

$$\begin{aligned}&\{X(t): t=1,\ldots,N\}:\\&\widetilde{X}(t)=(X(t),X(t+1),\ldots,X(t+M-1)),\\&t=1,\ldots,N',\quad N'=N-M+1\end{aligned} \qquad (۱۸)$$

بردارهای بالا به‌عنوان ورودی شبکه عصبی درنظر گرفته می‌شوند. پس ازآموزش شبکه عصبی، الگوریتم ارائه شده در بخش (۵-۲) به داده‌های فوق اعمال می‌گردد. به علت اینکه هر یک از داده‌های سری زمانی در مکان‌های ۱ تا M از پنجره گرفته شده می‌توانند قرار بگیرند، لذا هر یک از داده‌ها در تعداد Mتا از داده‌های اعمال شده به شبکه عصبی حضور دارند. برای اینکه سری بازسازی گردد، از داده مذکور در هر یک از M پنجره میانگین‌گیری شده و میانگین در سری بازسازی شده قرار می‌گیرد. برای شناسایی داده‌های نامناسب، تفاضل سری بازسازی شده با سری اصلی، همانند روش NR از یک فیلتر گوسی عبور داده می‌شود و بدین‌ترتیب داده‌هایی که از آستانه بزرگتر هستند، با سری بازسازی شده جایگزین می‌گردند. بدین‌ترتیب سری تصحیح شده بدست می‌آید. ولی به دلیل اینکه داده‌های با خطای بسیار بزرگ، یادگیری شبکه عصبی و در نتیجه دقت تصحیح را تحت تاثیر قرار می‌دهند، یک بار روش با آستانه بزرگتری برای فیلتر گوسی اعمال شده و برای داده‌های تصحیح شده، مجددا شبکه عصبی آموزش داده می‌شود و سپس روش مذکور بدان اعمال می‌گردد و این بار تفاضل سری بازسازی شده با سری اولیه داده شده از فیلتر گوسی با آستانه کوچکتر عبور داده شده و داده‌ها تصحیح می‌گردند.

پس از تصحیح داده‌ها، الگوهایی که به‌صورت تکرار شونده به‌عنوان داده نامناسب تشخیص داده شده‌اند، از میان داده‌های نامناسب تشخیص داده شده حذف می‌گردند. بدین‌منظور ابتدا تفاضل سری بازسازی شده با سری اصلی به‌عنوان سری خطا تشکیل می‌شود:

$$err(i)=X(i)-X_{coorected}(i) \qquad \text{i=1,2,…,L} \qquad (۱۹)$$

که L برابر با طول سری است.

اگر برای داده نامناسبی، ازمیان داده‌های ۲۴ یا ۴۸ ساعت قبل یا بعد ازآن داده نامناسب دیگری موجود باشد و قدرمطلق تفاضل دامنه سری خطا در این نقاط، بین دو آستانه‌ای که توسط فرد خبره تعیین می‌شوند، قرار گیرد هر دو داده مذکور از مجموعه داده‌های نامناسب تشخیص داده شده حذف می‌گردند.

$$Thre1<\left|err(i)-err(i+k)\right|<Thre2 \qquad (۲۰)$$
$$k=\pm 24,\pm 48$$

٦- پیاده‌سازی روش پیشنهادی روی داده‌های بار شبکه برق تهران

روش پیشنهادی بر روی داده‌های بار شبکه برق تهران در سال‌های ۱۳۸۱ تا ۱۳۸۳ اعمال گردید. ابتدا سری بین صفر و یک نرمالیزه شده و اندازه پنجره (M) برابر با ۷ و N مورد استفاده در E-AANN برابر با ۱۰۰ قرار داده شده است. تعداد نرون‌های شبکه عصبی اول برابر با ۷-۱۶-۳-۱۶-۷ و برای شبکه عصبی دوم برابر با ۷-۱۱-۳-۱۱-۷ می‌باشد. آستانه برای اجرای اول روش، برابر با ۴ و برای اجرای دوم برابر با ۲ و آستانه‌های مورداستفاده برای حذف داده‌های نامناسب متناوب، برابر با ۰/۰۵ و ۰/۴ قرار داده شده است. برای این داده‌ها، در شکل‌های (۳) و (۴) عملکرد روش E-AANN در مقایسه با روش NR و SSA برای برخی روزها آمده است

روش جدید در مجموع، عملکرد بهتری دارد و اعمال آن به صورت on-line سرراست‌تر و دارای قابلیت اطمینان بیشتری نسبت به دو روش دیگر است وتصحیح‌های بکار رفته نیز با الگوهای بار سازگاری بیشتری دارد. داده‌های بار تصحیح شده با استفاده ازنرم‌افزار پیش‌بینی بار پژوهشگاه نیرو که با استفاده از شبکه عصبی و سیستم خبره فازی، میزان بار را پیش‌بینی می‌کند، برای سال ۱۳۸۳ و با استفاده از سال‌های ۱۳۸۱ و ۱۳۸۲ به‌عنوان داده‌های آموزش پیش‌بینی گردید که نتایج در جدول (۲) نشان داده شده است.

۷- نتیجه‌گیری

اشکالات ثبت و اتفاقات ناخواسته در سیستم قدرت بر کیفیت داده‌های بار و در نتیجه بر کیفیت پیش‌بینی بار تاثیر می‌گذارند. در این مقاله، مروری بر روش‌های مانده هنجار شده و آنالیز طیف تکین برای شناسایی و تصحیح داده‌های نامناسب بار مورد استفاده در پیش‌بینی بار انجام گرفت و سپس برای این منظور، روش جدیدی مبتنی بر شبکه‌های عصبی خود انجمنی

(AANN) ارائـه گـردید. نـتایج پـیاده‌سـازی روش جدیـد برای داده‌های بار برق تهران در سال‌های ۱۳۸۱ تا ۱۳۸۳ در مقایسه با روش مانـده هنجارشده و آنالیز طیف تکین ارایه شده است. روش پیشنهادی به‌خوبی بیشتر خطاها را پیدا کرده و تصحیحی سازگار با الگوهای بار موجود ارائه کرده و دقت پیش‌بینی را بالاتر می‌برد و در مجمـوع روش SSA بـه دلـیل سادگی و دقت قابل قبول و روش AANN بـه دلـیل دقـت بهتـر و حفـظ عملکرد در حالت on-line، روش‌هـای خوبـی بـرای شناسایی و تصحیح داده‌های نامناسب بار می‌باشند.

۸- مراجع

[۱] سـعیده برقـی‌نـیا، پـویا انصاری مهر، ناصر وفادار، "پیشنهاد روشی جهـت شناسـایی و تـصحیح داده‌های نامناسب بار شبکه سراسری ایران مورد استفاده در بر آورد بار کوتاه‌مدت"، پانزدهمین کنفرانس بین‌المللی برق، آبان ۱۳۷۹.

[۲] وحـید وحـیدی‌نـسب، سـعیده برقی‌نیا، پویا انصاری مهر، " کاربرد بـرازش منحنی در تلفیق با روش مانده هنجار شده برای شناسایی و تـصحیح داده‌هـای نامناسب بار"، بیستمین کنفرانس بین المللی مهندسی برق، آبان ۱۳۸۴.

[3] P. Ansarimehr, S. Barghinia, Z. Miesepasi, H. Habibi, " Identification and Modification of Improper Load Data Used in Short Term Load Forecasting", IEEE 2005 PowerTech Conference.

[۴] عبدالحـسین وهابـی، بابـک نجـار اعرابـی، کـارو لـوکس، پـویا انصاری‌مهر، سعیده برقی‌نیا"، کاربرد آنالیز طیف تکین در شناسایی وتـصحیح داده‌هـای نامناسـب بـار مـورد استفاده در پیش‌بینی بار کـوتاه‌مـدت"، بیست و یکمین کنفرانس بین المللی مهندسی برق، آبان ۱۳۸۵.

[5] A. Gholipour, C. Lucas, M. Shafiee, B.N. Araabi, "Extracting the Main Patterns of Natural Time Series for Long Term Prediction", Journal of Aymospheric and Solar Terrestrial Physics, Elsevier, 2005.

[6] R. Vautard, P. Yiou, M. Ghil, "Singular Spectrum Analysis: A Toolkit for Short Noisy Chaotic Signals", Physica D, 58: 95-126, 1992.

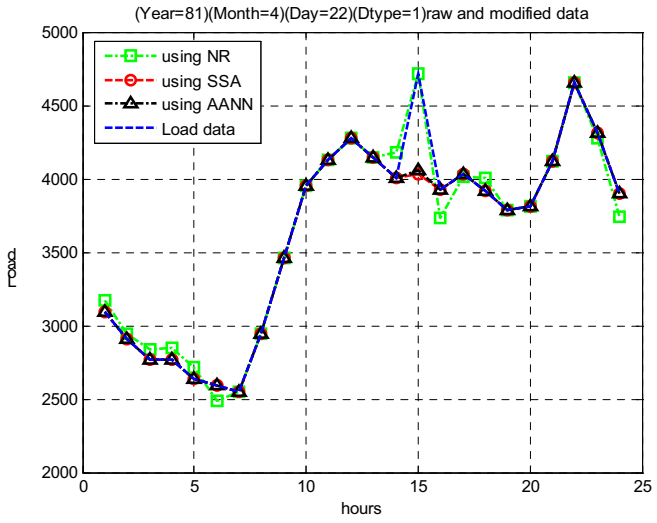
[7] M. T. Harandi and B. N. Araabi, "Impulse Noise Removal Based on Long Range Correlation in an Image", ICEICS 2003, IEEE

[8] M.A. Kramer, "Autoassociative Neural Networks", Computers in Chemical Engineering, vol. 16, no 4, pp. 313-328, 1992.

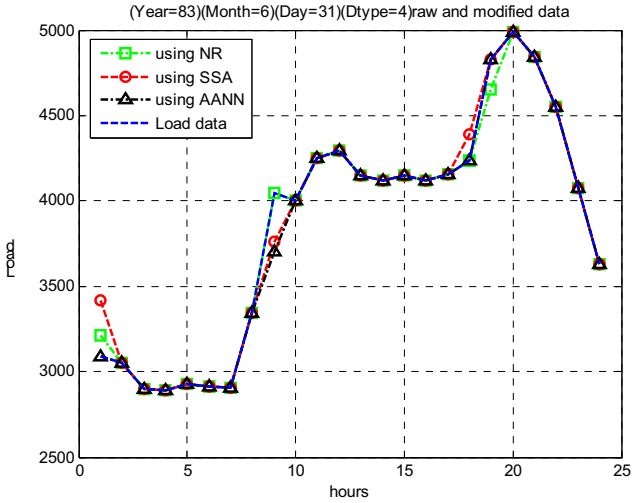
[9] Massieh Najafi, "Use of AutoAssociative Neural Networks for Sensor Diagnostics", MSc thesis, Texas A&M University, December2003

جدول(۲): :نتایج پیش‌بینی بار برق تهران در سال ۸۳ با استفاده از روشهای مختلف تصحیح داده‌ها

	Saturdays	Work days	Thursdays	Fridays	Special days	Total mean
Without Correction	4.56	3.24	4.04	4.61	2.70	3.63
NR	4.51	3.20	4.00	4.53	2.69	3.60
SSA	4.51	3.11	3.99	4.37	2.68	3.51
AANN	4.50	2.80	4.03	4.39	2.49	3.35



شکل(۳): مقایسه منحنی‌های تصحیح شده بار در سال ۸۱ با سه روش مذکور



شکل(۴): مقایسه منحنی‌های تصحیح شده بار در سال ۸۳ با سه روش مذکور