خوشه‌بندی سری‌های زمانی برای سهم‌بندی الگوهای مختلف مصرف از بار فیدر

رضا اسلامی[1]

[1]استادیار دانشکده مهندسی برق دانشگاه صنعتی سهند تبریز، eslami@sut.ac.ir

چکیده – داده‌ها و الگوهای قابل استخراج از بین آنها، از مهمترین شاخص‌های دنیای اطلاعات هستند و داده‌کاوی یکی ازپرکاربردترین روش‌هایی است که می‌تواند الگوهای مفید در داده‌ها را با حداقل دخالت کاربران شناسایی کند. خوشه‌بندی به عنوان یکی از روش‌های داده‌کاوی، مناسب ترین روش برای کار با داده‌ها و استخراج الگوهای موجود در آنها می‌باشد. امروزه با اندازه‌گیری هوشمند، امکان پردازش غیر متمرکز داده ها و نیز مخابرات وسیع برای تحویل اطلاعات به صورت آنی مهیا شده تا به وسیله تحلیل اطلاعات، انرژی مورد نیاز شبکه توزیع پیش‌بینی شود.

در پژوهش حاضر، ابتدا کاربرد و روش‌های مختلف خوشه بندی در سری‌های زمانی مورد ارزیابی قرار خواهد گرفت و همچنین نحوه‌ی جمع‌آوری داده‌های مصرف کنندگان مورد بررسی قرار می‌گیرد تا با روش‌های مختلف و متنوعی که وجود دارد الگوی بار مشترکین مختلف با توجه به سوابق آنها در سیستم، استخراج گردد. یک روش بهینه‌ی خوشه‌بندی مطرح می‌شود و سپس با داشتن الگوهای بارهای هر خوشه، سهم آنها از بار فیدر بدست می‌آید.

کلید واژه- الگوهای بار، خوشه‌بندی، داده‌کاوی، سری‌های زمانی و فیدر.

۱- مقدمه

معرفی و توسعه بازارهای رقابتی برق به سرعت در تعامل بین تامین‌کننده و مصرف‌کننده تغییریافته است [۲،۱]. در حال حاضر، تفکیک هر کدام از مولفه‌های بار شبکه‌های توزیع، یکی از نیازهای مبرم صنعت برق بوده و نتایج برگرفته از آنها می‌تواند در زمینه‌های مختلف اعم از تأمین الگوی بار الکتریکی، پیش بینی بار در زمان اوج مصرف، تعیین ظرفیت مانور شبکه-های توزیع و از همه مهم‌تردرجهت تدوین برنامه‌های مدیریت مصرف در بخش های گوناگون به کار گرفته شوند.

داده‌کاوی یکی از مهمترین این روش‌ها است که به وسیله آن الگوهای مفید در داده‌ها با حداقل دخالت کاربران شناخته می‌شوند و اطلاعاتی را در اختیار کاربران و تحلیلگران قرار می-دهد تا براساس آنها تصمیمات مهم و حیاتی را اتخاذ کنند. داده-کاوی به معنای استخراج دانش از حجم عظیم داده‌هاست و به عنوان مهمترین مرحله در فرایند کشف دانش معرفی شده است؛ به تعبیر دیگر داده‌کاوی به عنوان فرایند کشف و استخراج الگوهای پنهان از حجم بالایی از داده‌ها تعریف می‌شود[۴-۲]. امروزه داده‌کاوی را یک گام اساسی در فرآیند کشف دانش در پایگاه های داده می‌دانند [۵-۷]. استفاده روزافزون از داده‌های زمانی، به ویژه داده‌های سری زمانی، تلاش‌های مختلف تحقیق و توسعه در زمینه داده‌کاوی را آغاز کرده است.

خوشه‌بندی که یکی از روش‌های داده‌کاوی می باشد، از لحاظ فنی ،یک یادگیری بدون ناظر است که شامل یافتن خوشه‌هایی است که حداکثر واریانس را با یکدیگر دارند، در حالی که گروه‌های زمان‌بندی شده در هر خوشه دارای حداقل واریانس با یکدیگر هستند[۸]. معمولا تامین کنندگان بزرگ برق از خوشه بندی برای دسته بندی کردن مشتری‌هایی که دارای ویژگی های الکتریکی مشابه هستند استفاده می‌کند [۹].

این مقاله با مقایسه دو روش خوشه بندی k-means و SOM با استفاده از شاخص DB که نسبت فاصله بین خوشه‌ای (یعنی میانگین فاصله همه الگوهای در خوشه به مرکز خوشه) تقسیم بر فاصله بین خوشه‌ای (یعنی فاصله بین مراکز خوشه‌ای مختلف) بیان می‌کند به ارزیابی این روش ها برای دسته بندی مشترکین برق یک فیدر توزیع می‌پردازد و سهم خوشه‌ها مختلف مصرف را از بار فیدر به دست می‌آورد.

۲- خوشه بندیK-means

K-means از یک فرایند تکرار استفاده می‌کند که مشتریان یا بارهای مصرفی را به گروه‌هایی براساس فاصله‌بین خود و مرکز خوشه اختصاص می‌دهد. در ابتدا، مراکز خوشه‌ای به صورت تصادفی در مجموعه داده‌های نمونه انتخاب می‌شوند. سپس فاصله بین مشتری نمونه و مراکز خوشه محاسبه می‌شود تا مشتریان یا بار مصرفی به هر خوشه با کمترین فاصله اختصاص داده شود. سپس مرکز خوشه براساس افزودن یک مشتری یا بار جدید دوباره محاسبه می‌شود.

معادله (۱) الگوریتمk-means را که مجموعه‌ای از مشاهدات $x1, x2, ..., xn$ را ارائه می‌دهد، که در آن هر یک از مشاهدات یک بردار dimensional با تعداد ابعاد d است. خوشه بندیk-means با هدف تقسیم‌بندی n مشاهدات به $k \le n$ مجموعه $S = \{S1, S2, ..., Sk\}$ است به طوری که مجموع مربع اختلاف از میانگین (یعنی واریانس) برای هر خوشه حداقل شود[۱۰]. بنابراین تعریف می‌شود:

$$j = \sum_{j=1}^{K} \sum_{n \in S_j} \| X_n - \mu_j \|^2 \qquad (۱)$$

۳- خوشه بندی SOM

نقشه‌های خود سازمان دهنده یا SOM گاهی اوقات به عنوان نقشه های Kohonen به عنوان اصول شبکه های عصبی مطرح می‌شوند اما می‌توانند به عنوان یک روش خوشه‌بندی در نظر گرفته شوند [۱۱].

SOM فرآیند شبکه عصبی را اعمال می‌کند که یادگیری بی‌نظیری را برای تقسیم داده‌ها بکار می برد. در ساختار شبکه-های مستطیلی یا شش ضلعی معمولا از گره برای جداسازی داده‌ها استفاده می‌شود. هر گره شش ضلعی توسط یک بردار وزن تعریف شده است که شامل مجموعه‌ای از ابعاد مختلف بسته به بردار ورودی است. فرآیند نقشه برداری با به کارگیری وزن‌ها با مقادیر تصادفی در هر گره آغاز می‌شود. همانطور که شبکه پیشرفت می‌کند، هر بردار ورودی با وزن هر گره مقایسه می‌شود و گره با بیشترین شباهت، به نام بهترین واحد تطبیقی BMU[1] نامیده می‌شود، اختصاص می‌یابد. سپس وزن‌هادرBMU و گره‌ها براساس بردار ورودی تنظیم می‌شوند. این فرآیند تا زمانی که همه بردارهای ورودی در درون خوشه ها طبقه بندی شوند، تکرار می‌گردد [۱۲].

۴- مشخصات فیدر

مجموعه داده‌های فیدر مورد مطالعه، توسط گروه یکپارچه سازی سیستم‌های توزیع به عنوان بخشی از مطالعه در مورد ضریب نفوذ بالا PV خورشیدی در شبکه توزیع، جمع آوری شده است[۱۳].فیدر مورد نظر از میان۲۴ فیدر طبقه بندی شده‌ی آزمایشگاه ملی شمال غربی اقیانوس آرام (PNNL) انتخاب شده است[۱۴].این فیدر با نام اختصاصی4-R1-12.47 نماینده‌یک منطقه حومه شهری در سواحل غربی کشور امریکا می‌باشد و عمدتاً از خانه‌هایتک خانواده و بارهای سنگین تجاری تشکیل شده است. شبکه توزیع این فیدر رو باز نیست و ۱۰۰٪ زیر زمین می‌باشد. اکثریت بارها در نزدیکی پست قرار دارند. داده‌ها شامل بارساعتی در طول یک سال از۵۰ ترانسفورماتور تجاری و مسکونی هستند و در سال ۲۰۰۹جمع آوری شده‌اند. بارهای ترانسفورماتور های تجاری به تفکیک هر فاز مورد ارزیابی قرار می‌گیرند.در جدول ۱ مشخصات فیدر ارائه شده است.

جدول ۱: مشخصات فیدر

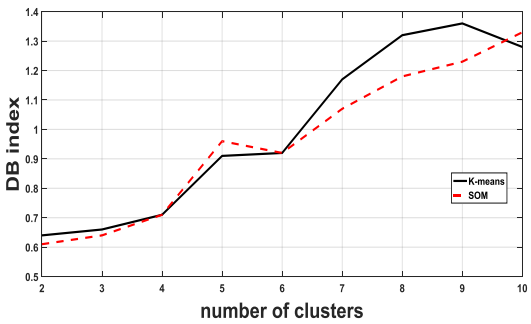
نقطه انشعاب	۳۰۲
ولتاژ(KV)	۱۲/۴۷
بار (KW)	۵۳۰۰
تنظیم‌کنندههای ولتاژ	۰
ریکلوزر	۰
ترانسفورماتورهای مسکونی	۳۸
ترانسفورماتورهای تجاری	۱۲
ترانسفورماتورهای صنعتی	۰
ترانسفورماتورهایتجاری	۰

شبیه سازی در این مقاله با استفاده از برنامه matlab R216a استفاده شده است.

۵- نتایج شبیه سازی

۵-۱- ترانسفورماتورهای مسکونی

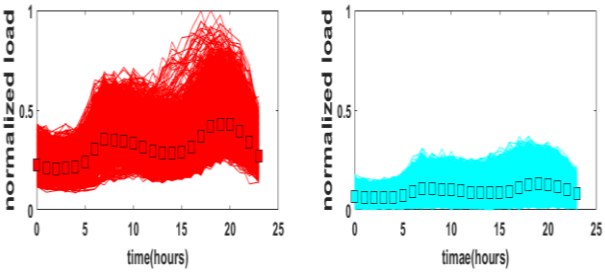
از طریق شاخص DB مطابق شکل به بررسی دو روش SOM و k-means پرداخت شد و شکل۱ عملکرد بهتر SOM را در زمانی که تعداد خوشه‌ها برابر با دو می‌باشند را نشان می‌دهد.



[1] Best Matching Unit

شکل۱: شاخص DB برای پروفایل‌های مسکونی

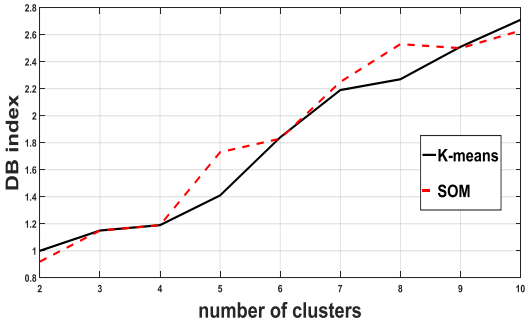
بنابراین نحوه‌ی خوشه‌بندی پروفایل‌های روزانه در طول یک سال ترانسفورماتورهای مطابق شکل ۲ خواهد بود.



شکل۲: خوشه بندی ترانسفورماتورهای مسکونی

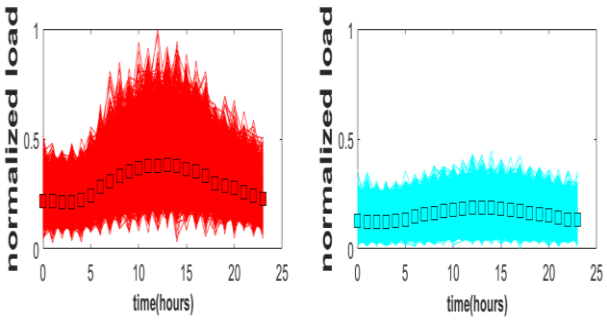
۵-۲- ترانسفورماتورهای تجاری

همچنین برای ترانسفورماتورهای تجاری شاخص DB مطابق شکل ۳ بـه بررسـی دو روش SOM و k-means پرداخـت شـد و شکل۳ عملکرد بهتر SOM را در زمانی که تعداد خوشـه‌ها برابـر با دو می‌باشند را نشان می‌دهد.



شکل۳: شاخص DB برای پروفایل‌های تجاری

بنابراین نحوه خوشه‌بندی پروفایل های روزانه در طـول یـک سال ترانسفورماتورهای مطابق شکل ۴ می‌باشد.



شکل۴: خوشه بندی ترانسفورماتورهای تجاری

۵-۳- سهم خوشه ها از بار فیدر

در این بخش سهم خوشه های بدست آمده از بار فیدر بـرای سال ۲۰۰۹ بدست می‌آید .جدول ۲ میانگین درصد سهم بار هـر خوشه از بار فیدر در ماه‌های سال را نشـان مـی دهـد کـه نشـان دهنده‌ی تغییرات در مصرف خوشه ها در ماه‌های مختلف دارد.

جدول ۲: میانگین درصد سهم بار خوشه‌ها از بار فیدر

ماه	مسکونی-خوشه ۱	مسکونی-خوشه ۲	تجاری-خوشه ۱	تجاری-خوشه ۱
ژانویه	32.944	21.497	27.282	18.275
فوریه	32.409	21.607	26.614	19.368
مارس	30.796	20.531	31.495	17.177
آوریل	29.6511	19.7483	32.686	17.914
مه	29.829	19.5483	32.10	18.522
ژوئن	30.43	19.51	33.05	16.993
ژوئیه	32.77	20.35	30.29	16.58
اوت	30.11	19.95	33.96	15.96
سپتامبر	29.8	19.89	33.21	17.09
اکتبر	28.21	18.76	35.1	17.92
نوامبر	30.98	20.54	30.16	18.30
دسامبر	35.79	21.48	24.16	18.54

۶- نتیجه گیری

در این مقاله یا استفاده از خوشه‌بندی کـه یکـی از روش‌هـا داده‌کاوی برای تحلیل داده‌ها با حجم زیاد مـی‌باشـد بـه بررسـی الگوهای بار یک فیدر شبکه توزیع پرداخت شد. دو روش خوشـه بندی k-means و SOM مورد ارزیابی قرار گرفـت و خوشـه‌هـای بهینه برای هر کلاس مصـرف بدسـت آمـد. در پایـان سـهم هـر خوشه از بار فیدر به تفکیک هر ماه حاصل شد که نشان دهنده‌ی تغییرات در میزان مصرف خوشه‌ها بود.

مراجع

[1] L. Park, S. Jeong, J. Kim, and S. Cho, "Joint Geometric Unsupervised Learning and Truthful Auction for Local Energy Market," IEEE Transactions on Industrial Electronics, vol. 66, pp. 1499-1508, 2019.

[2] M. D. Somma, G. Graditi, and P. Siano, "Optimal Bidding Strategy for a DER Aggregator in the Day-Ahead Market in the Presence of Demand Flexibility," IEEE Transactions on Industrial Electronics, vol. 66, pp. 1509-1519, 2019.

[3] Han, J.; Kamber, M. . 2006 "Data Mining, Concepts and Techniques", Morgan Kaufmannpublisher, second edition.

[4] Frawley, W.j.; Piatetsky-Shapiro, G.; Matheus, C.J; 1992. "Knowledge Discovery in databases:an Overview", AI Magazine,13, 3, 57-70.

[5] Jiawei Han, Micheline Kamber, Data Mining: Concepts and Techniques, London: Academic Press, 5, 2001.

[6] S. Kumar and M. Singh, "Big data analytics for healthcare industry: impact, applications, and tools," Big Data Mining and Analytics, vol. 2, pp. 48-57, 2019.

[7] C. Kong, G. Luo, L. Tian, and X. Cao, "Disseminating authorized content via data analysis in opportunistic social networks," Big Data Mining and Analytics, vol. 2, pp. 12-24, 2019

[8] Vasimalla, Kumar. "A survey on time series data mining." International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE) 2 (2014): 170-179

[13] Hoke, Anderson, et al. "Steady-state analysis of maximum photovoltaic penetration levels on typical distribution feeders." IEEE Transactions on Sustainable Energy 4.2 (2013): 350-357.
[14] Schneider, Kevin P., et al. Modern grid initiative distribution taxonomy final report. No. PNNL-18035. Pacific Northwest National Laboratory (PNNL), Richland, WA (US), 2008.
[9] V. Figueiredo, F. Rodrigues, Z. Vale, and J. B. Gouveia, “An Electric Energy Consumer Characterization Framework Based on Data Mining Techniques,”IEEE Transactions on Power Systems, vol. 20, no. 2, pp. 596–602, May 2005
[10] McLoughlin, Fintan. "Characterising domestic electricity demand for customer load profile segmentation." (2013).
[11] M. Halkidi, Y. Batistakis, and M. Vazirgiannis, “On Clustering Validation Techniques,” Journal of Intelligent Information Systems, pp. 107–145, 2001Gffg
[12] MathWorks, “Neural Network Toolbox Users Guide,” no. Version 4, 2002Uky