



(12)

CERERE DE BREVET DE INVENȚIE

(21) Nr. cerere: a 2021 00069

(22) Data de depozit: 23/02/2021

(41) Data publicării cererii:
30/08/2022 BOPI nr. 8/2022

(71) Solicitant:
• UNIVERSITATEA POLITEHNICA DIN
BUCUREȘTI, SPLAIUL INDEPENDENȚEI
NR.313, SECTOR 6, BUCUREȘTI, B, RO

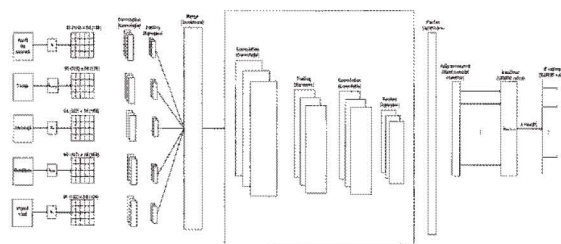
(72) Inventatori:
• BOICEA ADRIAN-VALENTIN,
CALEA FLOREASCA, NR.91-111, BL.F1,
SC.2, ET.2, AP.17, SECTOR 1,
BUCUREȘTI, B, RO

(54) **METODĂ DE PREDICȚIE A CONSUMULUI DE ENERGIE
ELECTRICĂ ȘI DE GAZE NATURALE, BAZATĂ PE BIG DATA
ȘI REȚELE NEURONALE CONVOLUȚIONALE CU INTRĂRI
MULTIPLE**

(57) Rezumat:

Invenția se referă la un procedeu de predicție a consumului de energie electrică sau de gaze naturale, bazat pe utilizarea Rețelelor Neuronale Convoluționale și a Big Data. Procedeu, conform invenției, folosește ca date de intrare/antrenare serii temporale de valori de consum energetic, preț, precum și de temperatură, irradianță, viteză a vântului, umiditate. Procedeu se poate utiliza la nivel de sistem de transport, de distribuție sau la nivel de gospodărie. Totodată, procedeu poate fi utilizat și la detectarea unor potențiale probleme de securitate ale sistemului de transport sau de distribuție, cuprinzând două etape: citirea datelor și, respectiv, obținerea predicției cu ajutorul Rețelei Neuronale Convoluționale cu intrări multiple. Datele de intrare se pot utiliza fie direct sub formă de serii temporale, fie se pot transforma în imagini.

Revendicări: 6
Figuri: 1



Cu începere de la data publicării cererii de brevet, cererea asigură, în mod provizoriu, solicitantului, protecția conferită potrivit dispozițiilor art.32 din Legea nr.64/1991, cu excepția cazurilor în care cererea de brevet de invenție a fost respinsă, retrasă sau considerată ca fiind retrasă. Întinderea protecției conferite de cererea de brevet de invenție este determinată de revendicările conținute în cererea publicată în conformitate cu art.23 alin.(1) - (3).



Titlul invenției: Metodă de predicție a consumului de energie electrică și de gaze naturale, bazată pe Big Data și rețele neuronale convoluționale cu intrări multiple

Această invenție se referă la un **procedeu de predicție a consumului de energie electrică și de gaze naturale**, la nivel de sistem energetic național utilizând date de consum stocate anterior. După cum se va putea observa, absolut toate conceptele prezentate în această descriere pot fi utilizate și în cazul unei rețele de distribuție de energie electrică, a unei rețele de transport de gaze naturale sau a unei rețele de distribuție de gaze naturale. Totodată, predicția de consum de gaze naturale sau de energie electrică, bazată pe aceste metode și concepte poate fi aplicată și la nivelul unei gospodării.

Predicția se realizează în cazul de față printr-o rețea neuronală convoluțională (RNC) cu intrări multiple. Citirea datelor se face fie prin transformarea în imagini a profilelor de consum și a variației variabilelor exogene, fie direct, prin preluarea seriilor de timp dintr-un fișier csv.

În prezent, există un număr restrâns de astfel de invenții iar acestea folosesc RNC cu o unică intrare, precum în [1,2,3,4].

În brevetul [1], predicția se face pe ziua următoare, utilizându-se ca model de antrenare o hartă de caracteristici bidimensională, rezultată în urma concatenării mai multor matrici care conțin consumul istoric și variabilele meteo exogene. O astfel de abordare în cazul Big Data, presupune un timp de calcul lung și poate conduce la o pierdere de caracteristici a datelor pe baza cărora se generează această hartă. Datele de intrare în brevetul [1] sunt organizate doar sub formă de serii de timp, nu și sub formă de imagini care sunt sugestive la detectarea unor anomalii în cazul volumelor mari de date folosite la antrenarea RNC.

Prezenta invenție realizează și codificarea datelor de intrare în imagini. Conceptele descrise în brevetul [1] sunt aplicate doar unor consumatori rezidențiali și se referă doar la energie electrică, nu și la gaze naturale. Prezenta invenție poate fi utilizată la nivel de sistem de transport, distribuție sau gospodărie atât în cazul gazelor naturale cât și al energiei electrice. Testarea rezultatelor predicției se face în brevetul [1] prin metode statistice, bazate pe coeficient de încredere. Dezavantajul în cazul de față este că această metodă nu e directă, așa cum este comparația cu un consum istoric deja realizat, utilizată în prezentul brevet.

În brevetul [2], datele de intrare sunt obținute prin prelucrarea datelor de consum ce rezultă din utilizarea simultană a mai multor RNC. Aceasta conduce la un timp de calcul lung. Un alt dezavantaj important este acela că generarea eșantionului de testare este făcută pe criterii arbitrare. Prezenta invenție utilizează un singur tip de rețea și anume o RNC cu intrări multiple ceea ce scurtează semnificativ timpul de calcul iar prognoza fiind făcută pe un an, calibrarea parametrilor rețelei din prezenta invenție se face prin comparația cu consumul istoric realizat pe anul anterior. Restul conceptelor prezentate în brevetul [2], sunt foarte asemănătoare celor din brevetul [1].

În brevetul [3], datele de antrenare pentru RNC sunt rezultate din aplicarea tehnicii LOTUS ce presupune un pas în plus de calcul. Tranziția de la zilele nelucrătoare la cele lucrătoare (și invers) se face prin hot coding și nu printr-o metodă directă. În cazul prezentei

invenții, ordonarea minutelor cu consum prognozat se face direct de la 1 la 525600 (numărul de minute dintr-un an). Cu excepția testării rezultatelor predicției, restul conceptelor prezentate sunt asemănătoare brevetelor de invenție [1,2].

În brevetul [4], deși setul de antrenare este de tip multisursă, antrenarea nu se bazează pe Big Data iar predicția este făcută pe ora următoare și nu pe termen mediu. În cazul prezentei invenții, antrenarea se face cu Big Data ceea ce înseamnă acuratețe îmbunătățită iar predicția se face pe un an. Un alt dezavantaj al soluției prezentate în brevetul [4] este acela că, pe lângă RNC, se mai folosește o altă rețea neuronală artificială, ceea ce conduce din nou la un timp de calcul mai lung, spre deosebire de prezentul brevet. Restul conceptelor prezentate în brevetul [4], sunt foarte asemănătoare celor din brevetele de invenție [1,2,3].

Niciunul dintre brevetele de invenție [1,2,3,4] nu realizează o prognoză la minut a consumului de energie electrică și nu prezintă concepte care să poată fi aplicate la nivel de transport, de distribuție sau gospodărie (ci sunt limitate doar la clienți rezidențiali). Ca atare, niciunul dintre brevetele [1,2,3,4] nu poate fi utilizat la detectarea unor potențiale probleme de securitate ale sistemului de transport sau de distribuție. Alte dezavantaje importante ale brevetelor [1,2,3,4] este că niciunul dintre acestea nu prezintă o modalitate de codificare a datelor de intrare în imagini și nu pot face o predicție a consumului de gaze naturale (la nivel de sistem de transport, distribuție sau gospodărie).

Pe de altă parte, prezenta invenție folosește o RNC cu intrări multiple, evitând astfel calculul matricial sau combinarea diverselor tipuri de date de consum cu variabile exogene ce poate conduce la o pierdere de caracteristici. În cazul prezentei invenții, predicția se poate face la nivel de transport, distribuție sau gospodărie, pe termen mediu (1 an), fie la gaze naturale, fie la energie electrică. Tot spre deosebire de brevetele de invenție [1,2,3,4], prezenta invenție poate fi utilizată de asemenea la detectarea unor potențiale probleme de securitate ale sistemului.

O variantă anterioară a prezentei invenții presupune o RNC cu o unică intrare [5]. Prin această intrare se preia o serie de timp ce constituie consumurile orare de energie electrică iar temperatura (care este unica variabilă exogenă considerată) este preluată printr-o altă rețea neuronală artificială, conectată la stratul de ieșire a RNC. Predicția în cazul acestei RNC se face pe o zi.

Problema pe care o rezolvă prezenta invenție este, așadar, într-o primă variantă constructivă, îmbunătățirea acurateții predicției pe termen mediu (1 an) a consumului de energie electrică sau de gaze naturale prin implementarea unei metode ce presupune includerea simultană a unor variabile exogene ce influențează consumul în mod direct. Într-o altă variantă constructivă, o altă problemă pe care invenția o rezolvă este aceea a implementării unei metode ce poate fi utilizată și la detectarea potențialelor breșe/probleme de securitate ale sistemului (de transport sau de distribuție).

Astfel, **avantajele invenției față de stadiul actual al tehnicii** sunt următoarele:

1. Predicția pe termen mediu (1 an) a consumului de energie electrică și de gaze naturale, bazată pe RNC cu intrări multiple și Big Data ceea ce conferă automat o acuratețe sporită față de restul metodelor prin faptul că facilitează includerea simultană a mai multor mărimi exogene

2. Prognoză la minut, pe 1 an, ceea ce facilitează utilizarea acestei metode și la detectarea potențialelor probleme de securitate ale sistemului, nu numai la predicție
3. Codarea datelor de consum în imagini, necesară în anumite aplicații cu caracter special
4. Posibilitatea de includere simultană a mai multor variabile exogene (legate de fenomene meteo sau de prețul istoric al energiei electrice sau gazelor naturale) ceea ce îmbunătățește acuratețea predicției
5. Determinarea pe cale directă a preciziei predicției, prin compararea cu un consum istoric, deja realizat
6. Posibilitatea utilizării metodei de predicție la nivel de sistem de transport, de distribuție sau la nivel de gospodărie (atât la gaze naturale cât și la energie electrică)
7. În caz de necesitate, posibilitatea utilizării metodei de predicție, separat, în funcție de tipul de generare a energiei electrice (eolian, solar, combustibil fosil etc.)
8. Posibilitatea utilizării metodei de predicție și cu o cantitate de date de intrare mai mică decât volumul de date presupus de conceptul Big Data
9. Posibilitatea utilizării metodei de predicție și pentru prognoze efectuate pe unități de timp mai mari de 1 minut.

S-a optat în cazul de față pentru RNC în detrimentul celorlalte metode de predicție pentru faptul că acest tip de rețea prezintă capacități importante de învățare profundă, asigură viteză bună de calcul și acuratețe înaltă a predicției, fiind compatibil cu conceptul de Big Data [2]. Aceasta deoarece posedă capacitatea de a extrage prin diferitele straturi convoluționale caracteristicile esențiale ale unui model caracterizat de volume mari de date [2].

Big Data nu presupune în toate domeniile același ordin de mărime în ceea ce privește volumul de date. Spre exemplu în energetică, ordinul acesta de mărime are o limită inferioară de tipul GB [6].

Procedeul din prezenta invenție presupune următoarele etape: I. citirea datelor; II. obținerea predicției cu ajutorul RNC cu intrări multiple pe baza propagării înainte a datelor și a propagării înapoi care constituie algoritmul de antrenare al rețelei.

I. Citirea datelor

Algoritmul propus se bazează pe RNC cu intrări multiple. Aceasta deoarece predicția de consum energetic se bazează și pe variabile exogene, reprezentate de temperatură, umiditate, viteza vântului, iradianță. Atât variabilele exogene istorice cât și consumul istoric reprezintă serii de timp. Fiecare dintre acestea, mai departe, vor reprezenta intrările multiple în RNC. O altă variabilă exogenă ce poate fi luată în discuție este prețul istoric al energiei electrice sau al gazului natural.

Citirea seriilor de timp se poate face fie prin transformarea acestora în imagini, fie direct sub formă de vectori de valori numerice, unidimensionali.

Deși prima variantă necesită o operație în plus, transformarea seriilor de timp în imagini poate prezenta o serie de avantaje precum [7, 8, 9]:

- stocarea datelor într-un format diferit de cel numeric, necesară în anumite aplicații;
- obținerea mai multor informații legate de comportamentul consumatorilor și facilitarea clusterizării acestora, clusterizare care îmbunătățește acuratețea predicției realizată cu o rețea neuronală artificială (dacă se optează pentru această soluție, va crește timpul de calcul);
- în caz de necesitate, posibilitatea unei compresii mai eficiente a imaginilor decât a seriilor temporale prin faptul că algoritmi de compresie, în cazul imaginilor, pot utiliza mai eficient relațiile statistice existente între datele care caracterizează imaginile;
- o mai bună structurare a datelor de intrare care în situația de față nu provin din aceeași categorie, eficientizându-se astfel operația propriu-zisă de predicție.

Citirea seriilor de timp, în cazul transformării acestora în imagini, se va face cu ajutorul coeficientului de scalare $k \in [0; 1]$, după cum urmează:

- a) se calculează maximul seriei temporale iar apoi restul valorilor seriei se va exprima în funcție de produsul din formula (1):

$$v_{crt} = \max(s_t) \cdot k \quad (1)$$

unde:

v_{crt} – valoarea curentă din seria temporală ce trebuie reprezentată;

$\max(s_t)$ – valoarea maximă din seria de timp.

În funcție de numărul, n , de ani cu care se decide antrenarea rețelei, maximul seriei curente poate reprezenta o medie aritmetică a maximelor acestor ani, după cum se poate observa în (2):

$$\max(s_t) = \frac{\max(s_t)_1 + \max(s_t)_2 + \max(s_t)_3 + \dots + \max(s_t)_n}{n} \quad (2)$$

- b) Astfel, în funcție de valorile fracționare ale lui k , vor rezulta pătrate în diferite niveluri de gri. Dacă valoarea lui k este mai apropiată de 1, atunci va rezulta un pătrat a cărui culoare va fi mai apropiată de negru, altfel dacă valoarea lui k este mai apropiată de 0, atunci va rezulta un pătrat a cărui culoare va fi mai apropiată de alb. Secvența aceasta de pătrate, caracterizată de diferite niveluri de gri, va constitui imaginea ce va servi drept date de intrare pentru RNC.

Latura acestor pătrate va avea o lungime de 20 de pixeli, ceea ce înseamnă aproximativ 0.528 cm. Ținând cont de faptul că un ecran de înaltă definiție are o rezoluție de 1920 x 1080 pixeli, vizualizarea unei astfel de imagini va presupune un număr maxim de 96 de pătrate pe orizontală și de 54 pe verticală. Atunci când rețeaua neuronală este antrenată cu Big Data, poate fi utilă folosirea unui ecran de tip 4 K, cu rezoluția 3840 x 2160 de pixeli (sau chiar a unui ecran 8 K). În cazul ecranului de tip 4 K, va rezulta un număr maxim 192 de pătrate pe orizontală, respectiv de 108 pe verticală. Indiferent de situație, trebuie aleasă reprezentarea în tonuri de gri pe 32 biți care poate reda maxim 4294967296 nuanțe.

Bineînțeles, se poate opta și pentru alte culori, diferite de tonurile de gri. Seria de valori numerice ale coeficientului de scalare va fi reprezentată și în această situație de diferitele intensități ale culorii alese. Aceasta se va face în funcție de temperatura culorii.

Procedeele descrise anterior este valabil pentru absolut toate seriile temporale, mai puțin pentru temperatură care poate fi caracterizată și de valori negative. În acest caz, se va calcula modulul fiecărei valori negative din serie și se va raporta la maximumul seriei. După determinarea lui k_{temp} pentru fiecare dintre valorile negative, acesta se va înmulți cu -1.

Formula (2) poate fi înlocuită și cu formula mediei ponderate. Astfel, consumul se poate pondera de exemplu cu temperatura, cunoscut fiind faptul că acest consum energetic tinde să crească atunci când temperatura scade sau atunci când temperatura crește foarte mult. Dacă se optează pentru o astfel de mediere, atunci calculul maximumului nu mai este corelat cu evenimente complet aleatoare precum pandemiile sau breșele de securitate care schimbă în mod radical profilul consumului. Din acest motiv, media aritmetică este mai relevantă într-o situație de acest gen.

Variabilele exogene se pot media la nivel național conform formulei (2) pentru fiecare unitate de timp, în funcție de rata de eșantionare utilizată de stațiile meteorologice.

Seriile temporale corepunzătoare consumului și variabilelor exogene trebuie să aibă aceeași dimensiune. Prognoza se face pe un an, la minut, pentru a avea o cunoaștere mai bună a momentului de timp de pe parcursul anului când poate apărea o posibilă breșă de securitate în sistem. Această breșă de securitate poate induce valori aberante de consum, consecutive, pe perioade de timp relativ lungi.

Astfel, vor rezulta serii de timp de dimensiunea 525600, egale așadar cu numărul de minute dintr-un an.

Transformarea seriilor temporale în imagini este mult mai sugestivă atunci când prognoza se face pe un an întreg, la minut. Detectarea valorilor aberante de consum ce pot fi induse de breșe de securitate, devine greoaie în cazul seriilor de timp de dimensiuni foarte mari, așa cum este și cazul de față. În situația utilizării imaginilor, aceste valori aberante sunt evidențiate imediat.

Totodată, atunci când diferența dintre prognoză și consumul propriu-zis realizat devine mai mare de un anumit prag (determinat prin încercări succesive), aceasta poate fi rezultatul unei potențiale probleme/breșe de securitate.

Cea de a doua variantă de citire, directă a seriilor temporale, se va face în mod clasic dintr-un fișier de tip csv.

Datele necesare citirii se pot prelua, de exemplu, de pe site-ul operatorului de transport energie electrică (în cazul predicției de consum energie electrică la nivel de sistem de transport).

Deci, transformarea seriilor temporale în imagini este mai sugestivă dar presupune un timp de calcul mai lung în raport cu citirea directă, în ceea ce privește obținerea predicției.

II. Obținerea predicției, arhitectura RNC și descrierea algoritmului de antrenare

În toate cazurile de tipuri de citire, algoritmul de antrenare a RNC, din punct de vedere principal, este identic. În ceea ce privește arhitectura RNC, în cazul citirii directe, se va renunța la transformarea în imagini a seriilor de timp, la stratul convoluțional dedicat acestora și la eventualul strat de agregare (Pooling), trecându-se direct la blocul de combinare (Merge), conform Fig. 1. Restul arhitecturii rețelei rămâne identic.

Transferul datelor de la intrare către ieșirea din RNC se face, în ordine, prin operațiile de convoluție, agregare și prin stratul complet conectat (Fully Connected Layer). Fiecare strat de convoluție prezintă o serie de filtre antrenabile (kernel) pentru filtrarea rezultatelor stratului anterior. Filtrarea este urmată de o funcție de activare pentru obținerea hărții caracteristicilor (feature map). Funcția de activare este o funcție ReLU (Rectified Liniar Unit) și este descrisă în (3):

$$f_{act}(x) = \max\{0, x\} \quad (3)$$

Pentru ambele tipuri de date de intrare (imagini sau serii de timp), dacă se dorește o acuratețe crescută și nu neapărat viteză de calcul, se renunță la straturile de agregare. Agregarea conduce implicit la o pierdere de caracteristici.

În cazul folosirii clusterizării, imaginile de la intrarea în rețea vor avea zone de instensități ale culorilor clar delimitate. Aceasta înseamnă că este utilă folosirea unor filtre de dimensiune 3 care pot identifica mai ușor marginile acestor zone iar apoi a unor filtre de dimensiune 5 pentru a determina conținutul propriu-zis al imaginii. Conform Fig.1., la intrare, înainte de combinare, nu se recomandă mai mult de 1, maxim 2 straturi convoluționale. După blocul de combinare din aceeași figură, nu se recomandă mai mult de 3 straturi convoluționale. Pe măsură ce straturile de convoluție se înmulțesc, acuratețea tinde să scadă [10].

Nu există o formulă în baza căreia să se determine numărul straturilor convoluționale pentru niciunul dintre cele două tipuri de citiri. În funcție de volumul de date cu care se decide a se face antrenarea rețelei, prin încercări succesive, se poate determina acest număr sub forma unei soluții de compromis între timpul de calcul și acuratețea predicției. Se recomandă ca predicția să se facă pe un an din trecut, pentru care se cunoaște deja consumul real. Astfel, acuratețea predicției se determină prin compararea predicției cu acest consum realizat și totodată, astfel, se face și calibrarea parametrilor rețelei. Acuratețea predicției se determină, deci, prin calculul abaterii absolute dintre consumul prognozat și cel realizat.

Așadar, în determinarea numărului de straturi convoluționale, validarea rezultatelor joacă un rol deosebit de important, aceasta fiind obținută în momentul în care abaterea absolută devine minimă. Acest prag de minimalitate va fi decis de utilizator.

În cele ce urmează, va descrie algoritmul de antrenare al RNC.

Algoritmul de antrenare aplicat rețelei din Fig. 1. este compus din două etape: a. inferență (propagare înainte) respectiv b. propagare înapoi.

a. Propagarea înainte presupune operațiile de convoluție, agregare și de utilizare a funcției de activare pentru obținerea hărții caracteristicilor. Aceste operații sunt redate în formula (4) [10]:

$$y_j^{s+1} = f(\sum_{i \in H_j} y_i^s * kern_{ij}^{s+1} + b_j^{s+1}) \quad (4)$$

unde:

y_j^{s+1} -rezultatul operației de convoluție de la etapa de calcul curentă (corespunzătoare stratului $s+1$)

H_j -setul de hărți de intrare

y_i^s - rezultatul operației de convoluție de la etapa de calcul anterioară

$kern_{ij}^{s+1}$ -ponderea aflată pe linia i și coloana j din fiecare filtru (kernel)

b_j^{s+1} -bias de tip aditiv, adăugat fiecărei hărți obținute pe baza funcției de activare

Dacă se optează pentru utilizarea stratului de agregare, aceasta va provoca o reducere dimensională a rezultatului convoluției, conform formulei (5) [10]:

$$y_j^{s+1} = B_j^{s+1} \cdot pool(y_j^s) + b_j^{s+1} \quad (5)$$

unde:

B_j^{s+1} -bias de tip multiplicativ

$pool(y_j^s)$ -funcție de reducere dimensională (agregare) a rezultatului convoluției

Ca funcție de activare pentru stratul de ieșire, se poate folosi funcția logistică, descrisă în formula (6):

$$f(x) = \frac{1}{1+e^{-x}} \quad (6)$$

unde:

x -un vector cu dimensiunea 525600 care poate lua valori între 0 și 1.

Tipul acesta de funcție a fost ales deoarece este util la rețelele neuronale folosite la predicții. Dacă modelul nu converge, se poate utiliza și funcția *softmax* care reprezintă o generalizare a funcției logistice.

b. Propagarea înapoi

Această operație are rolul de a actualiza parametrii rețelei în așa fel încât ecartul dintre ieșirea reală, obținută prin Metoda Gradientului, și cea dorită să fie minim. Acest ecart este redat prin funcția de entropie încrucișată din formula (7):

$$Pierd_i = -\log\left(\frac{e^{p_{ji}}}{\sum_{j=1}^{525600} e^{p_j}}\right) \quad (7)$$

unde:

p_{ji} -probabilitatea ca valoarea curentă i să ia valoarea prezisă j .

Pentru evitarea utilizării unei alte rețele neuronale artificiale, conectate la stratul de ieșire al RNC, rețea care ar trebui să trateze probleme legate de tranziția de la zilele lucrătoare la cele nelucrătoare și invers, la stratul de ieșire complet conectat se impune ca încărcările prezise la minut să se facă în ordine de la minutul 1 la minutul 525600.

În cazul în care abaterea absolută nu scade la mai puțin de 5 %, se va face antrenarea acestei RNC cu intrări multiple cu ADAM (Adaptive Moment Estimation) care furnizează rezultate bune mai ales în cazul citirii directe de serii de timp.

Ca domenii de utilizare a prezentei invenții se menționează:

- metodele de predicție consum, în cazul ambelor tipuri de citire, pot fi utilizate la evidențierea unor potențiale probleme/breșe de securitate ale sistemului energetic de transport sau de distribuție;
- metodele de predicție consum, în cazul ambelor tipuri de citire pot fi folosite direct și la predicția consumului de energie electrică, în cazul unei gospodării dotate cu smart-meter-uri care furnizează cantități importante de date (Big Data);
- metodele de predicție consum, în cazul ambelor tipuri de citire pot fi folosite direct și într-o rețea de distribuție de energie electrică, nu numai într-o rețea de transport a energiei electrice;
- metodele de predicție consum, în cazul ambelor tipuri de citire, pot fi folosite direct și în cazul transportului și distribuției de gaze naturale. Într-o atare situație, consumul de gaz metan este influențat direct de temperatură și de preț. Astfel, se consideră aici ca variabile exogene doar temperatura și prețul istoric al gazelor naturale;
- metodele de predicție consum, în cazul ambelor tipuri de citire, pot fi folosite direct și la nivelul unei gospodării, pentru estimarea consumului de gaze naturale;
- metodele de predicție de consum, în cazul ambelor tipuri de citire, se pot utiliza și separat, în funcție de tipul generării (eolian, solar, combustibil fosil etc.).

Pentru îmbunătățirea acurateții predicției, pe măsura trecerii timpului, este necesară îmbogățirea setului de date de antrenare, prin adăugarea automată la setul de antrenare a anului curent, după ce acesta s-a încheiat.

În cazul ambelor tipuri de citire, pentru creșterea vitezei de calcul, datele de intrare pot consta doar din consum (fără variabile exogene), din moment ce consumul este istoric și deja influențat de aceste variabile.

Totodată în cazul ambelor tipuri de citire, setul de date de antrenare este utilizat 100 % pentru antrenare iar testarea rezultatelor se face prin compararea predicției cu consumul realizat din anul istoric pentru care s-a făcut predicția (datele de consum din anul istoric sunt disponibile pe website-ul operatorului de transport de energie electrică). Astfel, 100 % din datele destinate testării, se vor utiliza în acest scop. Aceasta spre deosebire de situația cea mai des întâlnită în literatura de specialitate în care din totalul datelor utilizate, 80 % sunt folosite la antrenare iar 20 % la testare.

Procedeul descris anterior prezintă un important potențial de a fi implementat și produs industrial, având în vedere interesul sporit manifestat în prezent pentru problemele legate de predicția precisă a consumului de energie electrică sau de gaze naturale precum și de potențialele probleme de securitate pe care le implică operarea sistemului de transport sau distribuție de energie electrică sau gaze naturale. Procedeul poate fi utilizat, totodată, și la nivel rezidențial.

Bibliografie

- [1] C. Lilin *et al.*, “Resident load probability prediction deep learning method considering microclimate and user mode,” CN Patent CN111275571A, Jun. 12, 2020.
- [2] C. Fengming *et al.*, “A kind of method of the intelligent building microgrid load prediction based on deep learning,” CN Patent CN107423839A, Dec. 1, 2017.
- [3] D. Zhuofu *et al.*, “Short-Term Load Forecasting Method based on information fusion convolutional neural networks model,” CN Patent CN109543901A, Mar. 29, 2019.
- [4] F. Haiwei *et al.*, “Load prediction method, system and equipment based on multi-source data and hybrid neural network,” CN Patent CN111695731A, Sep. 22, 2020.
- [5] A. M. Tudose, D. O. Sidea, I. I. Picioroaga, V. A. Boicea and C. Bulac, “A CNN Based Model for Short-Term Load Forecasting: A Real Case Study on the Romanian Power System,” *2020 55th International Universities Power Engineering Conference (UPEC)*, Torino, Italy, 2020, pp. 1-6.
- [6] H. Akhavan-Hejazi and H. Mohsenian-Rad, “Power systems big data analytics: An assessment of paradigm shift barriers and prospects,” *Energy Reports*, vol. 4, pp. 91-100, Feb. 2018.
- [7] K. Park and S. Son, “A Novel Load Image Profile-Based Electricity Load Clustering Methodology,” in *IEEE Access*, vol. 7, pp. 59048-59058, 2019.
- [8] The University of Iowa, “Digital Image processing,” 2000. [Online]. Disponibil: <http://user.engineering.uiowa.edu/~dip/lecture/DataCompression.html> (accesat Dec. 21, 2020).
- [9] A. Shahzadeh, A. Khosravi and S. Nahavandi, “Improving load forecast accuracy by clustering consumers using smart meter data,” *2015 International Joint Conference on Neural Networks (IJCNN)*, Killarney, 2015, pp. 1-7.
- [10] Y. Sun *et al.*, “Multi-Input Convolutional Neural Network for Flower Grading,” *Journal of Electrical and Computer Engineering*, vol. 2017, Aug. 31, 2017.

Revendicări

1. Metodă (bazată pe RNC cu intrări multiple) de realizare a predicției consumului la nivel de sistem energetic național, pe un an, la minut, caracterizată prin aceea că utilizează variabile exogene precum temperatura, iradianța, viteza vântului, umiditatea și prețul energiei (în cazul predicției consumului de energie electrică) sau a temperaturii și prețului gazelor naturale (în cazul predicției consumului de gaze naturale).

2. Metodă de predicție conform revendicării 1, caracterizată prin aceea că, într-o primă variantă constructivă (citirea directă a seriilor de timp), metoda constă din următorii pași:

Pasul 1: Generarea seriilor de timp, rezultate din Big Data, corespunzătoare consumului și variabilelor exogene;

Pasul 2: Propagarea datelor către blocul de combinare

Pasul 3: Propagarea datelor prin inferență către straturile convoluționale și de agregare

Pasul 4: Propagarea datelor către blocul de aplatizare și cel complet conectat pentru determinarea valorilor prognozate pentru fiecare minut de pe parcursul anului.

3. Metodă de predicție conform revendicării 1, caracterizată prin aceea că, într-o a doua variantă constructivă (transformarea seriilor de timp în imagini), metoda constă din următorii pași:

Pasul 1: Generarea seriilor de timp, rezultate din Big Data, corespunzătoare consumului și variabilelor exogene;

Pasul 2: Transformarea seriilor de timp în imagini, pe baza coeficientului de scalare k după cum urmează:

a) Se determină maximul seriei temporale iar apoi restul valorilor seriei se va exprima conform relației $v_{crt} = \max(s_t) \cdot k$

Atunci când n este numărul de ani cu care se decide antrenarea rețelei, maximul seriei curente reprezintă o medie aritmetică a maximelor acestor ani.

b) În funcție de valorile fracționare ale lui k , vor rezulta pătrate în diferite niveluri de gri. Dacă valoarea lui k este mai apropiată de 1, atunci va rezulta un pătrat a cărui culoare va fi mai apropiată de negru, altfel dacă valoarea lui k este mai apropiată de 0, atunci va rezulta un pătrat a cărui culoare va fi mai apropiată de alb. Secvența aceasta, caracterizată de diferite niveluri de gri, va constitui imaginea ce va servi drept date de intrare pentru RNC.

Procedeul descris la punctele a) și b) este valabil pentru absolut toate seriile temporale. Temperatura poate fi caracterizată și de valori negative. Pentru acestea, se va calcula modulul fiecărei valori negative din serie și apoi se va raporta la maximul seriei. După determinarea

coeficientului de scalare pentru temperatură, k_{temp} , pentru fiecare dintre valorile negative, acesta se va înmulți cu -1.

Pasul 3: Transmisia fiecărei imagini, prin inferență, în parte către stratul de convoluție și eventualul strat de agregare

Pasul 4: Propagarea datelor către blocul de combinare

Pasul 5: Propagarea datelor prin inferență către straturile convoluționale și de agregare, respectiv către blocul de aplatizare și cel complet conectat pentru determinarea valorilor prognozate pentru fiecare minut de pe parcursul anului.

4. Metodă de transformare a profilului de consum în imagini, conform revendicării 3, caracterizată prin aceea că fiecare nivel de încărcare corespunde unui pătrat cu latura de lungime 20 de pixeli. Pătratele pot fi reprezentate în diverse tonuri de gri sau utilizând diverse alte culori, de diferite tonuri, în funcție de valoarea încărcării.
5. În cazul ambelor tipuri de citire, setul de date de antrenare este utilizat 100 % pentru antrenare iar testarea rezultatelor se face prin compararea predicției cu consumul realizat din anul istoric pentru care s-a făcut predicția. Astfel, 100 % din datele destinate testării, se vor utiliza în acest scop.
6. În cazul în care abaterea absolută dintre consumul prognozat și cel realizat nu scade la mai puțin de 5 %, se va face antrenarea acestei RNC cu intrări multiple cu ADAM (Adaptive Moment Estimation).

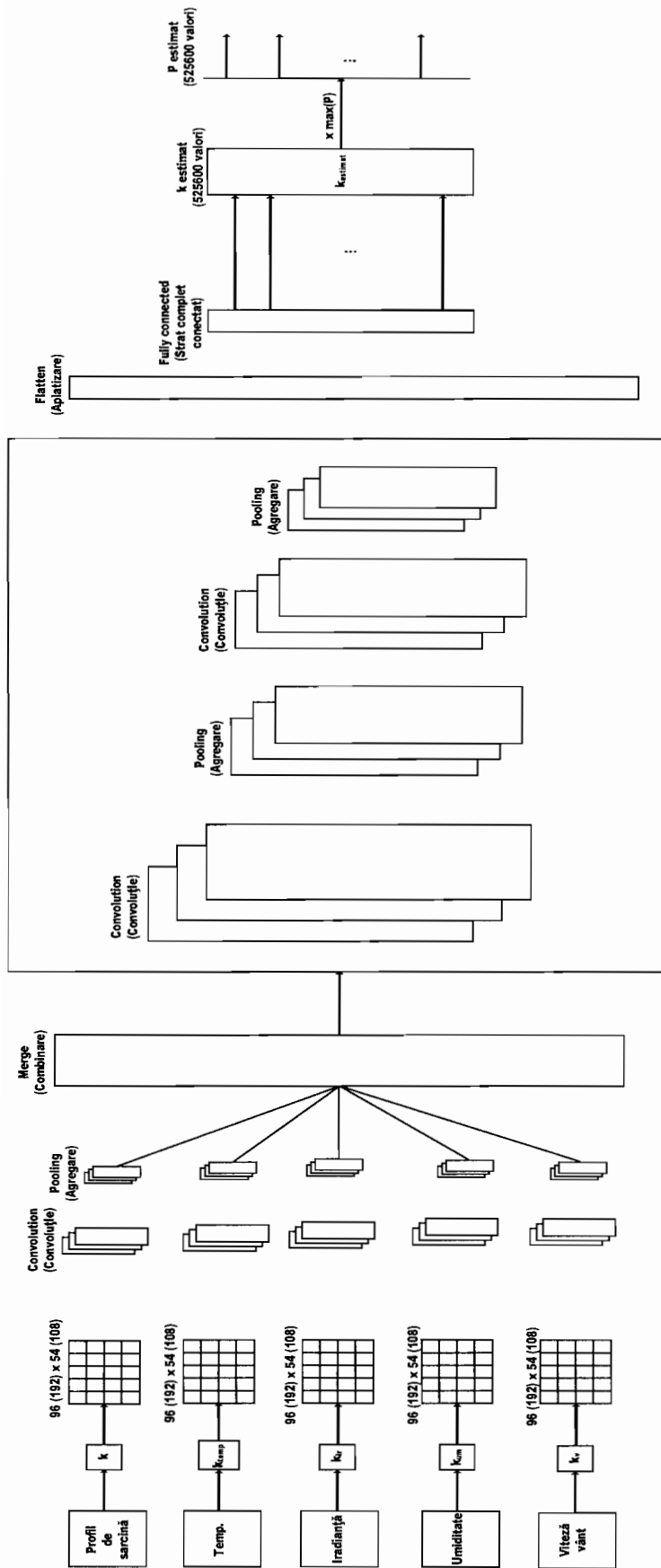


Fig. 1. Arhitectura RNC cu intrări multiple