

## Развитие на методика за групиране на денонощни товари графици

Никола Кибритев, Людмил Михайлов

**Grouping of daily load schedules:** *There is a need of methodology for grouping daily load schedules having similar characteristics. Such methodology was developed and tested in real environment and with real load data in 2011. Daily load schedules are grouped using statistical processing and the results of the grouping are visualized. The research is automated using MATLAB®. This article analyses the methodology from 2011 and exposes some weaknesses. Furthermore, it presents a development of the methodology by achieving better results. Other issues are discussed and conclusions are made.*

**Keywords:** *energy modeling, energy efficiency, energy management, forecasting, methodology*

### Въведение

В [2] е разгледана методика за групиране на денонощни товари графици (ДТГ). Предложената методика обособява групи от дни, за които товарите графици притежават сходни статистически характеристики. Направеното предположение, че методиката ще може да се използва за всякакъв вид енергийни товари не се потвърждава в някои случаи, тъй като са установени някои недостатъци:

- Не се разглежда хипотезата, когато не се изпълнява критерия на Фишер и се преминава към критерия на Стюдънт чрез преизчисляване на степените на свобода посредством [1], дали някои потенциални членове на групата, които са пропускани, вече заемат по-точно своето място в потенциалната група;
- Представените резултати от проверката обхващат само един месец – не може категорично да се докаже точната работа на методиката;
- Методиката групира дни, но не разглежда случаите, в които един ден може да принадлежи към повече от една група и как следва да се процедира в тези случаи;
- Методиката не разглежда случаите, когато се появяват аномалии – денонощни товари графици с различна форма, но еднакви в статистически смисъл попадат в една и съща група;
- Методиката не предлага начин за филтриране на дните и окончателно образуване на групи, в които да фигурират товари графици, които попадат в доверителния интервал на групата.

Редно е да се отбележи, че групирането само по себе си отразява вероятността един ден да бъде в една група с друг ден. След проведени тестове се оказва, че един ден може да принадлежи към повече от една група и е необходимо да се намери инструмент за филтриране, чрез който да се определи точно към коя група следва да се причисли съответния товаровграфик.

Целта на настоящото изследване е да се провери актуалността на предложената методика за групиране [2], като същата се подложи на тестове с цел повишаване на нейната работоспособност.

### Изложение

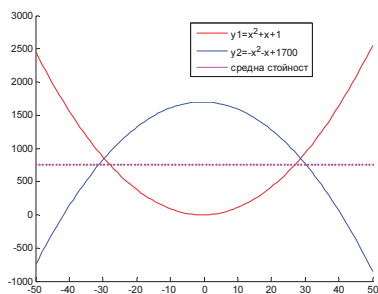
Предлага се подобряване работата на методиката от [2], като тя се допълва и преработва по следния начин:

1. Сортират се ТГ по избран признак за определен период.
2. За всеки ТГ се изчислява средния активен товар, коригираната дисперсия и степените на свобода.
3. Съставят се възможните двойки ТГ за избрания период – всеки ден с всеки.

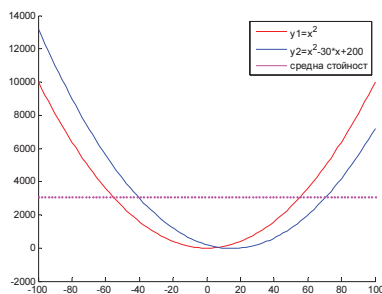
4. За всяка двойка ТГ се изчислява общото средноквадратично отклонение и  $t$ -критерия на Стюдънт.
5. За всяка двойка ТГ се проверява хипотезата за равенство на дисперсиите на разхода на електрическа енергия, посредством изчисляване на критерия на Фишер  $F$ , като в числителя е по-голямата от двете дисперсии.
6. От [1] се получава критичната стойност на критерия  $F$  на Фишер.
7. Сравняват се изчисленият  $F$  и критичната стойност на  $F - F_{cr}$ , като при:
  - a.  $F < F_{cr}$  следва, че за двата ТГ има статистическо равенство и за проверка за равенство на средните стойности на разхода на електрическа енергия може да се използва  $t$ -критерия на Стюдънт;
  - b. При  $F > F_{cr}$ , се изчисляват новите степени на свобода на двойката ТГ и се прилага отново  $t$ -критерия на Стюдънт, но за новите степени на свобода.
8. Изчисляват се коригираните средноквадратични отклонения за всички дни.
9. Изчисляват се степените на свобода на средноквадратичното отклонение.
10. Определя се общото средноквадратично отклонение за всички двойки дни.
11. Изчислява се  $t$ -критерия на Стюдънт.
12. Проверява се критичната стойност  $t_{cr}$  на  $t$ -критерия на Стюдънт при доверителна вероятност  $\gamma=0,95$  и степени на свобода  $k$  [1]. Сравняват се двете стойности за  $t$ -критерия – изчислената и критичната и се извършва преценка. Ако изчисленият  $t$ -критерий е по-голям от  $t_{cr}$ , то разходите на електрическа енергия за двойката ТГ се различават статистически значимо и товарите им графици не попадат в една група. В обратния случай двойката потенциално принадлежи към една и съща група.
13. Извършва се класиране на ТГ – на първо място се класира този ТГ, който потенциално се групира с най-много други ТГ.
14. Образуват се потенциалните групи.
15. Извършва се филтриране на групите посредством итеративен подход.
16. Извършва се визуализация на получените групи.

Описаните по-горе слабости на методиката [2] се демонстрирани на фиг. 1 и фиг.2. На **Error! Reference source not found.** се илюстрира случая, когато две функции имат близки средни стойности, огледални са една на друга, дисперсиите им също са много близки като стойности, както и средноквадратичните им отклонения. От казаното следва, че прилагайки методиката [2], тези два ТГ биха попаднали в една група, което ще доведе до изкривяване на групата и до резултати без практическо значение. За илюстрирания опростен случай се приемат две функции  $y_1 = x^2 + x + 1$  и  $y_2 = -x^2 - x + 1700$ , които са разгледани в интервала от -50 до 50. От фиг. 1 ясно се вижда, че ако това са реални ТГ, то те не би следвало да принадлежат към една и съща група. Стойностите на статистическите показатели на тези два ТГ са съответно: *Средна стойност:  $y_{1cp} = 851.1684$ ,  $y_{2cp} = 849.8316$ ; Средноквадратично отклонение:  $y_{1crk} = y_{2crk} = 764.6916$ ; Дисперсия:  $y_{1duc} = y_{2duc} = 5.8475e+05$ ; Изчислен  $t$ -критерий на Стюдънт: 0.0175.* Следва също така, че общото средноквадратично отклонение и изчисления  $t$ -критерий на Стюдънт ще бъдат равни и ТГ ще попаднат в една и съща група. На **Error! Reference source not found.** е показан друг случай, когато биха възникнали аномалии при групирането. Двата ТГ са изместени във времето един спрямо друг. За илюстрирания опростен случай се приемат две функции  $y_3 = x^2$  и  $y_4 = x^2 - 30x$ , които са разгледани в интервала от -100 до 100. От **Error! Reference source not found.** ясно се вижда, че ако това бяха реални ТГ, то те не би следвало да принадлежат към една и съща група, най-малкото защото има изместване във времето между тях и евентуалните заявки за доставка на съответния енергиен носител биха били също изместени, което ще доведе до съответния небаланс. Стойностите на статистическите показатели на тези два ТГ са

съответно: Средна стойност:  $y_{3cp} = 3400$  ,  $y_{4cp} = 3600$ ; Средноквадратично отклонение:  $y_{3сркв} = 3056$ ,  $y_{4сркв} = 3526$ ; Дисперсия:  $y_{3дис} = 9.3423e+06$ ,  $y_{4дис} = 1.2434e+07$ ; Изчислен  $t$ -критерий на Стюдънт: 0.6061.

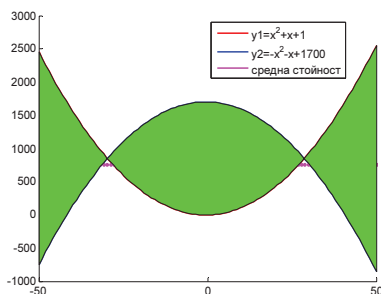


Фиг.1. Случайна близкисредностности, когато функциите са огледални една на друга

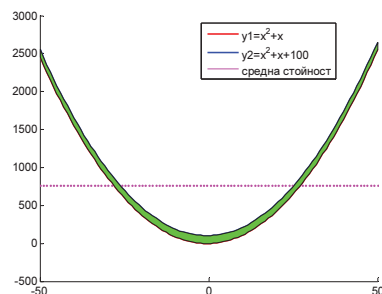


Фиг.2. Случайна близкисредностности, когато функциите са изместени във времето една спрямо друга

След проучване на предложените методи за намиране на сходства между криви [3], се предлага допълнителна обработка на ТГ с цел да се минимизира вероятността някой от посочените на Error! Reference source not found. и Error! Reference source not found. случаи да се появи. За целта се дефинира **разстояние между два ТГ**. Разстоянието между два ТГ представлява абсолютната стойност на разликата между моментните стойности на товарите за определен момент. Дефиницията е визуализирана на Фиг.3 и Фиг.4. Площта в зелено представлява отражение на разстоянието между ТГ. Колкото по-голяма площ има затворена между ТГ, толкова е по-голямо разстоянието между тях и толкова по-малко вероятно е те да принадлежат към една и съща група ТГ.



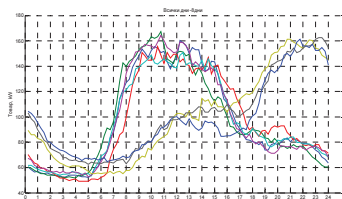
Фиг.3. Разстояние между ТГ, които не са в една група – площта в зелено



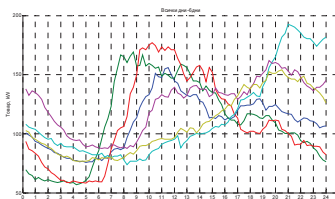
Фиг.4. Разстояние между ТГ, които са в една група – площта в зелено

На Фиг.5, Фиг.6 и фиг.7 са показани случаи на аномалии при групиране на ТГ на реален обект. На фигурите ясно се вижда огледалния характер на ТГ, които са попаднали в една група. Показаните аномалии са получени когато на обработка за групиране са подложени голям брой дни от няколко последователни години, без да се извършва предварително разделение на работни, почивни и друг тип дни. В илюстрираните случаи почивни дни със среден товар, идентичен с група работни дни попадат в една група. Въпреки, че средната стойност е близка, тези ТГ съществено се различават по моментните стойности на този товар и от там и

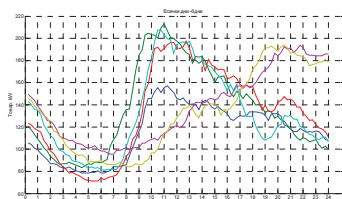
формата им е съществено различна. Случаят, когато товарите графици са изместени във времето се получава, когато се извършва смяната на часовото време, промяна в смените на работа на обекта и др. На базата на показаното до тук, можем да се направи следния извод, който е и част от предложените подобрения на методиката за групиране на ТГ: необходимо е да се добави допълнително филтриране на ТГ по параметър, различен от средната стойност, средноквадратичното отклонение и дисперсията на случайната величина.



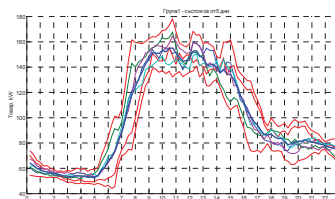
Фиг.5. Представяна Аномалия 1



Фиг.6. Представяна Аномалия 2



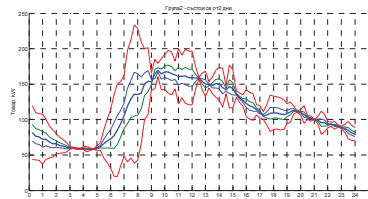
Фиг.7. Представяна Аномалия 3



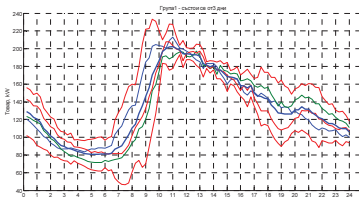
Фиг.8. Неутрализирана Аномалия 1

От представените фигури с аномалии се представя на отделна фигура водача на групата, един от членовете на групата и една от аномалиите. Илюстрира се разстоянието между тях посредством методиката, показана на Фиг.3 и Фиг.4. Избират се няколко ТГ, които очевидно принадлежат на групата и тези, които представляват аномалията. Постановката се представя посредством графиките на фиг.5, фиг.6 и фиг. 7. Разглеждат се 3 аномалии: Аномалия 1 (фиг.5): ТГ от групата: 25.06.2012 – понеделник, 12.10.2012 – петък, 06.10.2011 – четвъртък, 21.06.2012 – четвъртък, 22.06.2012 – петък; ТГ аномалии: 16.10.2011 – неделя, 10.10.2010 – неделя, 24.10.2010 – неделя; Аномалия 2 (фиг. 6): ТГ от групата: 03.12.2011 – събота, 23.10.2012 – вторник, 14.10.2011 – петък; ТГ аномалии: 06.03.2011 – неделя, 15.01.2011 – събота, 18.11.2012 – неделя; Аномалия 3 (фиг.7): ТГ от групата: 26.11.2011 – събота, 16.12.2011 – петък, 14.03.2011 – понеделник, 25.03.2011 – петък; ТГ аномалии – 27.02.2011 – неделя, 16.01.2011 – неделя.

На показаните фиг.8, фиг. 9 и фиг.10 резултати, аномалията е неутрализирана и ТГ са групирани коректно, посредством измерване на разстоянието между ТГ.



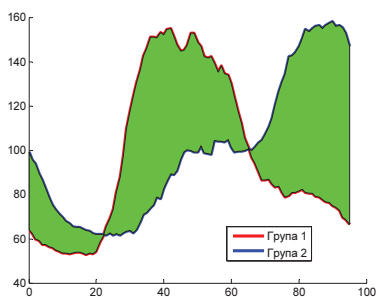
Фиг.9. Неутрализиране на Аномалия 2



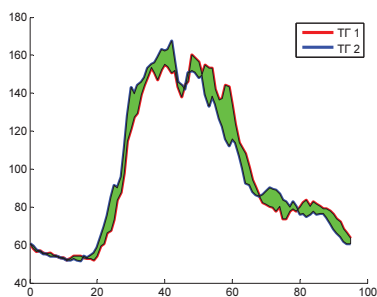
Фиг.10. Неутрализиране на Аномалия 3

Илюстрират се разликите в разстоянията между ТГ в група 1 и тези в група 2, като за сравнение се показват разстоянията между дни от група 1 – със зелен цвят.

Разликата е очевидна (Error! Reference source not found.). По статистически показатели ТГ от група 1



Фиг.11. ТГ в различни групи



Фиг.12. ТГ в една и съща група

и група 2 би трябвало да принадлежат към една и съща група. Прилагайки методиката за пресмятане на средното разстояние между ТГ, могат да се отделят в различни групи и да се осигури прецизност на системата за групиране на ТГ.

**За Група 1 и Група 2:** Средна стойност: Група1<sub>ср</sub> = 96.7, Група2<sub>ср</sub> = 99.2; Средноквадратично отклонение: Група1<sub>сркв</sub> = 35.7, Група2<sub>сркв</sub> = 31.7; Дисперсия: Група1<sub>дис</sub> = 1.2776e+03, Група2<sub>дис</sub> = 1.0067e+03; Изчислен t-критерий на Стюдънт: 0.7154.

**За ТГ1 и ТГ 2:** Средна стойност: ТГ1 = 96.8, ТГ2 = 96.7; Средноквадратично отклонение: ТГ1<sub>сркв</sub> = 37.3, ТГ2<sub>сркв</sub> = 37.1; Дисперсия: ТГ1<sub>дис</sub> = 1.3942e+03, ТГ2<sub>дис</sub> = 1.3773e+03; Изчислен t-критерий на Стюдънт: 0.0378.

Въпреки различията в стойностите за изчисления t-критерий и при двата случая ТГ попадат в една и съща група. Прилагайки метода на измерване на разстоянието между тях, може да се вземе правилно решение относно групирането на тези ТГ. Изчислявайки средното разстояние между ТГ се получава:

- Средно разстояние между Група 1 и Група 2 – 41.89;
- Средно разстояние между ТГ1 и ТГ2 – 7.07.

Съотнасяйки получения резултат към средната стойност на фактическата група – 96.7 се получава отклонение от 43.3% за случая на аномалия и 7.3% за случая на попадане в една група. За да се интерпретира резултата, нека се разгледа Фиг.3 и Фиг.4. В първия случай средното разстояние между ТГ е 1157, като  $y1_{ср} = 851.1684$ ,  $y2_{ср} = 849.8316$ . Видно е, че е налице сериозно отклонение на средното разстояние от средната стойност – над 100%. Във втория случай средното разстояние е 100, като  $y3_{ср} = 3400$ ,  $y4_{ср} = 3600$ . Видно е, че отклонението на средното разстояние от средната стойност на ТГ е около 3%.

Възниква въпросът, как се преценява какво отклонение да се допусне по време на групирането на ТГ. Това следва да се извърши след индивидуална преценка, според целта на групирането и степента на точност, която се търси, както и според характеристиките на разглеждания обект. По-ниска стойност на допустимото отклонение ще доведе до повече на брой групи, по-тесен доверителен интервал и по-малко членове на групите. В този случай ТГ-водач на групата, ще се отклонява в по-малка степен от отделните ТГ в групата. Обратно, допускайки по-голям процент на отклонение на средното разстояние от ТГ-водач на групата, ще се получават по-малко на брой групи, по-широк доверителен интервал и повече членове на групите.

За Аномалия2 и Аномалия3 може да се проведат същите изследвания и ще се получат идентични резултати.

От горе изложеното се достига до извода, че когато се работи с голям брой ТГ, които следва да се групират, може да се допусне по-нисък процент на отклонение, а когато се работи с по-малък брой ТГ е целесъобразно да се избере по-висок процент

на отклонение. Следва да бъде ясно, че всичко това зависи и от характерът на ТГ и типа на наблюдаваните обекти. Колкото по-непостоянен и непредсказуем е един обект, толкова по-трудно би се извършило групиране. В този случай следва да се започне с по-висок допуск за отклонение и да се намалява до достигане на оптимално групиране, като моментът на оптималност е строго индивидуален за всеки обект и зависи от неговото поведение в енергиен смисъл.

От илюстрираните примери е видно, че аномалиите се откриват лесно чрез оглед – те са силно различими в постановките с които работим. На основа на това следва да се приеме за оптимален процент на отклонение на средната стойност на ТГ-водач от средното разстояние между конкретния ТГ и ТГ-водач за конкретния обект тази стойност, за която не възникват аномалии при групирането, т.е. групите са чисти.

### **Заключение**

1. Предложена е методика за групиране на товари графици от всякакъв тип, посредством сравнителен анализ на статистическите им показатели – групиране, класиране и филтриране.

2. Описани са откритите слабости на методиката и са предложени инструменти за тяхното неутрализиране и по този начин значително подобряване на качеството на получаваните резултати. Предложено е разширяване на алгоритъма за групиране и е извършено тестване с голям брой реални данни от обекти със жилищно, стопанско и смесено предназначение.

3. Предложена е блокова схема за програмна реализация и автоматизация на методиката. Предложен е начин за визуализация на резултатите от методиката. Предложен е алгоритъм за финално групиране на дните отговарящи на критериите на методиката.

4. С оглед на резултатите следва, че:

- Методиката за групиране на денонощни товари графици е приложима в реални условия;
- Методиката подлежи на достъпна програмна реализация и симулация посредством MATLAB®;
- Методиката представлява инструмент за подпомагане на изследването на факторите, които влияят върху денонощните товари графици.

### **Литература**

[1] Статистически методи за изследване и оптимизиране на селскостопанска техника, Митков Ат., Д. М, София 1993, Земиздат.

[2] Групиране на денонощни товари графици, Кибритев Н., Научни трудове на русенския университет, 2011, том 50.

[3] CurveMatching, TimeWarping, andLightFields: New AlgorithmsforComputingSimilaritybetweenCurves, Efrat A., Q. F., SureshVenkatasubramanian AT&T Labs – Research

### **За контакти:**

Маг. инж. Никола Алексиев Кибритев, Русенски университет „Ангел Кънчев“, Катедра „Електроснабдяване и електрообзавеждане“, тел.:0898-713415, e-mail: [nkibritev@uni-ruse.bg](mailto:nkibritev@uni-ruse.bg)

**Докладът е рецензиран.**