IOP Conf. Series: Journal of Physics: Conf. Series **1065** (2018) 202005 doi:10.1088/1742-6596/1065/20/202005

Data analysis of energy surveys by preference aggregation method

M A Borisova, S V Muravyov
Division of Automation & Robotics
National Research Tomsk Polytechnic University
Pr. Lenina, 30, Tomsk, 634050, Russian Federation

E-mail: muravyov@tpu.ru

Abstract. Systematic conducting energy surveys of power networks to identify a possible energy loss is the basis for energy saving and energy efficiency. In the paper, a preference aggregation based method is proposed to arrange the big data of instrumental investigations of auxiliary power consumption by substations of electric power distribution networks. This approach provides a compact integral ordinal scale estimate of substations' energy consumption for own needs, which can provide a rapid decision making and visualization. Application of the proposed method for processing energy survey data of real backbone electric grids of Russia is discussed.

1. Introduction

The main task of conducting energy surveys is the calculation of electricity losses in electrical substations. International recommendations in energy efficiency require that in the process of electricity transmission in electric networks, losses should not exceed 4 % [1], however, the actual losses of main electric networks as a rule are not less than 10 %. Significant component of actual losses is the consumption of electric power for own needs by electrical substations [2,3].

Traditionally, an analysis of energy audit results is the work with a large amount of unstructured data [4] that is difficult to systematize and fully take into account. The paper objective is to solve the problem on the base of preference aggregation approach, which will enable efficient data compression preventing loss of useful information and provide clear visualization of energy audit results.

2. Preference aggregation

Let a set $\Lambda = \{\lambda_1, \lambda_2, ..., \lambda_m\}$ of m rankings of n objects of a set $A = \{a_1, a_2, ..., a_n\}$ be given. Each ranking is in the form of a chain and specifies a preference relation $\lambda_k = (a_1 \succ a_2 ... \sim a_s \sim a_t \succ ... \sim a_n)$ over the set A. The preference relation λ is a union of two relations: a strict preference relation ρ, i.e. $a_i \succ a_j$, and indifference (deemed as equivalence) relation τ, i.e. $a_i \sim a_j$, that is $\lambda = \rho \cup \tau$. We shall refer to the set of rankings Λ as the *preference profile* for given m and n.

To aggregate m preferences specified over a set of n objects means to determine a unique preference relation β called the *consensus ranking*, which provides the best compromise among the rankings of the initial profile. The treatment of the concept "best compromise" is defined by the *preference aggregation rule* used to find the consensus ranking.

In this work we will use the *Kemeny rule* [5,6] consisting in determination of such linear order (Kemeny ranking) β of objects that the distance $D(\beta, \Lambda)$ (defined in terms of the number of pairwise inconsistencies between the rankings) from β to the rankings of the initial profile Λ is minimal for all possible linear orders (permutations) of the objects.

The Kemeny rule assumes the existence of a nonunique consensus ranking: the number N of optimal solutions found by this method can exceed 10^7 even for small $m = 4$ and $n = 15$ [7]. To implement a convolution of the set of optimal solutions $\mathrm{B} = \{\beta_1, \beta_2, \ldots, \beta_N\}$ to a single final consensus ranking β_{fin} we will use the following condition: if a number of realations $a_i \succ a_j$ equals to a number of relations $a_i \prec a_j$ in all consensus rankings, then $a_i \sim a_j$ in the final consensus ranking β_{fin}; otherwise in the final consensus ranking β_{fin} there will be included one of relations $a_i \succ a_j$ and $a_i \prec a_j$, which is encountered more frequently in the optimal solutions.

To determine the Kemeny rankings we will use the recursive algorithm of our own design RECURSALL, implementing the branch-and-bound technique, allowing to find all possible Kemeny rankings for a given initial preference profile [8].

3. Decomposition of preference profile

When applying the Kemeny rule, it should be taken into account that the problem of finding the consensus ranking is *NP*-complete, i.e. having an exponential growth of the solution time as a function of the dimension $n = |A|$ of the problem [8]. Notice that, at problem dimension $n \le 20$ suitable for practical application, the RECURSALL algorithm allows to find all exact solutions within a reasonable time about several milliseconds. In situations where $n > 20$, one should resort to partitioning the set A into disjoint subsets A_i, i.e. $A = A_1 \cup A_2 \cup \ldots \cup A_k$, $\bigcap_{i=1}^{k} A_i \neq \varnothing$, where $|A_i| \le 20$, $i = 1, \ldots, k$.

The operation of partitioning the set A results in the decomposition of the preference profile Λ, thereby transforming its structure from linear to two-level one, and under multiple repetition of this operation the structure becomes hierarchical one. Consensus rankings β_i found over the set A_i will be included into the profile of higher hierarchy level, for which a consensus ranking can also be found. The process continues until the highest hierarchy level is reached.

4. Data analysis of real energy surveys

The Kemeny rule and the RECURSALL algorithm were used as the basis for the method for analyzing energy survey data of Backbone Electric Grids (BEGs) of Russia with help of preference aggregation technique. The Unified National Electric Grid of Russia includes eight BEGs, which in their turn consist of the Enterprises of Backbone Electric Grids (EBEGs), each of which unites a large number of *substations*.

The initial data for the method were contained in extensive tables of the values of the standard auxiliary expenses (SAE) of the substations [9]. The data were obtained during the energy survey of the BEGs of Russia and provided by the public company “Federal Grid Company of Unified Energy System”. The structure of the SAE contains 9 main components: heating of buildings (λ_1), lighting of buildings (λ_2), lighting of the territory (λ_3), cooling of the transformers (λ_4), heating of existing equipment (λ_5), charging devices (λ_6), communication equipment and telemechanics (λ_7), ventilation and air conditioning of buildings (λ_8) and other expenses (λ_9).

Since in many cases the number of substations n exceeded the upper permissible limit of 20 (see Section 3), we divided the substation sets into subsets (clusters) based on their close geographic location. This was justified by the need to analyze the consumption of resources by substations operating under similar climatic conditions [10].

The main stages of the method are as follows:

1) shaping the set substations $A = \{a_1, a_2, \ldots, a_n\}$, the SAE of which should be analyzed;
2) forming the preference profile $\Lambda = \{\lambda_1, \ldots, \lambda_m\}$, consisting of m rankings of n substations for each of the components (attribute) SAE; pair of substations a_i and a_j will be in a binary relation $a_i \succ a_j$ or $a_i \prec a_j$, or $a_i \sim a_j$ by the attribute λ_k, if the corresponding pairs of SAE values v_i and v_j consist in a binary relation $v_i < v_j$ or $v_i > v_j$, or $v_i = v_j$ by the same attribute λ_k respectively;
3) finding the consensus rankings $\mathrm{B} = \{\beta_1, \ldots, \beta_N\}$ for the profile Λ by the rule of Kemeny;
4) determination of the final consensus ranking β_{fin} by convolution condition (see Section 2).

In order to demonstrate an application of the proposed method to real data, let us consider a set of substations located in Arzamas area: Arzamasskaya (a_1), Bobylskaya (a_2), Lukyanovskaya (a_3) and Luch-500 (a_4). This set of substations belongs to the EBEG "Nizhegorodskoye" along with other areas located around the settlements Nizhny Novgorod, Poretskoe and Saransk. Near the cities of Arzamas and Saransk there are located four substations, near Nizhny Novgorod are 11 substations, near Poretskoe five substations. The values of the SAE for all attributes of four substations in the Arzamas area are reduced in Table 1.

Table 1. The SAE of four substations in the Arzamas area, 10^3 kWh.

Substation, $n = 4$	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9
a_1	449.28	9.73	95.92	1689.2	777.63	132.80	52.50	3.25	4.15
a_2	351.07	8.23	10.00	137.4	104.55	44.16	106.16	2.67	4.43
a_3	20.553	5.512	6.00	225.24	26.118	44.16	0	14.11	20.88
a_4	915.46	21.22	24.0	1159.8	370.8	132.8	52.5	65.32	329.78

Figure 1 shows the location of the set of substations $A = \{a_1, a_2, a_2, a_4\}$ in the Arzamas area, the preference profile $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_9\}$ constructed for the set A, and the resulting consensus ranking $\beta_{\mathrm{fin}} = \{a_3 \succ a_2 \succ a_1 \sim a_4\}$ found for the preference profile Λ using the RECURSALL algorithm. Figure 1 also provides a convenient visualization of the profile Λ, in which the position (rank) r_i of the element a_i in the ranking λ_j is represented by a saturation of green colour, where a less intense colour corresponds to the more preferable position of the element a_i.

The result of application of the preference aggregation method is the identification of two substations a_1 (Arzamasskaya) and a_4 (Luch-500) in the Arzamas area, for which it is necessary to undertake measures to reduce the SAE.

5. Conclusion

It is suggested in the paper a method based on the preference aggregation for the analysis and visualization of energy survey data of main electric grids. The method makes it possible to identify sources of economically inefficient expenditure of energy resources and unjustified energy losses, to provide compression of large volumes of energy survey data without loss of essential information. The proposed method can be a convenient and promising tool for organizations engaged in energy consulting.

Acknowledgment

This work was supported by the Russian Science Foundation, Project # 18-19-00203, in part of the method of preference aggregation development and by the Ministry of Education and Science of Russian Federation, basic part of the state task "Science", project 2.5760.2017/8.9 in part of real energy surveys data analysis.

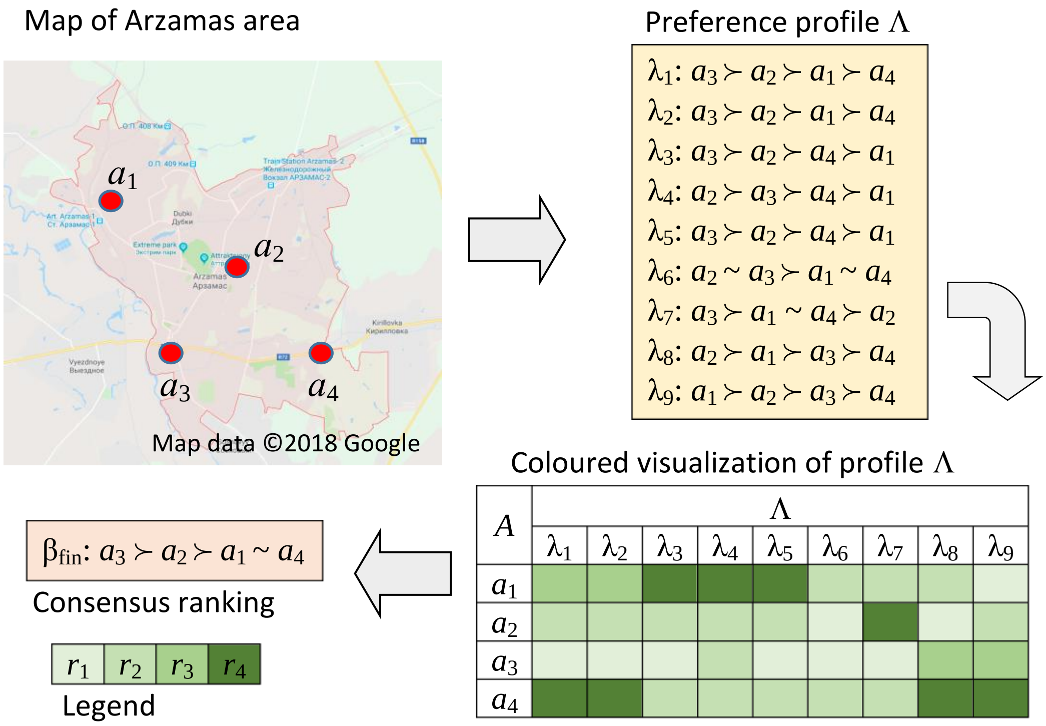

Figure 1. Example of set of substations $A = \{a_1, a_2, a_2, a_4\}$; corresponding preference profile $\Lambda = \{\lambda_1, \lambda_2, ..., \lambda_9\}$; and the final consensus ranking β_{fin} of the substations. Position r_i (rank) of element a_i in a ranking λ_j is shown by the bloom as in legend.

References

[1] International Energy Agency Energy Technology Perspectives <http://www.iea.org/media/ translations/ eer_ru.pdf > (last accessed 25.02.18)
[2] EU Energy Efficiency Directive (2012/27/EU) Guidebook for Strong Implementation, 2013, The Coalition for Energy Savings, Brussels, Belgium, <http://eedguidebook.energycoalition.eu/ images/PDF/EED.pdf> (last accessed 01.03.18)
[3] Eyupoglu C and Aydi M A 2015 *Procedia Soc. Behav. Sci.* **195** 1966-70
[4] Sommer K D, Sarge S, Spitzer P and Kessel R 2012 *Proc. XX IMEKO World Congress Metrology for Green Growth (Republic of Korea, Busan)* vol 1 (IMEKO) 1692-95
[5] Kemeny J G and Snell J L *Mathematical Models in the Social Sciences*, Ginn, New York, 1962
[6] Muravyov S V 2007 *Metrol. Meas. Syst.* **13** (1) 9-24
[7] Muravyov S V 2014 *Measurement* **51** 328-34
[8] Muravyov S V 2013 *Measurement* **46** (8) 2927-35
[9] Federal Grid Company of Unified Energy System <http://www.fsk-ees.ru/eng> (last accessed 05.02.18)
[10] Abeysinghe S, Wu J, Sooriyabandara M, Abeysekera M, Xu T and Wang Ch 2017 *Appl. Energy* **210** 1101-12