Paper accepted for presentation at 2009 IEEE Bucharest Power Tech Conference, June 28th - July 2nd, Bucharest, Romania

Clustering Techniques for Energy Losses Evaluation in Distribution Networks

G. Cartina, *Member, IEEE*, G. Grigoras, and E. C. Bobric, *Member, IEEE*

Abstract--The size and complexity of primary distribution networks make it impossible to carry out an exhaustive analysis of each individual feeder, so a sampling approach may be adopted to evaluate the energy losses for whole distribution networks. Clustering techniques were applied to several databases, obtaining particular characteristics of the feeders groups. For every group (cluster), the feeder that represents the best the group characteristics and the energy losses was identified.

***Index Terms*--clustering techniques, power losses, radial distribution networks, representative radial network.**

I. Introduction

THE evaluation and the minimization of power losses are important issues during planning and operation, with important technical and economical implications.

The determination of the energy losses depends of a number of parameters and variables that derive the design criteria and the operating conditions of the distribution networks. Thus, the distribution networks have a wide range of variables, such as network type (rural or urban), nominal voltage, length, installed transformation capacity, number of transformation points, the circuit type (overhead, underground, mixed), load being served, etc, [2], [5].

In this paper, we refer to the radial urban and rural distribution network.

Even if a radial network (RN) has more power losses compared with other, it does not imply that it is operating out of the normal condition. It may have more length, may be more loaded, may have more transformation points, etc, presenting different constructive or operative characteristics.

A policy for the reduction of losses can contain short- and long-term actions. Some short-term measures are the following:

- Identification of the weakest areas in distribution network and improving them;
- Reduction of the length of the distribution networks by relocation of distribution substation/installations of additional transformers;
- Installation of shunt capacitors for the improvement of the power factor, especially for urban distribution networks.

G. Cartina is with the Department of Power Systems, “Gh. Asachi” Technical University, Iasi, Romania (e-mail:gcartina@ee.tuaiasi.ro).

G. Grigoras is with the Department of Power Systems, “Gh. Asachi” Technical University, Iasi, Romania (e-mail:ggrigor@ee.tuaiasi.ro).

E.C. Bobric is with the Department of Electrical Engineering, “Stefan cel Mare” University, Suceava, Romania (e-mail:crengutab@ee.tuaiasi.ro).

978-1-4244-2235-7/09/$25.00 ©2009 IEEE

If the energy losses for a feeder of a distribution substation can be determined exactly, it is impossible to carry out an exhaustive analysis of each individual feeder within an electric company, with thousands of feeders, thousands of transformers, thousands of consumers, etc. In this case a sampling approach may be adopted to evaluate the energy losses for whole distribution network. Clustering techniques were applied to several databases, obtaining particular characteristics of the groups. For every group, the feeder that represents the best the group characteristics and the power energy was identified.

The obtained results demonstrate the ability of the proposed method to overcome difficult aspects encountered in process operation and planning problems of distribution networks.

II. Clustering analysis

Clustering analysis is the organization of a collection of objects (usually represented as a vector of measurements) into clusters (groups) based on similarity. It is a wonderful exploratory technique to help us understand the clumping structure of the data. There are two major methods of clustering: hierarchical clustering and k-means clustering, [2], [4]-[6].

Hierarchical clustering is subdivided into agglomerative methods, which proceed by series of fusions of n objects into groups, and divisive methods, which separate n objects successively into finer groupings. Agglomerative techniques are more commonly used. Hierarchical clustering may be represented by a two dimensional diagram, known as dendogram which illustrates the fusions or divisions made at each successive stage of analysis. An example of such a dendogram is given below in Fig. 1.

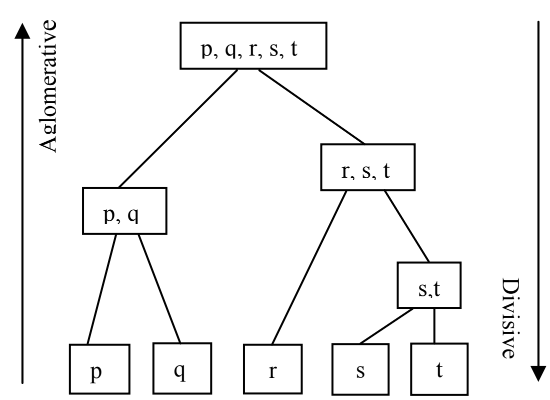

Fig. 1. Example of dendogram.

Hierarchical clustering is appropriate for small tables, up to several hundred rows. Differences between methods arise because of the different ways of defining distance (or similarity) between clusters.

III. Fuzzy Modeling of Load

Each study requires different types of consumption data. The values of nodal hourly power consumption are needed for load flow studies and the calculus of power/energy losses. The load distribution on a feeder at the time of peak load is obtained by applying appropriate ratios corresponding to measurement time of each transformer on feeder. This method assumes that the load profile on individual distribution transformers follows the same patterns as the load profile of the feeder, [8].

However, in distribution networks, typically, there are very few available real-time measurements of load values. Two primary variables can be considered for modeling the loads in distribution substations in this case: the loading factor *kI* (%) and power factor *cosφ*, so that the fuzzy representation of the active and reactive powers results from relations, [2], [5]:

$$P = \frac{kI}{100} \cdot S_n \cdot \cos\varphi ; \quad Q = P \cdot \tan\varphi \qquad (1)$$

where S_n is the nominal power of the distribution transformer from the distribution substations.

Thus, the hourly loading factor of a particular distribution transformer can be employed to approximate the nodal load. Also, because most utilities have no historical records of feeders, linguistic terms usually used by dispatchers are used, to describe the uncertain hourly loading factor. These linguistic terms are defined in function by the loading of the transformers at the peak load. Each loading level represented by a linguistic variable is described by a fuzzy variable and its trapezoidal membership function. In this paper, the loading factor *kI* and the power factor *cosφ* were divided into five linguistic categories (L.C.) with the trapezoidal membership function, Table I.

TABLE I
LINGUISTIC CATEGORIES OF THE KI AND COS φ

Linguistic Categories		x: kI (%)	x: cos φ	Linguistic Categories		x: kI (%)	x: cos φ
VS (Very Small)	x_1	10	0.75	M (Medium)	x_3	55	0.87
	x_2	10	0.77		x_4	65	0.89
	x_3	15	0.79	H (High)	x_1	55	0.87
	x_4	25	0.81		x_2	65	0.89
S (Small)	x_1	15	0.79		x_3	75	0.91
	x_2	25	0.81		x_4	85	0.93
	x_3	35	0.83	VH (Very High)	x_1	75	0.91
	x_4	45	0.85		x_2	85	0.93
M (Medium)	x_1	35	0.83		x_3	95	0.95
	x_2	45	0.85		x_4	95	0.97

IV. Clustering Techniques in Urban Distribution Networks

The primary characteristics data were updated and prepared for the grouping process, including the selection of the variables for a total of 44 urban feeders belonging to the electrical utility.

The feeders are divided in groups (8, in this case), based on the primary characteristics (the total length and the installed power of the transformers), using a statistical clustering method (centroid method), Fig. 3.

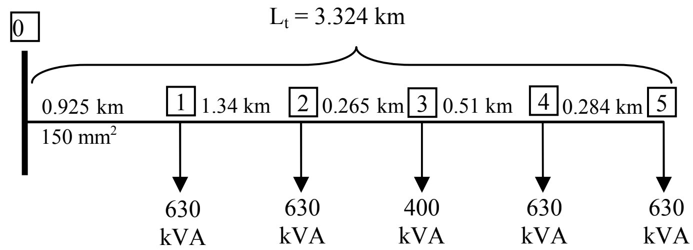

Fig. 2. Simplified representation of an urban feeder.

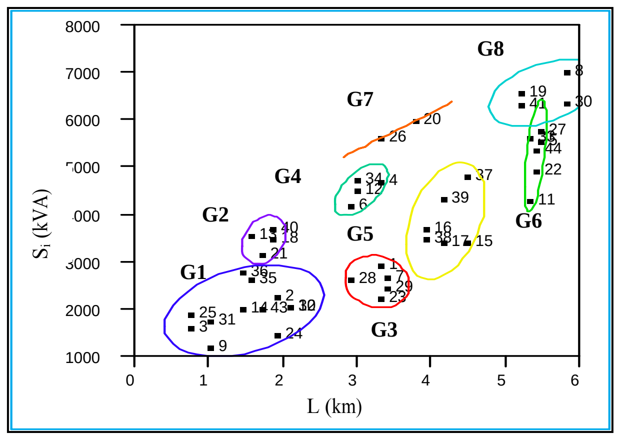

Fig. 3. Result of the grouping for the urban feeders.

The installed power of the feeder, in this case, is defined as the sum of the nominal power of the transformers served by the respective feeder:

$$S_i^f = \sum_{j=1}^{N_T} S_{nj}^f ; \qquad f = 1, \dots, N_f \qquad (2)$$

where: S_i^f – installed power of the feeder *f*, S_{nj}^f - nominal power of a transformer *j* served by the feeder *f*, N_T – total number of the transformers served by the feeder *d*, and N_f – total number of feeders in the database (N_f = 44, in this case).

The average values and the standard deviations corresponding to the total length (*L*) and installed power (S_i) for every group of feeders are calculated with the relations:

$$m_L^k = \frac{\sum_{f=1}^{N_f^k} L_f^k}{N_f^k}; \; m_{S_i}^k = \frac{\sum_{f=1}^{N_f^k} S_{if}^k}{N_f^k}; \quad k = 1, \dots, N_G \qquad (3)$$

$$\sigma_L^k = \sqrt{\frac{\sum_{f=1}^{N_f^k} (L_f^k - m_L^k)}{N_f^k}}; \quad \sigma_{S_i}^k = \sqrt{\frac{\sum_{f=1}^{N_f^k} (S_{if}^k - m_{S_i}^k)}{N_f^k}} \qquad (4)$$

where: L_f^k – the total length of the feeder *f* from the group *k*, S_{if}^k – the installed power of the feeder *f*, the group *k*, N_G – the total number of the feeder groups, and N_f^k – the total number of feeders from group *k*.

After calculating the statistical values (average, m and standard deviation, σ) for the primary characteristics of the feeder groups, the representative feeder (RF) for every group was defined as the real feeder with characteristics nearest to the average of the group. This representative feeder must characterize the best the peculiarities of the group, with respect to the primary characteristics (length and installed power, for urban feeders), and finally, the energy losses.

The average values (*m*) and the standard deviations (*σ*) corresponding to the total length (*L*) and the installed power (S_i) for the groups of urban feeders, and the characteristics of the representative feeders (L_{RF}, $S_{i\ RF}$) are presented in Table 2, for the considered urban distribution network.

TABLE II
THE CHARACTERISTICS OF THE GROUPS AND REPRESENTATIVE FEEDERS (RF)

Group	RF	L (km)		L_{RF} (km)	S_i (kVA)		$S_{i\ RF}$ (kVA)
		m_L	σ_L		m_{Si}	σ_{Si}	
G1	32	1.492	0.501	1.461	1975	452.33	2000
G2	13	1.765	0.139	1.585	3447.5	239.08	3550
G3	7	3.289	0.202	3.430	2590	259.33	2690
G4	12	3.077	0.169	3.022	4535	255.93	4520
G5	17	4.214	0.248	4.180	3870	575.95	3430
G6	22	5.427	0.069	5.380	5477.5	239.08	5431
G7	20	3.568	0.344	3.325	5805	275.77	5610
G8	19	5.537	0.343	5.24	6577.5	327.65	6580

To compare the RF with medium feeder of each group, from the viewpoint of power/energy losses, a fuzzy typical load profile (FTLP) has been considered, Fig. 4, [2].

The energy losses, as a function of the loading level, are calculated for every feeder of the groups. The average values of the energy losses were determined for every group, and every loading level. The crisp values of the energy losses corresponding to the distribution feeders (1, 7, 23, 28, and 29) from the *G3*, Fig. 3, considering linguistic load categories, at the peak load are presented in the Table III.

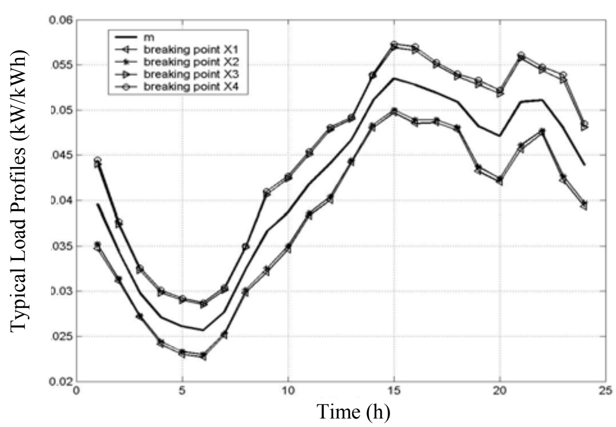

Fig. 4. Typical Load Profile versus its breaking points [2].

TABLE III
THE ENERGY LOSSES OF THE FEEDERS FROM G3 GROUP AS FUNCTION OF LOADING LEVELS, TABLE I

L.C.	ΔW_{G3} (kWh)						δΔW (%)
	1	**7(RF)**	23	28	29	**m_{G3}**	
VS	254.40	**240.48**	214.20	270.96	226.80	**241.37**	0.37
S	321.45	**299.88**	264.75	330.17	280.73	**299.40**	-0.16
M	421.37	**391.48**	342.73	422.06	363.86	**388.30**	-0.82
H	616.26	**570.45**	495.15	601.87	526.39	**562.02**	-1.50
VH	853.00	**787.35**	673.95	819.90	723.45	**771.53**	-2.05

The average of energy losses (for the group *G3*) and the representative feeder (*RF* = 7) are presented in the same table. The errors were calculated with the relation:

$$\delta\Delta W = (\Delta W_m - \Delta W_{RF})/\Delta W_m \qquad (5)$$

where: ΔW_G - the crisp value of the energy losses for the average of every group, and ΔW_{RF} - the crisp value of the energy losses for the representative feeder from every group.

V. CLUSTERING TECHNIQUES IN THE RURAL DISTRIBUTION NETWORKS

In rural distribution networks the same methodology can be used as in case of urban networks.

In this section only the clustering process is presented, which is different from the case of urban distribution networks.

For the rural distribution network the data were updated and prepared for the clustering process, including a total of the 43 rural distribution feeders of a distribution 20 kV network. The following data were considered: total length, axis length, and installed power.

The rural feeders are divided in 7 groups, Fig. 5, according to the primary characteristics (total length, axis length, and installed power), using a statistical clustering method (Ward method), [1].

The simplified representation of a rural distribution network is presented in Fig. 6, where the bold line marks the axis of the rural distribution network.

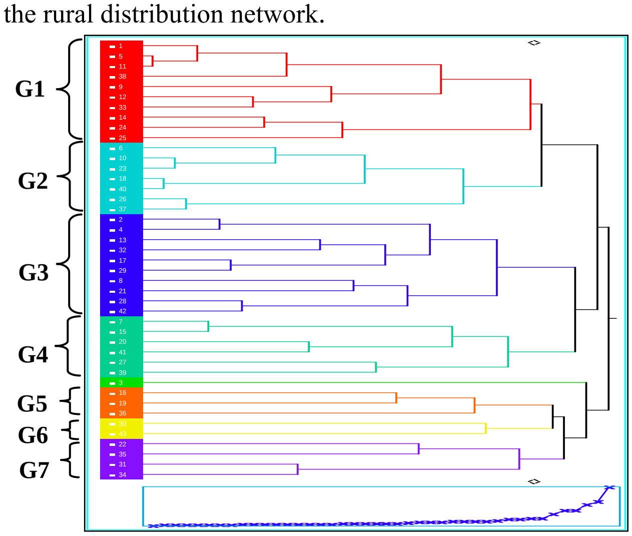

Fig. 5. The dendogram of the clustering process.

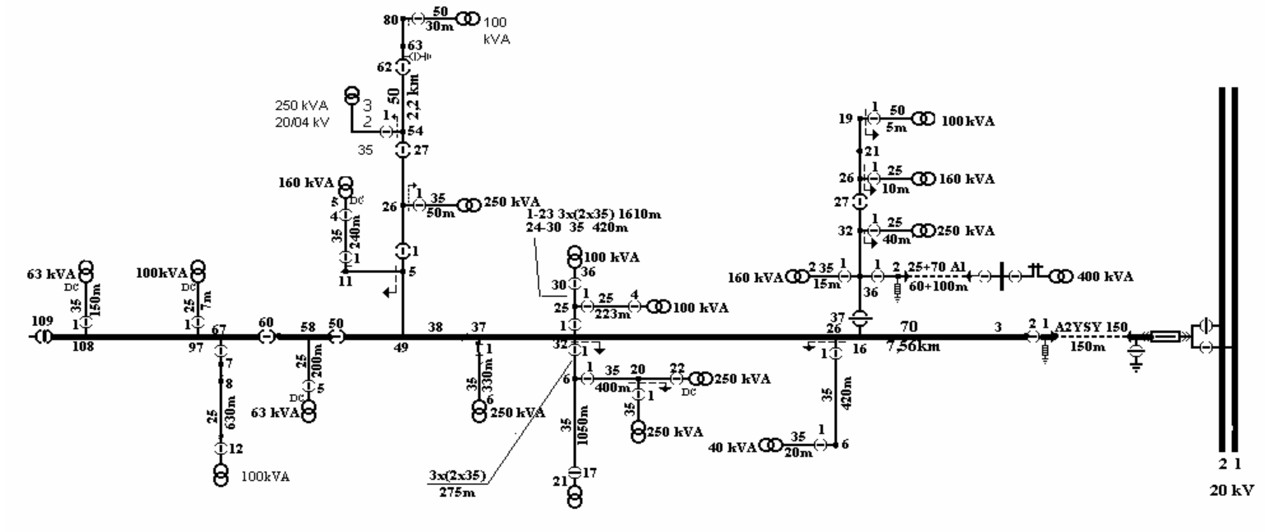

Fig. 6. The simplified representation of a rural distribution network.

TABLE IV
THE AVERAGE AND THE STANDARD DEVIATIONS FOR LT, LAX, AND SI

Group	No. NR	Lt (km)		Lax (km)		Si (kVA)	
		m_{Lt}	σ_{Lt}	m_{Lax}	m_{Lt}	σ_{Lt}	m_{Lax}
G1	10	6.92	4.19	4.50	2.28	2101.60	1663.95
G2	7	19.12	5.57	9.04	2.68	2405.86	639.81
G3	10	24.08	10.08	11.45	2.24	5377.60	1015.34
G4	6	17.23	6.86	7.80	3.04	10126.83	1891.82
G5	3	67.65	16.84	7.48	1.90	8572.00	1058.46
G6	2	85.48	25.62	20.70	7.78	11262.50	3483.03
G7	4	57.34	22.93	21.90	8.53	5023.50	3318.62

The average values and the standard deviations corresponding to the total length (L_t), axe length (L_{ax}) and the installed power (S_i) are indicated in Table IV.

The power losses of each feeder from the obtained groups were calculated, for several loading levels at the peak load, Table V.

TABLE V
LINGUISTIC CATEGORIES FOR THE POWER LOSSES OF THE GROUPS, (KW)

Group	Linguistic Categories				
	VS	S	M	H	VH
G1	12.77	14.23	17.02	20.67	29.66
G2	25.28	28.03	33.29	40.14	57.22
G3	46.20	53.01	65.96	83.00	125.35
G4	60.49	76.31	106.76	146.81	225.01
G5	107.26	122.87	153.08	193.07	295.72
G6	111.95	139.32	192.41	263.15	447.99
G7	73.86	86.99	111.89	144.96	233.14

VI. CONCLUSIONS

An exact calculation of the power/energy losses in the case of the distribution networks is impossible, however using clustering techniques and representative radial networks for every group can simplify enormously this task, maintaining a good precision.

The proposed methodology is proven as a useful tool for evaluation of power/energy losses and can be applied to different loading levels of distribution urban and rural networks.

VII. REFERENCES

[1] G. Cartina, Y.H. Song, G. Grigoras, *Optimal Operation and Planning of Power Systems*, Iasi: Casa de Editura VENUS, 2003.

[2] G. Cartina, G. Grigoras, E.C. Bobric, *Tehnici de clustering in modelarea fuzzy. Aplicatii in electroenergetica*, Iasi: Casa de Editura VENUS, 2005.

[3] G. Grigoras, G. Cartina, "Load Analysis in Distribution Systems using Power Fuzzy Correlation Models", in *Proc. of the 26th International Conference on Fundamentals of Electrotechnics and Circuit Theory*, Gliwice-Ustron, Poland, 2003.

[4] G. Grigoras, G. Cartina, "An improved method for power losses determination in rural distribution networks", in *Proc. of Electrical and Power Engineering International Conference*, Iasi, Romania, pp. 23-30, 2008.

[5] *JMP Statistics and Graphics Guide*, SAS Institute Inc., 1995.

[6] H. M. Khodr, J. Molea, I. Garcia, C. Hidalgo, P. C. Paiva, J. M. Yusta, A. J. Urdaneta, "Standard Levels of Energy Losses in Primary Distribution Circuits for SCADA Applications", *IEEE Trans. On Power Systems*, vol. 17, pp. 615 – 620, 2002.

[7] J. T. Saraiava, V. Miranda, L. Pinto, "Impact on some planning decisions from a fuzzy modeling of power systems", *IEEE Trans. on Power Systems*, vol. 9, pp. 819 – 825, 1994.

[8] S. Stalmans, "Loss Reduction in Distribution Networks", [Online]. Available: http://www.pbworld.com/news_events.

VIII. Biographies

Gheorghe Cartina (Member IEEE) received the M. SC. and Ph. D. degrees in Electrical Engineering from Technical University "Gh. Asachi" of Iasi, Romania, in 1964 and 1972 respectively. He is currently professor in the Department of Power Systems of Electrical Engineering Faculty at the same university. His research interests include especially to monitoring and optimal control of power systems.

Gheorghe Grigoras received the M. SC. and Ph. D. degrees in Electrical Engineering from Technical University "Gh. Asachi" of Iasi, Romania, in 2000 and 2005 respectively. He is currently lecturer in the Department of Power Systems of Electrical Engineering Faculty at the same university. His main areas of interest are analysis, planning, and optimization of power systems.

Elena-Crenguta Bobric (Member IEEE) received the M. SC. and Ph. D. degrees in Electrical Engineering from Technical University "Gh. Asachi" of Iasi, Romania, in 1995 and 2006 respectively. He is currently lecturer in the Electrical Engineering Faculty at the University "Stefan cel Mare" of Suceava. His research interests are the fuzzy technique applications in power systems.