A NEW APPROACH TO GENERALISATION OF SETTLEMENTS

Dr hab. Mieczysław Sirko
Department of Cartography
Marie Curie-Skłodowska University
Akademicka 19
20-033 Lublin, Poland

The purpose of this paper is to present, as objectively as possible, selection principles of settlements on small scale geographical maps of general and thematical character.

In objective selection of settlements on maps it must be assumed that studies should be based on a complete set of settlements, irrespective of their characteristics. It is also important that the applied research procedure allows us to select settlements on maps both „manually" and automatically. Automatization of the selection process aims at finding the most effective way to complete the undertaken work.

I have assumed that a settlement can be determined by a number of features. It is highly probable to obtain objective results when the features characterising the objects studied appear most distinctly. Realising the importance of this stage of work. I found it proper to take into consideration in my studies the following variables; the number of the settlement inhabitants, administrative significance, urbanisation level, economical role, accessible road and train communication, historical and touristic importance, location in relation to the river network. The selected features are divided into 54 classes, which allowed unequivocal determination of the characteristics of the settlements studied.

The area of studies (East-Central Poland) is characterised by various settlement forms and contains 3,988 settlements.

Having started to gather source - materials, I tried to select them so that they contained a full settlement image. This condition is satisfied by maps on the scale 1: 100 000.

The studied area was divided into 38 dcm^2 in order to take into account the features of the particular areas differing in their settlement character. This division was made on 1:300 000 map of Poland.

Of the 9 adopted variables characterising a settlement only two - the number of inhabitants and economical role - are of quantitative character. The other 7 concern qualitative features. This in turn necessitated the expression of the studied features by means of indices. Every feature was divided into several classes. Every successive class of settlements, according to the adopted variables, was by one order lower in relation to the settlements included into a higher class. It can thus be stated that the classification system of the variables reflects the settlement ranks, and that the distinction hierarchy of the classes has been preserved.

Measurements constitute the basis for obtaining values concerning the particular settlement sets, which correspond to definite classes of the objects studied. This operation comprises among other things numbering. If so, the next procedure can be done. Each class of the studied feature receives its serial number, namely, the lowest class is numbered one, and the highest class receives the number corresponding to the number of the selected settlement groups for the given variable. As a result, the number of the inhabitants of a settlement, divided into 10 classes, is numbered from 1 to 10, and the location of the settlements in relation to the river network from 1 to 5, in which 5 classes have been distinguished. Using such a procedure for determining index values, the particular classes of both mentioned features are not equivalent. As at this stage of work the real value of both variables cannot be determined it is advisable to compare the indices of the features adopted for studies. Therefore, normalisation should have been carried out, through dividing the numbers denoting the particular classes by the total number of the classes of the feature studied. This resulted in obtaining index values which were in the interval from one, for the most important class within the feature studied, to zero for the least important class. It can be expressed by the formula:

$$W = \frac{k-1}{n-1}$$

where „W" is the index for the specific class of the feature studied, „k" is the class No. of the feature concerned, „n" is the total number of the classes of the feature studied.

Next, each of the 3,988 settlements had to be included into a definite class in regard to each of the 9 variables (of economic, historical, touristic etc. importance), i.e. to determine the value of index „W" for the particular features. As a result a set of data characterising all settlements was obtained.

So far I have been concerned with determination of index values for the particular features, but I have not explained which of the variables adopted for studies are more and less important for correct selection of settlements on maps.

The purpose of this part of the paper is an attempt at presenting in an accurate mathematical form the rank of the variables adopted as selection criteria of settlements on maps. The point here is to show whether, e.g., the number of the inhabitants of a settlement is an important feature of the problem dealt with. If so/no, then what is its significance in relation to the other variables, e.g. that of historical importance. I have decided to solve this problem by factor analysis. Characterised by mathematical accuracy it will allow elimination of subjectivity in evaluation of each feature. Accordingly, the role of „W" indices is to be determined here.

There is rich literature concerning the methods of factor analysis, in which algorithms, their possible application as well as limits and conditions are discussed. Due to that a number of errors and mistakes, particularly in interpretation of the obtained results, can be avoided and the author is simultaneously dispensed from algorithm description. The final stage of factor analysis is calculation of the sum of factor loads which determine the importance (denoted further as R^2) of each of the variables in the factor space.

The calculated values (R^2) of the particular variables caused division of the features into two groups. One comprised variables being the most significant for selection of settlements to be put on the map. They were; economic ($R^2 = 0.97$) and administrative ($R^2 = 0.91$) significance of the settlement, urbanisation level ($R^2 = 0.90$) and number of inhabitants ($R^2 = 0.80$).

The other group comprised variables of a lower R^2 value concerning historical significance of the settlement ($R^2 = 0.60$), road communication access ($R^2 = 0.60$), touristic significance ($R^2 = 0.58$), railways access ($R^2 = 0.50$) and location in relation to the river network ($R^2 = 0.29$). The obtained results concernthe whole area studied which is characterised by a big differentiation of the settlement network. Therefore, the variable values (R^2) had to be calculated for the particular squares into which the studied area was divided.

Then 9 individual indices (W_1) were calculated for each feature of every settlement. For this purpose index „W" resulting from coding the particular classes of each feature had to be multiplied by the value R^2 corresponding to a given variable.

Thus

$$W_{1i} = W_{ji}\, R^2_i$$

where „W_{1i}" is the individual index (W_1) for the i-th variable, „W_{ji}" is the index „W" for the settlement „j" and variable „i", R^2_i is the rank of the i-th variable.

Since R^2 determines the rank of a given feature, the obtained value of the individual index

(W_1) also expresses its rank for a definite settlement in relation to this variable.

Knowing the values of all 9 individual indices characterising the given settlement in relation to all variables and then summing them up, I obtained a new summary index which expressed complexly the role of every settlement in relation to other settlements of the studied area.

It has been explained so far that the rank of every settlement can be determined quantitatively according to the summary index. However, the fact that a given settlement will be placed on the map being made does not depend only on its rank. One of the most important factors is the map scale. With its decrease the generalisation role increases. One of the possible solutions taking into account the scale on selection of settlements is application of F. Topfer's square root formula;

$$nb = na\sqrt{\frac{Ma}{Mb}}$$

where „nb" is the number of settlements on the map being made, „na" is the number of settlements on the source map, „Ma" is the scale denominator of the source map, „Mb" is the scale denominator of the map being made.

Calculation by using the above formula were performed separately for each of the squares into which the area was divided.

Selection of settlements is also determined by the distance between them on the map scale. Therefore, the criterion of minimal distance between the settlements should be taken into consideration. In this paper I have adopted 2 mm as the smallest admissible distance between settlements on the map scale. However, it should be stressed that it is not the matter of the visual aspect of a geographical map. The author does not discuss problems connected with features of the cartographical mark; its dimensions, shapes and the like. Settlement marks have been treated laboratorially as dot elements. Accordingly, only distribution of an abstracted picture not corresponding to real, graphical map can be considered.

The method used in this paper allows selection of settlements both on general and thematical geographical maps. Selection of settlements for general geographical maps depends on their rank. The settlement rank with regard to the features adopted for studies is determined by the value of the summary index. A bigger or smaller number of settlements selected for a map, however, depends on the results of calculations made by using the root square formula with consideration of the minimal distance coefficient.

A separate problem is selection of settlements for thematical maps. Variables and the calculated individual indices are very helpful in selection of settlements for preparing any thematical map. Depending on the map theme, every feature treated individually or in connection with others can be

the basis to determine which settlements should be placed on the thematical map being prepared. The theme and purpose of a map finally determine the main features. For example, for communication maps settlements being important road and railway line junctions should be taken into consideration, i.e. such whose values of the individual indices are the highest for the above mentioned features. Thus, in dependence of the map purpose, settlements can be selected for maps of various scales.

The presented selection procedure of settlements is relatively simple but it is time-consuming work. It is therefore advisable to automatize this process. Selection of settlements can be effectively accomplished by means of digital computers using analysis programme for summary index values, in which all settlements should be rejected in the selection process, which do not meet the condition of the adopted minimal distance and the value of their summary index is lower than that of the neighbouring settlements. The number of the settlements which should be introduced into the map being prepared is determined by the results of calculations made from the square root formula. In this way the needed set of settlements can be obtained in dependence of the scale of the map being prepared.

If we assume that 1:100 000 map fully illustrates the network of 3,988 settlements (100 %), 1:300 000 map will contain 2,294 (57,55 %) settlements, 1:500 000 one - 1,631 (40,90 %), 1:1 000 000 one - 1,012 (25,38 %), and 1:2 000 000 one - 603 (15,12 %). The obtained data served to obtain automatically a cartographic picture of settlements.

The purpose of this paper was not to form an instruction for making concrete maps but to work out methodological bases for settlement selection. In the light of the obtained results this goal seems to have been reached. The purposed method is as much objective as much the features adopted for studies allow a full characteristics of settlements and as much objectively the evaluation of the features with regard to their rank in selection of settlements was made.

Although the presented paper concerns generalisation of settlements on maps, its scope is much wider. It shows a new approach to the selection procedure of settlements on maps, which, after appropriate adaptation, might be successfully used for generalisation of other elements of the map contents.