DEVELOPING INTELLIGENT GEODEMOGRAPHIC TARGETING SYSTEMS

Stan Openshaw

WORKING PAPER 94/13

SCHOOL OF GEOGRAPHY • UNIVERSITY OF LEEDS

Why?

The market for geodemographics is currently believed to be worth around £44m per annum. Indeed there has been a resurgence of interest in geodemographics due to the 1991 census and the appearance of major new clients; in particular the energy utilities. However, it still appears that most of the current systems are not greatly different from what might be termed first generation geodemographic products of the late 1970s. They can be criticised as being essentially 1960s technology being run on 1990s hardware. The products themselves have improved with time, particularly notable is the much higher accuracy of the postcode-census linkage, vastly better graphics and mapping systems with some basic GIS functionality being added, the incorporation of non-census data to enhance the marketing relevant of the 1991 Census, some minor improvements in the classification methods used, and a much greater recognition of the need for customisation of segmentation systems. So whilst there is no doubt that some progress has been made with a gradual evolution of geodemographic technology, the basic view argument here is that in general geodemographics still remains remarkably "dumb" and fails to exploit the new opportunities for sophistication and improved performance that are available.

It is easy to abandon geodemographics on the grounds that lifestyle marketing systems based on individual data are so obviously superior. However, this is by no means proven. Individual level marketing systems have their own sets of problems that need to be resolved and still await the emergence of smarter and intelligent generations of database mining and targeting tools better able to handle the wealth of data on consumer behaviour and lifestyles than those that now exists (Openshaw, 1994). A further important consideration is that in the mid 1990s there is no longer any need to be exclusive, when it comes to the choice of targeting technologies. The sensible marketeer should not rely only on one approach or only one geodemographic system but seek to utilise the full range of opportunities that exist, allowing each application to define what is best for itself. There is no longer any need for prescription or blinkered vision in the emerging era of computational marketing; an era in which "smart" marketing machines could in theory search, evaluate, and then suggest whatever is found to be "best" and "safest" for particular needs in the context of a specific application (Openshaw, 1992, 1994).

Building marketing machines is not easy and at presently fully functional systems do not yet exist. Here a start is made by designing, building, and then demonstrating via case studies a purely Geodemographic Targeting Machine (a GTM/1). There is a major opportunity to take a quantum leap forward in geodemographics by becoming more intelligent and smarter in how we apply the technology. This is partly a matter of improving on what is already standard practice by both developing better algorithms based on artificial intelligence (AI) tools and making good use of compute intensive statistical methods, but it is also a matter of becoming smarter by explicitly seeking to include more of what we already know about the sociospatial structure of residential neighbourhoods into the geodemographic targeting process.

Intelligent Knowledge Based Systems (IKBS) was at one time synonymous only with expert systems, but there are many different paths to intelligent systems, including the incorporation of meta-knowledge in analysis systems and designing systems that behave in a smarter fashion. Currently we seem to use geodemographics in an uniquely dumb manner. For instance, improvements in the precision of a geodemographic system is widely seen as a function of one or more of the following

a. a better classification method,
b. a move to smaller postcode units from census enumeration districts,
c. the incorporation of multi-source non-census data, and
d. the use of GIS decision support tools.

However, whilst laudable, item (a) still ignores 50 years or more of accumulated knowledge about the structure and nature of residential patterns and processes. Somehow, better results are viewed as a purely an algorithmic matter. Item (b) may actually result in more precise geographic descriptions (due to smaller areas) but in worse targeting (neighbourhood effects are lost and smaller number problems become much more serious affecting data reliability). Item (c) may also cause many problems due to the mixing of data with different levels of error, spatial and temporal resolution, and uncertainty; for example, the disaggregation of census enumeration district data by some approximation to unit postcodes, or the aggregation of a non-census data source for 1994 or 1993 from postcodes to approximate census eds for use with census data for April 1991. Maybe these problems matter, maybe they do not. The problem is knowing when they do and when they don't, or else of attempting to handle the classification uncertainty that they cause. Item (d) is essentially a distraction, an often expensive visualisation and window dressing device. GIS and maps are good communication tools but they cannot discriminate between the good, the bad, and the rubbish; all may look equally appealing when mapped! GIS technology should instead be used to improve the spatial targeting process rather than decorate the results of procedures which are otherwise largely GIS free.

The paper seeks to build as an academic research project a prototype of an intelligent geodemographic target marketing machine. The challenge is to develop a much improved smarter and more intelligent geodemographic targeting system. There are a number of basic design criteria:

1) it should be better than any of the alternatives,
2) it should be tailored to a particular application,
3) it should incorporate knowledge where relevant,
4) it should handle rather than ignore the problems,
5) offers a means of achieving the best possible levels of performance within the limitations of the technology, and
6) be safe with self diagnosis of failure.

There are two key components of any geodemographic system: the classification and the geographic targeting. Improved versions of both are discussed in Sections 2 and 3. However, this is not by itself sufficient to create an intelligent geodemographic systems which needs a number of additional components to handle the various problems and provide some means of ensuring safe application. Section 4 describes the construction of a geodemographic targeting machine that meets the basic design criteria. Section 5 demonstrates its application on case studies and section 6 discusses further developments.

2 An intelligent census classifier

2.1 Census classification problems and uncertainties

So far all commercial geodemographic systems are partly, or largely, or wholly based on census data. The census classification process is now well understood with over two decades of experience (Openshaw, 1993). However, what is much less well appreciated is that the techniques that are used are fairly crude and fail to include any of the existing knowledge about the sociospatial structure of residential areas and lack a means of handling the known geographical problems associated with spatial classification.

Some of the problems are endemic and insoluble. For instance, the results of any classification exercise depend on a number of entirely subjective (albeit informed) decisions regarding the choice of variables, the number of clusters, selection of method, interpretation of results, and even the composition of data to be used. There is no easy way of optimising the whole process because there is no global objective function that could be used. Indeed, a seemingly good descriptive classification in terms of census variables may turn out to be much poorer in a geodemographic application than a much poorer census classification. Performance in a targeting context depends on a not yet understood relationship between the nature of the particular application and the characteristics of the classification being used. Maybe it does not matter, maybe it does; the problem is that the processes are not well understood and may never be. At one time, the end-user had no option but to have faith, now it should be possible to do more by recognising and then handling the various sources of uncertainty and problem that exist.

There are a number of major sources of uncertainty in any classification. They include:

1) precisely which cluster best represents an enumeration district or postcode,
2) what label best describes a cluster and to what extent its members match the cluster profile,
3) whether the cluster label provides a good description of the constituent households or just a small minority of them,
4) whether or not an enumeration district or postcode can reasonably belong to more than one cluster,
5) whether the classification adequately handles local patterns as well as offering a good global discrimination,
6) what is the best number of clusters to use,
7) whether a good description of census data has any relevant for marketing, and
8) how to incorporate in a consistent manner 'real' neighbourhood effects into the classification.

Many of these potential problems might only matter in terms of certain applications, but there is no way of knowing. Likewise, there need be no simple relationship between the importance of these problems and the quality of the classification. Tuning and spatial resolution may seem very important, but it may not be the weakest link. Likewise, improved classifiers might well be worthy of academic attention, but in a marketing context they may not matter much or at all. The secret is to become less naive and more aware of the problems and then develop a technology to handle rather than ignore them

2.2 Developing more intelligent spatial classifiers

In general there are three ways of becoming smarter in developing a spatial classification: improve the classification method, incorporate knowledge into the classification, and improve the application of the results in a targeting context.

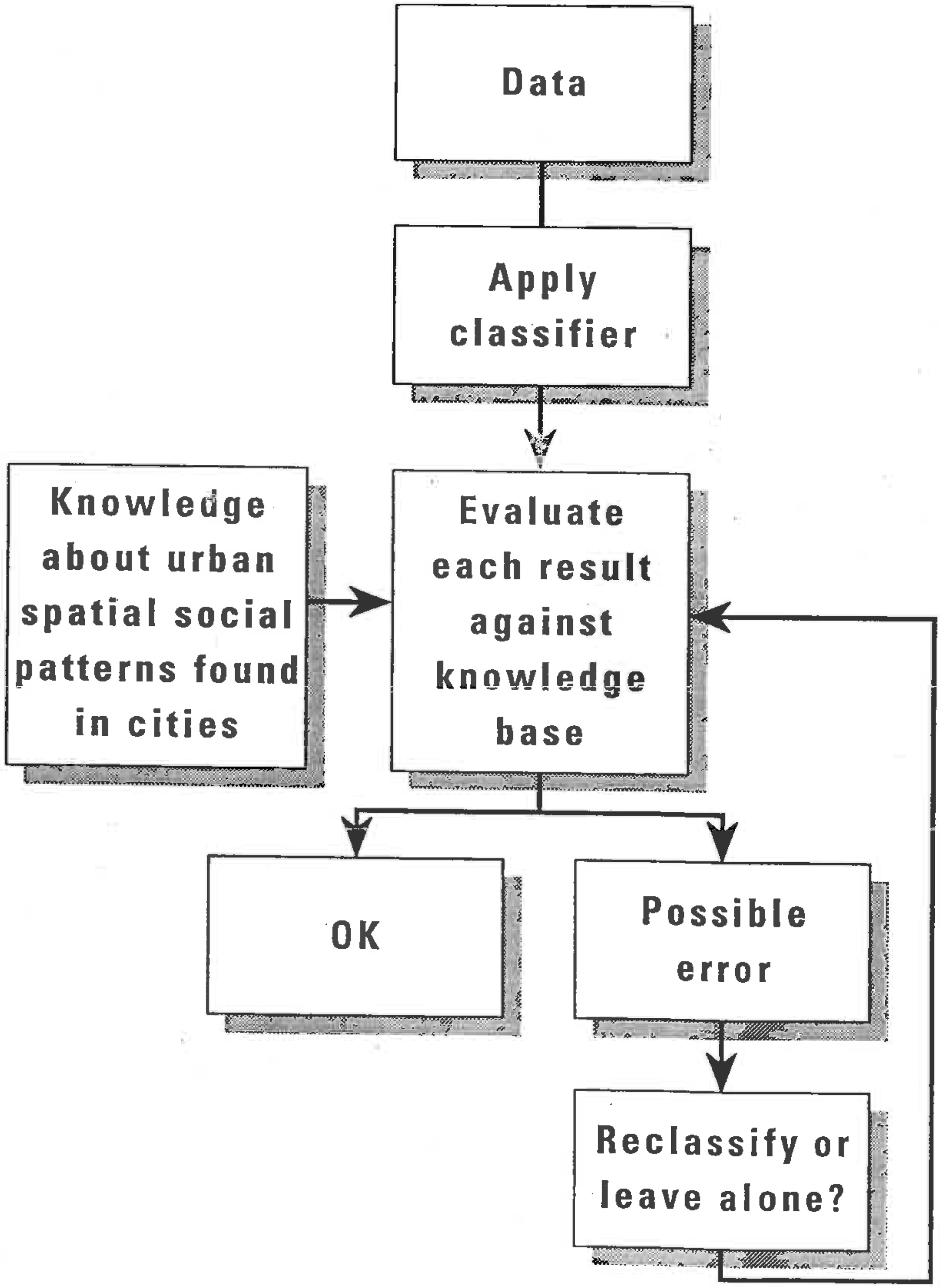
It is not clear at present which is likely to be the best strategy, although there does seem to be limits on how much improvement in a classification can be gained purely by algorithmic development. There is also the problem that a poor classification with for instance 2 × M clusters will almost certainly always outperform an optimal classification with only M cluster, although measuring and assessing performance is not easy. However, despite problems neglect is not the answer and it would seem that at the very least the classifications used should be based on good practice (Openshaw and Wymer, 1994).

Openshaw (1994) argues that one obvious way forward is to seek to become more intelligent and incorporate geographic knowledge directly into the classification algorithms. There are various ways of doing this. Perhaps the simplest is to exploit the large numbers of degrees of freedom that exist whenever a hundred and fifty thousand (or more) areas are to be classified into a 100 or less groups. There are many different good solutions and the opportunity must exist to bias the results to maximise sensibility in terms of external criteria. Figure 1 outlines one possibility, there are other options that could be investigated as a means of incorporating external knowledge into the classification process.

2.3 New methods of spatial classification

Geographers observed long ago that nearby areas tend to be similar, but are not always so. Why not incorporate this principle in the census classification process. Currently, census classifications have no notion of geographical structure (i.e. contiguity) in them and attempt to optimise a global function based on a common cluster morphology in high dimensional spaces based on minimum variance hyperspheres. This is somewhat rigid and artificial.

Figure 1: A Knowledge based Classification of Residential Areas



Openshaw (1994b) outlines a neurocomputing classifier adapted to handle census data that allows cluster structure to emerge in a bottom-up manner, rather than be imposed top-down. It also handles size variation in areas and gives most emphasis to areas for which the data are most reliable; most methods treat all areas equally and are prone to small number effects. This neural net approach has been extended to handle the notion that nearby areas may well be similar. It is considered important that a spatial classification should provide a good local description as well as attain a useful global level of discrimination. Unfortunately there is no notion of a good local description in a conventional approach. The Kohonen self-organising map is extended to handle geographic space. This can be achieved in a number of different ways: (1) update neurons within two different neighbourhoods simultaneously (the usual neuron 'distance' metric and in a k^{th} nearest neighbour geography metric), and (2) the use of multiple nearby areas such that the spatial averaging this entails diminishes with iteration time. This is the subject of an ongoing research project funded by the ESRC.

Experiments so far indicate that 'good' results can be obtained, albeit after a massive amount of compute time. For example, classifications based on both principles required 4 days each on a fast workstation. Clearly this is an area where parallel supercomputing is needed if much more progress is to be made and these new methods widely used. Also, there remains the difficult task of knowing how to measure 'goodness' in a classification that is relevant in a geodemographic context. It is quite apparent that minimising a statistical measure of classification quality may not mean much, but what does? Performance is application dependent.

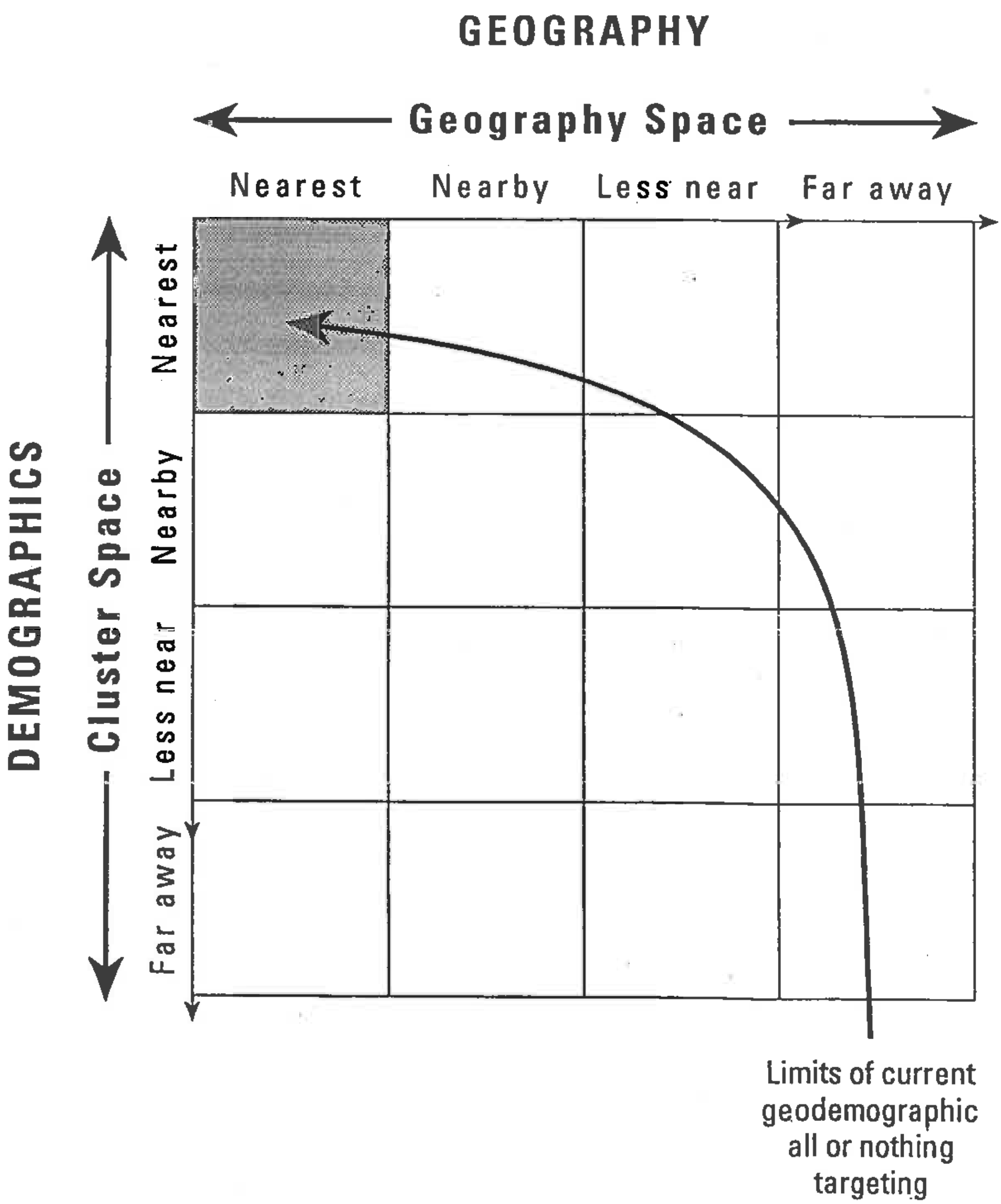
3 A fuzzy neighbourhood targeteer

3.1 Incorporating neighbourhood

Geodemographics rests heavily on what is essentially an ecological fallacy; namely, that people living in areas with similar features will (a) have these features, (b) behave in a similar way, and (c) behave in a similar way to others in the same cluster even if they have different characteristics. This can also be termed a spatial neighbourhood effect, whereby people behave in a similar manner to their "neighbours" even if they are different. The spatial extent of this effect is probably highly variable, context specific (e.g. voting behaviour is different from retail behaviour) and is also related to socioeconomics (Harrop, Health, Openshaw, 1991). In geodemographics, the spatial neighbourhood is the census enumeration district (ed) (or postcode) which vary greatly in size, shape, and social homogeneity. This aspect is usually overlooked and, indeed, the representation of neighbourhood effects in geodemographic systems is by no means consistent, yet it is an important part of geodemographic targeting. In fact it is worth emphasising all census enumeration districts (and also postcodes consisting of more than one household) are not homogenous. Most classification labels when applied to these small areas will be accurate often only for a minority (sometimes a very small minority) of the residents living there. This is a geographical fact of life. A classification of enumeration districts provides a representation of neighbourhood effects but only at that scale. Postcodes are more homogenous but maybe they are too small for real neighbourhood effects to operate. As a result an important feature of geodemographics is being very poorly handled in current systems. An intelligent system would seek to incorporate rather than ignore these effects.

Openshaw (1989, 1989b) argues that there are two major sources of uncertainty that need to be taken into account rather than ignored in geodemographics: see Figure 2. There is uncertainty as to which cluster an ed (or postcode) belongs too, some may well be near to several different clusters and the all or nothing nature of the classification hides this from view. This uncertainty also reflects the non-homogenous natures of the areas. However, there is also uncertainty in the geographic space. This reflects possible error in the postcode-ed linkage, but also uncertainty as to the scale of any application specific neighbourhood effects in a particular locality. These effects interact in that areas very near to other areas and belonging to different clusters may well either be mis-classified or characterised by more than cluster profile. It is argued that a fuzzy geodemographic targeting system would be worth considering if there is empirical evidence to support the existence of significant amounts of fuzzyness.

Figure 2: Fuzziness in Geographic & Cluster Spaces



3.2 Fuzzy geodemographics

It is of some interest, therefore, to note that Openshaw, Blake, and Wymer (1994) report the results of a fuzzy analysis of a 1991 census classification. Table 1 describes some of the findings. It is immediately apparent that quite large numbers of census eds often exist outside a specified target cluster but are nearby in both geographic and cluster spaces. For example, prosperous pensioners are quite strongly clustered geographically; most eds being within about 300m of each other (see first column in Table 1). However, there is almost as many within a similar geographic distance to those areas which are classified as something else but have similar characteristics (see second and third columns). These are, clearly, prime prospects that are missed by conventional geodemographics. It also makes good urban geographic sense as such areas can be seen in most British cities.

The principal problem with fuzzy geodemographics is explaining it, determining optimal estimates of the fuzzy parameters relating to both the geography and cluster spaces, and simply handling the much large amounts of data that are now needed. For example, a UK wide geodemographic system installed on a pc would require about 10MB of diskspace, a fuzzy geodemographic system based on eds would require about 100MB and 1000MB if based on postcodes. It also needs accurate grid-references for eds and postcodes and the latter may well have to wait for the Ordnance Survey's Address Point Product.

3.3 Fine tuning the fuzzy logic

The results in Table 1 can be taken further by the simple expedient of looking more closely at the composition of the eds (or postcodes) assigned to the various cells of the table. Table 1, column 1, shows the geographic distance distribution of nearest neighbours. This can be expanded to include a closer look at absolute distances from the centroid of the target clusters. It might be inferred that there is fuzzyness here also and that those eds (or postcodes) far away in either or both distance metrics might best be removed from the targeting. In this way, it is possible to become much more discriminating and thus more intelligent. Why use all the eds in a target cluster anyway since it is only reasonable that

Table 1 Fuzzyness in a geodemographic cluster

geographic distance (m)	Similarity space distance units		
	0	0.25	0.50
0	(2)	0	0
100	(349)	164	128
200	(701)	350	324
300	(319)	229	269
400	(116)	152	188
500	(64)	88	133
750	(96)	149	208

Notes:

(1) **cluster is prosperous pensioners in retirement areas with 1800 eds Figures in brackets is the nearest neighbour distance distributions for these eds**

(2) **similarly space units are measured in terms of distance to parent cluster**

some, perhaps many, should be excluded in particular areas for some applications. Details of the fuzzy tuning is a matter for a subsequent paper.

4 A Geodemographic Target Marketing Machine (GTM/1)

4.1 Basic design

Neither the development of a better census data classifier or a fuzzy targeteer will have much impact unless they can be built into an easy to use system that covers the entire geodemographic targeting process. The complexity and sophisticated analysis technology is hidden behind an interface to provide a system no more difficult to apply than a conventional product. Figure 3 outlines the general structure of a Geodemographic Targeting Machine that handles the entire process. Note that a number of previously undefined grey areas are now filled in. The GTM/1 involves four major steps:

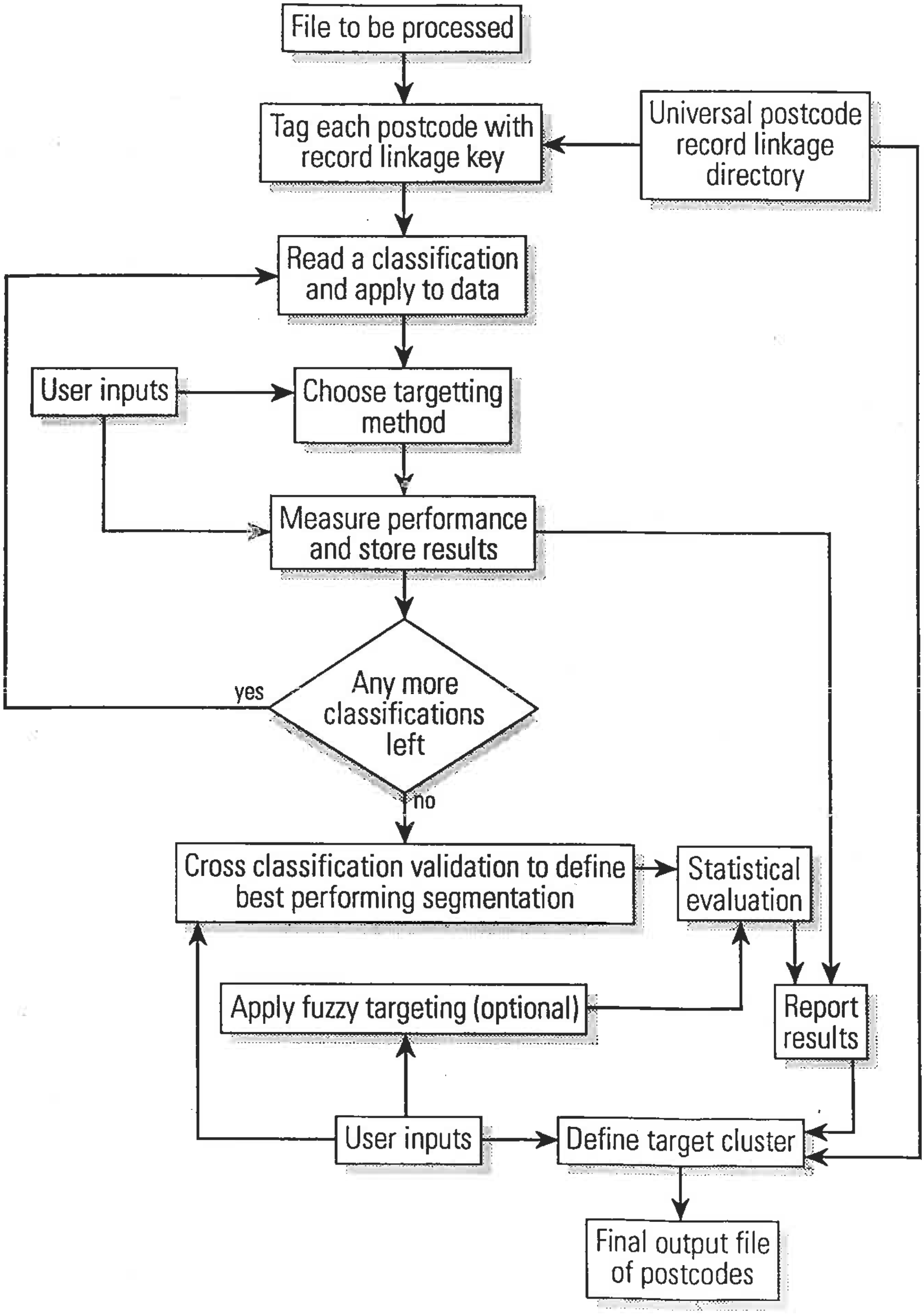
<u>step 1</u> Test the segmentation system to ensure that it is working reasonably well on the application. Failure or poor performance might be due to data processing errors in creating the data to be analysed, a response that is mainly random, or a segmentation system that fails to perform much better than a random classification at a similar level of resolution.

<u>step 2</u> Find the best number of clusters and the best segmentation system to use, excluding from consideration any poor performing clusters.

<u>step 3</u> Test the selected targeting to determine whether the observed results are better than you would find if the response was random.

<u>step 4</u> Investigate fine tuning the targeting by identifying and exploring the fuzzyness in the profiling provided by the selected system.

Figure 3 A Geodemographic Targeting Machine



4.2 Multiple Classification systems

The key idea is very simple. Instead of using only one geodemographic cluster system offering two or three levels of resolution, a database containing a large number of different classifications is examined. The logic here is as follows. Even if the world's best classification has been produced it still may not be optimal or best when applied to a particular dataset for a specific applications. Of course the results of applying intelligent census classifiers should provide substantially improved results but there can be no guarantee that this will always happen, performance may well be application dependent. The task of providing a first rate description of the structure of Britain's residential areas is quite different from that of providing an optimal segmentation of brown sauce consumption or of the clients of company 'X'. Inappropriate variables might have been used or there could be too few or even too many clusters. The task of producing a high quality classification with k clusters now needs to be repeated probably for a large number of values between 5 and at least 1,000. The falling cost of computing makes this quite feasible. Once the concept of a single global, all purpose classification is declared redundant then it becomes important to examine a range of many different classifications, each with a different number of clusters in it. It is probably also desirable to include in this process additional sets of classification systems corresponding to both different clustering methods (both dumb and intelligent) and different sets of variables.

The underlying assumption is that if performance depends on the resolution (i.e. number of clusters) the quality, and the basis of the classification used in some undefined and little understood fashion, then why not explore 50 or 100 different systems to find whatever is best in a given context. Once multiple systems become available so it becomes important to at least consider the prospect of achieving better segmentation purely by evaluating multiple different cluster solutions to permit each application to find whatever is best for it. This approach becomes economic once computers are fast enough to both create and then use multiple classification systems in a fully automated manner, and also to evaluate them in a fairly rigorous way.

4.3 Measuring performance and cross segmentation validation

Once there are multiple different classification systems to investigate so it becomes desirable to be able to find the best one to use. This is not easy because there is very little research that can be used to define clusters that provide a good and reliable profile of customers that is robust to data uncertainty, relatively free of bias, and thus offers a sound basis for a subsequent mailout. As the numbers of clusters increase the segmentation may well become more accurate but less reliable and robust. However, what constitutes an optimal number depends on the data being processed. For a very large file (say more than a million postcodes) then 500 or even 1000 cluster solutions might yield very good and reliable results; but for a very small file of several thousand postcodes then 10 or 20 clusters might be too many for reliable results. The key criteria of accuracy, robustness, bias, and uncertainty also become important as response rates fall. It is very easy for small number effects to dominate; for example, the apparent best performing clusters may be those with the smallest denominators and there comes a point when the best performers could just as well be zero performers; for example, if due to uncertainty their performance is not significantly different from zero or random. In practice this means that the improvement in performance gained by using cluster systems with increasing numbers of clusters could, at some quite unpredictable point, be totally offset by the increase in uncertainty due to small number effects. The problem is that there is no theoretical way of predicting this point, or of even knowing whether any fixed number of clusters is too few or too many for a particular application. The number of responders is clearly one critical factor but so too is their spread over the classification. Evaluation is, therefore, an extremely important topic in geodemographics that has been almost totally neglected for far too long. It needs to be an integral part of any GTM design.

The principal source of uncertainty of interest here is that resulting from the use of a particular client file representing as it does a particular instance of response to a mailout or promotion or offer which may well have not been random or, indeed, selected in any sampling related manner. It is only to be expected that a different database (i.e. the same promotion 3 months later or a different promotion) will give different results and there is probably nothing much that can be done about this data dependency. However, it is possible to assess the statistical accuracy of a particular segmentation on a specified dataset. There are four types of error and uncertainty that are applicable here: uncertainty due to the limited set of observations, systematic error due to classification defects, prediction error, and uncertainty about which classification to use.

Fortunately, there are various compute intensive statistical methods that can be used to estimate the effects of these problems and thus provide a practical means for handling them. The statistical bootstrap offers a simple but extremely effective means of measuring the accuracy of different segmentations of the same database, see Efron and Tibshirani (1993) for details of the method. For example, if you estimate that 50% of the clients in a database live in 10 different clusters, then bootstrap estimates can be made of the accuracy of this quantity. Likewise, you can use it to define confidence intervals for cluster response rates and then ignore those indistinguishable from random or user defined inadequate levels of performance. There are no analytical expressions you can use to make these estimates hence why the bootstrap is so useful. It is also non parametric. However, there are three difficulties: it is compute intensive requiring bootstrap samples of 100 or more and with 500,000 responders, this can involve examining 100 segmentations each of 500,000 cases; the results are specific to the data at hand; and it is necessary to devise a measure of geodemographic targeting performance that is relevant.

Various performance measures can be defined: for instance, define a maximum mailout size and then seek the best possible result in terms of response, the uncertainty mainly being restricted to the level of response that is likely to be achieved; identify best in a minim bias sense performing clusters; a combination of (2) and (1) which requires multiple bootstraps, and maximise a profit function with some estimate of its confidence intervals

This process has to be performed for each classification separately and then the results examined to identify the best. The definition of best might be variously: the largest unbiased value of a performance indicator (for example for a fixed mailout the level of response), that segmentation with the smallest standard deviation or that with the smallest probability of being a failure (i.e. less than a particular minimum value). There are a number of other possibilities and the final decisions need to incorporate the end-user's experience and requirements; maybe it should not be a purely automated decision although it could be.

4.4 Statistical evaluations to find the best segmentation

A final question concerns assessing whether or not the targeting is meaningful and worthwhile. This can be measured by measuring the probability that a given segmentation might be no better than could be achieved by chance alone. There are two ways of assessing this:

Test (1) comparing the observed result against a large sample of random segmentations of the same data, and

Test (2) comparing the observed result against what might be expected to occur if the response file was random.

Test (1) is really a measure of the performance of the segmentation. Poor performance would be due to too many clusters or too few, or because the data being segmented is random. Test (2) is really only a measure of whether the observed segmentation is different from random, or some other benchmark. In both cases a Monte Carlo significance test procedure can be used to evaluate the null hypothesis of randomness. This testing is only used as a guide and is a useful means of ensuring at least a degree of safety. It can be applied at both the moment a classification is input and also after the best has been chosen. However, randomness is a fairly limited benchmark and it may be more useful to use instead a model as a benchmark, or to assess whether a particular desired level of performance can be achieved.

A final consideration is how to find the best solution from those that pass all the preceding tests. Typically performance, as measured by a user selected option, does not dramatically vary once a certain plateau level of performance is reached. Some cluster systems will have dropped out; others will have been rejected because performance is too poor or too uncertain, or the confidence intervals are too large. However, this will usually still require that a final single solution is selected from several survivors. Occam's razor provides a useful rule of thumb principle here. In a geodemographic context it can be argued that the best solution might be that cluster system which is based on the fewest parent clusters but still yields a response rate (or some other performance measure) not significantly worse than many of the other candidate solutions with more clusters in them. The rationale is that parsimony is always worthwhile. Further, a cluster system with fewer clusters may be more robust in a mail out because it offers a slightly greater shotgun effect and also a better degree of insulation against the various sources of spatial data uncertainty. Statistical significance could also be based on a validation dataset set aside at the beginning of the profiling exercise. However, in both cases the test of significance depends on a bootstrapped statistic designed to assess whether the results are worse.

5 Case studies

5.1 Analysis of the response to a newspaper offer

A national newspaper advert resulted in a 14,000 response file. The aim now is to define a target market of 5 million people that can be used for a follow-up direct mail campaign. In this study the 1991 census population distribution is used to define the response rate, producing an average of 0.2%. The question now is how well can an intelligent geodemographic cope with this quest to find from the sparse response reasonably good performing areas.

The GTM/1.0 is set running with the constraint that any clusters with a response rate significantly poorer than average should be excluded. A selection of results are shown in Table 2. The run through 42 alternative segmentations systems offering varying levels of resolution took 11 hours on a workstation. The first aspect to note is that there are quite large numbers of deleted clusters where the response rate is worse than average or insignificantly different from average. Second, that response rates reach a series of plateaus but with the best results being produced by one of the smart classifications in which the classification was constrained to provide a local fit as well as a global one. It is noted that even with this low level of response all the results were significantly different from random, providing a comforting degree of re-assurance. The recommended classification according to the previous rule of thumb was the intelligent classification denoted by an asterisk in Table 2.

5.2 **Analysis of a client database**

In the second case study a client database of about 250,000 is analysed to identify an appropriate segmentation. The aim is to define those areas than contain 30% of the clients so that direct mail can be used to try and increase penetration of those areas where fairly large concentrations of clients exist.

The GTM/1 was constrained to exclude below average performing clusters. Some of key results are shown in Table 3, they took 14 hours on a workstation to produce. Again large numbers of clusters are deleted as having levels of response either significantly below average or indistinguishable from average. This is most useful because it provides a means of discounting unreliable clusters where data uncertainty effects are strong. It is useful also because it allows the data to define the level of detail in the segmentation that is most relevant and likely to yield robust and reliable results.

The second interesting discovery is that the level of performance continues to increase with the number of clusters, although it does slow down it has not yet levelled out. This would

Table 2 GTM/1 results for case study 1

Number of clusters	Deleted clusters	Mean response	Range of values Low	Upper
10	7	.084	.082	.086
20	12	.098	.096	.100
30	22	.107	.105	.100
40	29	.112	.110	.114
50	33	.130	.127	.132
60	38	.137	.134	.139
70	48	.122	.120	.125
80	54	.125	.123	.127
90	62	.136	.134	.139
100	71	.142	.140	.144
200	139	.138	.135	.141
300	213	.140	.137	.142
400	303	.141	.138	.143
500	371	.143	.141	.145
750	555	.146	.144	.149
900	703	.148	.146	.151
64[1]	47	.127	.125	.129
64*	44	.153	.150	.157
64	44	.140	.138	.143
64	43	.125	.123	.128
64[2]	44	.123	.121	.126

Notes

1 four sets of 64 cluster intelligent classifications

2 a self-organising neural net based system

Table 3 GTM/1 results for study 2

Number of clusters	Deleted clusters	Mean response	Range of values lower	upper
10	5	.677	.669	.684
20	12	.834	.821	.846
30*	21	.884	.874	.892
40	28	.838	.823	.853
50	35	.935	.925	.945
60	43	.931	.920	.942
70	48	.985	.975	.996
80	55	.999	.983	1.014
90	63	1.029	1.017	1.040
100	68	1.002	.994	1.011
200	146	1.022	1.012	1.034
300	219	1.134	1.154	1.117
400	299	1.156	1.173	1.141
500	379	1.152	1.165	1.141
750	581	1.232	1.252	1.211
999	684	1.261	1.281	1.241
64[1]	41	1.000	.996	1.015
64	42	1.000	.991	1.012
64	43	.985	.975	.996
64	42	.995	.986	1.005
64[2]	43	.972	.962	.982

suggest that where large amounts of data are involved, geodemographic systems with perhaps a few thousand clusters might be of interest now that the GTM can handle small number effects quite well. It is again noted that for all the classifications examined here, the results were significantly different from random and thus it can be concluded that the segmentation is working well. Finally, the intelligent spatial classifications on this occasion offered only small levels of improvement. The rule of thumb selection is for the 30 cluster solution, although virtually any could be used and the decision might well reflect other considerations such as size of desired target group.

5.3 **Analysis of a response**

The entire client database of 250,000 was mailed and a response of 25,000 achieved. The aim is to find a segmentation containing significantly above average response rates so that a second mailing can be more focused. Whilst it might appear that an average response of about 10% is extremely good, nevertheless, carefully targeted mailing to existing client databases can sometimes achieve 20 to 30 percent response rates. The purpose here is to identify a segmentation that will achieve a response rate of at least 15% (1.5 times greater than average) as the basis for a second mail out. A summary of the results is shown in Table 4. The run took 10 hours on a workstation.

Performance is now much more erratic. The segmentation systems with small numbers of clusters are totally deleted. Also there are now some indications that some of the results are becoming unreliable especially with the larger number of clusters; in particular the segmentations based on the 100, and 900 seem to be producing random results with many others starting to show slight signs of distress, see Table 5. The intelligent spatial classifications are quite safe on this occasion and yield good results and the preference is for one of these to be chosen.

Table 4 GTM/1 results for case study 3

Number of clusters	Deleted clusters	Mean response	Range of values low	upper
10	10	-	-	-
20	18	-	-	-
30	28	-	-	-
40	36	20.6	19.5	21.8
50	45	18.6	18.2	19.0
60	54	18.7	18.3	19.1
70	59	19.1	18.6	19.6
80	68	18.8	18.4	19.1
90	72	18.7	18.2	19.3
100	81	18.6	18.2	19.0
200	137	21.1	20.7	21.6
300	207	20.5	20.1	20.8
400	244	21.5	20.9	22.2
500	292	21.6	20.8	22.4
750	400	21.6	21.1	22.2
900	439	21.8	21.2	22.4
64[1]	58	18.8	18.4	19.1
64*	61	19.3	18.9	19.8
64	59	18.6	18.3	19.0
64	59	19.2	18.8	19.6
64[2]	60	19.0	18.6	19.3

Table 5 The probability that the target profile yields non-random results

Number of clusters	probability
10	-
20	-
30	-
40	.76
50	.42
60	1.0
70	.95
80	.57
90	.17
100	.04
200	.35
300	.21
400	.20
500	.10
750	.16
900	.03
64[1]	1.0
64	1.0
64	1.0
64	1.0
64[2]	1.0

6 Future developments

The current mark 1.0 system seems to work extremely well. The strategy of devising a more intelligent targeting process by a computational route seems to work, provided the computing costs are affordable. The intelligence in the system results both from its use of advanced classification and targeting tools and more especially from its ability to adapt to the task in hand. It becomes intelligent by evaluating more of the alternatives, incorporates knowledge about residential neighbourhood effects in the context of a particular application, and its use of statistical techniques to identify the best results. In theory this should lead to an unbeatable combination of geodemographic, statistical, and artificial intelligence technologies that is flexible enough to provide a framework that can expand as more advanced and powerful computer hardware becomes available. It also provides a degree of safety and some promise that the best attainable results will be achieved in the context of a specific application. There are also various ways that the technology can be extended and customised to suit particular needs.

In Figure 2, GTM version 1.0, is a sequential process. This has some attractions, principally it is easy to implement and it can be seen as an evolution of current practice that makes it more understandable. However, its sequential nature is also a restriction because there is no opportunity for the fuzzy analysis to drive the targeting. For instance, it is possible that the best results at the end of the process may be achieved via a relatively poor performing classification that would be filtered out by the sequential process. This totally fuzzy targetteer would be possible but would require about 100 times more compute power, so version 2.0 is perhaps best left for next year and parallel workstation platform, even though it is useful to note that the GTM concept itself is flexible and might be expected to improve by exploiting the new opportunities as they become available. It is a future proof, extendible technology. Meanwhile as an add-on that is compatible with any or all of the existing geodemographic systems already in the market place, the GTM/1 has the potential to breathe new life and performance into geodemographics.

Acknowledgements

Some of the research reported here was supported by an ESRC Research Grant.

References

Efron, B., Tibshirani, R J, 1993, An introduction to the Bootstrap Chapman and Hall, London

Harrop, M., Health, A., Openshaw, S., 1991, 'Does neighbourhood influence voting behaviour and why?' in I Crewe, P. Norris., D. Broughton (eds) British Elections and Parties Yearbook 1991 Harvester Wheatsheaf, London p103-120

Openshaw, S., 1989, 'Making geodemographics more sophisticated', J. of the Market Research Society 31 111-131

Openshaw, S., 1989B, 'Learning to live with errors in spatial databases', in M Goodchild and S Gopal (eds) The accuracy of spatial databases Taylor and Francis, London 264-276

Openshaw, S., 1992, 'A review of the opportunities and problems in applying neurocomputing methods to marketing applications' J. of Targeting, Measurement and Analysis for Marketing 1, 170-186

Openshaw, S., 1993, 'Special Classification', in B Leventhal, C Moy, J Griffin (eds) An introductory guide to the 1991 Census NTC Publications, Henley 69-82

Openshaw, S., 1994, 'Developing smart and intelligent target marketing systems', J. of Targeting, Measurement and Analysis for Marketing (forthcoming)

Openshaw, S, 1994, 'Neuroclassification of spatial data', in B C Hewitson, R G Crane (eds) Neural Nets: applications in geography Kluwer Academic Publishers, Boston 53-70

Openshaw, S., Wymer, C., 1994, 'Classification and regionalisation', in S Openshaw (ed) Census User's Handbook Longmans, London (forthcoming)

Openshaw, S., Blake, M., Wymer, C., 1994, 'Using neurocomputing methods to classify Britain's residential areas' Paper presented at GIS-RUK94, Leicester University; School of Geography, Working Paper (forthcoming)

Produced By
School of Geography
University of Leeds
Leeds LS2 9JT
From Whom Copies May Be Ordered

June 1994