Aggregation Techniques for Statistical Confidentiality

Josep Domingo-Ferrer[1] and Vicenç Torra[2]

[1] Dept. Enginyeria Informàtica i Matemàtiques,
Universitat Rovira i Virgili,
Autovia de Salou, s/n, E-43006 Tarragona (Catalunya, Spain),
e-mail:jdomingo@etse.urv.es, http://www.etse.urv.es/~jdomingo
[2] Institut d'Investigació en Intel.ligència Artificial - CSIC,
Campus UAB s/n, E-08193 Bellaterra (Catalunya, Spain)
e-mail: vtorra@iiia.csic.es, http://www.iiia.csic.es/~vtorra

Abstract. This chapter describes microaggregation, a technique for statistical confidentiality that uses aggregation operators. We describe the goals of statistical confidentiality and its application to continuous and categorical data. We show the application of the method to a small publicly available data set. The chapter finishes by reviewing some of the practical problems of the application of microaggregation to statistical disclosure control.

1 Introduction

National Statistical Offices disseminate information that allows researchers, media and general public to perform their own analyses and studies. This information is collected beforehand from individual and corporate respondents. Although released information should be as detailed as possible from the users' viewpoint, dissemination is in conflict with respondents' privacy. Disclosure risk is defined as the risk of re-identification of particular respondents. That is, if some sensitive and confidential data that have been released are subsequently linked to a particular respondent, confidentiality is lost.

Statistical offices must avoid disclosure to protect the confidentiality of the information provided by respondents. The two usual approaches to disclosure protection (*i.e.* re-identification avoidance) are either to partially suppress data (data suppression or sampling methods) or to perturb them before publication. In this way, the disclosure risk decreases. However, data have to maintain the so-called analytical validity [20], *i.e.* similar results should be obtained when the same statistical analysis is performed on disclosure-protected data and on the original confidential data. Therefore, as [14] point out, "Statistical offices are confronted with the problem of ensuring that the risk of a breach of confidentiality (a disclosure) is acceptably low, while at the same time preserving as much as possible the information content of the data to be released."

Statistical Disclosure Control (SDC) is the discipline that seeks to modify statistical data so that they can be published without giving away the identity of any respondent behind the data. SDC methods for microdata[1] (sets of records, each containing information about an individual respondent such as a person, household, ...) are usually known as *masking methods*. At present, there is a wide range of masking methods for microdata [6]. For example: additive noise, global recoding, record swapping, microaggregation, resampling, PRAM, etc). This chapter focuses on microaggregation, a perturbative method based on aggregation operators.

In addition to masking methods, other tools are required by Statistical Offices to ensure data protection. We underline the following:

- Measures to assess the level of information loss due to a masking method (information loss measures).
- Measures to assess the disclosure risk associated to the masked data (disclosure risk measures).

Information loss measures quantify the data suppression/perturbation introduced by an SDC method in the masked dataset. Disclosure risk measures evaluate the re-identification risk inherent to the masked dataset. Both measures are useful to make the decisions and trade-offs when releasing the data. For perturbative masking (which does not involve data suppression) information loss measures are based on the differences between some computations on the original data and on the masked data, *e.g.* differences between correlations, differences between contingency tables, distance between original and masked records. Disclosure risk measures for perturbative masking are based on re-identification algorithms.

Re-identification algorithms take as input an original microdata set and a masked version of it and try to link masked records with the original ones. For example, record linkage [15] is used to link records in separate files that relate to the same respondent. Re-identification algorithms in [12,20,15] are based on the presence of a set of common variables in both files. The more records are correctly linked, the more re-identification and the higher the disclosure risk. Algorithms have been designed so that not only exact matches are considered but also partial matches. This is so because, as pointed out in [20], "the normal situation in record linkage is that identifiers in pairs of records that are truly matches disagree by small or large amounts and that different combinations of the non-unique, error-filled identifiers need to be used in correctly matching different pairs of records". In particular, this is the case with masked data.

In this chapter, we consider the use of aggregation methods for statistical confidentiality. The structure of this chapter is as follows. In Section 2 the general approach to microaggregation is described. Section 3 discusses

[1] National Statistical Offices also release *tabular data* (tables containing aggregated data), but these will not be dealt with in this paper.

microaggregation for quantitative variables. Section 4 deals with microaggregation for qualitative variables. Section 5 contains the chapter conclusions, as well as some comments on open research issues.

2 Microaggregation

Microaggregation is a perturbative masking method and as such consists on the replacement of the variable values of each respondent by perturbed values. In the case of microaggregation, perturbed values correspond to the aggregation of the values of some similar respondents.

Microaggregation relies on a confidentiality rule that is very often applied by Statistical Offices: publication of a data vector is not allowed if it corresponds to a group of less than k respondents. To allow for publication, at least k respondents should share the same values. To satisfy this condition, groups of at least k similar respondents are formed and the individual value for each respondent is replaced by the average value of the group. According to this, microaggregation is defined by the following two steps:

1. Partition of the respondents into a set of disjoint groups so that the number of respondents in each group is at least k.
2. For each respondents, the value of each variable is replaced by the aggregation of the corresponding values of all respondents in the same group.

According to this, formalization of the method is based on the notion of "similarity" / "distance" and the use of an aggregation operator:

1. Let $X = \{x_1, x_2, \cdots, x_n\}$ be a set of given data for n respondents for which p variables are observed. Thus, x_i is a p-dimensional vector $x_i = (x_{i1}, \cdots, x_{ip})$. Let $d(x, y)$ be a p-dimensional distance. Then, if we denote a partition of X into g groups by $G = \{G_1, \cdots, G_g\}$, with $|G_i|$ being the cardinality of groups G_i and $\hat{x}_i$ being the average data vector of G_i, the optimal partition G for microaggregation is the one that minimizes:

$$\sum_{i=1}^{g} \sum_{x_j \in G_i} d(x_j, \hat{x}_i) \tag{1}$$

   while satisfying $|G_i| \geq k$ for all $G_i \in G$.
2. Given an optimal solution G, x_j is replaced by $\hat{x}_i$ if x_j belongs to group G_i ($\hat{x}_i$ is the average of the elements in G_i).

Note that usual clustering methods (e.g. k-means) do not directly apply to this problem because they do not consider group cardinality contraints.

Exactly solving the above minimization problem has been shown to be NP-hard when data are numerical and $d(\cdot, \cdot)$ is the squared Euclidean distance [13]. This is enough to understand why existing practical methods are of heuristic nature [5]. We distinguish below two types of heuristics, depending on whether they deal with univariate or multivariate data:

Univariate microaggregation: A single variable is considered in the masking process (i.e., $p = 1$) or, if several ones are considered, they are masked in successive and independent processes. Fixed-size microaggregation and variable-size microaggregation methods exist. The former yield partitions where all groups except perhaps one have the same size k. The remaining group can contain more than k elements. In variable-size microaggregation the only restriction is that all groups have at least k elements. In this case, several groups can exist with more than k elements. Figure 1 illustrates the advantages of variable-sized groups. If fixed-size microaggregation with $k = 3$ is used, we obtain a partition of the data into three groups, which looks rather unnatural for the data distribution given. On the other hand, if variable-sized groups are allowed then the five data on the left can be kept in a single group and the four data on the right in another group; such a variable-size grouping yields more homogeneous groups, which results in lower information loss. The following approaches have been considered for group building:

- Elements are sorted either in ascending or descending order. Groups are defined with k consecutive elements. This approach is followed in [2].
- Genetic algorithms [10,5]. The fitness function is proportional to Expression (1).
- Modified Ward's algorithm [10,5]. Ward's hierarchical agglomerative clustering has been modified to accomodate group cardinality constraints.

Multivariate microaggregation: In this case, several variables are considered and masked simultaneously (not necessarily all variables at a time, see Section 3). This is, a single partition is used to mask several variables. Two alternatives can be considered:

- *Projected data.* A new variable is defined as the projection of the multivariate data into a single axis. The optimal partition is computed on the new variable. Aggregated values are computed using that partition and the original values.
- *Unprojected data.* Clustering is performed directly on the multivariate data. Modified Ward's algorithm can again be used here.

3 The numerical case

When variables are numerical the usual distance $d(\cdot, \cdot)$ in Expression (1) is the squared Euclidean distance. Therefore, the function to minimize is the within-groups sum of squares and, thus, Expression (1) becomes:

$$\sum_{i=1}^{g} \sum_{x_j \in G_i} ||x_j - \hat{x}_i||^2 \tag{2}$$

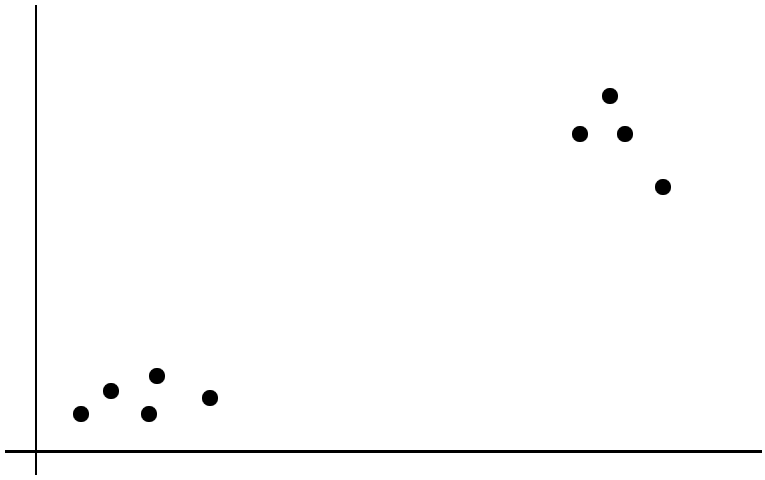

Fig. 1. Variable-sized groups versus fixed-sized groups

At present [10,5,2], the aggregation of the values of all respondents in the same group is computed using the arithmetic mean. There are several results worth mentioning when the arithmetic mean is used as $\hat{x}_i$ in Expression (2):

- Modified Ward's algorithm is attractive in this case because the underlying Ward hierarchical clustering chooses at each step two groups for joining which minimize the increase in the within-group sum of squares.
- The use of heuristics is theoretically justified, because it has been shown in [13] that minimizing Expression (2) with $\hat{x}_i$ being the arithmetical mean is NP-hard.
- It is reported in [7] that multivariate microaggregation on unprojected data taking two or three variables at a time (rather than all variables) is the microaggregation algorithm offering the best tradeoff between information loss and disclosure risk.

It has to be said that aggregation operators other than the arithmetical mean could also be applied, the most straightforward being the OWA operator [21,22]. This is so because this operator can model low influence of extreme values[2]. This is an attractive property in statistical confidentiality, where no individual value should dominate (*i.e.* contribute too much to) the group aggregated value published in the masked dataset.

[2] An OWA operator with a non-decreasing fuzzy quantifier Q such that its derivative is large for values near to 0.5 and small for values near 0 and 1. This is, a fuzzy quantifier corresponding to "about one half".

4 The categorical case

Two aspects have to be addressed in the categorical case: the definition of a distance and the definition of an aggregation procedure. The degree of difficulty of both definitions is similar.

In general, there exist two main approaches to deal with ordinal scales and define operators on these scales. On the one hand there are sets of methods that rely on an underlying semantics defined for each term (e.g. an interval or a fuzzy set). Operators are then defined according to this semantics. On the other hand there exist methods that try to circumvent the explicit definition of a semantics by directly operating over the terms. The latter approach can be implemented by first establishing the desired properties of the operator and then considering all possible operators that satisfy these properties (*e.g.*, t-conorms in [11]); another implementation possibility is first assuming a set of basic operators in the ordinal scale and then building complex operators from the basic ones. This second possibility is the one chosen in [3] to define the convex combination operator (assumption of a negation function on the ordinal scale) and in [8] to define aggregation (assumption of a dual pair of t-norm / t-conorm on the ordinal scale). However, the basic operator is sometimes associated with an implicit semantics. This is the case when negation is considered.

In order to define qualitative aggregation for statistical disclosure control, we consider two approaches: (i) define operations based on user-defined semantics and (ii) define operations based on semantics inferred from basic operators. In particular, we consider the negation function as the basic operator. As shown below, negation implicitly induces a semantics [16] from which aggregation operators and distances can be defined. We do not consider the definition of aggregation operators from properties because aggregation is not associative and the definition of a non-associative m-dimensional operator on a n-dimensional domain through tables is far from practical: n^m values are required (and consistency has to be checked).

We take negation as our basic operator instead of other alternatives (*e.g.*, t-norms in [8]) because the meaning of negation is clear for non-experienced users. As [4] and [18] point out, the negation of a term can be understood as its antonym. From a practical point of view, classical negation (an involutive operator over the domain) restricts the semantics too much because it implies equal informativeness for all terms. To overcome this difficulty, we use the negation introduced in [16]. Both negations are reviewed below.

Classical negation functions are defined over an ordinal scale $L = \{l_0, \cdots, l_n\}$ (with $<$ defined such that $l_0 < \cdots < l_n$) as functions from L to L that satisfy the following two properties [11,9]:

N1: if $l < l'$ then $N(l) > N(l')$ for all l,l' in L
N2: $N(N(l)) = l$ for all l in L

For a given L, conditions N1 and N2 determine the negation function N. This is formally stated by the following proposition proven in [1]:

Proposition 1. *For each set of ordered linguistic labels* $L = \{l_0, \cdots, l_n\}$ *there exists only one negation function that satisfies conditions N1 and N2. This negation is of the form:* $N(l_i) = l_{n-i}$ *for all* l_i *in* L.

This proposition implies that each term in the pair $< l_i, l_{n-i} >$ carries the same information or, in other words, that l_i and l_{n-i} are equally informative. Moreover, this proposition shows that as soon as conditions N1 and N2 are fulfilled, implicit assumptions are made for the underlying semantics of the labels.

However, equal informativeness is not adequate in some situations because the *amount* of information should be different for different labels. To overcome equal informativeness, [16] introduced a new negation function over linguistic labels. The new negation is a function from L to parts of L (*i.e.*, $\wp(L)$). [16] also introduced a method to induce a semantics from that negation. Here semantics is understood as a mapping from the set of labels L into intervals in $[0, 1]$.

Definition 1. [16] A function N from L to $\wp(L)$ is a negation function if it satisfies:

C0: $N(l)$ is a non-empty interval in L
C1: if $l < l'$ then $\max N(l) \geq \min N(l')$ for all l,l' in L
C2: if $l \in N(l')$, then $l' \in N(l)$

In Definition 1, C1 and C2 are generalizations, respectively, of N1 and N2, and C0 is a technical condition.

Note that the negation defined here is, following [4], an inner operator. An alternative definition would be to consider a membership function for each l_i and the negation of each membership function. This would be an external operator following [4]. Both definitions of negation correspond to the idea of antonyms.

A semantics is a mapping of each label into subintervals of the unit interval. Thus, assuming that there exists a negation function N_I in $[0, 1]$ that is the counterpart of the function N in L, it is possible to define a semantics consistent with N_I. In other words, for each label l_i an interval $I(l_i) \subseteq [0, 1]$ is defined, where the set of intervals constitute a partition of $[0, 1]$ consistent with N_I. Consistence is defined as follows (see Definition 3.8 in [16]):

- The negation of all the elements in the interval of label l_i is in the intervals of the negation of l_i. This is,

$$\text{for all } l_i \in L, \text{for all } x \in I(l_i), N_I(x) \in I^*(N(l_i))$$

  where $I^*(\Lambda)$ is defined as $I^*(\Lambda) = \cup_{l \in \Lambda} I(l)$.

- No label in $N(l_i)$ is "superfluous". This condition needs only to be formulated for the extreme labels in $N(l_i)$: if $N(l_i) = \{l_{i_0}, \cdots, l_{i_n}\}$, then:
  1. there exists a value $x \in I(l_i)$ such that $N_I(x) \in I_R(l_{i_0})$
  2. there exists a value $x \in I(l_i)$ such that $N_I(x) \in I_L(l_{i_n})$

  where I_R and I_L mean intervals open in the upper limit and the lower limit, respectively.

An example of semantics that is consistent with the usual negation in the unit interval $N_I(x) = 1 - x$ is the following one given in [16]:

Proposition 2. *Let N be a negation function from L to $\wp(L)$ according to Definition 1. Then, the following expression defines a semantics for L into $[0,1]$ consistent with $N_I(x) = 1 - x$.*

$$I(l_i) = \Big[\frac{\sum_{l<l_i} |N(l)|}{\sum_{l \in L} |N(l)|}, \frac{\sum_{l \leq l_i} |N(l)|}{\sum_{l \in L} |N(l)|}\Big]$$

where $|\cdot|$ is the cardinality operator.

When the negation function N is such that $|N(l)| = 1$ for all $l \in L$, the intervals reduce to the one with maximal neutrality (this is, maximal similarity on their imprecision). In other words, all intervals have the same measure.

Using the mapping of Proposition 2 it is possible to define a distance and an aggregation operator over L by considering the intervals induced by N. An example is given after the next definition.

Definition 2. Let L be an ordinal scale and N be its negation function. Then the distance $d : L \times L \to [0,1]$ is defined as

$$d(l_i, l_j) = (center(I(l_i)) - center(I(l_j)))^2$$

where $center(\cdot)$ is the center of the interval (if $I(l_i) = [min, max]$ we define $center(I(l_i)) = (min + max)/2$). Similarly, an m-dimensional aggregation operator $\mathbb{C} : L^m \to L$ is defined as

$$\mathbb{C}(a_1, \cdots, a_m) = I^{-1}((1/m) \sum_{i=1}^{m} center(I(a_i)))$$

where $I^{-1}(x) = l$ if $x \in I(l)$.

4.1 Example

This approach has been applied to mask 21 records of the *American Housing Survey 1993* (data publicly available from the U. S. Bureau of the Census

through the *Data Extraction System* [19]). Two variables have been considered: *BUILT* (year the household structure was built) and *DEGREE* (long term average heating degree days). Both are defined on an ordinal scale.

Variable *BUILT* has an explicit semantics: its range is $(01-09, 80-93)$ where 09 corresponds to years earlier than or equal to 1919, 08 to the interval $(1920-1929)$, 07 to the interval $(1930-1939)$, and in a similar way for categories $06, 05, 04, 03$ and 02; 01 corresponds to year 1979, and categories from $80-93$ correspond to years $1980-1993$. Distance and aggregation are computed for this variable according to these intervals and using expressions in Definition 2.

The range of variable *DEGREE* is {*coldest, cold, cool, mild, mixed, hot*}. In this case, distance and aggregation have been defined using negation functions. We have used the one given in Table 1. Note that the term *mild* overlaps with *hot* and *mixed* as its negation is defined as {*cool, cold, coldest*}. This term is not considered when building the intervals. Using Proposition 2, the negation in Table 1 induces the intervals given in Table 2 (a graphical interpretation is given in Figure 2).

The original records to be masked and the masked records are given in Table 3 (left and right columns, respectively). Computing the information loss of this method for the records in Table 3 with a distance-based information loss (based on the distance between the old and the new record [17]) yields an average loss of 0.144 (the average is computed over the set of records).

$N(coldest) = \{hot\}$	$N(mixed) = \{cool, cold\}$
$N(cold) = \{hot, mixed\}$	$N(hot) = \{cold, coldest\}$
$N(cool) = \{mixed\}$	$N(mild) = \{cool, cold, coldest\}$

Table 1. Negation function for variable *DEGREE*

$I(coldest) = [7/8, 8/8]$	$I(mixed) = [2/8, 4/8]$
$I(cold) = [5/8, 7/8]$	$I(hot) = [0/8, 2/8]$
$I(cool) = [4,8, 5/8]$	$I(mild) = [0/8, 4/8]$

Table 2. Induced intervals for variable *DEGREE*

5 Conclusions and open research issues

This chapter has described microaggregation and its application to disclosure protection of continuous and categorical variables. While quantitative

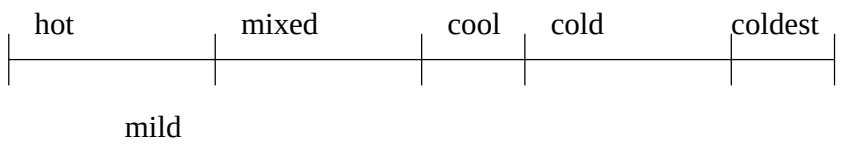

Fig. 2. Intervals induced for variable *DEGREE*

Built	Degree	Built'	Degree'
80	mild	81	mixed
81	cool	81	mixed
81	cool	81	mixed
84	mild	84	cool
84	cold	84	cool
84	cold	84	cool
84	cool	85	cool
85	mixed	85	cool
85	cool	85	cool
85	cool	86	mixed
86	cool	86	mixed
86	mild	86	mixed
86	mild	86	mixed
86	mild	86	mixed
87	cool	86	mixed
87	coldest	88	cold
88	cool	88	cold
89	cool	88	cold
92	cold	92	cold
92	cold	92	cold
93	cold	92	cold

Table 3. Original variables BUILT and DEGREE and their masked versions. Groups displayed in this table have been built with $k = 3$.

microaggregation has been a research topic for about ten years, qualitative microaggregation is quite new. Due to this, the research problems in both areas are quite different.

For continuous variables, the squared Euclidean distance is the natural similarity measure to use and the arithmetic mean is the most straightforward aggregation operator. With these choices, it has very recently been shown that exact optimal microaggregation of continuous variables is NP-hard. Practical microaggregation methods such as those described in [5] are heuristic. An open research issue is to design heuristics which guarantee that the information loss they cause is at most $p\%$ percent above the minimal information loss, where p is an input parameter. A second open research issue is to find the microaggregation heuristic that achieves an optimal tradeoff

between information loss and disclosure risk. Indeed, in [7] it is shown that multivariate microaggregation on unprojected data taking only two or three variables at a time is a good option; it remains to be seen which is the best option. A third open issue is whether aggregation operators alternative to the arithmetic mean (*e.g.* OWA) could improve the aforementioned tradeoff and/or reduce the complexity of exact optimal microaggregation.

In ordinal scales, research should be focused on the definition of methods for microaggregation. However, not all existing aggregation operators are suitable. Note that operations for statistical confidentiality should have a clear semantics and be defined in an easy way (the smaller the number of parameters, the best), because their potential users are seldom experts in soft computing methods. Moreover, operators have to be tested to determine their information loss and the disclosure risk when data is released. It is clear that methods with high information loss and disclosure risk are of little interest.

In this work, we have proposed the use of negation functions to define a semantics for ordinal scales when such a semantics does not exist beforehand. Further work is needed to exhaustive analyze its implications for information loss and disclosure risk of the resulting methods.

Acknowledgments

Partial support of the European Commission under project IST-2000-25069 "CASC" and of the U.S. Census Bureau under contracts no. OBLIG-2000-29158-0-0 and OBLIG-2000-29144-0-0 is gratefully acknowledged.

References

1. Agustí, J., Esteve, F., Garcia, P., Godo, L., Sierra., C., (1991), Combining multiple-valued logics in modular expert systems, in *Proc. 7th Conference on Uncertainty in AI*, Los Angeles, July.
2. Defays, D., Nanopoulos, P., (1993), Panels of enterprises and confidentiality: the small aggregates method, in *Proc. of the 1992 Symposium on Design and Analysis of Longitudinal Surveys*, Ottawa: Statistics Canada, 195-204.
3. Delgado, M., Verdegay, J. L., Vila, M. A., (1993), On aggregation operations of linguistic labels, *Int. J. of Intel. Syst.*, 8:351-370.
4. De Soto, A. R., Trillas, E., (1999), On antonym and negate in fuzzy logic, *Int. J. of Intel. Syst.*, 14:295-303.
5. Domingo-Ferrer, J., Mateo-Sanz, J. M., (2000), Practical data-oriented microaggregation for statistical disclosure control, *IEEE T. on Knowledge and Data Engineering* (to appear).
6. Domingo-Ferrer, J., Torra, V., (2000), Disclosure methods and information loss for microdata, in *Confidentiality, Disclosure and Data Access: Theory and Practical Applications for Statistical Agencies*, North-Holland, forthcoming.

7. Domingo-Ferrer, J., Mateo-Sanz, J. M.,(2001), An empirical comparison of SDC methods for continuous microdata in terms of information loss and disclosure risk, in *Second Eurostat-UN/ECE Joint Work Session on Statistical Data Confidentiality*, Skopje, Macedonia, March.
8. Godo, L., Torra, V., (2000), On aggregation operators for ordinal qualitative information, *IEEE T. on Fuzzy Systems*, 8:143-154.
9. Herrera, F., Herrera-Viedma, E., Verdegay, J. L., (1995), A sequential selection process in group decision making with a linguistic assessment approach, *Int. J. Information Science*, 80:223-239.
10. Mateo-Sanz, J. M., Domingo-Ferrer, J., (1998), A comparative study of microaggregation methods, *Qüestiió*, 22:511-526.
11. Mayor, G., Torrens, J., (1993), On a class of operators for expert systems, *Int. J. of Intel. Syst.*, 8:771-778.
12. Newcombe, H. B., Kennedy, J. M., Axford, S. J., James, A. P., (1959), Automatic linkage of vital records, *Science*, 130:954-959.
13. Oganian, A., Domingo-Ferrer, J., (2001), On the complexity of microaggregation, in *Second Eurostat-UN/ECE Joint Work Session on Statistical Data Confidentiality*, Skopje, Macedonia, March.
14. Pannekoek, J., Willenborg, L., (1999), Preface to the special issue on statistical disclosure control, *Netherlands Official Statistics*, 14:4-5.
15. Robinson-Cox, J. F., (1998), A record-linkage approach to imputation of missing data: analyzing tag retention in a tag-recapture experiment, *J. of Agricultural, Biological and Environmental Statistics*, 3:48-61.
16. Torra, V., (1996), Negation functions based semantics for ordered linguistic labels, *Int. J. of Intel. Syst.*, 11:975-988.
17. Torra, V., (2000), *On information loss measures for categorical variables*, Deliverable 3 of contract no. OBLIG-2000-29144-0-0 "Optimizing the Tradeoff Between Information Loss and Disclosure Risk for Categorical Microdata", submitted to U. S. Bureau of the Census.
18. Torra, V., (2001), Aggregation of Linguistic Labels when Semantics is Based on Antonyms, , *Int. J. of Intel. Syst.*, (in press).
19. U. S. Bureau of the Census (2000), *The Data Extraction System*, http://www.census.gov/DES/www/welcome.html
20. Winkler, W.E., (1995), Matching and record linkage, in *Business Survey Methods*, New York: Wiley, 355-384.
21. Yager, R. R., (1988), On ordered weighted averaging aggregation operators in multi-criteria decision making, *IEEE Trans. on SMC*, 18:183-190.
22. Yager, R. R., (1996), Quantifier guided aggregation using OWA operators, *Int. J. of Intel. Syst.*, 11:49-73.