საქართველოს მეცნიერებათა ეროვნული აკადემიის მოამბე, ტ. 3, №2, 2009
BULLETIN OF THE GEORGIAN NATIONAL ACADEMY OF SCIENCES, vol. 3, no. 2, 2009

Cybernetics

Constructive Method for the Recognition of Patterns

Zurab Kipshidze*, Alexander Chaduneli*, Gulaber Ananiashvili*, George Ghlonti**

* *N. Muskhelishvili Institute of Computational Mathematics, Tbilisi*

(Presented by Academy Member M. Salukvadze)

ABSTRACT. On the basis of Coding theory universal methods for recognition of patterns and clustering are elaborated, providing an adoption of realization of patterns to the set of target class, i.e. recognition of patterns and division of the set of realization into clusters. *© 2009 Bull. Georg. Natl. Acad. Sci.*

Key words: *realization of patterns, code, coding structure, decoding, syndrome, factorization, neighbor classes, primitive element, clusters.*

1. An algorithm for the recognition of patterns

The recognition of patterns represents two-stage extrapolation, the first stage of which consists in statistical analysis of recognition space (channel, environment), establishment of a standard object, and definition of the set $a_i \in A_N$ of possible deviations from it in *n*-dimensional vector space V^n over the field *GF*(*q*) where *q* is the prime number.

On the basis of those deviations $a_i \in A_N$ the structure of correction code is built [2] by means of decoding operator *H*, presented as a linear matrix of the range *r*n*. The operator *H* provides factorization of the space V^n into adjacent *N* non-compact classes corresponding to the kernel *G* of a zero subspace *H* and realizations a_i as class leaders (creators).

$$(a_i + G)H = s_i \, (i = 0, N-1),$$

where a_i is *n*-dimensional vector belonging to the space of vectors *V*.

$$a_i = \alpha_{i1}, \alpha_{i2}, ..., \alpha_{in},$$

where $a_{ij} \in \{0,1,...,q-1\}$ are the elements of *GF*(*q*) field. $S_i = a_i H$ is a syndrome of vector a_i of length *r*, on the basis of which the adjacent class with the leader a_i is defined.

Second stage of extrapolation is the recognition process itself, when localization of the deviation a_i, thus correction of deviation and object perception is effected, leading to restitution of the initial state of the system. Let us note that some *i*-th (*i* = 0, *N*-1) element from adjacent classes is selected for target class. It represents one possible realization of the object under recognition.

Suppose that during the recognition process some realization a_i is obtained for recognition such that it does not belong to *i*-th standard class, i.e. $a_j \neq a_i$. In this case the distance between classes, i.e. the difference

$$a_j - a_i = e$$

should be defined, and *e* should be added to a_j giving some standard realization

$$a_j + e = a_i.$$

Thus the problem is solved – an object is recognized, it is represented by some realization of *i*-th standard class.

Below an algorithm for construction of operator *H* is presented.

Suppose the set of vectors $a_i \in A_N$ is given where $A_N \subset V^n (i = \overline{0, N-1})$ and $0 < N \leq q^n$. Let *W* be the correlation of the type $w = a^i - a^j (0 < j \leq N)$, and *H* be a linear mapping in the space V^n

$$WH \supset Q = (0,0,...,). \tag{1}$$

The latter formula in matrix representation will assume the form:

$$W(j)H \supset m^j, \tag{2}$$

where m^j is *j*-th line of *H* matrix.

To obtain the operator *H* it is necessary for any given natural $j \leq n$ and $i < q$ to construct the following:

a) A set of vectors $w = (w_1,..., w_2)$ where $W_i(j)$ with their *j*-th components equal to *i* $(w_j = i)$, and with zero for all other components to the right of *j*-th:

$$w_\sigma = 0 \ (\sigma > j);$$

b) A set $W_i^*(j) = i^* W_i(j) + e^j$,

where e^j is a unitary vector of *n* length such that its *j*-th digit is equal to 1 and i^* is a solution of the equation

$$ix \equiv -1(\text{mod}\, q);$$

c) A set $w_i(j) = \bigcup_{i=1}^{q-1} w_i^*(j)$, which gives $\bigcup_{i=1}^{q-1} (i(W(j))H \supset \theta$ or otherwise $W_i(j)H \supset m^j$, where m^j is *j*-th line of *n*r* -dimensional *H* matrix.

The last equation leads us to the conclusion that for m^j line of matrix *H* any vector may be chosen from the space $V^n \setminus W^*(j)H$.

The transformation matrix *H* may be given in idempotent form, when in the case of zero columns omitted we obtain

$$H = \left[\frac{B}{E_r}\right].$$

The construction rule for this matrix is based on successive selection of lines m^j with $N \leq r$ as necessary and $\log_q g < r$ as sufficient conditions, where *r* is the range of the matrix *H*, and $g = \max\{g(j)S\}$, and *g*(*j*)-*w*(i) is the amount of vectors.

For recognition of patterns use of arithmetic codes with so-called check by module is also advisable. In this case realizations are presented as numbers. Module, minimal from the point of view of realization, is defined, which will be used for code generation. Module should be chosen in such a way that each deviation (realization) have its adjacent class. In this case conformity of realizations (patterns) to standard class, which is one of the residual classes, should be guaranteed.

1. Algorithm for Clustering

Suppose that the realizations under recognition are the members of the field *FG*(*q*) the non-zero members of which are the powers of the primary element *α*. Suppose the elements $\alpha^{i_1}, \alpha^{i_2}, ... \alpha^{i_n}$ are associated in one cluster. These elements should be the roots of the equation $x^{q^m} - x = 0$ and on the other hand, they are the roots of the equation of the degree *k*, which might be presented in the following way:

All a_{ij} roots, satisfying this equation, are the realizations of patterns and are located in the same cluster. The equation for other clusters is composed in the same way.

If all the elements of the field are realizations to be recognized and it is necessary to distinguish them in two classes, an equation is formed only for one cluster. Another cluster will be composed by the remaining roots of the equation $x^{q^m} - x = 0$. The same for any kind of decomposition in case that it is impossible to form an extra equation. In case when none of the members of the field $GF(q^m)$ represents the realization of patterns, corresponding equations for clusters should be written for the existing realizations.

Consider for example a situation, when the realizations that should be recognized are the vectors of the length *n*=4 from the field *GF*(2). The Galois field $GF(2^4)$ consisting of 16 elements is obtained by the module $x^4 + x + 1$ of the remnant of a polynomial. This polynomial is conversant to the field *GF*(2) and it is the minimal function of the primitive α member of the field.

Suppose α is the class of remnants which includes *x*. Then α should be the root of the polynomial $x^4 + x + 1$ and the primitive element of the field. 15 non-zero elements of this field should be represented as follows:

$$\alpha^0 = 1 = (0001);$$
$$\alpha^1 = \alpha = (0010);$$
$$\alpha^2 = \alpha^2 = (0100);$$
$$\alpha^3 = \alpha^3 = (1000);$$
$$\alpha^4 = \alpha + 1 = (0011);$$
$$\alpha^5 = \alpha^2 + \alpha = (0110);$$
$$\alpha^6 = \alpha^3 + \alpha^2 = (1100);$$
$$\alpha^7 = \alpha^3 + \alpha + 1 = (1011);$$
$$\alpha^8 = \alpha^2 + 1 = (0101);$$
$$\alpha^9 = \alpha^3 + \alpha = (1010);$$
$$\alpha^{10} = \alpha^2 + \alpha + 1 = (0111);$$
$$\alpha^{11} = \alpha^3 + \alpha^2 + \alpha = (1110);$$
$$\alpha^{12} = \alpha^3 + \alpha^2 + \alpha + 1 = (1111);$$
$$\alpha^{13} = \alpha^3 + \alpha^2 + 1 = (1101);$$
$$\alpha^{14} = \alpha^3 + 1 = (1001);$$
$$\alpha^{15} = 1 = \alpha^0.$$

Suppose that 5 elements, in particular $\alpha, \alpha^3, \alpha^5, \alpha^6 \text{and} \alpha^8$, should be associated in one class. The corresponding equation should be represented as follows:

$$(x-\alpha)(x-\alpha^3)(x-\alpha^5)(x-\alpha^6)(x-\alpha^8) = 0.$$

Multiplication of members and use of the table presented above will lead to the following:

$$x^5 - \alpha^8 x^4 + \alpha^8 x^3 + \alpha^8 x^{14} - \alpha^8 = 0.$$

Patterns presented by the elements $\alpha, \alpha^3, \alpha^5, \alpha^6 \text{and} \alpha^8$ should be certainly grouped in one cluster, as they are roots of the above equation.

კიბერნეტიკა

სახეთა ამოცნობის კონსტრუქციული მეთოდი

ზ. ყიფშიძე*, ა. ჩადუნელი*, გ. ანანიაშვილი*, გ. ღლონტი*

* *ნ. მუსხელიშვილის გამოთვლითი მათემატიკის ინსტიტუტი*

(წარმოდგენილია აკადემიკოს მ. სალუქვაძის მიერ)

კოდირების თეორიის გამოყენებით დამუშავებულია სახეთა ამოცნობისა და კლასტერირების უნივერსალური მეთოდები, რომლებიც უზრუნველყოფენ სახეთა რეალიზაციების მიკუთვნებას ეტალონური კლასის სიმრავლესთან, ე.ი. სახეთა ამოცნობას და რეალიზაციათა სიმრავლის ნებისმიერად დაყოფას კლასტერებად.

REFERENCES

1. *Z. Kipshidze, G. Ananiashvili* (2003), Informatsiis teoria, kodireba da sinergetika [Theory of Informatics, Coding and Synergy], Tbilisi (in Georgian).
2. *U. Peterson, E. Ueldon* (1976), Kody, ispravlyayushchie oshibki, M. (in Russian).

Received December, 2008