STATISTICAL CLASSIFICATION IN MONITORING SYSTEMS

Alexander A. Grusho, Nick A. Grusho and Elena E. Timonina
Institute of Informatics Problems,
Federal Research Center "Computer Science and Control"
of the Russian Academy of Sciences
Vavilova 44-2,
119333, Moscow, Russia
Email: grusho@yandex.ru

KEYWORDS
statistical decision functions, bans of probability measures, secure architectures, intelligent monitoring

ABSTRACT
The paper is devoted to the statistical classification problems. Repeated classification in control and monitoring systems is complicated by nonzero mistakes of traditional statistical decisions. At repeated applications of rules of statistical classification small probabilities of mistakes generate a large number of wrong decisions. At construction of monitoring systems of information security in computer systems wrong decisions are especially dangerous. Therefore for construction of secure architecture of control and monitoring systems it is necessary to look for nonconventional statistical decisions.

In finite set of words of finite length the ban is the word having zero probability of appearance. If the statistical criterion has the critical set consisting of only bans of supposed probability measure, the probability of wrong rejection of this measure is equal to zero. Therefore repeated application of such criterion won’t generate to false alarms in monitoring systems.

In the paper we consider a case of statistical classification when classes are defined by finite sets of probability distributions on a space of infinite sequences. We use bans to define decision functions and prove conditions when these decisions produce no mistakes.

INTRODUCTION

The paper is devoted to the statistical classification problems. Statistical classification of data is often used in different mathematical modeling problems. However repeated classification in control and monitoring systems is complicated by nonzero mistakes of traditional statistical decisions. At repeated applications of rules of statistical classification small probabilities of mistakes generate a large number of wrong decisions (Axelson, 1999).

At construction of monitoring systems of information security in computer systems the wrong decisions are especially dangerous. Monitoring and control systems are widely explored in different directions (Socolov et al., 2013). One of the most important model of monitoring systems is a stochastic one. Therefore for construction of secure architecture (Grusho et al., 2015a) of control and monitoring systems it is necessary to look for nonconventional statistical decisions. That is why we use a concept of a ban of probability measure in discrete probability space (Grusho and Timonina, 2011; Grusho et al., 2010).

In finite set of words of finite length the ban is the word having zero probability of appearance. If the statistical criterion has the critical set consisting of only bans of supposed probability measure, the probability of wrong rejection of this measure is equal to zero. Therefore repeated application of such criterion won’t generate to false alarms in monitoring systems (Grusho et al., 2013a).

There are certain special applications of statistical methods defined by bans. For example in (Denisov, 2015) the search of inserted functional relations in random sequences is investigated. The important problem is to determine bans of considered measures. This problem was solved with help of statistical simulation of the analyzed measures (Grusho et al., 2013b). Such simulation helps to get consistent estimation of the set of bans.

At research of such criteria it is proved that under certain conditions there is consistency meaning that power function tends to 1 on each alternative (Grusho et al., 2013a). Conditions when power function becomes equal to 1 for all alternatives on a finite step are found (Grusho et al., 2014, 2015b).

In this paper we present the generalization on a case of statistical classification when classes are defined by finite sets of probability distributions on a space of infinite sequences. The paper is structured as follows. Section 2 introduces definitions and previous results. In Section 3 the main results are proved. In Conclusion we shortly analyze future problems of construction of decision functions.

MATHEMATICAL MODEL. BASIC DEFINITIONS AND PREVIOUS RESULTS
Let $X = \{x_1, \ldots, x_m\}$ be a finite set, $X^\infty$ be a Cartesian product of $X$, $X^\infty$ be a set of all sequences where $i$-th element belongs to $X$. Define $\mathcal{A}$ be a $\sigma$-algebra on $X^\infty$, generated by cylindrical sets. $\mathcal{A}$ is also Borel $\sigma$-algebra in Tychonoff product $X^\infty$, where $X$ has a discrete topology (Bourbaki, 1968; Prokhorov and Rozanov, 1993).
On $(X^\infty, A)$ a probability measure $P$ is defined. Assume $P_n$ be a projection of $P$ on the first $n$ coordinates of sequences from $X^\infty$. It is clear that for every $B_n \subseteq X^n$

$$P_n(B_n) = P(B_n \times X^\infty).$$

(1)

Let $D_n$ be a support of measure $P_n$:

$$D_n = \{ \tau_n \in X^n, P_n(\tau_n) > 0 \}.$$

Denote

$$\Delta_n = D_n \times X^\infty.$$

The sequence $\Delta_n$, $n=1,2,...$, is nonincreasing and $\Delta(P) = \lim_{n \to \infty} \Delta_n = \bigcap_{n=1}^{\infty} \Delta_n$. (2)

The set $\Delta(P)$ is closed and it is a support of $P$. If $\tau_k \in X^k$, then $\tau_{k-1}$ is obtained from $\tau_k$ by dropping the last coordinate.

**Definition 1.** Ban (Grusho et al., 2014) of measure $P_n$ is a vector $\tau_k \in X^k$, $k \leq n$, such that

$$P_n(\tau_k \times X^{n-k}) = 0.$$

If

$$P_{k-1}(\tau_{k-1}) > 0,$$

then $\tau_k$ is the smallest ban (Grusho et al., 2014).

If $\tau_k$ is a ban of measure $P_n$ then for every $k \leq s \leq n$ and for every $\tau_s$ sequence starting with $\tau_k$ we have

$$P_s(\tau_s) = 0.$$

If there exists $\tau_n \in X^n$ such that $P_n(\tau_n) = 0$ then there exists the smallest ban. That is why further we say simply a ban of measure $P$.

Let on $(X^\infty, A)$ probability measures $P^{(1)}, ..., P^{(s)}$ be defined. As before we also define $P_n^{(i)}, D_n(P^{(i)}), \Delta_n(P^{(i)}), i = 1, 2, ..., s.$

Further under $\Delta_n^{(i)}$, $i = 1, 2, ..., s$, we will understand a set of the smallest bans of measure $P_n^{(i)}$, which have lengths equal to $n$.

Let’s construct a graph $G$ on vertices $P^{(1)}, ..., P^{(s)}$. Vertices $P^{(i)}$ and $P^{(j)}$ are connected by edge in graph $G$ if and only if $\Delta(P^{(i)}) \cap \Delta(P^{(j)}) \neq \emptyset$. Let $Q^{(1)}, ..., Q^{(r)}$ be sets of vertices in components of graph $G$, and for $i = 1, ..., r$, $V_i$ be a set of indexes of vertices, including in $Q^{(i)}$. Denote

$$\Delta(V_i) = \bigcup_{j \in V_i} \Delta(P^{(j)}), i = 1, r.$$

For $x \in X^\infty$ denote $x|_n$ be a vector which includes the first $n$ coordinates of the sequence $x$.

Then let’s consider a sequence of decision functions $d_n(x|_n) = i, i = 1, r, n = 1, 2, ...$

The basic problem considered in the paper is to find conditions when there exists such $N$ that for all $n \geq N$ we can determine such decision function $d_n(x|_n)$ by bans that

$$P^{(i)}_n(d_n(x|_n) = j) = 1,$$

where $i = 1, 2, ..., s$, $j = 1, ..., n$, $i \in V_j$.

**MATHEMATICAL RESULTS**

Let $P^{(i)}, Q^{(i)}$ and $V_i$ be defined as in Section 2. The solution for the basic problem is described in the next theorem.

**Theorem 1.** There exists a sequence of $d_n(x|_n)$, $n = 1, 2, ..., defined by bans, for which exists such $N$, that for every $n \geq N$ equations (3) are fulfilled if and only if

$$\Delta(V_i) \cap \Delta(V_j) = \emptyset,$$

for $i \neq j, i, j = 1, ..., r$.

The proof of the theorem 1 is based on several lemmas.

Let $\tau_k$ be the smallest ban of measure $P^{(i)}$. Then define $I_i(\tau_k)$ be the elementary cylindrical set in $X^\infty$, which is generated by the vector $\tau_k$.

**Lemma 1.** For every sequence $x \in I_i(\tau_k)$ it follows that $x \not\in \Delta(P^{(i)})$.

**Proof.** Suppose that there exists $x \in I_i(\tau_k)$ that belongs to $\Delta(P^{(i)})$. From formula (2) it follows that $x \in \Delta_n(P^{(i)})$ for every $n = 1, 2, ...$. By the definition of $\Delta_k(P^{(i)})$ the vector $x|_k$ defined by the first $k$ coordinates of $x$ belongs to the set $D_k(P^{(i)})$. Then

$$P_k(x|_k) > 0.$$

Besides

$$x|_k = \tau_k,$$

that contradicts to supposition. The lemma 1 is proved.

Let’s define the open set $S_i$:

$$S_i = \bigcup_{k=1}^{\infty} \bigcup_{\tau_k \in \Delta_k} I_i(\tau_k).$$

(4)

From lemma 1 it follows that

$$S_i \cap \Delta(P^{(i)}) = \emptyset.$$

**Lemma 2.** The set $S_i$ can be represented in the next form

$$S_i = X^\infty \setminus \Delta(P^{(i)}).$$

**Proof.** From

$$S_i \cap \Delta(P^{(i)}) = \emptyset$$

it follows that

$$S_i \subseteq X^\infty \setminus \Delta(P^{(i)}).$$

Let’s assume that

$$x \in X^\infty \setminus \Delta(P^{(i)}).$$

If $x \in X^\infty \setminus \Delta(P^{(i)})$ then

$$x \not\in \Delta(P^{(i)}) = \bigcap_{n=1}^{\infty} \Delta_n(P^{(i)}).$$

The sequence of sets $\{\Delta_n(P^{(i)})\}$ is not increasing. Then there exists $n$ such that for every $t \geq n$ we have $x \not\in \Delta_t(P^{(i)})$. That means that $P_t(x|_t) = 0$. Thus there exists the smallest ban $\tau_k$ such that $x \in I_i(\tau_k)$, so $x \in S_i$. Lemma is proved.
Lemma 3.
\[ \Delta(V_j) \cap \Delta(P^{(i)}) = \emptyset \]
if and only if
\[ \Delta(V_j) \subseteq S_i. \]

Proof. From the condition of lemma 3 it follows that
\[ \Delta(V_j) \subseteq X^\infty \setminus \Delta(P^{(i)}). \]
Then from lemma 2 \( \Delta(V_j) \subseteq S_i. \)
On the other hand if \( \Delta(V_j) \subseteq S_i, \) then
\[ \Delta(V_j) \subseteq X^\infty \setminus \Delta(P^{(i)}), \]
and it follows that
\[ \Delta(V_j) \cap \Delta(P^{(i)}) = \emptyset. \]
Lemma is proved.

Lemma 4. If
\[ \Delta(V_j) \cap \Delta(P^{(i)}) = \emptyset \]
then \( \exists N_i \) such that
\[ \Delta(V_j) \subseteq \bigcup_{k=1}^{N_i} \bigcup_{\pi_k \in \Lambda_k} I_i(\pi_k). \]

Proof. Tychonoff product \( X^\infty \) is a compact space (Bourbaki, 1968) and therefore from an every infinite cover of a compact by open sets it is possible to select a finite cover. The closed set \( \Delta(V_i) \) is a compact and \( \Delta(V_j) \subseteq S_i. \) That’s why due to definition (4) there exists \( N \) such that
\[ \Delta(V_j) \subseteq \bigcup_{i=1}^{N_i} \bigcup_{\pi_k \in \Lambda_k} I_i(\pi_k) = \sigma_{N_i}(i). \]

Lemma 4 is proved.

The set \( \sigma_{N_i}(i) \) is a cylindrical set. Therefore it can be represented in the next form
\[ \sigma_{N_i}(i) = C_{N_i}^{(i)} \times X^{N_i}, \]
where
\[ C_{N_i}^{(i)} \subseteq X^{N_i}. \]

Lemma 5. For every \( j = 1, ..., r, \)
\[ \Delta(V_j) = \bigcup_{n=1}^{\infty} \bigcup_{i \in V_j} \Delta_n(P^{(i)}). \]

Proof. For every \( i = 1, ..., s \)
\[ \Delta(P^{(i)}) = \bigcup_{n=1}^{\infty} \Delta_n(P^{(i)}). \]

Union in formula (6) for every \( j \) is finite and sequence \( \Delta_n(P^{(i)}) \) for every \( i \) is non increasing. Then
\[ \bigcup_{i \in V_j} \bigcup_{n=1}^{\infty} \Delta_n(P^{(i)}) = \bigcup_{n=1}^{\infty} \bigcup_{i \in V_j} \Delta_n(P^{(i)}). \]
The lemma 5 is proved.

Lemma 6. If \( \forall t \neq j, \)
\[ \Delta(V_t) \cap \Delta(V_j) = \emptyset, \]
then \( \exists N: \)
\[ \Delta_N(V_t) \cap \Delta_N(V_j) = \emptyset. \]

Proof. According to lemma 5
\[ \bigcap_{n=1}^{\infty} \bigcup_{i \in V_t} \Delta_n(P^{(i)}) \cap \bigcap_{n=1}^{\infty} \bigcup_{i \in V_j} \Delta_n(P^{(i)}) = \emptyset, \]
and
\[ \bigcap_{n=1}^{\infty} \{ \{ \bigcup_{i \in V_t} \Delta_n(P^{(i)}) \} \cap \{ \bigcup_{i \in V_j} \Delta_n(P^{(i)}) \} \} = \emptyset. \]
Due to compactness of Tychonoff product (Bourbaki, 1968) it follows that \( \exists N_{t,j} \) that for \( \forall N \geq N_{t,j} \)
\[ \bigcup_{n=1}^{N} \bigcup_{i \in V_t} \Delta_n(P^{(i)}) \cap \bigcup_{n=1}^{N} \bigcup_{i \in V_j} \Delta_n(P^{(i)}) = \emptyset. \]

From monotonicity to \( n \) of
\[ \bigcup_{i \in V_t} \Delta_n(P^{(i)}) \]
and
\[ \bigcup_{i \in V_j} \Delta_n(P^{(i)}) \]
it follows that
\[ \bigcup_{i \in V_t} \Delta_N(P^{(i)}) \cap \bigcup_{i \in V_j} \Delta_N(P^{(i)}) = \emptyset. \]
Lemma 6 is proved.

Corollary of Lemma 6. If true probability distribution of \( x \) is in \( Q^N \) then from lemma 4 it follows that supports of all measures \( P^{(i)}_{N_j}, j \notin V_i, \) are covered by bans of true measure. That is every vector from \( D_N(P^{(j)}) \) includes bans of true measure.

Let’s now prove the theorem 1. Let’s define the decision function \( d_N(\pi_N) \) satisfying to conditions of the theorem 1. Note that defined earlier sets \( C_{N(i)}^{(i)}, i = 1, 2, ..., s, \) are determined by bans. Then they are defined by the smallest bans. Let’s denote \( N = \max_{i=1,2,...,s} N^{(i)}. \)

Consider the next matrix \( M = \|A_i^{(i)}\| \) of size \( s \times N, \)
where element of \( i \)-th row and \( j \)-th column is the set of smallest bans of the measure \( P^{(i)} \) and of the length \( j. \)

For the observed random sequence \( x \) let’s determine vectors \( x|_1, ..., x|_N. \) For every one of these vectors compare its value with elements of certain column of matrix \( M. \) If any coincidence is found then let’s mark that row where the vector is an element of the set of the smallest bans. It is clear that the true probability distribution cannot produce any coincidence. It follows from the fact that in \( x \) there are no bans of the true distribution.
It follows that

\[ d_N(x|\tau_N) = j \]

when \( i \in V_j \).

Define the set

\[ B_N^{(j)} = \{ \tau_N : d_N(\tau_N) = j \} \]

and denote

\[ R_N^{(j)} = \bigcup_{i=1, i \neq j} \{ \tau_N : d_N(\tau_N) = i \}. \]

Then for \( P^{(i)} \in Q^{(j)} \)

\[ D_N(P^{(i)}) \subseteq B_N^{(j)}. \]

Due to the definition for all \( i = 1, 2, ..., s \), we have

\[ P_N^{(i)}(R_N^{(j)}) = 0. \]

Then due to definition

\[ R_N^{(j)} \cap D_N(P^{(i)}) = \emptyset. \]

That’s why

\[ R_N^{(j)} \times \{ \tau_N \} \subseteq \sigma(i). \]

As

\[ P_N^{(l)}(B_N^{(l)}) = 1 \]

for \( l \neq j \) and \( l \in V_i \), \( d(\tau_N) \) is correctly defined, then

\[ D_N(P^{(l)}) \cap D_N(P^{(i)}) = \emptyset \]

and

\[ D_N(P^{(i)}) \subseteq R_N^{(j)}. \]

Then for \( t \neq j \)

\[ \Delta_N(V_j) \cap \Delta_N(V_i) = \emptyset. \]

It follows that

\[ \Delta(V_j) \cap \Delta(V_i) = \emptyset. \]

The theorem 1 is proved.

From the theorem 1 we can get the main result of (Grusho et al., 2016).

**Theorem 2.** Let \( P^{(1)}, ..., P^{(s)} \) are such that for all \( i \), \( |Q^{(i)}| = 1 \), then there exists \( N \) and a function \( d_N(\tau_N) \) defined by bans such that for all \( i = 1, 2, ..., s, \)

\[ P_N^{(i)}(d_N(\tau_N) = i) = 1, \]

if and only if for all pairs \( i, j, i \neq j, i = 1, 2, ..., s, j = 1, 2, ..., s, \)

\[ \Delta(P^{(i)}) \cap \Delta(P^{(j)}) = \emptyset. \]

Under the considered conditions we can define decision functions \( g_n(x|n) \), which are determined by the conditions

\[ g_n(x|n) = i, \]

where \( i \) is the \( \min\{j\} \) such that

\[ x|n \in D_n(V_j). \]

It is possible to prove that there exists \( N \) when functions satisfy equation

\[ P_N^{(i)}(g_N(x|n) = j) = 1 \]

for all \( i \in V_j \).

But searching bans is equivalent to signature analysis, which proved to be quick.

It is interesting to compare these types of decision functions. Let’s define rooted trees containing admissible trajectories of random data \( x \in X^\infty \) produced under the probability distribution \( P^{(i)} \). The root of every tree means the first element of admissible sequence. Every infinite branch of the tree uniquely defines a sequence \( x \in X^\infty \) in the same way at it is usually done in \( m \)-arc tree, edges show possible ways of development of the random sequence. All admissible parts of branches of the length \( n \) define the set \( D_n(P^{(i)}) \) and \( X^n \setminus D_n(P^{(i)}) \) is a set of all bans of the length \( n \). The set \( D_n(P^{(i)}) \) can be represented by trees of the height \( n \). All bans of the length \( n \) also can be represented by trees.

If we use the smallest bans for classification we use reduced vectors which have lengths less or equal \( n \). We can see smallest bans on the \( D_n(P^{(i)}) \)-trees. When we use the smallest bans then the number of steps to check that the vector \( x|n \) does not belong to \( D_n(P^{(i)}) \) demands less or equal than \( n \) steps. That’s why we say about decision functions defined by bans.

No marks in the matrix \( M \) means that \( x|n \) may belong to \( D_n(P^{(i)}) \). But all vectors \( x|n \) have to possess bans of all other measures when \( n \geq N \), because \( x|n \) doesn’t belong to supports of measures in other components of the graph \( G \). If \( n < N \) it is possible to make a mistake. It means that probability distribution in another component may be admitted as a part of the true one.

**CONCLUSION**

Conditions under which the statistical classification of random infinite sequence is reduced to consideration of finite space of vectors of finite length \( N \) are found in the article. The offered approach allows to generalize the
found conditions on a case when some classes of probability measures contain an infinite number of measures. There is an open problem of constructive estimation of parameter $N$. If this problem was solved, from conditions on supports of measures in infinite spaces it would be possible to pass to conditions on supports of measures in finite spaces. In certain cases this problem is effectively solved. However the overall picture isn’t visible yet.

Acknowledgements

This work was supported by the Russian Foundation for Basic Research (grant 15-07-02053).

REFERENCES


AUTHOR BIOGRAPHIES

ALEXANDER A. GRUSHO, Professor (1993), Doctor of Science in physics and mathematics (1990). He is Head of laboratory in Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences and Professor of Moscow State University.

Research interests: probability theory and mathematical statistics, information security, discrete mathematics, computer sciences.

His email is grusho@yandex.ru.

NICK A. GRUSHO has graduated from the Moscow Technical University. He is Candidate of Science (PhD) in physics and mathematics. At present he works as senior scientist at Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences.

Research interests: probability theory and mathematical statistics, information security, simulation theory and practice, computer sciences.

His email is info@itake.ru.

ELENA E. TIMONINA has graduated from the Moscow Institute of Electronics and Mathematics and obtained the Candidate degree (PhD) in physics and mathematics (1974). She is Doctor in Technical Science (2005), Professor (2007). Now she works as leading scientist in Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences.

Research interests: probability theory and mathematical statistics, information security, cryptography, computer sciences.

Her email is eltimon@yandex.ru.