

**ОДЕССКИЙ НАЦИОНАЛЬНЫЙ  
ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ**

**Институт компьютерных систем**

**Кафедра системного программного обеспечения**

**ПРЕДСТАВЛЕНИЕ ИСХОДНЫХ ДАННЫХ В  
ЗАДАЧАХ НЕЙРОСЕТЕВОГО  
ПРОГНОЗИРОВАНИЯ**

**КРИСИЛОВ**

**Виктор Анатольевич**

**Тел.: (380-482) 22-78-40  
28-85-66**

**E-mail: VictorK@OL405.pasco.net**

**ЧУМИЧКИН**

**Константин Васильевич**

**Тел.: (380-482) 65-27-93**

**E-mail: k\_chum@ukr.net**

**КОНДРАТЮК**

**Александр Васильевич**

**Тел.: (380-48) 711-36-77**

**E-mail: alkond@ukr.net**

## ПРЕДСТАВЛЕНИЕ ИСХОДНЫХ ДАННЫХ В ЗАДАЧАХ НЕЙРОСЕТЕВОГО ПРОГНОЗИРОВАНИЯ

### Аннотация

В работе рассматриваются вопросы предварительных преобразований данных в задачах прогнозирования временных рядов посредством нейронных сетей. Сформулированы требования к предварительным преобразованиям, необходимые для уменьшения ошибки прогноза. Приведены практические результаты применения предложенных соображений при прогнозировании финансовых временных рядов.

### **V.A. KRISILOV, K.V. CHUMICHKIN, A.V. KONDRATUK**

Odessa national polytechnic university. Computer systems institute.  
VictorK@OL405.paco.net, k\_chum@ukr.net

## INPUT DATA REPRESENTATION IN TIME SERIES NEURAL BASED FORECASTING TASKS

### Abstract

Data preprocessing aspects in time series neural based forecasting tasks are discussed. Low range forecasting error achievement conditions are proposed. Discussed methods are applied to financial time series forecasting.

### Введение

Задача прогнозирования временных рядов (ВР) была и остается актуальной, особенно в последнее время, когда стали доступны мощные средства сбора и обработки информации. Широкое применение нейронных сетей (НС) в данной задаче обусловлено наличием в большинстве ВР сложных закономерностей, не обнаруживаемых линейными методами. Одним из самых важных этапов в решении задачи нейросетевого прогнозирования – это формирование обучающей выборки (ОВ). Именно от состава, полноты, качества ОВ существенно зависят время обучения НС и достоверность получаемых моделей. Поэтому целью данной работы является выявление критериев качества проведения предварительных преобразований на этапе построения ОВ.

### 1. Прогнозирование ВР с использованием нейронных сетей. Существующий подход к решению задачи

Временным рядом  $\{a(t)\}$  будем называть множество значений некоторой величины в последовательные моменты времени.

$$\{a(t)\} \equiv a(t_1); a(t_2); \dots; a(t_{i-1}); a(t_i); a(t_{i+1}); \dots$$

Прогнозирование временного ряда – вычисление величины его будущих значений либо характеристик, позволяющих определить эту величину, на основании анализа известных значений. Величина, подлежащая прогнозу, называется прогнозируемой величиной (ПВ).

При прогнозировании предполагается, что значение прогнозируемой величины зависит от каких-либо факторов, назовем их определяющими факторами, или признаками. Один из подходов к задаче прогнозирования основан на предположении зависимости ПВ от предыдущих значений ВР. Теоретическим обоснованием такого подхода является теорема Такенса [1].

Если временной ряд порождается динамической системой, т.е. значения  $\{a(t)\}$  есть произвольная функция состояния такой системы, существует такое число  $d$  (примерно равное эффективному числу степеней свободы данной динамической системы), что  $d$  предыдущих значений ВР однозначно определяют следующее значение.

Схему решения задачи прогнозирования можно представить в виде последовательности этапов (рис.1).

1. Этап предварительных преобразований (ПП).

На практике большинство прогнозируемых ВР порождаются сложными динамическими системами, с множеством степеней свободы, поэтому  $d$  для них велико. Кроме того, в самом ВР может присутствовать случайная составляющая. Поэтому на данном этапе выполняются предварительные преобразования исходных данных [2, 3] позволяющие уменьшить ошибку прогнозирования.

Предварительным преобразованием будем называть получение для момента времени  $t_i$  набора из определяющих факторов и соответствующего им значения ПВ (рис. 2). Определяющие факторы представляют собой некую функцию от прошлых (по отношению к моменту  $t_i$ ) значений ВР. Назовем набор определяющих факторов вектором признаков  $\{X\}$ , или описанием ситуации.

$$\{X\} \equiv x_1; x_2; \dots; x_n,$$

$$\{X\} = \Phi(a(t_i), a(t_{i-1}), \dots, a(t_{i-d}))$$

Описанию ситуации соответствует ПВ  $\{Y\}$ , которая может являться как самими будущими значениями ВР, так и характеристиками, позволяющими их определить с требуемой точностью.

$$\{Y\} \equiv y_1; y_2; \dots; y_m,$$

$$\{Y\} = \Psi(a(t_{i+1}), a(t_{i+2}), \dots, a(t_{i+k}))$$

После выполнения ПП для различных моментов времени  $t$ , ВР представляется в виде множества значений функции зависимости ПВ от определяющих факторов (или в виде множества наборов).

Далее из множества полученных наборов выделяются два непересекающихся подмножества (как правило, хронологически следующих одно за другим). Одно из них представляет собой обучающую выборку (ОВ), на которой будет выполняться обучение нейронной сети. Другое подмножество представляет собой контрольную выборку (КВ), которая не предъявляется нейронной сети в процессе обучения и используется для проверки качества прогноза. Таким образом, прогнозирование временного ряда сводится к задаче интерполяции функции многих переменных. Нейронная сеть используется для восстановления этой функции по множеству наборов, входящих в состав ОВ.

## 2. Этап структурного синтеза НС.

На данном этапе производится выбор архитектуры нейрона и структуры связей между нейронами.

## 3. Параметрический синтез НС.

Выполняется обучение НС. Как правило, используются методы градиентного спуска, в частности алгоритм обратного распространения ошибки и его модификации [4]. Следует отметить, что данный этап наиболее требователен к вычислительным ресурсам и занимает 50-90% времени решения задачи.

4. Проверка ошибки прогноза на КВ. Если значение ошибки находится в допустимых пределах, то задача считается решенной, и обученная НС используется для получения прогноза. В противном случае в зависимости от предполагаемой причины ошибки производится возврат к этапам 1, 2 или 3.

## 2. Этап предварительных преобразований

К сожалению, в работах большинства авторов основное внимание уделено выбору архитектуры и способа обучения НС, а этап ПП рассматривается лишь в контексте конкретной практической задачи, либо не рассматривается вообще. Как правило, приводится описание определенного типа ПП и результатов, полученных от его использования в той или иной области, а сравнительного анализа с другими типами ПП и критериев, по которым можно было бы их сравнить, не приводится.

Тем не менее ПП влияет на результат решения задачи прогнозирования не менее, чем структура и способ обучения НС, т.к. результат ПП является исходными данными для этих двух этапов. Поэтому целесообразно более подробно остановиться на этапе ПП и попытаться сформулировать и обосновать основные требования к ПП, необходимые для уменьшения ошибки прогноза.

Некоторые требования уже известны, однако они носят эмпирический характер. Авторы, используют ниже перечисленные требования неявно, выбирая такие типы ПП, которые уже удовлетворяют им [2, 3].

### 2.1. Требования к ПВ

Основное требование к ПВ – это возможность восстановления будущих значений ВР с требуемой точностью.

### 2.2. Понижение размерности пространства признаков

Использование в качестве предварительного преобразования сверток исходных данных позволяют описать ситуацию меньшим количеством признаков без потери или с допустимой потерей точности. Это приводит к сокращению времени обучения НС.

### 2.3. Уменьшение взаимозависимости между входами

Взаимозависимость входов может привести к снижению информативности описания ситуации, и, следовательно, к ухудшению качества обучения. Следует отметить, что использование сверток позволяет отчасти решить эту проблему, т.к. большинство методов сжатия информации основано на исключении избыточности.

## 3. Требования стационарности и непротиворечивости

Ввиду отсутствия формальных критериев оценки качества проведения ПП целесообразно ввести требования, необходимые для уменьшения ошибки прогноза, а также критерии их выполнения.

### 3.1. Условие стационарности входных значений

Как уже было отмечено, на этапе обучения НС восстанавливается целевую функцию по множеству наборов ОВ т.е. решает задачу интерполяции. На этапе использования обученной НС (получении прогноза) сеть будет использовать восстановленную зависимость для прогнозирования ПВ, т.е. решать задачу экстраполяции. Для корректного решения задачи экстраполяции как задачи интерполяции необходимо обеспечить стационарность ряда признаков  $\{X\}$ . Напомним, что стационарным называется временной ряд, распределение значений которого инвариантно относительно момента времени, для которого оно построено. Для характеристики стационарности будем использовать то, что для двух выборок, построенных в разные моменты времени закон распределения должен оставаться тем же. Т.е. величина

$$s = \int_{-\infty}^{\infty} \min[p_{OB}(\{X\}); p_{KB}(\{X\})] dX$$

должна равняться единице.

Но ввиду многомерности  $\{X\}$  и с учетом независимости признаков охарактеризуем стационарность каждого из компонент  $x_i$  (рис. 3).

$$s_i = \int_{-\infty}^{\infty} \min[p_{OB}(x_i); p_{KB}(x_i)] dx_i$$

Кроме того, полагая, что закон распределения равномерен на участке  $[m_i - \sigma_i; m_i + \sigma_i]$ , где  $m_i$  - среднее значение, а  $\sigma_i$  - стандартное отклонение  $i$ -го признака, получим

$$s_i = \begin{cases} \frac{\sigma_i^{OB} + \sigma_i^{KB} - |m_i^{OB} - m_i^{KB}|}{\max(\sigma_i^{OB}; \sigma_i^{KB})}, & \text{если } (\sigma_i^{OB} + \sigma_i^{KB} - |m_i^{OB} - m_i^{KB}|) \geq 0 \\ 0, & \text{если } (\sigma_i^{OB} + \sigma_i^{KB} - |m_i^{OB} - m_i^{KB}|) < 0 \end{cases}$$

Следует отметить, что предположение о равномерности закона распределения вносит некоторую погрешность в вычисление величины  $s_i$ , и для некоторых распределений со значительной асимметрией может оказаться недопустимым. Однако, как показала практика, такое упрощение вполне подходит в качестве оценки.

Общую характеристику стационарности получим усреднением по всем признакам.

$$s = \frac{1}{n} \sum_{i=1}^n s_i$$

Если для ОБ и КВ имеет место

$$s \approx 1,$$

будем считать, что условие стационарности выполняется.

### 3.2. Условие непротиворечивости наборов

Для корректного восстановления функциональной зависимости необходимо, чтобы наборы, входящие в состав как ОБ, так и КВ, не противоречили друг другу. Т.е. чтобы для двух наборов с одинаковым описанием ситуации значение ПВ не было разным [5]. Однако на практике вероятность того, что у двух наборов описание ситуации  $\{X\}$  будет одинаковым, весьма мала, т.к. количество наборов ограничено, а  $\{X\}$  – многомерная величина с большим или бесконечным множеством принимаемых значений. Поэтому будем считать одинаковыми векторы признаков  $\{X^{(1)}\}$  и  $\{X^{(2)}\}$ , для которых выполняется условие

$$\{X^{(1)}\} \in o(\{X^{(2)}\})$$

Границы окрестности выбираются так, чтобы это условие выполнялось для 50-100% наборов.

Тогда противоречивость двух наборов определим как

$$\Delta\{Y^{(1,2)}\} = \begin{cases} \frac{1}{m} \sum_{i=1}^m |y_i^{(1)} - y_i^{(2)}|, & \text{если } \{X^{(1)}\} \in o(\{X^{(2)}\}) \\ 0, & \text{иначе} \end{cases}$$

А общая противоречивость будет равна

$$\Delta = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \Delta\{Y^{(i,j)}\}$$

где  $N$  – общее количество наборов в ОБ и КВ.

Полученное значение общей противоречивости не должно превышать требуемой точности решения.

## 4. Практические результаты

Эксперименты предполагали исследование известных типов ПП [2, 3, 5] на соответствие требованиям стационарности и непротиворечивости и соотнесение полученных результатов с фактической ошибкой прогноза на КВ.

Исходными данными служил ВР курса валют GBP/USD на рынке Forex с 06.02.2001 по 03.03.2001 с интервалом 15 минут.

Параметры прогнозирующей системы выбирались в соответствии с опытом реальных валютных операций:

время описания ситуации: 24 часа, что соответствует глубине погружения  $d = 96$ ;

прогнозируемая величина: приращение курса за 4 часа;

выборки: было сформировано 1464 набора, 75% из них (1098 наборов) составляли ОБ, 25% (366 наборов) – КВ;

структура НС: трехслойный персептрон с 80 нейронами в скрытом слое;

алгоритм обучения: классический алгоритм обратного распространения ошибки с моментом 0.5 и параметром скорости обучения 0.1, обучение останавливалось по достижении ошибки НС на ОБ 0.05 от первоначальной.

Исследовались следующие типы ПП:

1. Не преобразованные значения.

Описание ситуации представляло собой 96 последних значений временного ряда.

2. Значения приращений.

В [2] предлагается использовать в качестве описания ситуации не сами значения ВР, а разности двух последовательных значений, именуемые приращениями.

3. Промежутки стабильности.

В [5] было отмечено, что в ряде случаев описание в виде приращений оказывается избыточным, и предлагалось описывать ситуацию в виде промежутков стабильности, т.е. участков ВР, где курс стабильно рос, падал, или оставался неизменным. Каждый из таких промежутков характеризуется парой значений приращение-длительность. Описание ситуации состояло из 14-ти таких промежутков.

#### 4. Вейвлет-преобразование.

В основе вейвлет-преобразования лежит свертка вида

$$c_i = \int_{-\infty}^{\infty} x(t) \cdot \varphi_i(t) dt,$$

где  $c_i$  - полученный коэффициент преобразования,

$\varphi_i(t)$  - базис преобразования.

Более подробно о вейвлетах написано в [6], а об их применении к анализу финансовых ВР – в [3].

Ситуацию описывали 28 коэффициентов вейвлет-преобразования.

Полученные результаты приведены в таблице.

Таблица.

Тип ПП	Критерии выполнения требований		Результаты обучения	
	Стационарность, $s$	Противоречивость, $\Delta$	Время обучения, итераций	Средняя ошибка на КВ
Не преобразованные значения	0.64	2.3	3148	5.8
Значения приращений	0.86	1.5	2352	4.05
Промежутки стабильности	0.82	0.56	1516	1.26
Вейвлет-преобразование	0.85	0.62	1479	1.4

Отметим, что на стационарных и непротиворечивых данных были достигнуты меньшие значения времени обучения НС и ошибки прогноза на КВ. На нестационарных и противоречивых данных, напротив, даже после большого числа циклов обучения, ошибка на КВ была значительна. Полученные данные подтвердили соответствие сформулированных критериев практике.

#### Выводы

С учетом сформулированных в работе требований схема решения задачи прогнозирования примет вид (рис. 4).

1. Выполнение ПП, оценка выполнения требований стационарности и непротиворечивости. Если требования выполняются, то переход к этапу 2, в противном случае – изменение параметров или выбор другого типа ПП.

2. Структурный синтез НС.

3. Обучение НС.

4. Проверка ошибки прогноза на КВ. Если значение ошибки находится в допустимых пределах, то задача считается решенной, и обученная НС используется для получения прогноза. В противном случае в зависимости от причины ошибки производится возврат к этапам 1, 2 или 3.

Это позволит сократить затраты времени на построение прогнозирующей системы, т.к. для проверки качества ПП не обязательно выполнять длительный и ресурсоемкий процесс обучения НС, достаточно проверить выполнение сформулированных требований.

#### Список литературы.

1. Sauer, T., Yorke, J.A., Casdagli, M. // Journal of Statistical Physics. Embedology – 1991. – Vol. 65. – P. 579–616.
2. Бестенс Д., ван ден Берг В., Вуд Д. Нейронные сети и финансовые рынки. Принятие решений в торговых операциях. – М.: ТВП, 1997.
3. Востров Г.Н., Любченко В.В., Полякова М.В. Моделирование временных рядов с использованием вейвлет-сетей. // Искусственный интеллект. – Донецк – 2000. – Вып. 3. – С. 207–214.
4. Уоссерман Ф., Нейрокомпьютерная техника: теория и практика. – М.: Мир, 1992.
5. Тарасенко Р.А. Метод анализа и повышения качества обучающих выборок нейронных сетей для прогнозирования временных рядов. // Диссертация на соискание ученой степени кандидата технических наук. – ОНПУ – 2002.
6. Daubechies I. Ten lectures on wavelets. – Philadelphia: SIAM, 1992.

## Существующая схема решения задачи прогнозирования

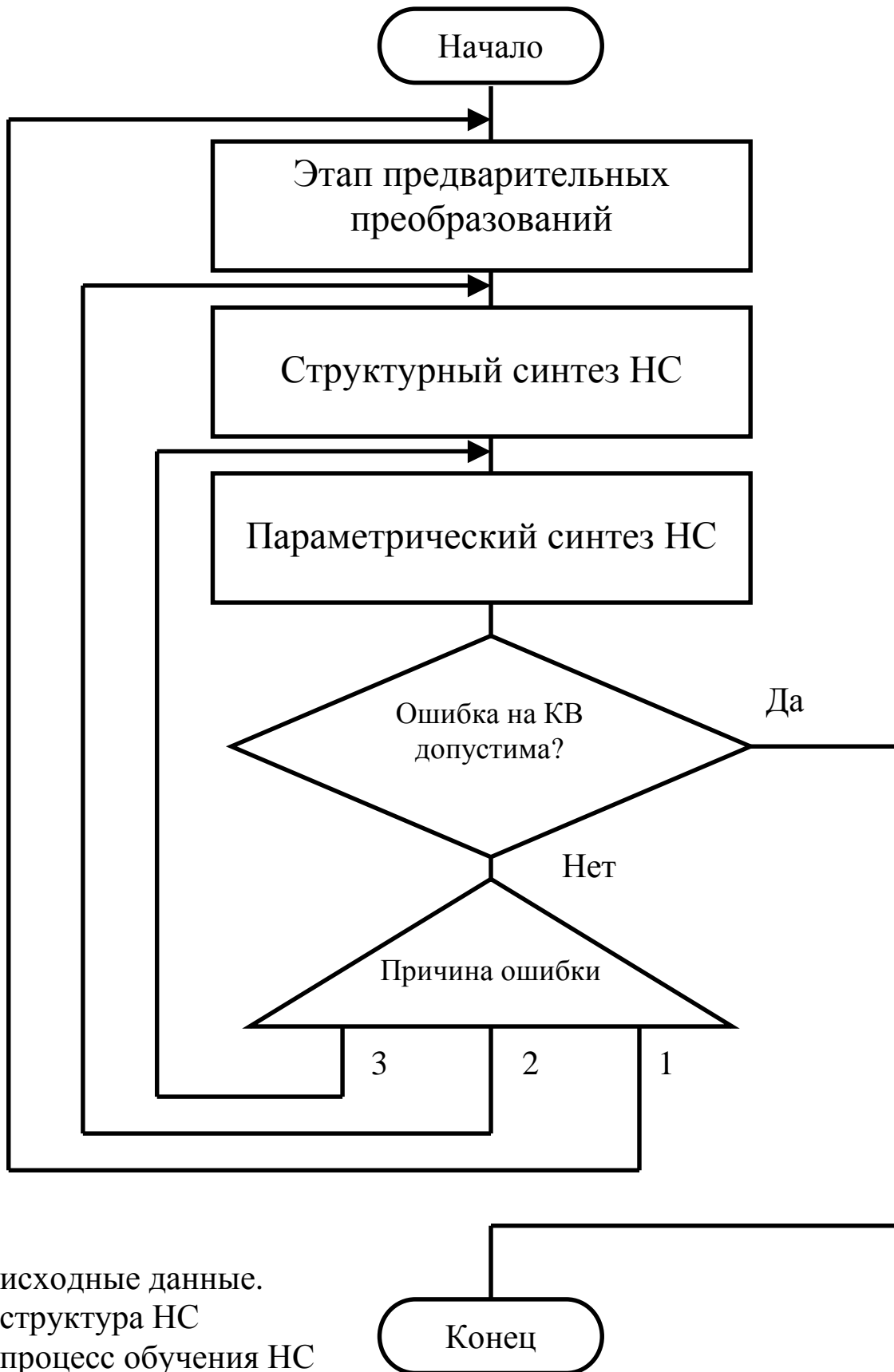
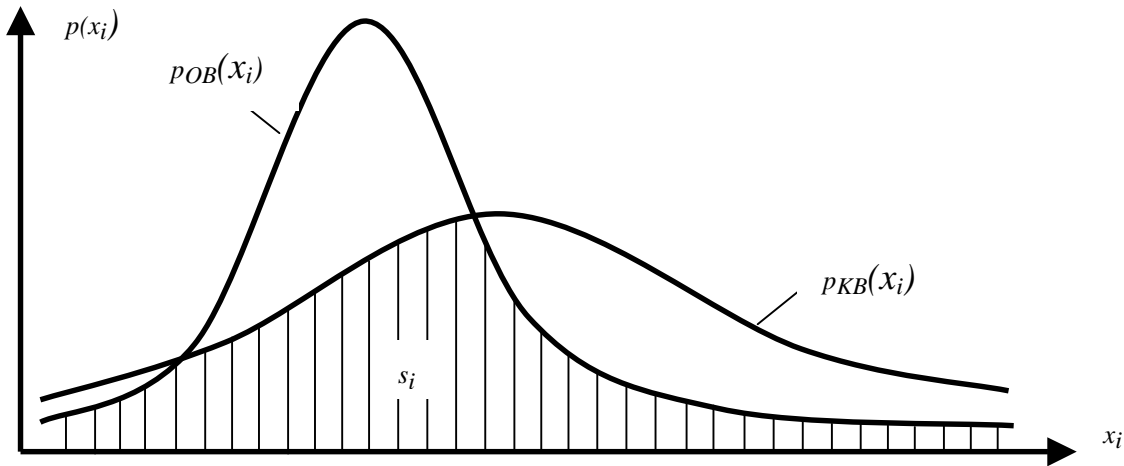


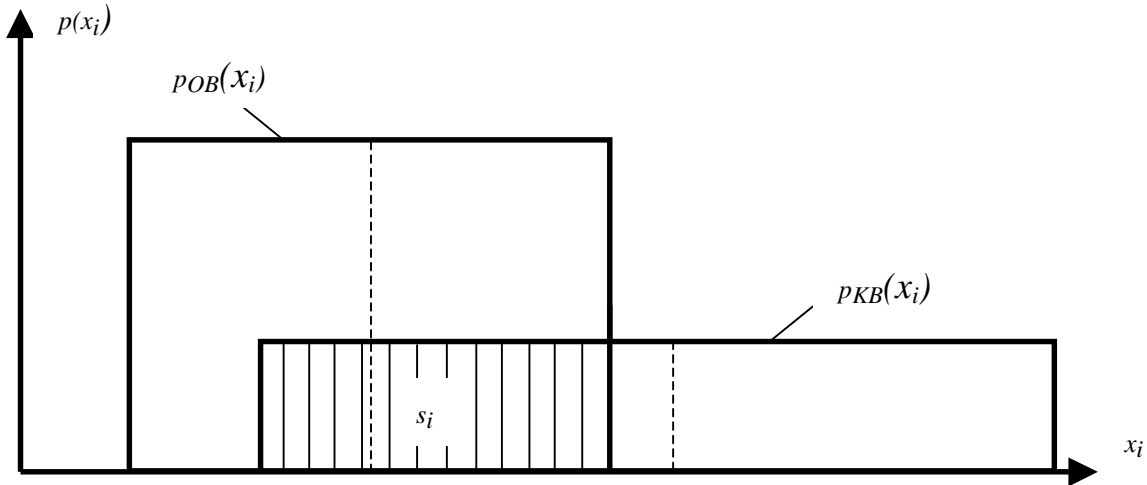
Рис. 1

## Условие стационарности входных значений

$$s = \int_{-\infty}^{\infty} \min[p_{OB}(\{X\}); p_{KB}(\{X\})] \{dX\}$$



$$s_i = \int_{-\infty}^{\infty} \min[p_{OB}(x_i); p_{KB}(x_i)] dx_i$$



$$s_i = \begin{cases} \frac{\sigma_i^{OB} + \sigma_i^{KB} - |m_i^{OB} - m_i^{KB}|}{\max(\sigma_i^{OB}; \sigma_i^{KB})}, & \text{если } (\sigma_i^{OB} + \sigma_i^{KB} - |m_i^{OB} - m_i^{KB}|) \geq 0 \\ 0, & \text{если } (\sigma_i^{OB} + \sigma_i^{KB} - |m_i^{OB} - m_i^{KB}|) < 0 \end{cases}$$

Стационарность: 
$$s = \frac{1}{n} \sum_{i=1}^n s_i$$

Рис. 3

## Предложенная схема решения задачи прогнозирования

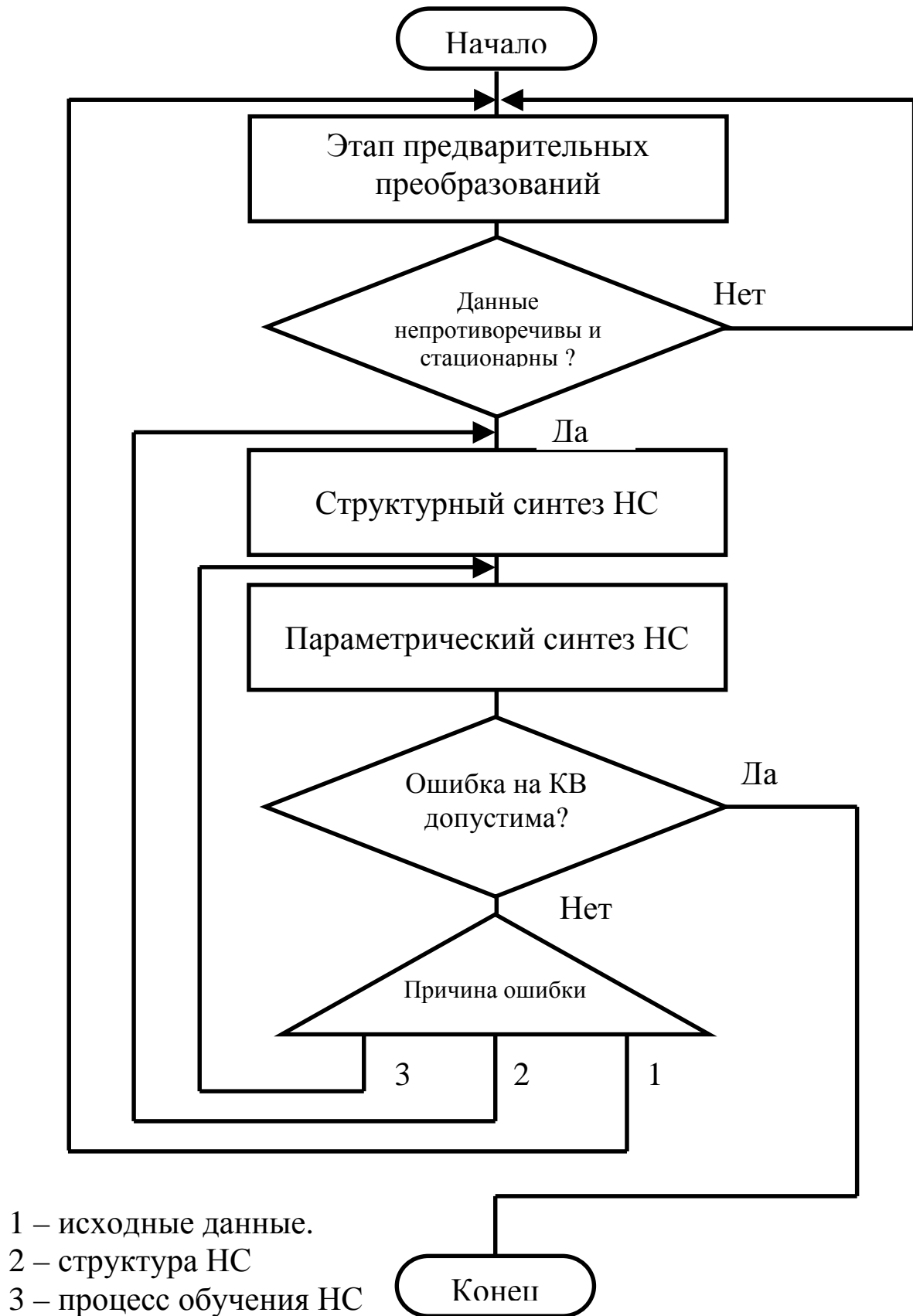
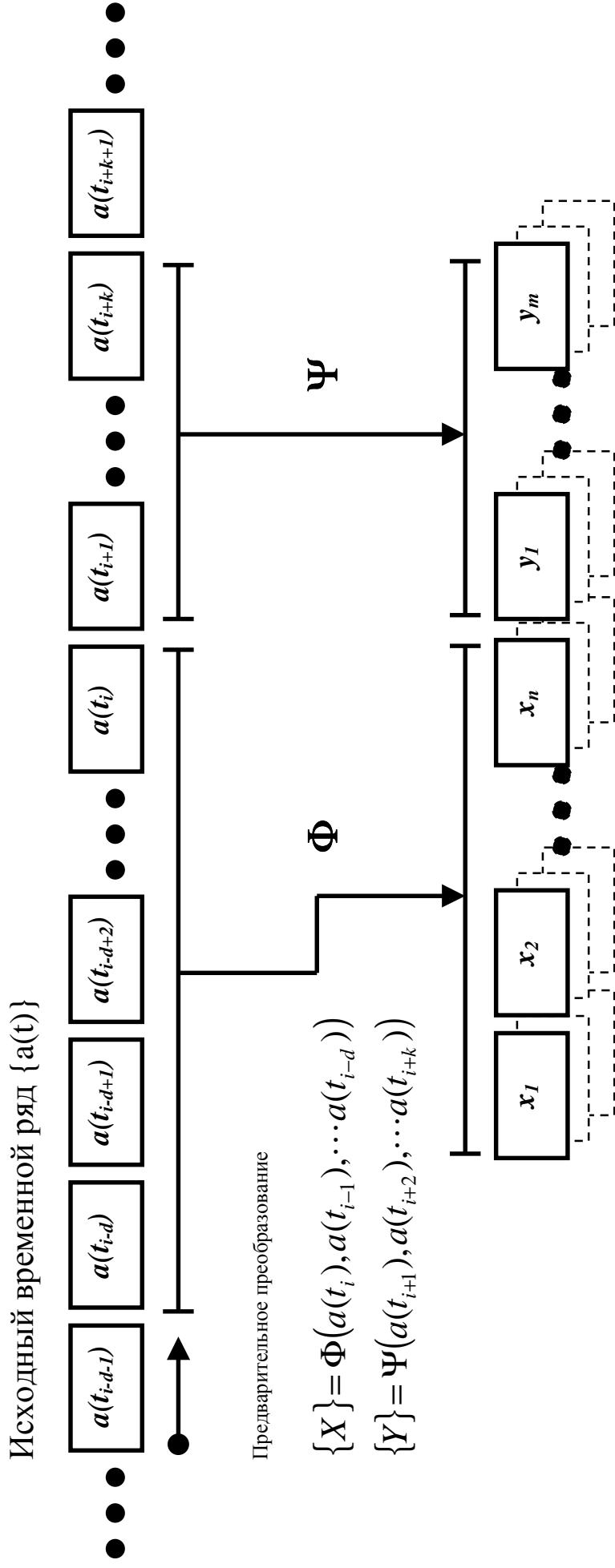


Рис. 4



## Этап предварительных преобразований



Множество формируемых наборов  $\{\{X\}, \{Y\}\}$

Рис. 2