

ЗАДАЧА ИДЕНТИФИКАЦИИ ПРИ СОЗДАНИИ ИНФОРМАЦИОННОЙ СИСТЕМЫ ДЛЯ МОНИТОРИНГОВЫХ ИССЛЕДОВАНИЙ

УДК 519.713:631.411.6

ШАРОНОВА Наталья Валерьевна

д.т.н., профессор, зав. каф. «Интеллектуальных компьютерных систем», Национальный технический университет «Харьковский политехнический институт»

Научные интересы: интеллектуальная обработка данных, искусственный интеллект

КОЗУЛЯ Мария Михайловна

магистр кафедры Компьютерного мониторинга и логистики, Национальный технический университет «Харьковский политехнический институт»

Научные интересы: интеллектуальная обработка данных, искусственный интеллект

СВЯТКИН Ярослав Владимирович

к.т.н., доцент кафедры Компьютерного мониторинга и логистики, Национальный технический университет «Харьковский политехнический институт»

Научные интересы: интеллектуальная обработка данных, искусственный интеллект

ПОСТАНОВКА ПРОБЛЕМЫ

В данной статье обсуждаются закономерности формирования знаний и интеллектуального анализа данных с целью предоставить пользователю возможность правильно ориентироваться в среде сбора, анализа, обработки и оценки информации для решения задач мониторинговых исследований в комплексных сферах деятельности, в частности экологической. Целью исследования явилось повышение эффективности и качества информационной технологии интеллектуальной обработки данных в автоматизированных информационных системах на основе информационных моделей и методов идентификации знаний.

В процессе работы над созданием современных компьютерных систем для решения задач, объединяющих различные виды деятельности, на первый план выдвигается проблема представления и извлечения знаний [1, 2]. О понятии «знание» много говорится в современной литературе, выберем в качестве рабочего определения следующее. Знание – это результат

адекватного отражения действительности человеком в виде представлений, понятий, теорий, суждений [1].

На уровне представления знаний в компьютерной системе отражены как отдельные элементы знаний, так и связи между ними. Уровень представления знаний отличается следующими особенностями: интерпретируемостью, наличием классифицирующих связей, наличием ситуативных отношений (например, одновременности хранения, совместимости знаний в машинной памяти). Для уровня знаний характерны такие признаки, как наличие специальных процедур: обобщение, наполнение имеющихся в системе знаний и т.д. Это становится особо необходимым при решении задач устойчивого развития, объединяющих в себе знания, по крайней мере, трех основных составляющих: социальной, экономической и экологической областей науки.

Актуальность этой темы определена перспективностью повышения эффективности обработки информации, исходящей из системы комплексного мониторинга. В работе проведен теоретический анализ обработки информации по состоянию сложных объектов

природно-техногенного содержания для комплексной оценки экологичности составляющих систем [3, 4, 5] и определения оптимальной модели принятия управленческого решения.

ЗАДАЧИ ИССЛЕДОВАНИЯ

В ходе теоретического анализа имеющихся результатов в этой области исследований в работе поставлены и решены такие задачи:

- 1) обоснование математических моделей компараторной идентификации характеристик природно-техногенных комплексов с определением их состояния и процессов в них по стабилизации нарушений;
- 2) формирование системы факторов регулирования экологичности систем как меры соответствия их природному гомеостазу;
- 3) возможности получения новых знаний о факторах дестабилизации качества объектов окружающей среды на примере исследования состояния техногенно-нагруженных почв.

ИССЛЕДОВАНИЕ ВОПРОСА И НАУЧНЫЕ РЕЗУЛЬТАТЫ

Согласно теории надежности [6], оптимальное управление в технической безопасности достигается при условии максимизации эколого-экономической эффективности (достижение экологического эффекта с минимальными затратами), что определяется стационарным экологическим состоянием объекта за счет аккумулирующей способности окружающей среды и возможности нейтрализации негативного влияния в ходе ее функционирования [5].

Таким образом, преобразование потенциально опасного стохастического потока информации на объект при экологической адаптации и самоорганизации систем обусловлено детерминированной функцией причинно-следственных связей в достижении устойчивого состояния при действии оператора H_w регулирования гомеостаза:

$$S_0 \xrightarrow{S_w} < S_1 \xrightarrow{q(x,t)} S_1^1 \xrightarrow{A_w} < S_2 \xrightarrow{D_w \rightarrow H_w} S_{\max} (\Delta S \rightarrow 0) \quad (1)$$

где S_w – стохастический оператор при действии естественных влияний и внешнего регулирования $F(t, u)$ в пределах природоохранных мер); $q(x, t)$ – связь между системами, вещественно-энергетический поток, A_w – оператор перехода стохастических функций в детерминированные, что определяет переход системы в новое стационарное состояние в результате самоорганизации вида; D_w – детерминированный оператор, который отвечает за функцию выхода, что приводит объект в равновесное экологическое состояние при условии стабилизации внешнего и внутреннего гомеостаза – оператор H_w схемы «вход – выход», который для системного объекта определяется достижением максимального энтропийного состояния S_{\max} и отсутствием дестабилизирующих явлений ($\Delta S \rightarrow 0$).

Модель системного образования по оценке экологической устойчивости рассматривается как кибернетический «черный ящик». Согласно выводам о возможностях существующих подходов по решению задач экологического качества [7, 8] предложено объединить метод бихевиористической идентификации – аппроксимации объективно существующей статистически переменной функциональной зависимости некоторым полиномом (априорная информация мониторинга), и метод интроспективного анализа – получение дополнительной информации по данным мониторинга «поведения» территориально-объектных систем. Оценка экологичности систем объекта исследования с включением процедуры риск-анализа включает:

- решение задач структурно-параметрической идентификации при синтезе модели, которая минимизирует несогласованность между выходом модели и реальной системы на множестве данных экспериментальной выборки (в классе аппроксимированных полиномов реализовано на практике экологических исследований состояния почв);
- процедуру проверки соответствия состояния, процессов в системах объекта требованиям безопасности на сравнении адекватности параметров модели комплексной оценки экологичности систем характеристикам природного гомеостаза – компараторная идентификация экологичности [3, 9].

Критерий оценки экологичности взаимодействия составляющих эколого-социально-экономического объекта при стабилизации социально-экономического развития представляет собой экологический компаратор (ЭК) – измеритель соответствия структуры, функциональных процессов в системах интегрального объекта естественному внутреннему гомеостазу.

Для установления уровня экологичности функционирующих природно-техногенных комплексов, территориально-объектных систем применение компараторной идентификации является более объективной оценкой, чем экспертное оценивание в баллах. Оценка экологичности принимает количественные значения 0 и 1 в виде меры соответствия, что позволяет объединить по энтропийному параметру ΔS изменения в состоянии систем и объекта и вероятности P нарушения связей в среде.

Согласно теории компараторной идентификации [7]: компаратор – измеритель соответствия системной структуры, функциональных процессов внутреннего гомеостаза интегрального объекта, представленного в виде кортежа X^s и X^t входного влияния X вида

$$F(A, X^{(s)}) = F(A, X^{(t)}), \quad s, t = \overline{1, k}, \quad s \neq t$$

Данный кортеж определяет вид термодинамического потока, представляя его как некоторое отношение R между составляющими системами в объекте согласно параметру модели A и оператору F (структура), реализуя внутренний гомеостаз и равновесие в системах. При учете однозначности вероятностно-энтропийных характеристик состояния систем оценка экологичности их определяется видом оператора F и параметров модели A , для которых несоответствие требованиям безопасности (ситуация неэквивалентности) приводит к неравенству типа

$$F(A, X^{(s)}) < (\leq) F(A, F^{(t)}), \quad s, t = \overline{1, k}, \quad s \neq t \quad (2)$$

При изучении территориально-объектных систем с общей оценкой поведения $X = (x_1, x_2, \dots, x_n)$, $x_i \in X, i = \overline{1, n}$ анализируются количественные изменения для отдельных составляющих, которые представляют как критерии качества, выраженные через частные критерии параметров измерения:

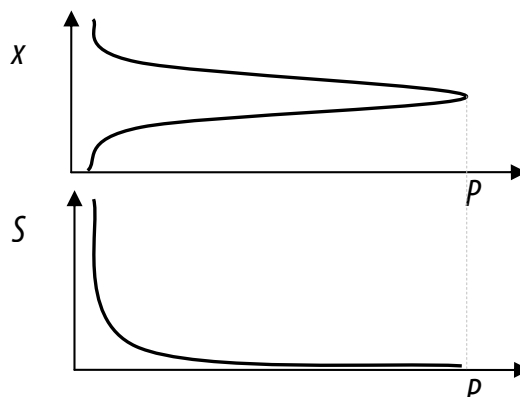
$$K(x_i) = \langle k_1(x_i), k_2(x_i), \dots, k_m(x_i) \rangle$$

Например, при исследовании состояния экосистемы рассматривают отдельно систему «почва» с целью измерения физических параметров ее состояния и химического состава как индикаторов экологичности и соответствия состоянию гомеостаза.

На основе анализа отдельных характеристик систем при определенных условиях состояния исследуемого объекта определяют вероятностную оценку реализованной структуры и ее соответствие требованиям экологической безопасности, предоставленных в виде $P: X \rightarrow V$, где P – оператор модели оценивания, а $V = P(X)$ – модель многофакторных оценок альтернатив решения ($v_i = P(x_i), i = \overline{1, n}$).

При комплексной оценке экологической безопасности предусмотрено представление данных о факторах состояния объекта, процессах внутреннего характера (в системах и между системами) и внешнего («объект – ОС») в виде $S, \Delta S, P, Risk$.

Для решения комплексной экологической задачи предлагается перейти от результатов анализа статистических наблюдений к рассмотрению формирования термодинамических потоков (процессов), которые позволяют удержать систему в состоянии равновесия ($\Delta S \rightarrow 0$) или уменьшать негативные влияния между системами за счет трансформационных преобразований ($S \rightarrow \max$), фиксировать дестабилизацию в объекте ($\Delta S \rightarrow \max$), учитывая однозначность вероятностно-энтропийных характеристик состояния систем и процессов (P, S и параметр x) (рис. 1).



а) состояние равновесия $\Delta S \rightarrow 0$

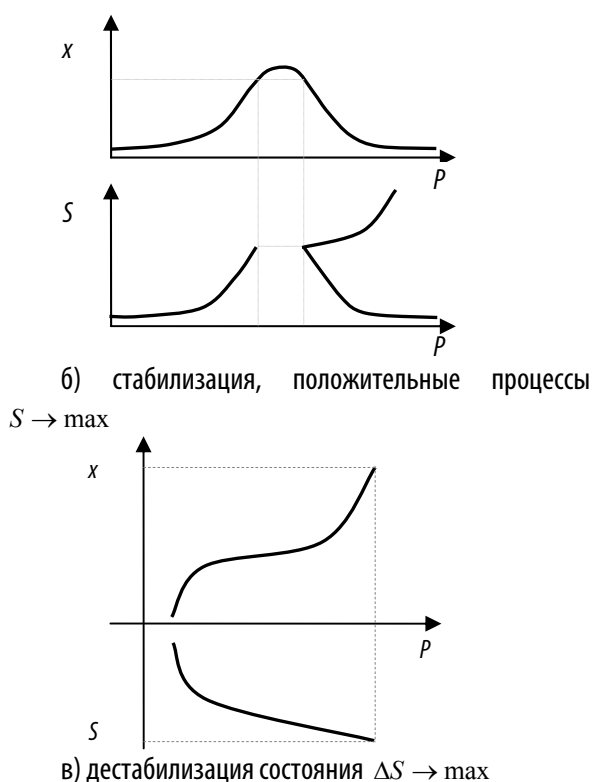


Рисунок 1 – Вероятностно-энтропийная оценка экологичности систем (P, S и параметр x)

При условии вероятностно-энтропийного оценивания экологичности с использованием элементов компараторной идентификации получают такие результаты:

– отрицательная оценка

$$k_{jn}(x_i) = \begin{cases} \max_{x_i \in X} k_j(x_i), & \text{при } k'_j(x_i) \rightarrow \min, \\ \min_{x_i \in X} k_j(x_i), & \text{при } k''_j(x_i) \rightarrow \max, \end{cases} \quad (3)$$

где $k'_j(x_i)$ – оценка фактора как вероятность стабилизационных процессов, P , положительный риск $Risk_n$; $k''_j(x_i)$ – оценка фактора по значению энтропии S или ее изменению ΔS согласно нормализованному показателю экологичности в виде индекса загрязнения (ИЗ);

– положительная оценка по принятию экологического решения:

$$k_{jn}(x_i) = \begin{cases} \max_{x_i \in X} k_j(x_i), & \text{при } k'_j(x_i) \rightarrow \max, \\ \min_{x_i \in X} k_j(x_i), & \text{при } k''_j(x_i) \rightarrow \min. \end{cases} \quad (4)$$

Функция состояния объекта определяется через естественные (A_1) и управляющие параметры (A_2). Функция $P(x_1, x'_1, x_2, x'_2, \dots, x_m, x'_m)$ – это предикат, который задается декартовым произведением $A_1 \times A_2$ как отображение произведения множества A_1, A_2, \dots, A_m на множество $\Sigma = \{0, |1\}$. Символы 0 и +1 соответствуют состоянию системы (значение энтропии или информации), при котором переменные $x_1, x'_1, x_2, x'_2, \dots, x_m, x'_m$ являются аргументами предиката (со штрихом – управляющие параметры):

$$P(x_1, x'_1, x_2, x'_2, \dots, x_m, x'_m) = K(f_1(x_1), f_2(x_2), \dots, f(x_m), f'(x'_1), f'(x'_2), \dots, f'(x'_m)) \quad (5)$$

$$, P(x_1, x'_1, x_2, x'_2, \dots) = t$$

Идентификация экологического соответствия территориально-объектной системы осуществляется экологическим компаратором состояния. Нарушение экологического гомеостаза объекта фиксируют в информационной системе мониторинга и представляют физическими данными $E_{x_1} M_{x_1}, E_{x_2} M_{x_2}, \dots, E_{x_n} M_{x_n}$, что находит отображение в виде исходных сигналов f_1, f_2, \dots, f_n по входам в компаратор ЭК (рис. 2) [3, 4].

Таким образом, определяется предикат ЭК вида:

$$P(E_{x_1} M_{x_1}, E_{x_2} M_{x_2}, \dots, E_{x_n} M_{x_n}) = \text{ЭК}(f_1(E_{x_1} M_{x_1}), f_2(E_{x_2} M_{x_2}), \dots, f_n(E_{x_n} M_{x_n})) \quad (6)$$

Внутреннее состояние системного объекта отображается зависимостями:

$$Y_1 = f_1(x_1, x'_1); Y_2 = f_2(x_2, x'_2) \quad (7)$$

При условии экологической приоритетности решения учитывают стабилизацию естественных параметров и факторы регулирующего управления, представляя предикатом

$$P(x_1, x'_1) = D(f(x_1), f(x'_1)) \quad (8)$$

Установленный экологический компаратор (рис. 2) реализует предикат

$K(Y_1, Y_2, \dots, Y_n) = t$, (9) который отвечает отношению « K - предикат компаратора» [7,8,3].

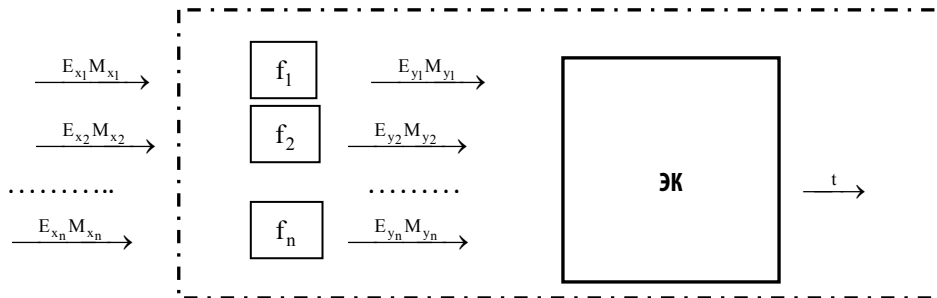


Рисунок 2 – Схема работы экологического компаратора

Если принять состояние систем при динамическом равновесии территориально-объектных систем X_1 и любое управляющее действие как влияние внешнего фактора на систему, то, реализуя анализ вида последовательности уравнений (5–8), предикат состояния исследованного системного объекта принимает вид (9).

Соответствие предиката установленным изменениям энтропии определяет состояние объекта, связанное с нарушением или сохранением упорядоченности между системами и устойчивостью экологического равновесия в целом:

$$D(Y_1, Y_1') = \begin{cases} -1, & \text{при } x_1 \neq x_1' \text{ - для энтропийного анализа процессов,} \\ \text{или} \\ 1, & \text{при } x_1 \neq x_1'; \\ 0, & \text{при } x_1 = x_1' \text{ или } x_1' = 0. \end{cases} \quad (10)$$

Таким образом учитывают уменьшение энтропии при увеличении управляющего влияния – упорядочение системы $dS < 0$; рост энтропии при увеличении управляющего влияния – неравновесное состояние $dS > 0$; неизменность энтропии – система в состоянии равновесия $dS = 0$.

Если $X_1 \geq X_1'$ или $X_1' = 0$, то система или остается в состоянии равновесия, или приходит к нему ($t = 0, -1$). При других условиях система выходит из равновесия $t = 1$:

	X_1'	A_2		
	0	1	2	3
X_1	0	0	1	1
A_1	1	0	0	1
2	0	-1	1	1
3	0	-1	-1	1

Состояние систем-системного объекта контролируется передачей сигналов, независимых от каналов внутреннего и внешнего подчинения систем и их элементов и объекта с внешней средой [3–5].

Реализация методики компараторной идентификации качества системных объектов рассмотрена на примере исследования соответствия экологическому состоянию почв ландшафтно-геохимических комплексов территории Змиевского района, которые находятся под влиянием промышленно-энергетического комплекса и соответственно загрязнены тяжелыми металлами (ТМ).

По методу компараторной идентификации (отношение (3–10)) определяем геохимическое подпространство A_1, A_2, \dots, A_m (m – размерность), для которого (x_1, x_2, \dots, x_m) составляющие факторов влияния в виде $x_1 \in A_1, x_2 \in A_2$, что представляют предметное про-

странство из катионных и анионных форм ТМ
 $A_1 = A_2 = \dots = A_m = U$:

$$U = \left\{ \underbrace{\text{Zn, Co, Ni, Pb, Sr, Cu}}_{x_1}, \underbrace{\text{Mo, Cr, V}}_{x_2} \right\}.$$

Для оценки результатов трансформации в миграционном потоке ТМ состояние системы предоставляют в виде двудольного графа. Влияние ТМ на сопредельные среды (в т.ч. и человека) учитывает снижение подвижности за счет самопроизвольных процессов образования ими нерастворимых соединений, накопление в почве и отсутствие в биоматериале, что задает множество возможных вариантов состояний. Безопасность поступления ТМ в объекты ОПС определяется посредством введения отношений: $\{x_1, x_2\}$ $m = 2$, $A_1 = \{\text{Zn, Co, Ni, Pb, Sr, Cu}\}$. $A_2 = \{\text{Mo, Cr, V}\}$, тогда $S = A_1 \times A_2$ есть множество пар вида (x_1, x_2) , для которых отношения формируются по значению энтропийного состояния, то есть анализа процессов изменений и самоорганизационных способностей системы «почва» вообще. Отношения для одного пространства являются однотипные, реализованные операциями: объединение – дизъюнкция $\vee \cup$ – или; пересечение – конъюнкция $\wedge \cap$ – и. Для анализируемого состояния допустимыми приняты те значения вероятности отклонения, количества которых от нормативно установленного ограничения находятся в пределах малого риска – 20%:

$$P(x_1 - x_n) = \begin{cases} 1, & \text{при } x \leq 0,2 \\ 0, & \text{при } x \geq 0,2 \end{cases}.$$

Энтропия при самопроизвольных процессах должна увеличиваться, что дает основание для идентификации состояния безопасности в связи с нейтрализацией миграционной подвижности загрязнителя в почве (табл. 1).

Превышение естественного содержания химических элементов в почве в связи с техногенным поступлением и «взаимодействие» их в почве с составляющими почв (сорбция) и между собой с образованием малоподвижных соединений с учетом объединения этих процессов для каждого элемента (конъюнкция $P' \cap Q$) имеет такой окончательный вид $P' = x^0 y^0 \vee x^0 y^1 \vee x^0 y^2 \vee x^0 y^3 \vee x^0 y^4$ согласно сформированной базе оценки их состояния:

Таблица 1

Определение состояния тяжелых металлов по энтропийной оценке их активности

Элемент	Энтропийное состояние	Компаратор	Cr	V
Zn	-2,265	1	1	1
Co	-1,619	0	1	1
Ni	-1,616	0	1	1
Pb	-2,042	1	1	1
Sr	-1,687	1	1	1
Cu	-1,844	1	1	1
Mo	-1,462	0	1	1
Cr	-2,639	1	1	1
V	-2,434	1	1	1

	Zn	Ni	Pb	Sr	Cu
Zn	1	1	1	1	1
Ni	0	0	0	0	0
Pb	1	1	1	1	1
Sr	1	1	1	1	1
Cu	0	0	0	0	0

По полученной совершенной дизъюнктивной нормальной форме (СДНФ): $P' = x^0 y^0 \vee x^0 y^1 \vee x^0 y^2 \vee x^0 y^3 \vee x^0 y^4$ получают такие данные

	Zn	Pb	Sr	Cr	V
Zn	0	0	0	1	1
Pb	0	0	0	1	1
Sr	0	0	0	1	1
Cr	1	1	1	0	0
V	1	1	1	0	0

Учитывая пересечения состояний Zn с соответствующим состоянием других тяжелых металлов, определено СКНФ вида

$$\begin{aligned} &x^0 y^3 \wedge x^0 y^4 \wedge x^1 y^3 \wedge x^1 y^4 \wedge \\ &\wedge x^2 y^3 \wedge x^2 y^4 \wedge x^3 y^0 \wedge x^3 y^1 \wedge. \end{aligned} \quad (11)$$

$$\wedge x^3 y^2 \wedge x^4 y^0 \wedge x^4 y^1 \wedge x^4 y^2$$

Самопроизвольные процессы накопления ТМ возможны при малой P отклонения от минимального их содержания, стабильность обеспечения «положительного состояния», определенного значительным отклонением от допустимого ограничения, объясняется присутствием элементов в почве и ходе указанных процессов:

$x^i \leq -2,3$ при $P = 0,1$ и $y^j \leq -2,99$ при $\alpha = 0,05, P = 0,95$. С учетом данных таблицы 1 и выражения (10) имеем такую оценку состояния элементов:

Zn	$x^0 = -2,78$	$1 \cdot 1 \wedge 1 \cdot 1$	1
Pb	$x^1 = -2,04$	$0 \cdot 1 \wedge 0 \cdot 1$	0
Sr	$x^2 = -2,02$	$0 \cdot 1 \wedge 0 \cdot 1$	0
Cr	$x^3 = -2,7$	$1 \cdot 1 \wedge 1 \cdot 1 \wedge 1 \cdot 1$	1
V	$x^4 = -2,46$	$1 \cdot 1 \wedge 1 \cdot 1 \wedge 1 \cdot 1$	1

На основе компараторной идентификации с учетом «состояние – процесс» благодаря применению СДНФ и СКНФ с учетом феноменологических знаний о поведении ТМ в объектах ОПС обосновывается оценка их опасности для состояния живых организмов и человека. Таким образом, вероятность нивелирования негативного влияния техногенного действия на экосистемы за счет процессов трансформации в миграционных потоках может быть установлена на основе построения интеллектуальных систем экологического мониторинга.

НАУЧНОЕ И ПРАКТИЧЕСКОЕ ЗНАЧЕНИЕ ПОЛУЧЕННЫХ РЕЗУЛЬТАТОВ

Обосновано введение компараторной идентификации состояния систем и процессов как критерия принятия управленческого решения в системе экологического мониторинга природно-техногенных объектов,

расширение возможностей по снижению риска для природных систем.

ВЫВОДЫ И ПЕРСПЕКТИВЫ ИССЛЕДОВАНИЙ

В работе предложена новая информационная основа для принятия управленческого решения на основе теоретико-практических исследований состояния сложных природно-техногенных объектов и введения компараторной идентификации, что позволило получить следующие результаты:

1) установлено общее описание функционирования природно-техногенных комплексов на основе идентификации состояния и процессов, нарушений и реализации новых состояний (1) (рис. 1);

2) определены условия эффективного сочетания факторов воздействия с целью регулирования экологичности систем с использованием положений компараторной идентификации характеристик состояния разнородных систем объекта исследования (3–10);

3) дана оценка состояния тяжелых металлов и изменений их состояния в антропогенно-нагруженных почвах, что позволило по результатам компараторной идентификации говорить о возможности снижения опасности для экосистем благодаря экологическому регулируемому воздействию трансформационных процессов при их участии (табл. 1, (11)) и новых подходах в экологическом управлении.

ЛИТЕРАТУРА:

- Gavrilova T. A. Izvlecheniye i strukturirovaniye znaniy dlya ekspertnykh sistem / T. A. Gavrilova, K. R. Chervinskaya. – М.: Radio i svyaz', 1992. – 200 s.
- Kandrashina Ye. YU. Predstavleniye znaniy o prostranstve i vremeni v sistemakh iskusstvennogo intellekta / Ye. YU. Kandrashina, L. V. Litvintseva, D. A. Pospelov. – М.: Nauka, 1989. – 328 s.
- Kozulya T. V. Modelirovaniye struktury i identifikatsiya sostoyaniya korporativnoy ekologicheskoy sistemy (KES) / T. V. Kozulya, N. V. Sharonova // Problemi informatsiy-nikh tekhnologiy. – 2007. – № 01 (001). – S. 178–187.
- Sharonova NV Entropy as Substratum of identifying the Corporative Ecological system (CES) condition / NV Sharonova, TV Kozulya // Vestnik Khersonskogo natsional'nogo tekhnicheskogo universiteta. – 2008. – № 2 (31). – S. 518–527.
- Kozulya T. V. Sistema pidtrimki priynyattya yekologichnogo rishennya v umovakh kontseptsii KES i novitnikh tekhnologiy yekologichnogo analizu / T. V. Kozulya, D. Í. Émel'yanova // Vestnik Khersonskogo natsional'nogo tekhnicheskogo universiteta. – 2010. – № 2 (38). – S. 285–293.
- Pampuro V. I. Strukturnaya informatsionnaya teoriya nadezhnosti sistem / V. I. Pampuro. – К.: Naukova dumka, 1992. – 324 s.
- Bondarenko M. F. O bionike intellekta / M. F. Bondarenko, YU. P. Shabanov-Kushnarenko // Bionika intellekta. – Khar'kov: KHNURE, 2004. – № 1 (61). – S. 3–14.
- Petrov K. E. Komparatornaya strukturo-parametricheskaya identifikatsiya modeley skalyarnogo mnogofaktornogo otsenivaniya / K. E. Petrov, V. V. Kryuchkovskiy. – Kherson: Oldi-plyus, 2009. – 294 s.
- Kozulya T. V. Opredeleniye struktury korporativnoy ekologicheskoy sistemy i identifikatsii yeye sostoyaniya / T. V. Kozulya, N. V. Sharonova // Sistemni doslidzhennya ta informatsiy-ni tekhnologii. – 2009. – № 3. – S. 117–129.