УДК 004.056.5:519.17 Кобозева А.А., Хорошко В.А.

К ВОПРОСУ ПОСТРОЕНИЯ МАТЕМАТИЧЕСКОЙ МОДЕЛИ СИСТЕМЫ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

1. Введение.

В современных условиях массового распространения средств электронной вычислительной техники, расширяющимися возможностями несанкционированных действий над информацией, необходимостью защиты не только государственной и военной, но и промышленной, коммерческой, финансовой тайн, защита информации в целом и защита информации в автоматизированных системах в частности становится все более сложной проблемой. Для ее решения каждая автоматизированная система, система информационных технологий глобальных компьютерных сетей (ГКС) (далее – Система) должна включать подсистему защиты информации (СЗИ). Для обеспечения автоматизированного учета возможности возникновения как естественных, так и искусственных каналов утечки информации, анализа результатов такой утечки, формализации процесса восстановления Системы после атаки, а также для решения многих других вопросов с привлечением современного математического аппарата необходимо создание адекватной математической модели Системы.

Большинство из построенных еще в прошлом веке математических моделей автоматизированных систем, обязанных своим существованием теории управления, основывались на «классическом» подходе, который строился на том, что положение объекта управления в пространстве признаков известно абсолютно. Однако наработанный здесь математический аппарат оказывался несостоятельным для описания объектов, которые плохо формализуются, обладают свойствами, плохо известными априори и изменяющимися в процессе функционирования. Таковой является и любая информационно-технологическая система (ИТ-система).

С конца прошлого века начал развиваться «неклассический» подход в теории управления, основывающийся на аналогиях архитектуры и целей функционирования сложных технических и биологических систем, являющихся естественными системами управления [1-6]. Подход при построении таких моделей до сих пор был ориентирован на нейросетевую элементную базу. Использование нейронных сетей при моделировании естественных систем управления, к сожалению, не обеспечивает всех выдвигаемых к моделям требований, обладающих в итоге рядом существенных недостатков [7,8]. В связи с этим в [9] предпринята попытка создания принципиально новой универсальной модели Системы, основывающейся на принципах функционирования нервной системы человека (НСЧ), без использования нейросетей. Основными математическими инструментами являются теория графов, матричный анализ и теория возмущений.

Цель настоящей работы – дальнейшее теоретическое обоснование процесса построения графово-матричной модели Системы, основные тезисы по формированию которой были выдвинуты в [9].

2. Аналогия между системой информационной безопасности и нервной системой человека

При создании модели Системы в [9] используется концепция ее разделения на управляющую и управляемую части [10]. При этом в качестве управляющей системы (УС) выступает СЗИ, моделируемая с использованием основных принципов функционирования НСЧ, являющаяся составной частью совокупной Системы, объекта управления (ОУ). Логика использования выбранной биологической системы вытекает из очевидной аналогии между НСЧ и системой информационной безопасности (рис.1). Следует учитывать, что процесс моделирования здесь носит комплексный характер и использует НСЧ, начиная с формы представления информации, программирования информационных полей и заканчивая

архитектурой ИТ-систем с встроенными механизмами обеспечения информационной безопасности и эволюционным протеканием процессов.

Моделирование защищенных информационных процессов основано на единстве представления информации в иерархии НСЧ, в которой сообщения передаются универсальным контейнером, определяемым структурированным информационным полем ДНК. Структурированный характер имеют распределеные информационные поля нейронных комплексов нервной системы, благодаря которым в НСЧ существуют адаптивные механизмы памяти, накапливающие жизненный опыт. Возможность реализации адаптивных механизмов памяти в искусственных информационных полях – основная предпосылка эволюции ИТ-систем. Программирование в НСЧ носит избыточный распределенный характер, что обеспечивает высокую функциональную устойчивость информационных процессов.

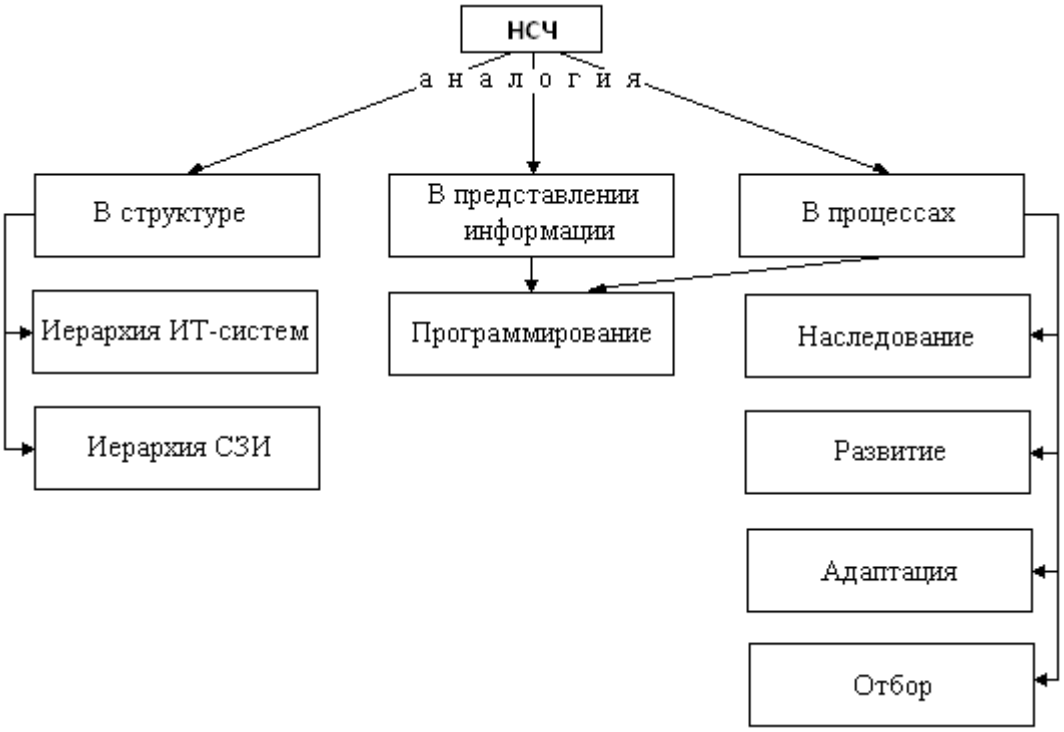


Рисунок 1. Аналогия между нервной системой человека и системой информационной безопасности

Отдельные искажения информации, с одной стороны, компенсируются избыточностью информационных полей, а с другой – позволяют реализовать механизм мутаций и эволюционные процессы развития и отбора. В частности, адаптивные процессы в информационных полях позволяют ИТ-системе развиваться и накапливать опыт в условиях расширения поля угроз, а наследование опыта в последующих реализациях системы сводится к передаче соответствующих информационных полей. Иерархия адаптивной системы информационной безопасности отражает разделение функций защиты на управляющую (проверка форм представления информации и т.д.) и управляемую, реализующую взаимодействие системы со средой (рис.2). Архитектурной особенностью НСЧ является внутренний характер механизмов защиты, реализуемый в иерархии ИТ-системы. При моделировании искусственных систем следует учитывать, что при реализации адаптивных механизмов НСЧ и информационных полей ее функции защиты информации должны быть внутренними функциями проектируемой системы.

Основные принципы организации и функционирования НСЧ, на базе которых формируется модель Системы:

1. *Автономность;*
2. *Дискретность;*

3. *Максимальная начальная приспособленность;*
4. *Минимум исходных данных.*

В качестве основной целевой функции рассматривается *выживание ОУ* [9].

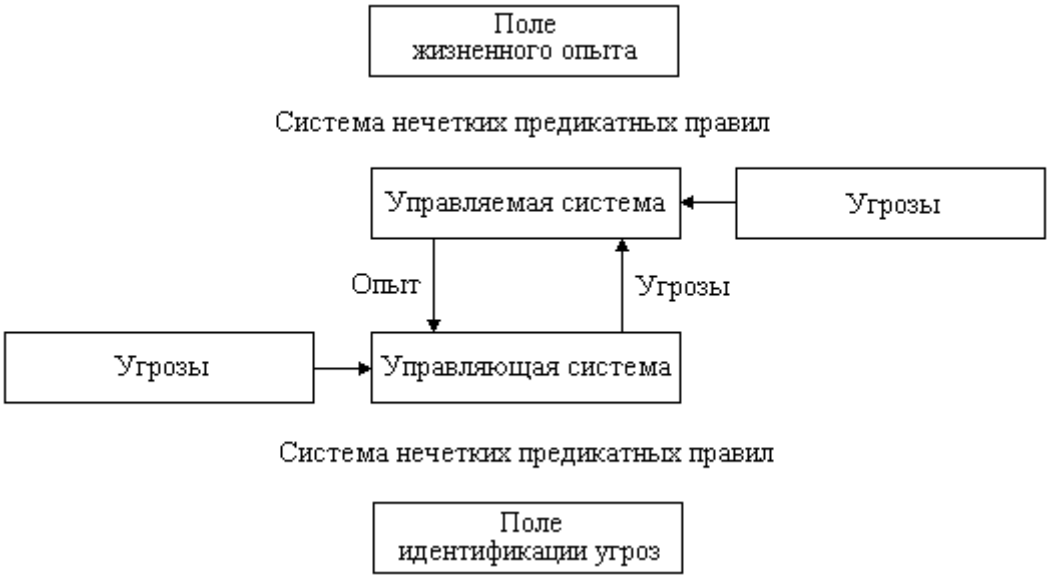


Рисунок 2. Иерархия адаптивной системы информационной безопасности.

3. Связь количества информации графа Системы и информации Системы

В качестве модели ОУ в [9] используется взвешенный неориентированный граф со структурным отношением «состоять из», что позволяет не только достаточно легко удовлетворить первым трем основным принципам НСЧ, перечисленным выше, но и обеспечить ее иерархию. Математическая локализация информации, циркулирующей в Системе, при выбранном способе построения модели осуществляется в возмущениях максимальных собственных значений и соответствующих собственных векторов матрицы смежности графа Системы. Возмущения происходят при установлении связи между СЗИ и непосредственно защищаемой информацией за счет появления нового ребра, соединяющего граф СЗИ с вершиной, соответствующей Системе в целом [9]; информация «вводится» в систему информационной защиты. В соответствии с теорией информации покажем, что такой «ввод» приводит к *увеличению количества информации графа-модели*. Для этого необходимо определить вероятностную схему графа (probability scheme). Этот вопрос не является тривиальным, поскольку до настоящего момента не существует единой вероятностной схемы для определения энтропии и количества информации произвольного графа. Долгое время традиционным в этом вопросе был взгляд, высказанный в [11], и развитый в [12], для простоты изложения которого введем некоторые понятия теории графов [13].

Два графа X_1 и X_2 называются *изоморфными*, если между их множествами вершин существует взаимно однозначное соответствие, сохраняющее смежность (изоморфизм является отношением эквивалентности).

Автоморфизмом помеченного графа X называется изоморфизм графа на себя. Каждый автоморфизм α графа X есть подстановка его множества вершин V, для которого $|V| = n$, сохраняющая смежность. Множество всех автоморфизмов графа образуют группу [12,13]. Будем обозначать группу автоморфизмов X как $G(X)$ (группа подстановок на множестве V). Для всякой вершины $v \in V$ *орбитой* этой вершины называется подмножество множества V, состоящее из всех таких элементов $w \in V$, что $\alpha v = w$ для некоторой подстановки

$\alpha \in G(X)$. Пусть A_i, $i=\overline{1,h}$ - различные орбиты $G(X)$. Тогда $A_i \cap A_j = \varnothing$, если $i \neq j$, а $\bigcup_{i=1}^{h} A_i = \{1,\ldots.n\}$, т.е. орбиты формируют разбиение множества $\{1,\ldots.n\}$ [12].

В [12] строится вероятностная схема P_x:

$$P_x = \begin{pmatrix} A_1, A_2, \ldots, A_h \\ p_1, p_2, \ldots, p_h \end{pmatrix}$$

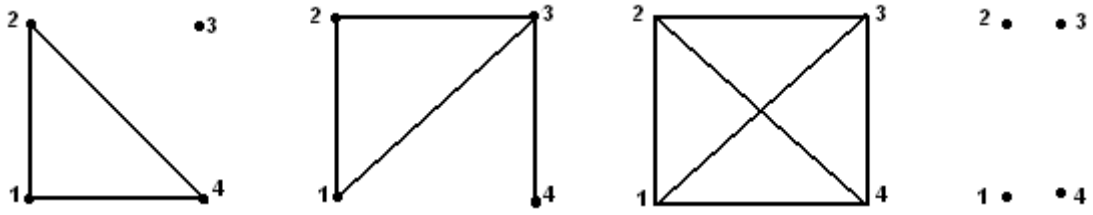
где $p_i = \dfrac{|A_i|}{n}$, $i=\overline{1,h}$. Тогда энтропия графа $I_g(X)$ определяется как энтропия P_x:

$$I_g(X) = -\sum_{i=1}^{h} p_i \log_2 p_i \,. \qquad (1)$$

Выражение (1) служит также для определения *меры сложности графа*.

Однако приведенная вероятностная схема для определения энтропии и сложности графа, основанная на использовании орбит, при подробном ее рассмотрении обладает серьезным недостатком, не отражая для некоторых графов реальную картину их сложности, а потому, очевидно, ее использование нежелательно для графовой модели Системы. Для наглядности изложения поясним это на примерах.

Для простоты рассмотрим графы, для которых $|V| = 4$ (рис.3).



а) $G(X) = \{\{1,2,3,4\}, \{2,1,3,4\}, \{4,2,3,1\}, \{1,4,3,2\}, \{2,4,3,1\}, \{4,1,3,2\}\}$,
$A_1 = \{1,2,4\}$,
$A_2 = \{3\}$

б) $G(X) = \{\{1,2,3,4\}, \{2,1,3,4\}\}$,
$A_1 = \{1,2\}$,
$A_2 = \{3\}$,
$A_3 = \{4\}$

в) $G(X) = \{мн-во\ всех\ подста-новок\ для\ \{1,2,3,4\}\}$,
$A_1 = \{1,2,3,4\}$

г) $G(X) = \{мн-во\ всех\ подста-новок\ для\ \{1,2,3,4\}\}$,
$A_1 = \{1,2,3,4\}$

Рисунок 3. Множества орбит групп автоморфизмов различных графов

В соответствии с (1) меры сложности (энтропии) графов, представленных на рис.3, соответственно равны:

а - $I_g(X) = 2 - \dfrac{3}{4}\log_2 3$; б - $I_g(X) = \dfrac{3}{2}$; в - $I_g(X) = 0$; г - $I_g(X) = 0$.

Как видно, сложность полного и нуль-графа равны нулю (положение не изменится при изменении мощности множества вершин), хотя очевидно, что сложность полного графа должна быть больше.

В литературе были предложены другие подходы при построении вероятностной схемы для определения энтропии графа, один из которых был описан в [14], использовался авторами при изучении молекулярных процессов, был развит в дальнейшем в [15] и состоит в следующем.

Пусть граф представляется некоторыми N элементами (например, вершинами, ребрами, расстояниями, кликами и т.д.), каждому элементу присвоен вес $w_i, i = \overline{1, N}$. Вероятность того, что случайно выбранный элемент i имеет вес w_i определяется как

$$p_i = \frac{w_i}{\sum_{i=1}^{N} w_i},$$

при этом $\sum_{i=1}^{N} p_i = 1$. Вероятностная схема графа имеет вид:

$$\begin{array}{ll} \textit{Элемент} & 1, \quad 2, \quad \ldots \quad N \\ \textit{Вес} & w_1, w_2, \ldots \quad w_N \\ \textit{Вероятность} & p_1, \ p_2, \ldots \quad p_N \end{array} \qquad (2)$$

и дает возможность определить целый ряд информационных индексов, учитывая (1).

Рассмотрим в качестве элементов графа его вершины. Пусть вес каждой вершины соответствует, например, ее степени. Здесь нулевая сложность нуль-графа будет отличаться от достаточно высокой сложности полного графа, каждая из которых будет получена из уравнения (1), но в соответствии с вероятностной схемой (2) (вероятность того, что случайно выбранная вершина i в полном графе из N вершин имеет степень a_i, равна $p_i = \frac{1}{N}$).

Шеннон определяет информацию как уменьшение энтропии системы относительно максимальной энтропии, которая может существовать в системы с таким же числом элементов:

$$I = H_{\max} - H \ . \qquad (3)$$

Используя (1), вычислим энтропию графа с общим весом $W = \sum_{i=1}^{N} w_i$ и весами вершин w_i для вероятностной схемы (2):

$$\begin{aligned} H = -\sum_{i=1}^{N} p_i \log_2 p_i = -\sum_{i=1}^{N} \frac{w_i}{W} \log_2 \frac{w_i}{W} = -\sum_{i=1}^{N} \frac{w_i}{W} \log_2 w_i + \sum_{i=1}^{N} \frac{w_i}{W} \log_2 W = \\ = \log_2 W - \frac{1}{W} \sum_{i=1}^{N} w_i \log_2 w_i \end{aligned} \qquad (4)$$

Как следует из (4), максимальное значение энтропии определяется как

$$H_{\max} = \log_2 W \qquad (5)$$

Подставляя (5) и (4) в (3), получим выражение для количества информации взвешенного графа:

$$I = \frac{1}{W}\sum_{i=1}^{N} w_i \log_2 w_i \,. \qquad (6)$$

Если в качестве веса вершины используется ее степень, то можно показать [16], что количество информации графа (сложность графа) может быть оценено в соответствии с формулой:

$$I = \sum_{i=1}^{N} a_i \log_2 a_i \;\;, \qquad (7)$$

полученной из (6) с использованием вероятностной схемы (2). Здесь a_i - степени вершин графа.

Поскольку «ввод» информации в СЗИ происходит при введении в граф Системы дополнительного ребра, т.е. непосредственно связано с изменением степеней вершин, воспользуемся для оценки изменения количества информации графа Системы при его окончательном формировании соотношением (7). Система в целом сконцентрирована в графе с единственной вершиной 1 (на рис.4 номера вершин отмечены красным цветом) нулевой степени. Количество информация такого графа в соответствии с (7) равно 0: пока о Системе неизвестно ничего, ее энтропия максимальна. Добавление ребра $<1,2>$ (на рис.4 ребро отмечено штриховой линией) вместе с подграфом, отвечающим СЗИ, вводя информацию в СЗИ, приведет к «появлению информации» и о самом графе: количество информации графа Система станет ненулевым в соответствии с (7). Таким образом, очевиден следующий вывод.

Вывод. При предложенном способе построения графово-матричной модели Системы [9], количество информации графа Системы будет определяться информацией, хранимой в Системе.

Замечание 1. Каждый последующий шаг детализации графовой модели Системы увеличивает количество информации графа, что очевидно говорит в пользу адекватности модели.

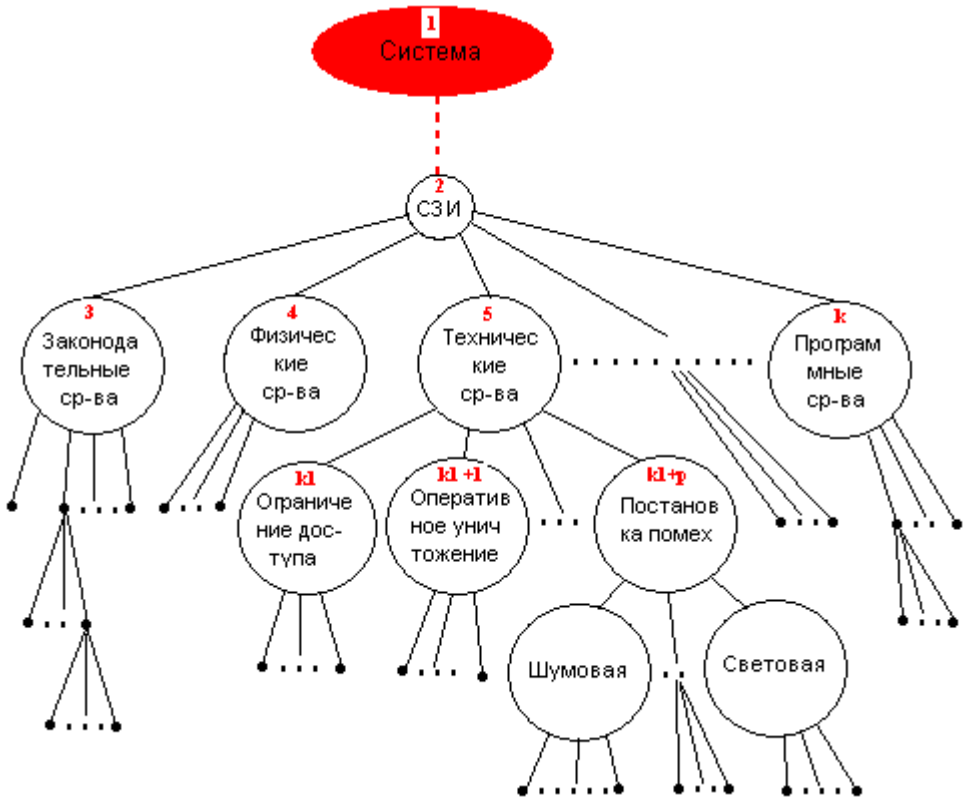


Рис.4. Помеченный граф Системы

Замечание 2. Можно показать, что введение информации за счет ребра $<1,2>$ приведет также к росту линейной сложности графа, выражаемой через линейную сложность его матрицы смежности, которая, в свою очередь, определяется количеством арифметических операций, необходимых для вычисления произведения

$$\overline{MS}\,x\,,$$

где $\overline{MS}$ - матрица смежности графа Системы, x - произвольный вектор соответствующей размерности [17].

Замечание 3. Введение информации в СЗИ увеличит значение связности (connectedness) графа Системы, определяемой как

$$Conn = \frac{2|E|}{|V|^2}$$

где E - множество ребер, что повлечет за собой возрастаное *реберной сложности* графа Системы [16].

4. Заключение

На основании проделанной работы можно утверждать, что предлагаемый новый подход к вопросу моделирования Системы в целом и СЗИ в частности, основанный на очевидной аналогии между НСЧ и системой информационной безопасности, использующий в качестве математических инструментов теорию графов, матричный анализ, теорию возмущений, является чрезвычайно перспективным.

Выбранный способ построения графово-матричной модели ИТ-системы обеспечивает ее согласованность с теорией информации.

Литература:

1. Жданов А.А. Об одном имитационном подходе к адаптивному управлению. Сб. «Вопросы кибернетики». Научный совет по комплексной проблеме «Кибернетика» РАН. М., 1996, с. 171-206.
2. Жданов А.А. Метод автономного адаптивного управления. – Известия Академии Наук. Теория и системы управления, 1999, № 5, с.127-134.
3. Жданов А.А., Крыжановский М.В., Преображенский М.Б. Нейронная адаптивная система управления. Труды международной конференции «Интеллектуальные и многопроцессорные системы» IMS’2002,с. 115-118.
4. Жданов А.А. Метод автономного адаптивного управления, его свойства и приложения. Перспективные информационные технологии и интеллектуальные системы, с.4-14.
5. Осовецкий Л.Г., Нестерук Г.Ф., Бормотов В.М. К вопросу иммунологии сложных информационных систем. Изв.вузов. Приборостроение. 2003, т.46, №7, с.34-40.
6. Нестерук Г.Ф., Осовецкий Л.Г., Нестерук Ф.Г. Адаптивная модель нейросетевых систем информационной безопасности. Перспективные информационные технологии и интеллектуальные системы, с.14-16.
7. Архипов А., Ишутин А. Применение моделей обнаружения аномалий для выявления атак // Четверта науково-технічна конференція. Правове, нормативне та метрологічне забеспечення системи захисту інформації в Україні. Тези доповідей. – 2006 .- с. 71-72.
8. Хорошко В.А., Терейковский И.А. Использование искусственных нейронных сетей в задачах распознавания атак на компьютерные системы. – Науково-технічний журнал «Захист інформації». – 2006. - № 3. – С. 57-65.

9. Кобозева А.А., Хорошко В.А. Модель системы защиты информации, основанная на принципах естественной системы управления. – Вісник ДУІКТ. – 2007. - №3. – С.
10. Лищук В.А. Математическая теория кровообращения. – М.: Медицина, 1991. – 256с.
11. Trucco E. On the Information Content of Graphs: Compound Symbols; Different States for each Point. Bull. Math. Biophys. 18, 1956, pp. 237-253.
12. Mowshowitz A. A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the University of Michigan, 1967.
13. Харари Ф. Теория графов. – М.: Мир, 1973. 300с.
14. Bonchev D., Trinajstic N. Chemical Information Theory, Structural Aspects. Intern. J. Quantum Chem. Symp. 1982. – 16. – pp. 463-480.
15. Bonchev D. Information-Theoretic Indices for Characterization of Chemical Structures. Research Studies Press, Chichester, UK, 1983.
16. Bonchev D., Buck G.A. Quantitative Measures of Network Complexity // Chapter 5 in “Complexity in Chemistry, Biology, and Ecology”. – Springer US. - 2005.
17. Neel D.L., Orrison M.E. The Linear Complexity of a Graph. Mathematics Subject Classification: 05C85, 68R10. - 2006.

Информационные технологии и компьютерная инженерия, №3,2007, С.164-171.

АНОТАЦІЯ

УДК 004.056.5: 519.17

Кобозєва А.А., Хорошко В.О. До питання побудови математичної моделі системи інформаційної безпеки.

Робота є наступним кроком на шляху теоретичного обгрунтування процесу побудови графово-матричної моделі системи інформаційної безпеки.

Літ. 17 назв., іл. 4.

Автор:

Кобозєва Алла Анатоліївна – канд.фіз.-мат.наук, доцент каф. прикладної математики Одеського національного політехнічного університету.