УДК 004.056

П.А. Глыбовский, А.П. Глухов, Ю.А. Пономарев, М.В. Шиленков

ПОДХОД К ОЦЕНИВАНИЮ И ПРОГНОЗИРОВАНИЮ УРОВНЯ ЗАЩИЩЕННОСТИ ИНФОРМАЦИОННЫХ И ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМ

Глыбовский П.А., Глухов А.П., Пономарев Ю.А., Шиленков М.В. **Подход к оцениванию и прогнозированию уровня защищенности информационных и телекоммуникационных систем.**

Аннотация. В статье рассматриваются подходы к долгосрочному прогнозированию количественных и качественных показателей подсистемы защиты информационных и телекоммуникационных систем, оценивается возможность их применения для анализа защищенности систем от несанкционированного доступа.

Ключевые слова: система оценки защищенности, политика безопасности, модель безопасности, подсистема защиты, методы самоорганизации.

Glybovsky P.A., Gluhov A.P., Ponomarev Yu.A., Shilenkov M.V. **Approach to the Evaluation and Prediction of the Level of Security of Information and Telecommunications Systems.**

Abstract. The article discusses approaches to long-term forecast of quantitative and qualitative indices of the security subsystem of information and telecommunication systems.The possibility of their use for the analysis of protection of systems against unauthorized access is assessed.

Keywords: system security assessment; security policy; security model; security subsystem; methods of self-organization.

1. Введение. В настоящее время при оценке уровня безопасности корпоративной информационной инфраструктуры применяются как методики отечественных и международных стандартов, ведомственных руководящих документов, так и методы прогнозирования ситуаций и оценки требуемых значений показателей уровня защищенности в совокупности с оценками экономической эффективности инвестиций в обеспечение безопасности и защиту информации [1,2].

К основным принципам создания системы оценки защищенности информационной инфраструктуры и анализа прогнозов показателей уровня защищенности относятся следующие:

- комплексное оценивание соответствия типовым требованиям руководящих документов Российской Федерации, международных стандартов ISO;
- количественное оценивание текущего уровня безопасности, задание допустимых уровней рисков;
- выявление и блокирование наиболее опасных уязвимостей до осуществления атак на уязвимые ресурсы;
- определение функциональных отношений и зон ответственности при взаимодействии подразделений и лиц по обеспечению информационной безопасности;

- обеспечение поддержания внедренного комплекса защиты, регулярные доработки организационно-распорядительной документации, модификация технологических процессов и модернизация технических средств защиты.

Следует отметить, что количественное оценивание подразумевает проведение инструментальных исследований информационной инфраструктуры на наличие уязвимостей [2].

2. Процесс оценки уровня защищенности ИТКС. Процесс анализа, контроля и оценки защищенности компонент информационных и телекоммуникационных систем (ИТКС) представляет собой процесс выполнения во времени некоторой совокупности отдельных процедур, связанных с получением информации об объекте оценки и проведением собственно требуемых оценок защищенности на предмет удовлетворения состояния соответствующей компоненты ИТКС каким-либо требованиям по защищенности (безопасности) данной компоненты (рисунок 1). Каждая из процедур выполняется по соответствующей методике, и в результате чего формируются показатели уровня защищенности, которые подвергаются прогнозированию.

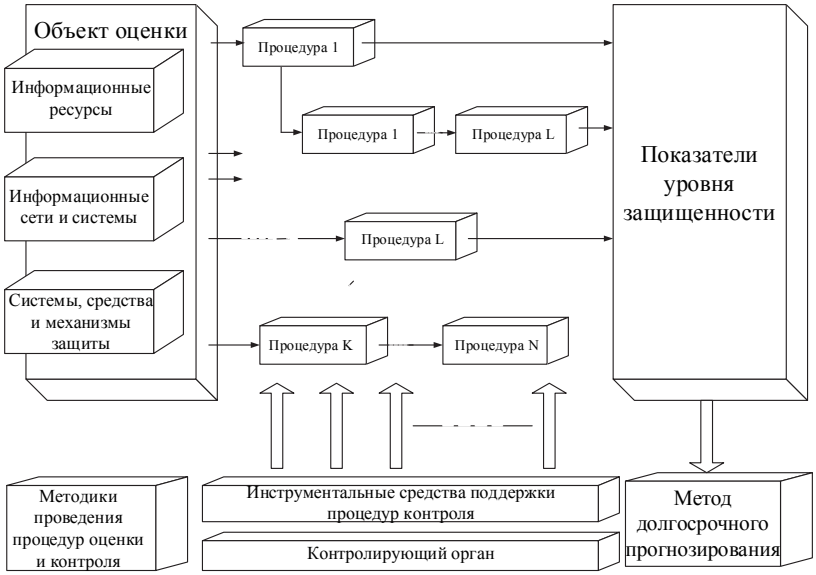

Рисунок 1 – Процесс оценки уровня защищенности ИТКС

Прогнозирующей моделью называют уравнение (или систему уравнений), описывающее предполагаемый ход моделируемого про-

цесса в будущем [3]. Соответствующие уравнениям зависимости, представленные в виде таблиц или графиков, являются прогнозом [4].

При использовании методов самоорганизации прогнозирующие модели строятся на основе опытных данных (по предыдущим наблюдениям процесса или, иначе, по его предыстории) [4]. Обычно такая информация представляется в таблицах через равные отрезки времени (дискретные интервалы или шаги). Задачей прогнозирования процесса (или временного ряда) является определение его систематической составляющей, называемой трендом [5,6]. Общие тенденции изменения временных рядов могут иметь регулярный (например, возрастающий) или циклический характер и описываться, соответственно, как полиномиальными, так и циклическими трендами. В ряде случаев тренд может являться функцией не только времени, но и ряда других наблюдаемых величин и параметров (при этом он применяется для построения так называемых нормативных прогнозов), флуктуации и нерегулярные составляющие временных рядов включают в себя ошибки измерения, а также отклонения от тренда, вызываемые воздействием не учитываемых, изменчивых факторов.

3. Критерии селекции. При синтезе трендов оптимальной сложности организуется перебор большого числа трендов-претендентов по критериям [7,8]:

- регулярности;
- минимума смещения;
- баланса переменных;
- комбинированным критериям.

Критерий регулярности. Проверочная последовательность используется в методе группового учета аргументов (МГУА) для выбора числа членов и степени уравнения регрессии: это позволяет получить наиболее регулярное решение (оптимальной сложности), малочувствительное к небольшим изменениям исходных данных. Такое решение дает небольшую ошибку на новых точках, в чем и состоит основная цель МГУА в задачах однократного (краткосрочного) прогноза и, иногда, идентификации. Величина критериев регулярности вычисляется по данным отдельной проверочной последовательности.

Разделение исходной выборки на две последовательности - обучающую и проверочную - производится по одному из следующих приемов:

Опытные точки ранжируются в ряд по дисперсии (квадрат расстояния от центра «тяжести» пространственной фигуры, охватывающей все представляющие точки):

$$D_i^2 = \sum_{i=1}^{n} \left(x_{ij} - \bar{x}_i \right)^2 , \qquad (1)$$

где $j=1,2,...,N$; N – общее число точек; n – число аргументов;

$$\bar{x}_i = \frac{1}{N} \sum_{j=1}^{N} x_{ij} .$$

В этом случае для определения коэффициентов частных описаний используются возможно более удаленные точки.

После ранжировки по дисперсии можно произвести дополнительное разделение точек на нечетные - четные. В таком случае первая и вторая части последовательности имеют близкие статистические свойства.

Если исходные данные снимаются в строгой временной последовательности (например, опытные данные физического поля или технологического процесса), то ранжировка точек в ряд недопустима [9]. В этом случае можно применить деление точек по четным и нечетным индексам [10]. После разделения по тому или другому способу первые N_A точек составляют обучающую последовательность, следующие N_B точек - проверочную.

В качестве критериев регулярности используются:
относительная среднеквадратичная ошибка:

$$\Delta^2(B) = \frac{\sum_{i=1}^{N_B} \left(q_i^* - q_i \right)^2}{\sum_{i=1}^{N_B} \left(q_i^* \right)^2} \to \min , \qquad (2)$$

коэффициент корреляции:

$$K_{qq}^* = \frac{\sum_{i=1}^{N_B} q_i^* q_i}{\sqrt{\sum_{i=1}^{N_B} (q_i^*)^2 \sum_{i=1}^{N_B} q_i^2}} \to 1 , \qquad (3)$$

индекс корреляции:

$$\rho = \sqrt{1 - \frac{\sum_{i=1}^{N_b} (q_i^* - q_i)^2}{\sum_{i=1}^{N_b} (q_i^* - \overline{q}_i^*)^2}} \to 1, \qquad (4)$$

где $\Delta(B)$ - относительная среднеквадратичная ошибка; q_i, q_i^* - полученное по модели и действительное значение выходной величины в i-й точке ($t = 1, 2, \ldots \ldots, N_B$); N_B - число точек проверочной последовательности:

$$\overline{q}_i^* = \frac{1}{N_B} \sum_{i=1}^{N_B} q_i^* .$$

Величины $\Delta^2(B)$, K_{qq^*}, ρ вычисляются на проверочной последовательности B. Чем меньше величина относительной среднеквадратичной ошибки или чем больше величина коэффициента корреляции или индекса корреляции, тем модель более регулярна.

Значения коэффициента корреляции и индекса корреляции в меньшей степени зависят от конкретного процесса и более подходят для сравнения результатов моделирования, чем значение среднеквадратической ошибки на проверочной последовательности.

Критерий минимума смещения. Модели, найденные по минимуму критерия регулярности, хорошо решают задачу краткосрочного прогноза при малом уровне шумов в исходных данных. При увеличении помех минимум критерия регулярности смещается влево, в сторону выбора более простых моделей, и при большой интенсивности шумов оптимальными оказываются линейные прогнозирующие модели.

Более устойчивыми к помехам являются различные формы критерия минимума смещения, так как, хотя при увеличении помех минимальное значение показателя смещения тоже увеличивается, минимум получается при одной и той же сложности модели, отвечающей истинной модели (это объясняет, почему такой критерий рекомендуется для решения задачи открытия законов). При малых шумах в исходных данных критерий регулярности и критерий несмещенности равноценны и приводят обычно к одинаковым результатам.

Общая, основная идея построения критерия минимума смещения состоит в том, что модели, получаемые при использовании различных частей таблицы исходных данных, должны по возможности мало отличаться друг от друга, а при полном отсутствии смещения -

совпадать. Это означает, что если исходные данные разбиты на две части, *A* и *B*, то модель, полученная на множестве точек *A*, должна возможно меньше отличаться от модели, полученной на множестве точек *B*.

Для расчета показателя смещения все имеющиеся экспериментальные данные делятся на две части *A* и *B*. Сначала первая последовательность данных является обучающей, а вторая проверочной. Уравнения регрессии (частные модели), получаемые при этом, обозначим $q(A) = f(x_i,x_j)$. Далее первая последовательность служит проверочной, а вторая — обучающей, что выражается уравнением регрессии (частной моделью) $q(B) = f(x_i,x_j)$. Близость этих уравнений оценивается по величине среднеквадратического расхождения их выходов, рассчитанных по всем табличным точкам:

$$n_{см}^2 = \frac{\sum_{i=1}^{N}\left[q_i(A) - q_i(B)\right]^2}{\sum_{i=1}^{N} q_{1табл}^2}. \qquad (5)$$

В многорядных алгоритмах МГУА эта величина, называемая показателем смещения, используется для выбора *F* лучших частных описаний, пропускаемых на последующие ряды селекции. В качестве суммарной характеристики смещения частных моделей ряда используется среднее значение показателей смещения самых несмещенных уравнений:

$$N_{см} = \frac{1}{F}\sum_{i=1}^{F} n_{см_i}^2. \qquad (6)$$

В некоторых программах вместо $N_{см}$ используется одно, лучшее значение $n^2_{смi}$ Ряды селекции наращиваются до тех пор, пока величина $N_{см}$ уменьшается.

В комбинированных алгоритмах МГУА значения $n^2_{см}$ рассчитываются для всех моделей, заданных в таблице постепенного усложнения, и в результате выбирается самая несмещенная модель.

В некоторых алгоритмах применяются следующие разновидности показателя смещения, основанные на анализе решений:

$$n_{см}^2 = \frac{1}{N}\sum_{i=1}^{N}[q_i(A) - q_i(B)]^2 \to 0, \tag{7}$$

$$n_{см}^2 = 2\frac{\sum_{i=1}^{N} q_i(A)q_i(B)}{\sum_{i=1}^{N}(q_i^2(A) + q_i^2(B))} \to 1. \tag{8}$$

В том случае, когда все особым образом нормированные аргументы частных моделей МГУА имеют примерно одинаковые максимальные и минимальные значения, для оценивания частных описаний можно воспользоваться критерием минимума смещения коэффициентов:

$$n_{см}^2 = \frac{\sum_{i=1}^{N}(a_i - b_i)^2}{\sum_{i=1}^{N}(a_i^2 - b_i^2)}. \tag{9}$$

Возможны и другие формы этого критерия, однако соответствующий показатель смещения, как функция параметров модели, должен быть безразмерным, симметричным и малочувствительным к различию в малых коэффициентах, например,

$$n_{см}^2 = 2\frac{\sum_{i=1}^{N} a_i b_i}{\sum_{i=1}^{N}(a_i^2 + b_i^2)} \to 1. \tag{10}$$

В ортогонализованных алгоритмах МГУА критерий минимума смещения коэффициентов упрощается. Действительно, в этом случае частные модели содержат всего один коэффициент, смещение которого может быть вычислено по формуле:

$$n_{см}^2 = \frac{(a-b)^2}{a^2 + b^2} \to 1. \tag{11}$$

Пусть $f(q_1(t), q_2(t), ..., q_S(t)) = 0$ – функция баланса (т. е. «закон», связывающий переменные $q_1(t), q_2(t), ..., q_S(t)$). Из множества всех прогнозирующих моделей для переменных *q(t)* нужно выбрать такую

систему, для которой на интервале экстраполяции (в районе точки прогноза) эта связь выполняется наилучшим образом. Нарушение баланса переменных можно характеризовать величинами небаланса

$$b_i = f(q_1(t), q_2(t),..., q_S(t)), \qquad (12)$$

которые рассчитываются для *f*, принадлежащих интервалу экстраполяции, по различным прогнозирующим моделям для каждой переменной *qi (t)*.

В результате выбирается наилучшая комбинация моделей. Их усложнение удобнее всего задавать в виде специальных таблиц постепенного усложнения.

Критерий баланса позволяет выбрать наилучший прогноз системы переменных набора так называемых «вееров» прогнозов (трендов) для каждой переменной.

В некоторых задачах функция связи между переменными ясна из физических соображений и использование ее в алгоритме является вполне естественным. В других случаях из физики процесса можно судить о существовании функциональной зависимости между переменными, однако вид ее неизвестен. При этом рекомендуется предварительно восстановить функцию баланса, используя один из алгоритмов МГУА на основе критерия минимума смещения.

В том случае, когда прогнозируемые переменные признаны несвязанными, возможно искусственное расширение множества переменных за счет дополнения их временными рядами, образованными из исходных по выбранным законам. Эти зависимости и выступают в дальнейшем в качестве функций баланса. При этом линейные функции баланса приводят к неоднозначности выбора модели, в то время как нелинейные – обеспечивают более надежную селекцию.

Если взаимосвязь между переменными существует, то на участке интерполяции могут быть построены несмещенные законы:

$$\begin{aligned} q_1 &= f_1(q_2, q_3, ..., q_S); \\ q_2 &= f_2(q_1, q_3, ..., q_S); \\ &\cdots \\ q_S &= f_S(q_1, q_2, ..., q_{S-1}); \end{aligned} \qquad (13)$$

Допустим, что функции (13), называемые прямыми, достаточно просты и допускают «обращение», т. е. выражение некоторых аргу-

ментов, входящих в правую часть, через остальные. Тогда из системы уравнений (13) можно получить:

$$\begin{aligned} q_1 &= f_{1обр}(q_2, q_3, ..., q_S); \\ q_2 &= f_{2обр}(q_1, q_3, ..., q_S); \\ &\dots \\ q_S &= f_{Sобр}(q_1, q_2, ..., q_{S-1}); \end{aligned}, \quad (14)$$

причем *t-я* обратная функция находится из прямой ($i \neq i$).

Из всех возможных вариантов построения обратной функции следует предпочесть тот, при котором она находится из самой несмещенной прямой (исходной) функции. Величины $b_{1i}=(f_1 - f_{1обр})^2{}_i$, $b_{2i}=(f_2 - f_{2обр})^2{}_i, \dots, b_{Si}=(f_S - f_{Sобр})^2{}_i$ - рассчитанные на интервале экстраполяции (t_i принадлежит интервалу прогнозирования) по прогнозирующим моделям (трендам) для каждой переменной, характеризуют небаланс системы прогнозов в точке, а величина:

$$B = \sum_{i=1}^{m} \left(\frac{\sum_i b_{ij}}{\sum_i f_i^2(i)} \right), \quad (15)$$

используется для выбора системы моделей.

Критерий баланса. Среди комбинаций трендов (по одному из «веера» каждой переменной), дающих малое, близкое к нулю значение критерия баланса переменных, обязательно находится одна комбинация, отвечающая действительному ходу процесса, так как закон, связывающий переменные, постоянен во времени. Однако среди таких комбинаций могут быть и ложные, не отвечающие ходу процесса, поэтому необходимо выделить «истинный» баланс среди нескольких «ложных». Для решения этой задачи рекомендуется использовать дополнительные ограничения (критерий «физической реализуемости прогноза» и селекцию модели по двум целесообразно выбранным критериям).

При реализации критерия баланса в виде «кольца разностей прямых и обратных функций» необходимо иметь в виду, что в случае тождественности f_j и $f_{jобр}$. Величина B не позволяет сделать однозначный выбор модели, поэтому рекомендуется переход к показателю:

$$B^N = \sum_{i=1}^{m} \left(\frac{\sum_i \left[(q_i(i) - f_i(i))^2 + (q_i(i) - f_{iобр}(i))^2 \right]}{\sum_i x_i^2(i)} \right). \qquad (16)$$

В случае дифференциальных прогнозирующих моделей выход за пределы интервала интерполяции при формировании критерия баланса осуществляется при многошаговом интегрировании уравнений динамики. Наиболее точные прогнозы получаются при использовании в критерии истинных функций баланса, определяемых из физических соображений.

Комбинированный критерий селекции, его разновидности и свойства. Выбор критерия отбора (селекции) определяется типом решаемой задачи, принадлежит автору модели и отражает его цели. По рекомендациям, приведенным в работе, можно целесообразно выбрать один из эвристических критериев. Однако в некоторых ситуациях требования к моделям оказываются противоречивыми и не могут быть выражены в виде одного минимизируемого критерия. Кроме того, иногда выбранный критерий селекции не обеспечивает единственности выбора структуры модели или же требует слишком больших вычислений, так что желательно сокращение объема перебора. В ряде случаев может оказаться недостаточной «плавность» критерия, и требуется предварительно сократить число оцениваемых моделей, чтобы попасть в область надежной селекции.

В том случае, когда необходимо согласование нескольких противоречивых свойств (качеств) модели, рекомендуется перейти к комбинированным критериям, которые обычно формируются из составляющих критериев *K(1)* и *K (2)*:

$$K_{комб} = \sqrt{\lambda K_1^2 + (1-\lambda) K_2^2}, 0 < \lambda < 1, \qquad (17)$$

и находить модели, не улучшаемые по обоим критериям, а при изменении $\lambda_{во}$ всем интервале *[0,1]* построить множество паретовских структур моделей.

Кроме согласования противоречивых требований к модели (фактически, задача многокритериальной оптимизации), применение комбинированных критериев позволяет обеспечить единственность выбора структуры в том случае, когда один из критериев приводит к многозначности модели. Оба критерия должны быть выбраны целесообразно, т. е. с учетом физического смысла решаемой задачи моде-

лирования. Модели, равноценные по одному из критериев, будут отличаться по другому, что и определяет однозначный выбор.

Комбинированный критерий, учитывающий исходные критерии, может формироваться и в виде:

$$K_{комб}=\sqrt{\lambda(\frac{K_1}{K_{1\max}})^2+(1-\lambda)(\frac{K_2}{K_{2\max}})^2},0<\lambda<1, \quad (18)$$

приведенном к безразмерным величинам, что необходимо в том случае, когда K_1 и K_2 имеют различные единицы измерения и различный физический смысл.

Комбинированные критерии селекции. Теория самоорганизации подчеркивает, что целесообразный выбор критериев селекции и порядка их использования принадлежит человеку. В качестве составляющих K_1 и K_2 в комбинированных критериях могут быть использованы;

$\Delta(B)$ - критерий регулярности, $n_{см}$ - критерий минимума смещения, B_i - различные виды критериев баланса переменных, i_2 - критерий устойчивости прогнозов.

Ориентировочно можно сформулировать следующие правила применения комбинированных критериев селекции [11]:

1. Для алгебраических (полиномиальных) моделей, не содержащих среди аргументов время (т. е. для моделей, не предназначенных для долгосрочного прогноза), следует применять комбинированный критерий «смещение плюс регулярность»:

$$K_1=\sqrt{n_{см}^2+\Delta^2(B)}\to\min. \quad (19)$$

2. Для гармонических и алгебраических моделей, использующих функции времени (т. е. предназначенных для долгосрочного прогноза), в случаях, когда известен или может быть открыт закон физического баланса переменных, рекомендуется критерий «смещение плюс баланс переменных» [12]:

$$K_2=\sqrt{n_{см}^2+b^2}\to\min. \quad (20)$$

3. Для дифференциальных моделей (в виде конечно-разностных уравнений), предназначенных для многократного долгосрочного про-

гноза при помощи пошагового интегрирования, рекомендуется комбинированный критерий «несмещенность плюс устойчивость прогноза»:

$$K_3 = \sqrt{n_{см}^2 + i^2(N)} \to \min . \qquad (21)$$

При использовании дифференциальных моделей результаты долгосрочного прогноза могут быть неудовлетворительными, если вычислительная процедура не обладает достаточной устойчивостью, т. е. приводит к накоплению ошибок. Существует два способа достижения устойчивых результатов:

усреднение некоторого числа лучших частных моделей на каждом ряду селекции или только в конце синтеза модели;

использование комбинированного критерия K_3, содержащего требования устойчивости. Величина $i_2\ (N)$ формируется как сумма квадратов разностей табличных значении выхода модели и рассчитанных в процессе многократного интегрирования уравнения динамики.

Алгоритмы с усреднением также позволяют получать более устойчивые модели; истинная же модель всегда устойчива [13].

Применяются также комбинированные критерии, учитывающие три или четыре составляющих, например, вида:

$$K_4 = \sqrt{n_{см}^2 + b^2 + i^2(N)} \to \min . \qquad (22)$$

Последний этап построения математических моделей заключается в пересчете (адаптации) их параметров, необходимом для минимизации какого-либо (обычно, комбинированного) критерия, и позволяющем улучшить точность прогноза и добиться более полного соответствия свойств процесса и модели.

Адаптация может производиться для одной наилучшей модели (после селекции) и для всех структур моделей (до селекции), так что отбор производится по улучшенным значениям критериев селекции.

Среди составляющих комбинированного критерия могут быть такие, которые характеризуют только структуру модели, независимо от ее коэффициентов (например, показатели смещения). При этом адаптация сводится к минимизации подбором коэффициентов модели только второй составляющей комбинированного критерия (если она является характеристикой и структуры, и параметров), что упрощает расчеты и допускает применение градиентных процедур и методов случайного поиска [14].

4. Заключение. Выводы по статье:

1. Метод самоорганизации моделей основан на применении внешних критериев. Критерий называется внешним, если его определение основано на новой информации, «свежих» точках, не использованных при синтезе модели выбора: регулярности, минимума смещения, баланса переменных и комбинированных. Структура модели оптимальной сложности соответствует минимуму внешнего критерия. Принцип самоорганизации состоит в том, что при постепенном усложнении моделей некоторые критерии (имеющие свойства «внешнего дополнения») проходят через свой минимум. При помощи перебора моделей находят этот минимум и указывают единственную модель оптимальной сложности.

2. Для прогноза показателей уровня защищенности АИТС разработан нечеткий групповой метод обработки данных. Данный прогноз имеет важное значение для оценки требуемых показателей уровня защищенности АИТС и планирования защищенности систем от несанкционированного доступа в будущем.

Литература

1. *Стоянова О.В., Бояринов Ю.Г.* Моделирование социально-экономических систем на основе самоорганизации нейро-нечетких сетей // Вестник МЭИ. 2014 №4.
2. *Zelinka I., Oplatkova Z., Nolle L.* Analytic programming – symbolic regression by means of arbitrary evolutionary algorithm // I. J. of Simulation. 2013. vol. 6, no. 9. pp. 44–56.
3. *Дрейпер Н., Смит Г.* Прикладной регрессионный анализ // М.: Издательский дом "Вильямс". 2007. 912 с.
4. *Anastasakis L., Mort N.* The development of selforganization techniques in modeling: A review of the Group Method of Data Handling (GMDH) // Research Report № 823. Sheffield, United Kingdom, The University of Sheffield. 2015.
5. *Орлов А.А.* Принципы построения архитектуры программной платформы для реализации алгоритмов метода группового учета аргументов (МГУА) // Управляющие системы и машины: Международный журнал. 2013. № 2. С. 65–71.
6. *Краснощеков П. С., Петров А. А.* Принципы построения моделей // М.: Фазис. 2010. 264 с.
7. *Booch G., Maksimchuk R.A., Engel M.W., Young B.J., Conallen J., Houston K.A.* Object Oriented Analysis and Design with Applications // NY. Addison_Wesley Professional. 2013. 720 p.
8. *Бояринов Ю.Г., Стоянова О.В., Дли М.И.* Нейро-нечеткий метод группового учета аргументов для поддержки принятия решений по управлению региональными социально-экономическими системами // М.: Физматлит. 2005.
9. *Larose D.T.* Discovering knowledge in Data: An Introduction to Data Mining // New Jersey: Wiley & Sons. 2005. 336 p.
10. *Olson D.L.* Advanced Data Mining Techniques // Berlin: Springer-Verlag. 2008. 180 p.
11. *Kordik P.* Fully Automated Knowledge Extraction using Group of Adaptive Models Evolution // PhD thesis. Prague. 2014. 150 p.
12. *Стрижов В.В., Крымова Е.А.* Методы выбора регрессионных моделей // М.: ВЦ РАН. 2010. Т. 60. С. 2.

13. *Guyon I., Gunn S.* Feature extraction: foundation and applications // Springer. 2011. 778 p.
14. *Efron B., Hastie T., Johnstone I., Tibshirani R.* Least angle regression // The Annals of Statistics. 2014. vol. 32. no. 2. pp. 407–499.

References

1. Stojanova O.V., Bojarinov Ju.G. [Modeling of socio-economic systems based on self-organizing neural-fuzzy networks]. *Vestnik MJeI – Bulletin of the MEI*. 2014. no. 4. (In Russ).
2. Zelinka I., Oplatkova Z., Nolle L. Analytic programming – symbolic regression by means of arbitrary evolutionary algorithm. *I. J. of Simulation*. 2013. vol. 6, no. 9. pp. 44–56.
3. Drejper N., Smit G. *Prikladnoj regressionnyj analiz* [Applied Regression Analysis]. M.: Publishing house "Williams". 2007. 912 p. (In Russ).
4. Anastasakis L., Mort N. The development of selforganization techniques in modeling: A review of the Group Method of Data Handling (GMDH). Research Report № 823. Sheffield, United Kingdom, The University of Sheffield. 2015.
5. Orlov A.A. [Architecture principles of a software platform for the implementation of algorithms group method of data handling (GMDH)]. *Upravljajushhie sistemy i mashiny: Mezhdunarodnyj zhurnal – Control systems and machines: An International Journal*. 2013. no. 2. pp. 65–71. (In Russ).
6. Krasnoshhekov P. S., Petrov A. A. *Principy postroenija modelej* [Principles of constructing models]. Phasis. 2010. 264 p. (In Russ).
7. Booch G., Maksimchuk R.A., Engel M.W., Young B.J., Conallen J., Houston K.A. Object Oriented Analysis and Design with Applications. NY. Addison_Wesley Professional. 2013. 720 p.
8. Boyarinov U.G., Stoyanova O.V., Dli M.I. *Nejro-nechetkij metod gruppovogo ucheta argumentov dlja podderzhki prinjatija reshenij po upravleniju regional'nymi social'no-jekonomicheskimi sistemami* [Neuro-fuzzy group method of data to support management decision-making regional socio-economic systems]. M.: Fizmatlit. 2005. (In Russ).
9. Larose D.T. Discovering knowledge in Data: An Introduction to Data Mining. New Jersey: Wiley & Sons. 2005. 336 p.
10. Olson D.L. Advanced Data Mining Techniques. Berlin: Springer-Verlag. 2008. 180 p.
11. Kordik P. Fully Automated Knowledge Extraction using Group of Adaptive Models Evolution. PhD thesis. Prague, 2014. 150 p.
12. Strizhov V.V., Krymova E.A. [Methods selection of regression models]. *Vychislitel'nyj centr RAN – Computing Centre of RAS*. 2010. Issue 60, p. 2. (In Russ).
13. Guyon I., Gunn S. Feature extraction: foundation and applications. Springer. 2011. 778 p.
14. Efron B., Hastie T., Johnstone I., Tibshirani R. Least angle regression. The Annals of Statistics. 2014. vol. 32. no. 2. pp. 407–499.

Глыбовский Павел Анатольевич — к-т техн. наук, доцент, доцент кафедры систем сбора и обработки информации, Военно-космическая академия имени А.Ф. Можайского. Область научных интересов: теория распознавания образов, теория информации. Число научных публикаций — 45. p_glybovsky@mail.ru; ул. Ждановская, д. 13, Санкт-Петербург, 197198; р.т.: +7(812) 237-19-60.

Glybovsky Pavel Anatolievich — Ph.D., associate professor, associate professor of system for collecting and processing information department, Mozhaisky Military Space Academy. Research interests: theory of pattern recognition, information theory. The number of publica-

tions — 45. p_glybovsky@mail.ru; 13, Zhdanovskaya street, St.-Petersburg, 197198, Russia; office phone: +7(812) 237-19-60.

Глухов Александр Петрович — к-т техн. наук, начальник департамента информационной безопасности, ОАО «РЖД». Область научных интересов: информационная безопасность, охрана объектов железнодорожного транспорта. Число научных публикаций — 40. gie76@yandex.ru; ул. Ждановская, д. 13, Санкт-Петербург, 197198; р.т.: +7(812) 347-96-87.

Gluhov Aleksandr Petrovich — Ph.D., head of department of information security, JSC «RZhD». Research interests: information security, protection of railways. The number of publications — 40. gie76@yandex.ru; 13, Zhdanovskaya street, St.-Petersburg, 197198, Russia; office phone: +7(812) 347-96-87.

Пономарев Юрий Александрович — к-т техн. наук, доцент, заместитель начальника кафедры систем сбора и обработки информации, Военно-космическая академия имени А.Ф. Можайского. Область научных интересов: методы анализа социально-психологических характеристик, теория нечетких множеств. Число научных публикаций — 35. yurij_1969_2011@mail.ru; ул. Ждановская, д. 13, Санкт-Петербург, 197198; р.т.: +7(812) 237-19-60.

Ponomarev Yuri Aleksandrovich — Ph.D., associate professor, deputy head of system for collecting and processing information department, Mozhaisky Military Space Academy. Research interests: methods of analysis of the socio-psychological characteristics, fuzzy set theory. The number of publications — 35. yurij_1969_2011@mail.ru; 13, Zhdanovskaya street, St.-Petersburg, 197198, Russia; office phone: +7(812) 237-19-60.

Шиленков Максим Викторович — к-т техн. наук, заместитель генерального директора, ЗАО "Эврика". Область научных интересов: информационные технологии, информационная безопасность. Число научных публикаций — 40. ema5@yandex.ru; ул. Ждановская, д. 13, Санкт-Петербург, 197198; р.т.: +7(911)2231431.

Shilenkov Maxim Viktorovich — Ph.D., deputy general director, JSC "Jevrika". Research interests: information technology, information security. The number of publications — 40. ema5@yandex.ru; 13, Zhdanovskaya street, St.-Petersburg, 197198, Russia; office phone: +7(911)2231431.

РЕФЕРАТ

Глыбовский П.А., Глухов А.П., Пономарев Ю.А., Шиленков М.В. **Подход к оцениванию и прогнозированию уровня защищенности информационных и телекоммуникационных систем.**

В настоящее время при оценке уровня безопасности корпоративной информационной инфраструктуры применяются методы прогнозирования ситуаций и оценки требуемых значений показателей уровня защищенности в совокупности с оценками экономической эффективности инвестиций в обеспечение безопасности и защиту информации.

Для повышения эффективности прогнозирования уровня защищенности информационных и телекоммуникационных систем необходимо выполнение прогнозирование совокупности отдельных процедур. Каждая из процедур выполняется по соответствующей методике, и в результате чего формируются показатели уровня защищенности, которые подвергаются прогнозированию.

Одной из особенностей, свойственных задачам математического моделирования уровня защищенности информационных и телекоммуникационных систем, является наличие большого числа факторов, влияющих на исследуемый показатель, и ограниченного объема наблюдений данных факторов, используемых при построении модели. Для решения задачи прогнозирования уровня защищенности был предложен метод группового учета аргументов (МГУА), позволяющий обеспечить приемлемое качество модели в условиях многофакторности и ограниченности объема обучающей выборки.

SUMMARY

Glybovsky P.A., Gluhov A.P., Ponomarev Yu.A., Shilenkov M.V. **Approach to the Evaluation and Prediction of the Level of Security of Information and Telecommunications Systems.**

Currently, in the evaluation of the safety of corporate information infrastructure, techniques are applied for forecasting situations and assessing the required indicator values for security level in conjunction with the assessment of the economic efficiency of investment in safety and protection of information.

To improve the efficiency of predicting the level of security of information and telecommunication systems it is necessary to perform forecast of the aggregate of separate procedures. Each of the procedures is performed by the corresponding method, resulting in the formation of indicators of the level of protection, which are subject to forecasting.

One of the features, peculiar to the problems of mathematical modeling of the level of protection of information and telecommunication systems, is the presence of a large number of factors influencing the studied parameters. At the same time, there is a limited amount of observations of these factors used in the construction of the model. To solve the problem of predicting the level of protection we propose a group method of data handling (GMDH), which allows one to ensure an acceptable quality of the model in the conditions of multivariate models and the limited amount of training sample.