

УДК 519.7÷681.3

**ПРИЛОЖЕНИЕ ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ В УПРАВЛЕНИИ
ТЕХНОЛОГИЧЕСКИМИ РЕЖИМАМИ НЕФТЕПРОДУКТОПРОВОДОВ**

С.Е.Кутуков

(Уфимский государственный нефтяной технический университет)

Применение современных технологий SCADA, телемеханики и АСУ в магистральном транспорте нефти и нефтепродуктов подготовило почву и поставило на повестку дня разработку систем функциональной диагностики технологических режимов нефте- и продуктопроводов, которые призваны повысить оперативность принятия решений, расширить аналитические возможности служб, эксплуатирующих разветвленные сети, отслеживать тенденции развития процессов в трубопроводах; в ясных для пользователя понятиях интерпретировать собранные системой SCADA параметры технологического процесса перекачки; в удобной для анализа и дальнейшего принятия решения форме представлять результаты диагностирования и за счёт применения превентивных мер, устраняющих причины осложнений технологических режимов, экономить до трети энергозатрат на линейной части.

Под термином «функциональная диагностика» подразумевается система оперативной диагностики осложнений режимов эксплуатации магистрального нефтепровода исключительно по технологическим параметрам перекачки, регистрируемых штатными средствами телемеханики без специальных тестовых воздействий. Классификация и анализ причин снижения пропускной способности и эффективности, технологической и экологической безопасности функционирования участка магистрального нефтепродуктопровода, объединенных общим термином «осложнения» технологических режимов выходит за рамки настоящей статьи, а всё их многообразие сведено к пяти категориям: газовые и водные скопления, утечки, смолпарафиновые отложения и неподвижные объекты в полости трубопровода.

С одной стороны, практически все осложнения технологических режимов магистральных нефтепроводов приводят к снижению пропускной способности (увеличению гидравлических сопротивлений). С другой стороны, количество регистрируемых параметров в традиционных системах телемеханики и АСУ весьма ограничено: давление (P), температура (T) и расход (Q). К ним можно ещё добавить время (τ). Регистрация всех параметров сопряжена с некоторой погрешностью, собственной ошибкой обладают и решения обратных задач в алгоритмах диагностирования, что накладывает определённую неточность, или в математических терминах – «нечёткость». Поэтому для получения количественных оценок осложнений технологических режимов МНП целесообразно использовать методологию «Мягких вычислений», позволяющую корректно обрабатывать нечёткую, частью противоречивую, весьма ограниченную информацию об изменении технологических параметров перекачки, как во времени, так и по длине участка трубопровода. Следовательно, достоверно определить вклад каждого вида осложнений при одновременном их возникновении весьма непросто, так как признаки их проявления схожи, а информация весьма ограничена как по количеству, так и по качеству.

Математическим аппаратом, оперирующим нечеткими понятиями является «Мягкие вычисления» (Soft Computing) [1], основоположником которых является Лутфи Заде – профессор Берклиевского университета, США. Данный термин, введенный в употребление в 1990 г., означает интеграцию идей нечеткой логики, нейронных сетей, генетических алгоритмов и вероятностных рассуждений с целью построения интеллектуальных систем обработки информации, оперирующих с неточными и нечеткими данными, неопределенностью и частичной истинностью фактов и обеспечивающих при этом нечувствительность к ошибкам, низкую стоимость решения и его лучшее согласие с реалиями окружающего мира.

История применения генетических алгоритмов начинается с работ Р.Холстиена [2], Де Джонга [265], Дж. Холланда [3] и Д. Гольдберга [4] 1970-75 гг., в которых впервые были продемонстрированы возможности метода для решения задач многопараметрической оптимизации. Но только в последние 5-7 лет появились работы Savic D.A., Walters G.A., (Великобритания) [5], Vitkovsky J.P., Simpson A.R., Murthy L.J. (Австралия) [6,7], Tang K.-W., Zhang F. (Канада) [8], предложившие использовать генетические алгоритмы в проектировании и оптимизации водопроводных сетей решением обратных задач трубной гидравлики (определении коэффициента гидравлических сопротивлений или шероховатости труб).

Главным достоинством рассматриваемого подхода заключается в возможности находить субоптимальные (пригодные) решения на всём множестве возможных решений без излишних вычислительных затрат «комбинаторного взрыва» возможных вариантов с ростом размерности задачи [9]. Количество же независимых переменных (генов @ G_i) достигает 200-500 при идентификации 5-7 осложнений на эксплуатационном участке нефтепровода протяженностью 400-500 км.

Как показывают исследования, выполненные разными авторами, эффективность применения генетических алгоритмов значительно возрастает в случае их комбинированного использования с другими методами теории искусственного интеллекта. К числу наиболее перспективных направлений в этой области относятся исследования в области "нечёткой логики". Для диагностики технологических режимов нефтепроводов наиболее перспективным является приложение методологии генетических алгоритмов в сочетании с нечёткими множествами для генерации целевой функции.

Генетические алгоритмы образуют широкий класс методов глобальной многопараметрической оптимизации, базирующихся на моделировании механизмов естественной эволюции и популяционной генетики. Терминология, принятая в данной области математики заимствована из генетики. В соответствии с наиболее общим определением, генетические

алгоритмы - это методы случайного глобального поиска, копирующие механизмы естественной биологической эволюции. Генетические алгоритмы оперируют с популяцией оценок потенциальных решений (индивидуумов), используя принцип "выживает наиболее приспособленный". На каждом шаге алгоритма, образуется новое множество приближений, создаваемое посредством процесса отбора индивидуумов согласно их уровню пригодности. Операндом генетического алгоритма является ген @G – закодированная оценка того или иного осложнения технологического режима, объединенный в «хромосому» (chromosome, string) - n-мерный вектор &S. Область значений всех генов @G унифицирована - [0...7], что обеспечивает возможность применения операторов ГА между ними. Размерность кода зависит от требуемой точности оценки диагностируемых параметров. Значение кода @G определяет количественную оценку осложнения, которая может быть пропорциональной или прогрессивной (логарифмической). Пропорциональное кодирование целесообразно применять для водных и газовоздушных скоплений. Прогрессивное - более предпочтительно для фаззификации утечек, смоло-парафиновых отложений, скоплений грата, деформаций труб и пр.

Операция кодирования осложнения, т.е. представление чёткой количественной оценки в нечёткую интерпретацию, называется фаззификацией (fuzzyfication). Обратная операция – определение количественной оценки того или иного осложнения по коду гена – дефаззификацией.

Основными генетическими операторами являются [10]:

- *репродукция* (reproduction)
- *кроссинговер* (crossing over);
- *мутация* (mutation);
- *селекция* (selection).

Оператор *мутация* имеет несколько практических реализаций:

1. Изменение кода случайно выбранного гена @G на зеркально противоположное из диапазона значений [0...7]:

$$M(@G) = (7 - @G); \quad (1)$$

2. То же, но случайным образом из всего диапазона значений:

$$M(@G) = \text{random} \{0...7\}; \quad (2)$$

3. Циклическое переписывание значений кодов соответствующих генов из геномов смежных участков трубопроводов:

$$M(@G)_i = M(@G)_{i+n} \quad (3)$$

где i – индекс участка трубопровода (генома);

n – размер цикла.

Этот тип мутации воспроизводит реальную ситуацию на трассе – движение скоплений, инородных предметов или отложений с потоком продукта.

После мутации хромосомы - кандидаты копируются в новую популяцию хромосом $\{S\}^{(\zeta + 1)}$, и весь процесс повторяется применением операторов репродукции, кроссинговера и мутации. Оператор мутации вносит дополнительное разнообразие в популяцию, накладывая стохастический шум на процесс эволюции, стимулируя тем самым исследование различных частей пространства поиска, что снижает в конечном итоге риск сходимости к локальным оптимумам. Обычно мутация представляет собой случайное изменение гена, вероятность которого довольно низкая ($P_{\text{мут}} \approx 0,001$).

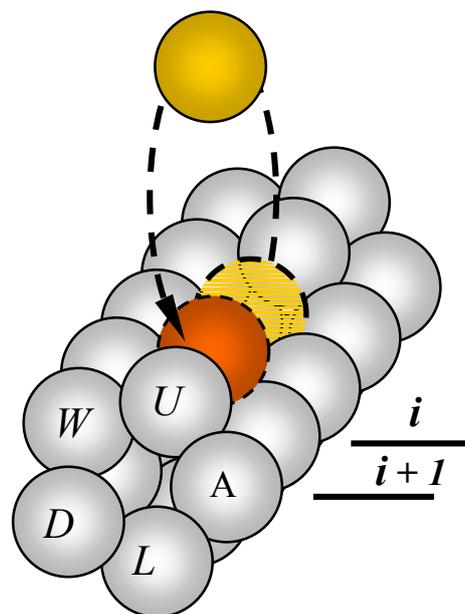


Рисунок 1 – Оператор циклической мутации

Разновидностью мутации является *инверсия* [10], при которой участок хромосомы покидает свое место в цепи хромосомы и, развернувшись на 180° , вновь занимает прежнее положение.

Кроссинговер - оператор рекомбинации двух различных хромосом из одной популяции (рисунок 2). Хромосомы скрещиваются, обмениваясь цепочками генов с вероятностью $P_{кр}$. Реализация оператора зависит от класса и размерности решаемой задачи. В простейшем случае одноточечного кроссинговера на практике применяется $P_{кр} = 0,6$.

Естественным развитием одноточечного кроссинговера являются схемы с несколькими точками (двухточечный, трехточечный и т.д.).

Предельным случаем является равномерный кроссинговер, когда каждая пара генов внутри 2-х хромосом обмениваются кодами в соответствии с определенной вероятностью.

В задаче диагностики осложнений технологических режимов необходимо различать кроссинговер в диагонали однотипных генов, т.е. описывающих один вид осложнений по длине трассы, и оператор между генами разных осложнений.

Т.е. оператор кроссинговера модифицирован с учетом декомпозиции хромосомы &S.

Более гибкой является схема, когда места сечений для кроссинговера выбираются отдельно для каждого участка &S и оператор осуществляется независимо в пределах каждой из хромосом.

Различные типы кроссинговера обладают общим положительным свойством: они контролируют баланс между дальнейшим использованием

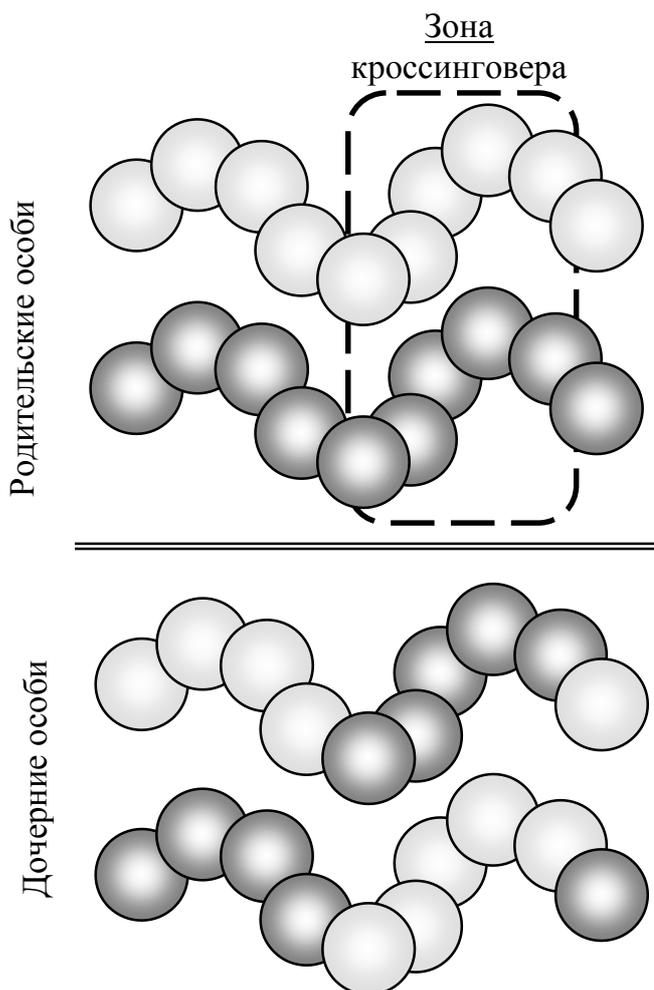


Рисунок 2 – Оператор двухточечного кроссинговера

уже найденных хороших подобластей пространства поиска и исследованием новых подобластей. Это достигается за счет сохранения общих блоков внутри хромосом - родителей и одновременного исследования новых областей в результате обмена частями строк (хромосом).

Кроссинговер и мутация играют разную роль. Кроссинговер, как правило, наиболее эффективен в начале поиска, в то время как мутация позволяет найти более точное решение в конце его [10]. Совместное использование операторов отбора и кроссинговера приводит к тому, что области пространства, обладающие лучшей в среднем оптимальностью, содержат больше членов популяции, чем другие. Таким образом, эволюция популяции направляется к областям, содержащим оптимум с большей вероятностью, чем другие.

Оператором *репродукции* в популяции выбирается пара хромосом, которые в процессе кроссинговера дают новое поколение, которое проверяется на приспособленность.

Селекция – ключевой оператор генетического алгоритма, реализующий принципы Ч.Дарвина «выживает наиболее приспособленный» («Survives the fittest»). В нашем случае, отбор производится на основании вероятностей $P(&S)$, вычисленных для каждого индивидуума популяции.

Селекцией из получившегося множества родительских и дочерних хромосом формируется новая популяция с исходным количеством индивидов, с которыми процедура повторяется до тех пор, пока значение функции пригодности лучшей хромосомы не останется в заданном количестве поколений N_p неизменной.

Для снижения затрат на адаптацию алгоритма в диагностике технологических режимов нефтепродуктопроводов целесообразно использовать такие зарекомендовавшие себя программные продукты, как расширение программного комплекса Genesis (GENEsYs, PARAGenesis, Gas), Genitor, SGA-C, Splicer, - пакет прикладных программ NASA, GALOPPS, Genie, Matlab – GAOT, PGA, GA Workbench, отличающиеся особенностями

практической реализации основных операторов генетического алгоритма: *репродукции* (reproduction); *кроссинговера* (crossing over); *мутации* (mutation) и *селекции* (selection) с преимущественным применением элитного отбора, циклической мутации, диагонального кроссинговера.

АЛГОРИТМ МНОГОПАРАМЕТРИЧЕСКОЙ ДИАГНОСТИКИ

На предварительном этапе выявляются осложнения, подлежащие оценке, из которых формируется структура хромосомы. Учитывая естественное ограничение, что дискретность пространства диагностического поиска не может превышать физического, обусловленного наличием датчиков по длине нефтепровода, каждому участку трассы между замерными пунктами поставлен в соответствие геном $\{@W, @D, @U, @A, @L\}$, состоящим из пяти генов:

$@W$ – скопления воды;

$@D$ – смолопарафиновые отложения

$@L$ – утечки.

$@A$ – газовые скопления;

$@U$ – неподвижные объекты

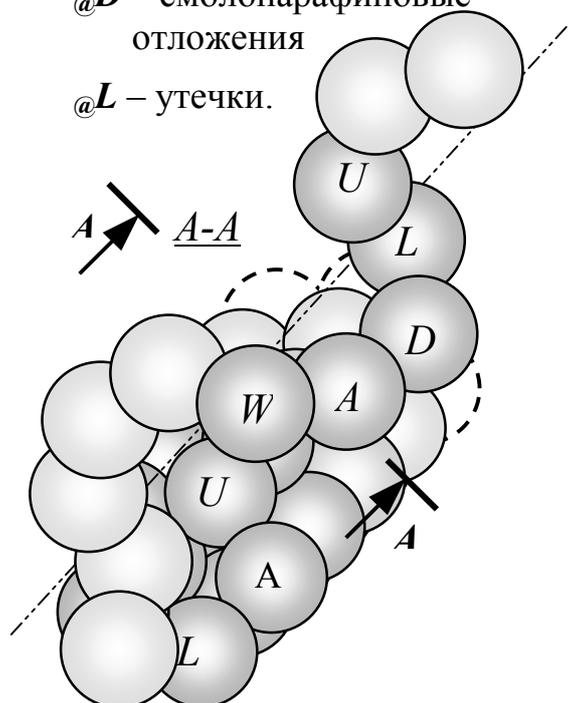


Рисунок 3 – Хромосома в алгоритме диагностики технологических режимов нефтепроводов

Представление строки кодов хромосомой, имеющей вид пространственной спирали удобно при объяснении разновидностей

оператора кроссинговера и мутации. Параллель генов @G, отвечающих за какой-либо один признак (осложнение) выглядит последовательным рядом на изображении такой спирали с торца (сечение А-А). Генотип, описывающий все осложнения одного участка трубопровода между замерными пунктами {@W, @D, @U, @A, @L} – участком спирали фиксированной длины на виде сбоку (рисунок 3).

Реализация генетического алгоритма в диагностике технологических режимов на трассе нефтепродуктопровода сводится к следующей последовательности операций (рисунок 4):

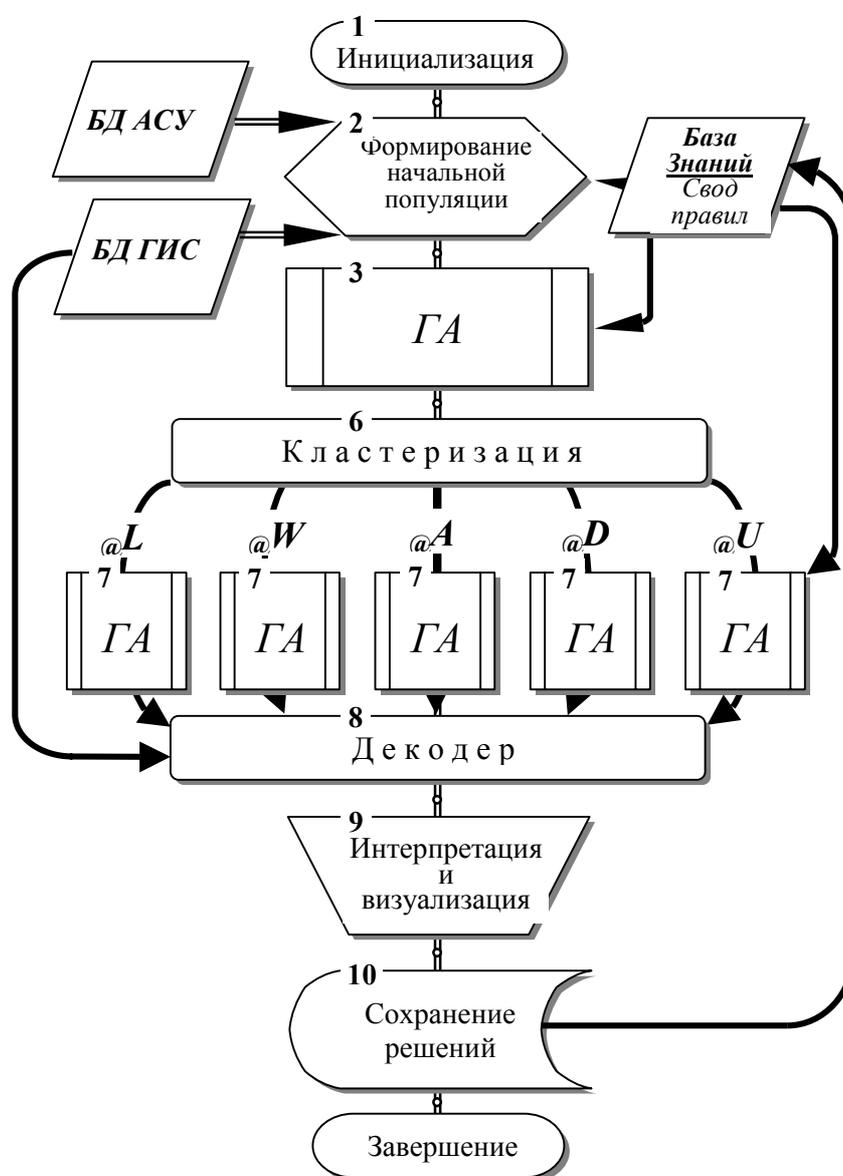


Рисунок 4 - - Блок-схема алгоритма диагностики

1. Оценивая условия существования каждого из внутритрубных образований на диагностируемом участке трубопровода в целом, определяется количество возможных осложнений k . После чего, для каждого участка трассы между замерными пунктами производится экспертная оценка максимально возможного уровня осложнения, которому ставится в соответствие максимальное значение кода - 7:

- для газовых скоплений – это длина всех нисходящих участков трубопровода;
- для водных скоплений – это длина всех восходящих участков трассы;
- для смолопарафиновых отложений рекомендовано принять предельную оценку толщины слоя - $\delta_{\max} = D/4$;
- для утечек, чтобы не дублировать существующие возможности, целесообразно брать значения на пределе чувствительности традиционных методов систем обнаружения утечек (3-5% от номинального расхода).

2. Генерируется начальная популяция $\{S\}^{(c=0)}$ из N индивидов. Количество особей в популяции должно быть не менее удвоенного произведения видов диагностируемых осложнений k на количество замерных пунктов по трассе эксплуатационного участка трубопровода ($m - 1$), из которых не менее 50% популяции – это результаты диагностики предыдущих периодов времени, которые извлекаются из Базы Знаний. Вторая часть индивидов формируются из условия максимального охвата всего пространства поиска решений алгоритмом диагностирования.

3. Формируется дочерняя популяция хромосом операциями кроссинговера и мутации, заложенными в используемом пакете прикладных программ. В случае выбора элитной хромосомы оператором кроссинговера или мутации она клонируется. Клоны полученных элитных хромосом, не участвуя в операциях кроссинговера или мутации, добавляются к дочерней популяции.

4. Определяется порог пригодности хромосом, по среднему значению функции пригодности:

$$\overline{fit} = \frac{\sum_1^N fit(\&S_i)}{N}. \quad (4)$$

5. На этапе селекции из родительской и дочерней популяций формируется новая из индивидов, годность которых больше средней по всей популяции:

$$fit(\&S_i) > \overline{fit}. \quad (5)$$

Цикл операторов пп.3-5 повторяется до тех пор, пока значение функции пригодности лучшей хромосомы не останется неизменной в заданном количестве поколений $Np = 15-30$.

6. Этап кластеризации – деление исходной популяции на группы $\{ @W \}$, $\{ @A \}$, $\{ @D \}$, $\{ @L \}$, $\{ @U \}$ по принципу «родственной» близости индивидов.

7. Применение операций генетического алгоритма в каждой группе с пропорциональным увеличением вероятности мутаций и уменьшения вероятности кроссинговера обеспечивает получение 5 решений, каждое из которых имеет своё максимальное в группе $\{ @W \}$, $\{ @A \}$, $\{ @D \}$, $\{ @L \}$, $\{ @U \}$ значение функции пригодности $fit(\&S^{max} | \{ @G_{ij} \})$, чем утоняются локальные максимумы функций пригодности по доминантному геному $\{ @G_{ij} \}$. Нормированные значения этих функций определяет вероятность адекватности интерпретации результатов диагностирования технологических режимов:

$$P(@G_i) = \frac{fit(\&S^{max} | \{ @G_i \})}{\sum_1^k fit(\&S^{max} | \{ @G_i \})}. \quad (6)$$

Окончательное заключение (право выбора варианта из предложенных альтернатив) остаётся за диспетчером.

8. Алгоритм дефаззификации хромосом зависит от принятых в п.1 вариантов кодирования:

- длина участков с газовыми скоплениями:

$$L_i^\Sigma = \frac{A_i}{7} \cdot \Delta L_i^{down}; \quad (7)$$

- длина участков с водными скоплениями:

$$L_i^\Sigma = \frac{W_i}{7} \cdot \Delta L_i^{up}; \quad (8)$$

- средняя по участку толщина слоя парафиновых отложений:

$$\delta_i^{cp} = \frac{D_i \cdot (D_i + 1)}{56} \cdot \frac{D}{4}; \quad (9)$$

- утечки: $Q_i^{leak} = 0,05 \cdot \frac{L_i \cdot (L_i + 1)}{56} \cdot Q;$ (10)

- осложнения, не зависящие от режима эксплуатации участка МНП:

$$\xi_i^{\text{оуз}} = \frac{\pi^2}{8} \cdot \frac{\zeta_i \cdot P_i^{\text{ном}} \cdot D^4}{\rho \cdot Q^2} \cdot U_i \cdot (U_i + 1), \quad (11)$$

где ΔL_i^{up} - общая длина восходящих сегментов на i -ом участке трубопровода между замерными пунктами;

ΔL_i^{down} - общая длина нисходящих сегментов на i -ом участке трубопровода между замерными пунктами;

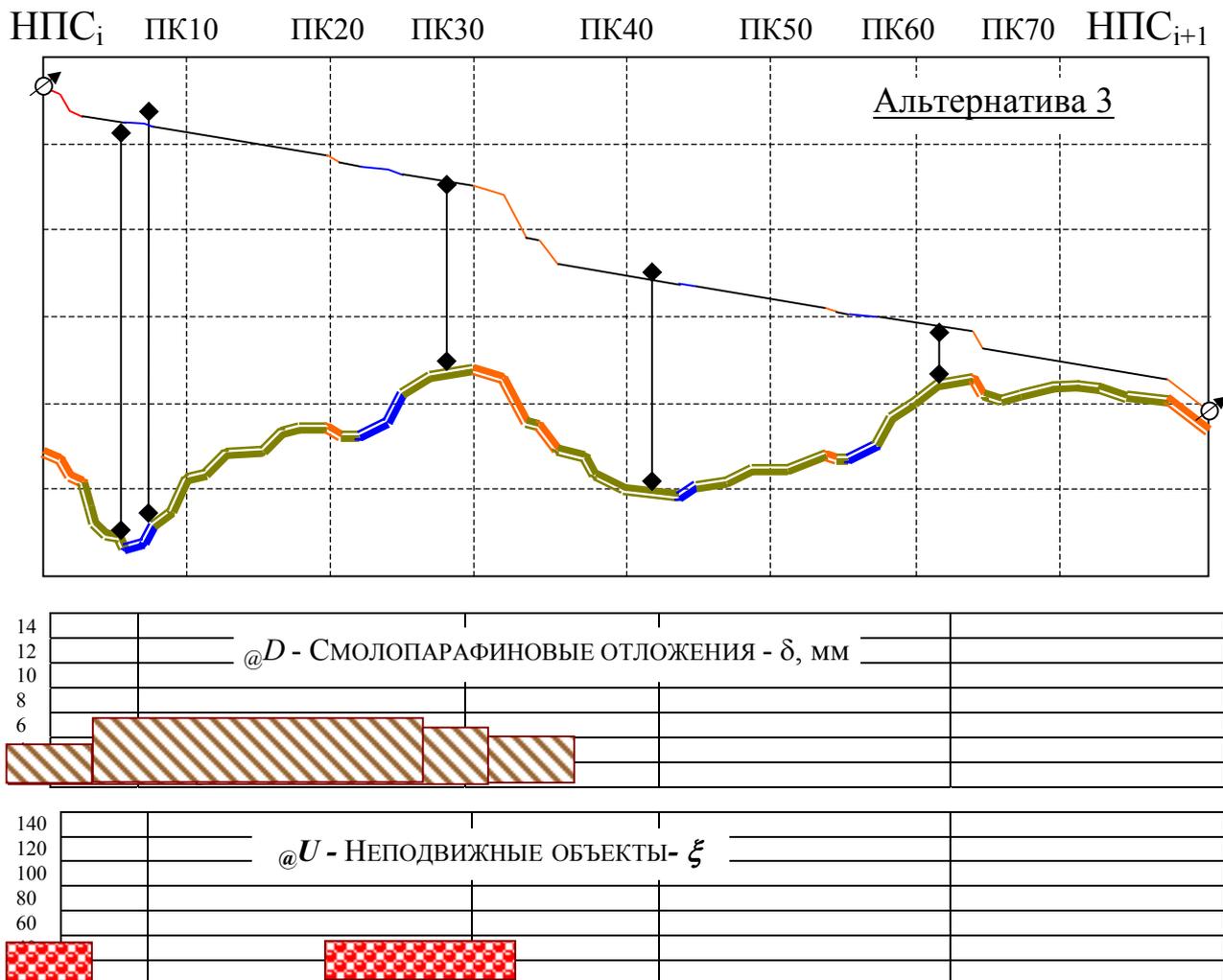
D – внутренний диаметр трубопровода;

$P_i^{\text{ном}}$ – номинал манометра на i -ом замерном пункте.

9. Несмотря на избыточное количество решающих правил для диагностики всех видов осложнений технологических режимов результаты диагностирования не могут быть интерпретированы однозначно. Можно лишь говорить о более или менее вероятном сочетании количественных оценок тех или иных осложнений. Поэтому результатом процедуры диагностирования является конечный набор альтернатив в сочетании с качественными и количественными оценками каждой из них. И только производственный опыт работы на конкретном объекте магистрального транспорта нефти или нефтепродукта может подсказать персоналу критерии окончательного выбора из предложенных альтернатив.

Один из вариантов визуализации результатов диагностирования

технологических режимов на профиле участка нефтепровода приведен на рисунке 5.



Альтернативы ситуаций на трассе МНП

№	Маркер	Доминирующий признак	Вероятность
1	@A	Газовые скопления	0,63-0,72
2	@W	Водные скопления	0,15-0,17
√3	@D	Смолопарафиновые отложения	0,12-0,13
4	@U	Неподвижные объекты	0,05
5	@L	Утечки	0,03

Рисунок 5 - Визуализация результатов диагностирования технологических режимов магистральных нефтепродуктопроводов

На профиле трассы синим цветом обозначены участки трубопровода с водными скоплениями, красным – с газовыми. Пьезометрическая линия с точками на концах вертикальных отрезков, показывающими давление в

гидростатических столбах перекачиваемого продукта на замерных пунктах, раскрашена в цвета определяющего её осложнения. Толщина слоя смолопарафиновых отложений и значения коэффициентов местных сопротивлений «неподвижных» объектов разнесены по отдельным шкалам, увязанными с делением нефтепровода на участки между смежными замерными пунктами. Список альтернатив интерпретации технологических параметров с условной вероятностью их адекватности реалиям ситуации на трассе, посчитанной по формуле (6) так же предлагается для анализа.

Необходимо отметить, что для каждой альтернативы из приведенного списка своя комбинация количественных оценок осложнений, а значит и свои значения изображенных параметров.

10. Сохраняя в Базе Знаний хромосомы – результаты диагностирования предыдущих периодов, обеспечивается обучение системы диагностики. Если период «памяти» достаточно большой, велика вероятность повторения ситуации на трассе. Использование в п.2 алгоритма цепочек кодов, близких к искомому решению значительно сокращают время расчётов и повышают достоверность полученных результатов.

Применение методологии “Soft Computing” в управлении систем магистрального транспорта нефти и нефтепродуктов принципиально меняет аналитические и прогностические возможности персонала, позволяет повысить оперативность принятия решений, отслеживать тенденции развития процессов в трубопроводах, в ясных для пользователя понятиях интерпретировать собранные системой SCADA параметры технологического процесса перекачки; в удобной для анализа и дальнейшего принятия решения форме представлять результаты диагностирования, что в конечном итоге является действенным средством повышения эффективности, промышленной и экологической безопасности эксплуатации изношенных нефтепроводов.

Литература:

1. Zadeh L.A. What is Soft Computer? // *Soft Computing*, 1997, № 1, p.1.
2. Holstien R.B. Artificial genetic adaptation in computer control systems, PhD thesis – University of Michigan, Ann Arbor, 1971.
- 3 De Jong K.A. Analysis of the behavior of class of genetic adaptive systems, PhD thesis – University of Michigan, Ann Arbor, 1975.
4. Goldberg D.E. *Genetic Algorithms in Search, Optimization and Machine Learning* – Addison-Wesley: Reading, M.A., 1989. – 412 p.
5. Savic D.A., Walters G.A. Genetic Algorithm for least-cost design for water-distribution networks. // *Journal of Water Resources Planning and Management*, ASCE, 123 (2), March-April, 1997. –pp.67-77.
6. Vitkovsky J.P., Simpson A.R., Lambert M. Transients for Calibration of Pipe Roughnesses Using Genetic Algorithms / Department of Civil and Environmental Engineering, the University of Adelaide, Australia, 2000. – 11 p.
7. Dandy G.C., Simpson A.R., Murthy L.J. An Improved Genetic Algorithm for Pipe Network Optimization // *Water Resources Research*, 1996. - № 32 (2), February. – p.p. 449-458.
8. Tang K.-W., Karney B., Pendlebury M., Zhang F. Inverse Transient Calibration of Water Distribution Systems Using Genetic Algorithms. /Conference on Computing and Control for the Water Industry (CCWI'99). – Exeter, U.K.,1999.
9. Захаров В.Н., Ульянов С.В. Нечёткие модели интеллектуальных промышленных регуляторов и систем управления. IV Имитационное моделирование. // *Техническая кибернетика*, № 5, 1994. – с.168-202.
10. Васильев В.И., Ильясов Б.Г. Интеллектуальные системы управления с использованием генетических алгоритмов. Учебное пособие. – Уфа:

УГАТУ, 1999. – 104 с.