

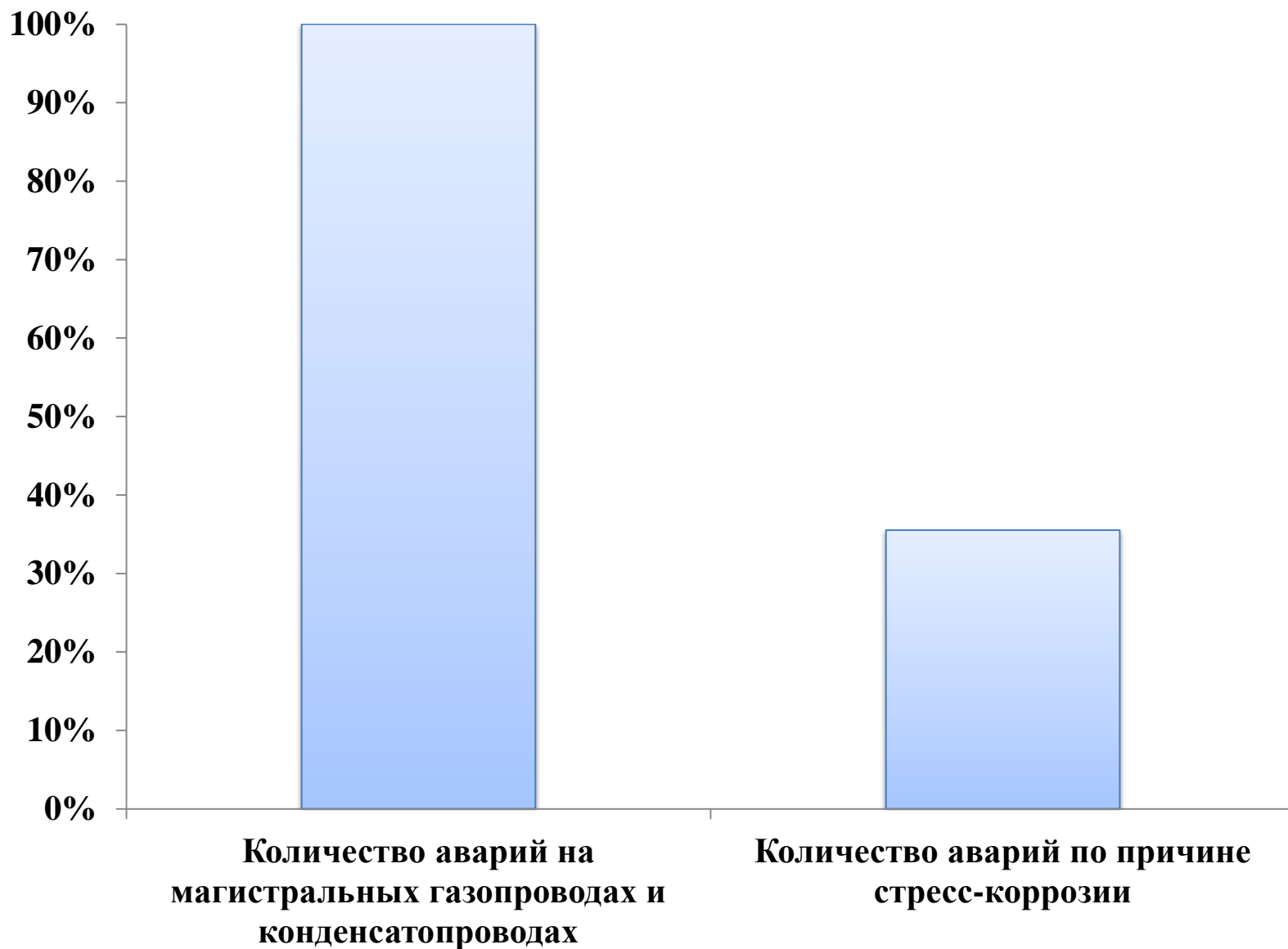
# Прогнозирование возможности реализации различных видов негативных событий на объектах магистральных газопроводов ПАО «Газпром»

*ведущий инженер отдела экспертизы нормативно-технической документации экспертно-аналитического Управления ООО «Газпром газнадзор»*

*Коваленко Сергей Владимирович*

2017 г.  
г. Москва

## Статистика стресс-коррозионных аварий за период с 1996 г. по 2016 г.



### Подготовительные этапы

### Состав работ на подготовительных этапах



# Распределение аварий, входящих в базу данных, по категориям и подкатегориям



### Внешние и внутренние признаки

✓к внешним (сопутствующим) признакам относятся нагрузки, воздействия, а также обстоятельства, в той или иной степени способствующие инициализации или развитию аварии, влияющие на физическое состояние объекта, его прочность, герметичность, несущую способность;

✓к внутренним (собственным) признакам относятся различные дефекты, повреждения как металла трубы, так и изоляционного покрытия, аномальные изменения механических свойств и т.д., определяющие динамику изменения физического состояния объекта, потерпевшего аварию.

### Случайные величины и случайные факторы

✓случайные величины (дискретные и непрерывные, приводимые к дискретному виду);

✓случайные факторы (качественные переменные, принимающие конечное число нечисловых значений уровней, обладающие рядом распределения но не имеющие числовых характеристик: среднего значения, дисперсии и т.д.).

## Вычисление недостающей информации (энтропии) и подготовка матрицы для отбора значимых признаков

Для I-й категории

Признак ( $x_1$ ):

Ряд распределения признака ( $x_1$ ):  $x_{11}, x_{12}, \dots, x_{1k}; p(x_{11}), p(x_{12}), \dots, p(x_{1k})$

Энтропия для признака ( $x_1$ ):  $H_I(x_1) = -\sum_{i=1}^k p(x_{1i}) \cdot \log_2 p(x_{1i})$

Признак ( $x_2$ ):

Ряд распределения признака ( $x_2$ ):  $x_{21}, x_{22}, \dots, x_{2l}; p(x_{21}), p(x_{22}), \dots, p(x_{2l})$

Энтропия для признака ( $x_2$ ):  $H_I(x_2) = -\sum_{i=1}^l p(x_{2i}) \cdot \log_2 p(x_{2i})$

Признак ( $x_N$ ):

Ряд распределения признака ( $x_N$ ):  $x_{N1}, x_{N2}, \dots, x_{Nm}; p(x_{N1}), p(x_{N2}), \dots, p(x_{Nm})$

Энтропия для признака ( $x_N$ ):  $H_I(x_N) = -\sum_{i=1}^m p(x_{Ni}) \cdot \log_2 p(x_{Ni})$

Аналогично для II-й и III-й категорий.

Полученные оценки энтропии для всех 3-х категорий:

Категории аварий	Энтропия ( $H$ ) по признакам $x_1 - x_N$			
	$x_1$	$x_2$	...	$x_N$
I	$H_I(x_1)$	$H_I(x_2)$	...	$H_I(x_N)$
II	$H_{II}(x_1)$	$H_{II}(x_2)$	...	$H_{II}(x_N)$
III	$H_{III}(x_1)$	$H_{III}(x_2)$	...	$H_{III}(x_N)$

## Выбор информативных признаков и категорий

Объединяем (суммируем по столбцам) оценки энтропии по одноименным признакам по всему массиву данных:

$$H_{\Sigma}(x_1) = H_I(x_1) + H_{II}(x_1) + H_{III}(x_1)$$

$$H_{\Sigma}(x_2) = H_I(x_2) + H_{II}(x_2) + H_{III}(x_2)$$

$$\dots$$

$$H_{\Sigma}(x_N) = H_I(x_N) + H_{II}(x_N) + H_{III}(x_N)$$

Выбираем для дальнейшего использования ряд признаков, обладающих наименьшими (по модулю) значениями энтропии (т.е. наименьшей недостающей информации).

Категории аварий	Энтропия ( $H$ ) по значимым признакам:			
	$x_1^*$	$x_2^*$	...	$x_N^*$
I	$H_I(x_1^*)$	$H_I(x_2^*)$	...	$H_I(x_N^*)$
II	$H_{II}(x_1^*)$	$H_{II}(x_2^*)$	...	$H_{II}(x_N^*)$
III	$H_{III}(x_1^*)$	$H_{III}(x_2^*)$	...	$H_{III}(x_N^*)$

\* Знак (\*) показывает, что какие-то признаки исключены на предыдущем этапе анализа.

Для установления типа (категории) аварий, в наибольшей степени обеспеченной достоверным прогнозом, суммируем по строкам оценки энтропии:

$$H_{\Sigma}^I = H_I(x_1^*) + H_I(x_2^*) + \dots + H_I(x_N^*)$$

$$H_{\Sigma}^{II} = H_{II}(x_1^*) + H_{II}(x_2^*) + \dots + H_{II}(x_N^*)$$

$$H_{\Sigma}^{III} = H_{III}(x_1^*) + H_{III}(x_2^*) + \dots + H_{III}(x_N^*)$$

Выбираем наименьшее значение суммарной энтропии, которое укажет категорию (т.е. вид) аварии, наиболее обеспеченную сведениями для прогноза.

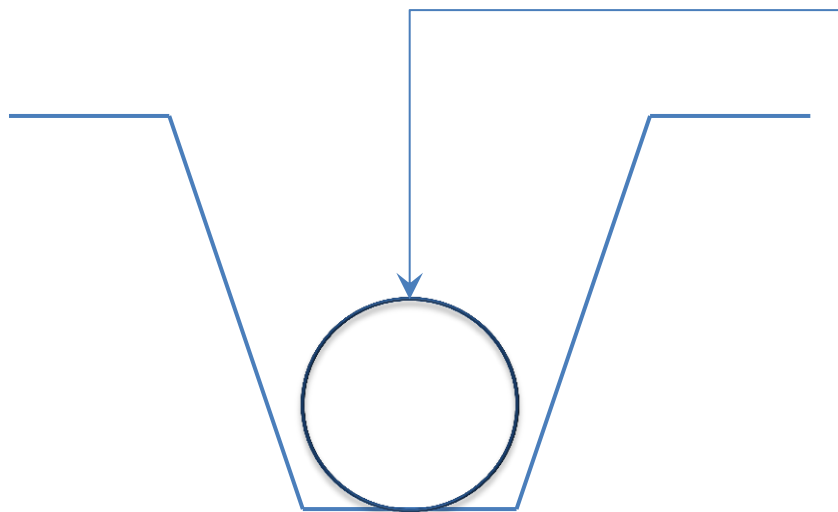
### Предупреждающая

**задача прогнозного определения вида (категории) возможной (гипотетической) аварии с использованием дополнительной информации о признаках, полученной шурфованием эксплуатируемого газопровода в конкретной его точке или иными методами, применяемыми при обследовании, например, воздушных переходов, площадочных трубопроводов, участков подводных переходов и т.д.**

### Подтверждающая

**методический инструмент в системе поддержки принятия решений, например, как поверочная задача отнесения случая реальной (произошедшей) аварии к той или иной категории, т.е. как задача подтверждения выводов специальной комиссии по расследованию аварии**

1. Данные тестового шурфования:



III-1		
№ п/п	Признаки	
	Наименование	Значение (уровень)
1	A	A <sup>o</sup>
2	B	B <sup>o</sup>
3	C	C <sup>o</sup>
...	...	...
N	L	L <sup>o</sup>

2. Возможные категории гипотетической аварии (полная группа несовместимых событий):

$$H_1, H_2, H_3$$

3. Априорные вероятности появления событий (гипотез):

$$P(H_1), P(H_2), P(H_3)$$

$$\sum_{i=1}^3 P(H_i) = 1$$

4. Совокупность  $N$  признаков, соответствующих каждому случаю аварий, зарегистрированных в БД.



1. Вычисление безусловной вероятности события  $V_{ABC\dots}$ :

$$P(V_{ABC\dots}) = \sum_{i=1}^n P(H_i) \cdot P(V_{ABC\dots}|H_i)$$

где:

$V_{ABC\dots}$  – событие, состоящее в реализации тестового вектора признаков  $\overrightarrow{V_{ABC\dots}}$ ;

$V_{ABC\dots}|H_i$  – событие, состоящее в том, что тестовый вектор признаков  $\overrightarrow{V_{ABC\dots}}$  отражает свойства категории аварий  $H_i$ ;

$P(V_{ABC\dots}|H_i)$  – вероятность события  $V_{ABC\dots}|H_i$ .

2. Вычисление вероятности события  $V_{ABC\dots}|H_i$  по критерию «попадания в категорию» хотя бы одного признака:

$$P(V_{ABC\dots}|H_i) = 1 - \left( (1 - P(A|H_i)) \cdot (1 - P(B|H_i)) \cdot \dots \right)$$

где:

$P(A|H_i), P(B|H_i), \dots$  – вероятности реализации независимых событий (значений признаков)  $A^o, B^o, \dots$  при условии выбора тестового вектора признаков из массива аварий  $H_i$ .

3. Вычисление вероятности реализации  $i$ -й категории аварии  $H_i$  в результате тестового шурфования:

$$P(H_i|V_{ABC\dots}) = \frac{P(H_i) \cdot P(V_{ABC\dots}|H_i)}{P(V_{ABC\dots})}$$

4. Сравнение и выбор максимального значения  $P(H_i|V_{ABC\dots})$ .

Построение «базового» распределения вероятностей по уровням каждого признака в каждой из 3-х рассматриваемых категорий

Признаки  $i$ -й категории

$A_i$		$B_i$		...
Значения (уровни)	Частота (вероятность)	Значения (уровни)	Частота (вероятность)	...
$A_{iB1}$	$P(A_{iB1})$	$B_{iB1}$	$P(B_{iB1})$	...
$A_{iB2}$	$P(A_{iB2})$	$B_{iB2}$	$P(B_{iB2})$	...
$A_{iB3}$	$P(A_{iB3})$	$B_{iB3}$	$P(B_{iB3})$	...
$A_{iB4}$	$P(A_{iB4})$	$B_{iB4}$	$P(B_{iB4})$	...
...	...	...	...	...

Присвоение совпадающим тестовым значениям/уровням ранее вычисленных ("базовых") значений вероятностей для каждой из 3-х рассматриваемых категорий:  
 При  $A_{iT} = A_{iB1}$   $P(A_{iT}) = P(A_{iB1})$ ,  $B_{iT} = B_{iB3}$   $P(B_{iT}) = P(B_{iB3})$ , ...

Переход к противоположным событиям для определения совместной вероятности независимых связанных (совместимых) событий:  
 $\bar{A}_{iT} = 1 - A_{iT}$ ,  $\bar{B}_{iT} = 1 - B_{iT}$

Определение для каждой из 3-х рассматриваемых категорий вероятности попадания в категорию хотя бы одного признака, т.е. значения (уровня) хотя бы одного признака:  
 $P(V_{AB...} | H_i) = 1 - P(\bar{A}_{iT} * \bar{B}_{iT}) = 1 - ((1 - P(A_{iT})) * (1 - P(B_{iT})) ...)$

Подготовка к вводу данных, приведение сведений БД к единому дискретному виду, вычисление базовых (условных) вероятностей по уровням признаков

$$X_1 \rightarrow P_1, \dots, X_n \rightarrow P_n$$

Тестовое шурфование с определением значений (уровней) признаков

$$A^\circ, B^\circ, C^\circ, \dots$$

Построение распределений вероятностей тестовых значений признаков по категориям:

$$P(A|H_1), P(A|H_2), P(A|H_3);$$

$$P(B|H_1), P(B|H_2), P(B|H_3);$$

$$P(C|H_1), P(C|H_2), P(C|H_3);$$

Установление наиболее вероятной физической природы аварии: сопоставление результатов расчета для

$$H_1; H_2; H_3 \text{ и выбор } P_{max}$$

нет

2-й этап расчета?

да

Корректировка исходных требований к массиву данных об авариях и составу «рабочего словаря» признаков

2-й этап решения задачи по массиву сведений для выбранной категории аварий ( $H_1, H_2$  или  $H_3$ )  
(при необходимости уточнения вида аварии)

Конец

1. Разработана методика прогнозирования вероятного типа (природы) гипотетической аварии магистрального газопровода на основе тестового шурфования газопровода с использованием аппарата распознавания образов, теории информации и с привлечением упорядоченного массива ретроспективных сведений о случаях аварий на магистральных газопроводах.
2. В предложенном подходе к прогнозированию возможного типа аварии отсутствует необходимость предварительного «обучения» системы распознавания, которое в процессе подготовки к решению задачи уже проведено, а система настроена на совокупности упорядоченных сведений о ранее произошедших случаях аварий.
3. Разработанная методика предназначена для прогнозной оценки наиболее вероятного типа («физической природы») гипотетической аварии на конкретном участке эксплуатируемого газопровода на основе данных, полученных с помощью пробного (одиночного или многократного) шурфования.

# Благодарю за внимание!

Коваленко Сергей Владимирович

ведущий инженер отдела экспертизы нормативно-технической документации экспертно-аналитического Управления ООО «Газпром газнадзор»  
тел.: (495) 355-98-71, email: [kovalenko@gaznadzor.gazprom.ru](mailto:kovalenko@gaznadzor.gazprom.ru)

Шапиро Владимир Дмитриевич

ведущий инженер отдела по анализу производственной безопасности экспертно-аналитического Управления ООО «Газпром газнадзор»  
тел.: (495) 355-98-74, email: [shapiro@gaznadzor.gazprom.ru](mailto:shapiro@gaznadzor.gazprom.ru)