УДК:681.5

М.Ю.ШЕСТОПАЛОВ, *проректор по научной работе, shestopalov 08@mail.ru*
Ю.А.КОРАБЛЕВ, *доцент, korablev@pochta.tvoe.tv*
Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И.Ульянова (Ленина)

M.U.SHESTOPALOV, *Vice Rector, shestopalov 08@mail.ru*
J.A.KORABLEV, *associate professor, korablev@pochta.tvoe.tv*
Saint Petersburg Electrotechnical University «LETI»

ДИАГНОСТИКА СОСТОЯНИЙ НЕПРЕРЫВНЫХ ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ НА ОСНОВЕ НЕЧЕТКОЙ ТОПОЛОГИЧЕСКОЙ ИДЕНТИФИКАЦИИ

В статье рассматриваются вопросы нечеткой идентификации математических моделей непрерывных физических и технологических процессов. Предлагается новый подход к построению диагностических моделей первого ранга неопределенности на основе нечеткой логики.

***Ключевые слова*:** нечеткая логика, моделирование, нечеткая топологическая идентификация, технологические процессы, сложные системы, диагностика.

DIAGNOSTICS OF CONTINUOUS TECHNOLOGICAL PROCESS BASED ON FUZZY TOPOLOGICAL IDENTIFICATION

In the article are observed the issues of fuzzy identification of continuous physical and technological processes simulation. Is suggested the new approach to construction of first rank uncertainty models diagnostics based on fuzzy logic.

***Key words*:** fuzzy logic, process simulation, fuzzy topological identification, technological process, complex systems, technical diagnosis.

Возникновение неисправностей, сбоев, отказов и, в конечном итоге, выход технологического оборудования из строя носит стохастический, трудно прогнозируемый характер, что особенно существенно для предприятий с непрерывными технологическими процессами. Поэтому задача диагностики состояний и коррекции неисправностей остается актуальной на протяжении всего хода производственного процесса. Одним из путей решения задачи повышения надежности производства является внедрение современных диагностических систем, ориентированных на обнаружение сбойных и предсбойных состояний, связанных со скрытыми дефектами, на ранних стадиях возникновения.

При этом особое внимание при разработке такого рода систем приходится уделять вопросам учета неопределенности технологической информации при принятии диагностических решений. Это связано как с неточностью или невозможностью измерений отдельных характеристик технологического процесса, так и с неоднозначностью и многовариантностью интерпретаций имеющихся знаний о диагностических свойствах таких сложных объектов, какими являются непрерывные технологические процессы.

Для решения задачи диагностирования в условиях неопределенности технологической информации необходима интеллектуализация процесса диагностики на основе привлечения теории нечетких множеств и

нечеткой логики, включая развитие и нестандартное использование на этой основе классических методов.

Примером такого подхода является использование обобщенной методологии, построенной по принципу эволюционного исследования сложных систем [2, 3] с привлечением классических статистических и нечетких технологий [1, 6, 7], позволяющего упорядочить процесс диагностики в виде поэтапной процедуры, организованной иерархически и решающей три слоя задач.

На нулевом уровне производится предварительное исследование диагностируемого технологического процесса, имеющего целью формирование его границ, определение вектора координат диагностической модели, т.е. получение модели нулевого ранга неопределенности. В зависимости от глубины знаний о процессе раскрытие начального уровня неопределенности реализуется либо на основе формализованных процедур (например, системный функционально-целевой и термодинамический анализ) либо привлечение экспертов в данной предметной области, использования их опыта и интуиции.

Первый уровень – экспериментальные исследования по сбору и систематизации измерительной информации для модели нулевого ранга неопределенности, обработка экспериментальной информации. В результате должны быть получены статистически представительные для дальнейших исследований исходные данные.

Второй уровень – нечеткая топологическая идентификация диагностической модели непрерывного технологического процесса, которая содержит три внутренних слоя задач:

1) анализ априорной информации;

2) проверка координат технологического процесса на изолированность (на связанность каждой координаты со всеми остальными);

3) анализ тесноты (наличия) связи между координатами процесса

Выходным результатом этого этапа является неориентированный нечеткий граф диагностической модели технологического процесса. Дополненная экспертными знаниями модель может быть доведена до модели первого ранга неопределенности (ориентированного графа). Но для целей диагностирования этот шаг не является обязательным, так как с точки зрения графового представления модели задача диагностики состоит в обнаружении, локализации и идентификации следующих видов отказов:

1) отсутствие вершины графа;

2) отсутствие связи между вершинами графа;

3) изменение силы (тесноты) связи.

Соответственно, ориентация связей между вершинами графа (причинно-следственный анализ) в задачах диагностики не является таким же важным этапом, как в задачах идентификации моделей непрерывных технологических процессов.

Представление диагностической модели непрерывного технологического процесса в виде нечеткого графа в рамках предлагаемой процедуры имеет свои особенности.

Традиционно нечеткий ориентированный граф определяется как граф $G = (X, U)$, у которого $X = \{x_i\}$, $i \in I = \{1,2,...n\}$ – четкое множество вершин, а $U = \{< \mu_U(x_i x_j)/(x_i x_j) >\}$ – нечеткое множество ребер, где $x_i x_j \in X$, $\mu_U(x_i x_j) \in [0,1]$ – значение функции принадлежности μ_U для ребра $(x_i x_j)$.

Здесь целесообразно рассматривать множество вершин графа также как нечеткое множество, т.е. $X = \{< \mu_U(x_i)/(x_j) >\}$, $i \in I = \{1,2,...n\}$ – нечеткое множество вершин и $\mu_x(x_i) \in [0,1]$.

Такое расширение определения нечеткого графа позволит описывать отказы типа выпадения (изоляции) вершин графа из модели.

Организационную последовательность проведения экспериментальных диагностических исследований можно представить в виде двух этапов.

I. Этап сбора экспериментальных данных для диагностики состояния непрерывного технологического процесса

На этом этапе реализуются следующие действия:

1) сбор экспериментальных данных для получения «пробных» выборок, необходимых

для исследования статистических свойств измеряемых сигналов;

2) первичная обработка данных с целью вычисления оптимальных периодов опроса и объема выборок;

3) сбор экспериментальных данных для получения «рабочих» выборок Т;

4) Нечеткая топологическая идентификация.

Топология объекта включает в себя множество координат и множество отношений между ними. Определение наличия этих отношений является целью нечеткой топологической идентификации. Таким образом, процесс построения топологии диагностической модели естественным образом разбивается на: 1) проверку координат на изолированность; 2) проверку координат на наличие связи.

Каждому типу проверки на втором этапе диагностических исследований соответствует своя теоретическая база, общим является использование в качестве основной информации экспериментальных данных, полученных на первом этапе, а также предположение о существовании корреляционной зависимости между исследуемыми координатами. Использование подобного предположения обуславливается тем, что в большинстве случаев координаты объекта являются случайными величинами и, следовательно, корреляционные критерии для проведения указанных выше проверок обладают наибольшей правдоподобностью и относительной легкостью реализации. В то же время, общность рассуждений не нарушается, если некоторые из координат процесса окажутся детерминированными, так как всякая детерминированная величина может рассматриваться как частный случай случайной.

Процедуры проверки координат на изолированность и наличие связи предполагают формирование матрицы корреляции R:

$$R = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nn} \end{pmatrix},$$

состоящей из парных коэффициентов корреляции (ПКК), учитывающих тесноту линейной корреляционной связи между всеми парами анализируемых координат с учетом их взаимного влияния и взаимодействия. Коэффициент корреляции между каждой парой координат X и Yr_{XY} рассчитывается по формуле

$$r_{XY} = \frac{N\sum_{i=1}^{N} X_i Y_i - \left(\sum_{i=1}^{N} X_i\right)\left(\sum_{i=1}^{N} Y_i\right)}{\sqrt{\left(N\sum_{i=1}^{N} X_i^2 - \left(\sum_{i=1}^{N} X_i\right)^2\right)\left(N\sum_{i=1}^{N} Y_i^2 - \left(\sum_{i=1}^{N} Y_i\right)^2\right)}}.$$

Для проверки координат на изолированность на основе матрицы корреляции рассчитывается для каждой координаты множественный коэффициент корреляции (МКК) $R_{0,1,2,\ldots,m}$, характеризующий степень тесноты связи между анализируемой координатой и всеми остальными независимыми переменными:

$$R_{0,1,2,\ldots,m} = \sqrt{1 - \frac{|R|}{|R_*|}},$$

где $|R|$ – определитель матрицы корреляции, $|R_*|$ – определитель матрицы корреляции, полученной после вычеркивания строки и столбца, представляющих связи исследуемой координаты.

Величина МКК лежит в пределах $0 \le R_{0,1,2,\ldots,m} \le 1$.

Для проверки на наличие связи на основе матрицы корреляции рассчитывается для каждой пары координат частный коэффициент корреляции (ЧКК) $r_{ij/km}$, характеризующий тесноту линейной корреляционной связи между парой анализируемых координат (x_i и x_j) без учета влияния на эту пару других переменных (x_k, x_m, и т.д.). Эти коэффициенты характеризуют так называемую чистую корреляцию. В матричном виде формулу для расчета частных коэффициентов корреляции можно записать следующим образом:

$$r_{ij/km} = \frac{-A_{ij}}{\sqrt{(A_{ii}A_{jj})}},$$

где A_{ii}, A_{jj}, A_{ij} – алгебраические дополнения соответствующих элементов матрицы парных корреляций R. Знак частному коэффициенту корреляции присваивается такой же, как и у парного коэффициента корреляции.

Величина ЧКК, как и МКК, лежит в пределах $0 \le r_{ij/km} \le 1$. При этом идеальными с точки зрения рассматриваемых процедур проверок являются крайние точки диапазона. При значении 0 для МКК делается вывод, что координата изолированная, значение ЧКК, равное 0, говорит об отсутствии связи между исследуемыми координатами. При единичных значениях ЧКК и МКК являются справедливыми альтернативные утверждения.

Однако оценки коэффициентов корреляции, вычисленные по ограниченной выборке, практически всегда отличается от нуля. Но из этого еще не следует, что коэффициент корреляции генеральной совокупности также отличен от нуля.

Если объем выборки небольшой, то найти выборочную ошибку этой величины достаточно сложно. Поэтому на практике можно использовать два основных подхода.

Первый подход базируется на применении методов математической статистики [4, 5]. В рамках этого подхода требуется оценить значимость выборочной величины коэффициента или, в соответствии с постановкой задач проверки статистических гипотез, проверить гипотезу о равенстве нулю соответствующего коэффициента корреляции. Если гипотеза H_0 о равенстве нулю коэффициента корреляции будет принята, то оценка коэффициента не значима, и делается вывод об изолированности координаты для МКК и отсутствии связи для ЧКК. Если гипотеза H_0 будет отвергнута, то выборочный коэффициент значим, справедливы альтернативные предположения. Проверка гипотезы о значимости оценки соответствующего коэффициента корреляции требует знания распределения этой случайной величины. К сожалению, распределение величины ПКК изучено только для частного случая, когда случайные величины, его образующие, распределены по нормальному закону.

Проверку статистических гипотез о значимости МКК и ЧКК проводят по-разному.

Для ЧКК в качестве критерия проверки нулевой гипотезы H_0 применяют случайную величину

$$t = \frac{|\,r_{ij/km}|\sqrt{n-k-2}}{\sqrt{1-r^2_{ij/km}}},$$

где k – число пассивных координат.

Если модуль коэффициента корреляции относительно далек от единицы, то величина t при справедливости нулевой гипотезы распределена по закону Стьюдента с $n-k-2$ степенями свободы, где n – объем выборки. Конкурирующая гипотеза H_1 соответствует утверждению, что значение $r_{ij/km}$ не равно нулю (больше или меньше нуля). Поэтому критическая область двусторонняя.

Проверка гипотезы H_0 о равенстве нулю ЧКК двумерной нормально распределенной случайной величины осуществляется в следующей последовательности:

• вычисляется значение статистики t;

• при уровне значимости α для двусторонней области определяется критическая точка распределения Стьюдента $t_{кр}(n-k-2;\alpha)$;

• сравнивается фактическое значение статистики t с критическим значением $t_{кр}(n-k-2;\alpha)$. Если $t < t_{кр}(n-k-2;\alpha)$, то нет оснований отвергнуть нулевую гипотезу, иначе гипотеза H_0 отвергается (коэффициент корреляции значим).

Когда модуль величины $r_{ij/km}$ близок к единице, распределение $r_{ij/km}$ отличается от распределения Стьюдента, так как значение $|r_{ij/km}|$ ограничено справа единицей. В этом случае применяют преобразование

$$y_{ij} = 0{,}5\ln[(1+|r_{ij/km}|)/(1-|r_{ij/km}|)].$$

Величина y_{ij} не имеет указанного ограничения, она при $n>10$ распределена приблизительно нормально с центром

$$\mu_1(r_{ij/km}) = 0{,}5\ln\,[(1+|r_{ij/km}|)/(1-|r_{ij/km}|)]+ \\ +0{,}5\,|r_{ij/km}|/(n-1)$$

и дисперсией

$$\mu_2(r_{ij/km}) = \sigma^2(r_{ij/km}) = 1/(n-3).$$

Если значение центрированной и нормированной величины

$$\left(y_{ij} - \mu_1(r_{ij/km})\right)/\sigma(r_{ij/km})$$

превышает значение квантили уровня $1-\alpha/2$ нормального распределения стандартизованной величины, то нулевая гипотеза отвергается.

Для МКК в качестве критерия проверки нулевой гипотезы H_0 применяют критерий Фишера

$$F = \frac{R^2_{0,1,2,\ldots,m}}{1-R^2_{0,1,2,\ldots,m}} \frac{n-m-1}{m},$$

где m – число координат.

Проверка гипотезы H_0 о равенстве нулю МКК проводится аналогично проверке равенства нулю ЧКК. Фактическое значение частного F-критерия сравнивается с табличным значением распределения Фишера при заданном уровне значимости α. Если фактическое значение превышает табличное, то включение координаты x_i в модель статистически оправданно. Если же фактическое значение меньше табличного, то нецелесообразно включение этой координаты в модель, координата изолирована.

Несомненным плюсом статистического подхода является возможность получения вероятностной оценки изолированности координаты или наличия связи как величины, равной $1-\alpha$.

Таким образом, реализация данного подхода находится фактически в русле классической логики: решение по каждой координате или связи принимается по принципу «нет/да». Но так как при этом вводится вероятностная оценка, то в полученной диагностической модели в форме нечеткого графа значения функций принадлежности для всех ребер и вершин будут равны $1-\alpha$ или α.

Второй подход базируется на применении методов нечеткой логики. В этом случае исходят не из необходимости знания законов распределения ПКК, а из качественной оценки тесноты связи между каждой координатой и всеми остальными (при проверке на изолированность) или между каждой парой координат (при проверке на наличие связи).

Для качественной оценки тесноты связи на основе МКК И ЧКК можно использовать следующую классификацию, часто применяемую на практике:

0.0 – нет связи (точно нет); 0.25 – слабая связь (скорее нет, чем есть); 0.5 – неопределенная связь (трудно отнести к слабой или сильной связи); 0.75 – возможная связь (скорее есть, чем нет); 1.0 – есть связь (точно есть).

Эта экспертное представление о трактовке тесноты связи позволяет построить механизм нечетких логических выводов для оценки нечеткой вероятности (значения функции принадлежности) наличия координаты и связи в нечетком графе, представляющем диагностическую модель технологического процесса.

В качестве входной лингвистической переменной R в процедуре проверки координаты на изолированность выступает МКК, а в процедуре проверки на наличие связи ЧКК. Каждая из этих переменных может принимать следующие нечеткие значения (термы): ноль (Н); меньше половины (МП); половина (П); больше половины (БП); единица (Е).

Начальное распределение термов входной переменной R может быть таким, как показано на рис.1.

Выходной лингвистической переменной P будет нечеткая вероятность наличия координаты при проверке на изолированность и тесноты связи при проверке наличия связи: нет связи (Н); слабая связь (СС); неопределенная связь (НС); возможная связь (ВС); есть связь (Е).

Начальное распределение термов выходной переменной P показано на рис.2.

База правил, связывающих входные и выходные лингвистические переменные, имеет вид:

Если R = Н, то P = Н;
Если R = МП, то P = СС;
Если R = П, то P = НС;
Если R = БП, то P = ВС;
Если R = Е, то P = Е.

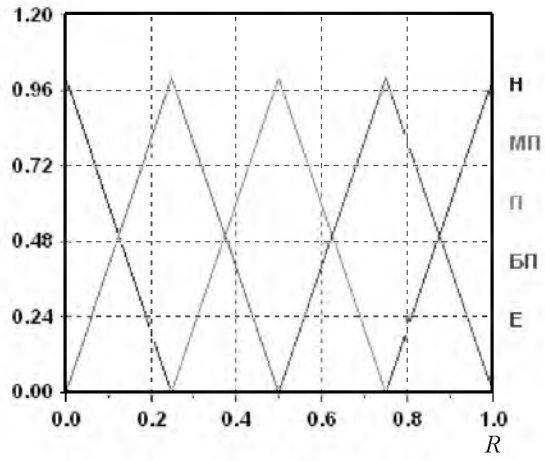


Рис.1

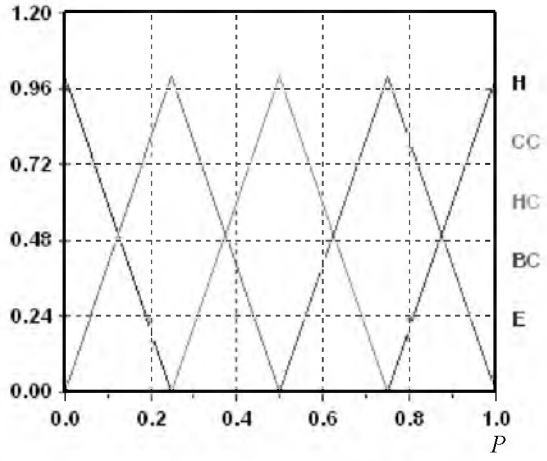


Рис.2

Механизм нечетких логических выводов может быть реализован по схеме MAX-MIN или MAX-PROD. В качестве метода дефазификации целесообразно использовать модифицированный метод центра тяжести.

На этапе настройки механизма нечетких логических выводов необходимо на основе полученных экспериментальных данных для нормального функционирования непрерывного технологического процесса с помощью изменения положения и формы термов лингвистических переменных добиться совпадения получаемой нечеткой топологической модели с экспертными представлениями о взаимосвязи между координатами. Возможна также вариация схем механизма нечетких логических выводов и методов дефазификации. При этом нечеткий граф должен получиться максимально близким к четкому (все степени принадлежности вершин и связей близки к 0 или 1).

После завершения настройки описанные алгоритмы готовы к работе в режиме диагностирования текущего состояния технологического процесса на основе нечеткой топологической идентификации диагностической модели в форме нечеткого графа.

ЛИТЕРАТУРА

1. *Алексеев А.А., Кораблев Ю.А., Шестопалов М.Ю.* Идентификация и диагностика систем. М., 2009.

2. *Вавилов А.А.* Структурный и параметрический синтез сложных систем. Л., 1979.

3. *Вавилов А.А., Имаев Д.Х.* Эволюционный синтез систем управления. Л., 1983.

4. *Елисеева И.И., Юзбашев М.М.* Общая теория статистики: Учебник / Под ред. И.И.Елисеевой. М., 2002.

5. *Кобзарь А.И.* Прикладная математическая статистика. М., 2006.

6. *Кораблев Ю.А., Шестопалов М.Ю.* Идентификация технологических процессов и управление ими на основе FUZZY- и NEURO-FUZZY-технологий: Учеб. пособие. СПб, 2007.

7. Прикладные нечеткие системы / К.Асаи, Д.Ватада, С.Иваи и др.; Под. ред. Т.Тэрано, К.Асаи, М.Сугено; пер. с япон. М., 1993.

REFERENCES

1. *Alekseev A., Korablev J., Shestopalov M.* Systems identification and technical diagnostics. Moscow, 2009.

2. *Vavilov A.* Structural and parametric synthesis of complex systems. Leningrad, 1979.

3. *Vavilov A., Imaev D.* Evolutional synthesis of control systems. Leningrad, 1983.

4. *Eliseeva I., Uzbashev M.* General statistics theory: Tutorial / Edited by I.Eliseeva. Moscow, 2002.

5. *Kobzar A.* Applied mathematical statistics. Moscow, 2006.

6. *Korablev J., Shestopalov M.* Identification of technological processes and control on the basis of FUZZY and NEURO-FUZZY technologies: Tutorial. Saint Petersburg, 2007.

7. Applied fuzzy systems / Edited by: T.Terano, K.Asai, M.Sugeno. Moscow, 1993.