

УДК 004.9

Тишин П.М. канд. физ-мат. наук, Маковецкий А.С.
Одесский национальный политехнический университет

ОПИСАНИЕ ЗАКОНОМЕРНОСТЕЙ ПРИ ДИАГНОСТИКЕ РАСПРЕДЕЛЕННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ, С ПРИМЕНЕНИЕМ МНОГОСОРТНОГО ЯЗЫКА ПРИКЛАДНОЙ ЛОГИКИ

Тишин П.М., Маковецкий О.С. Опис закономірностей при діагностиці розподілених інформаційних систем, із застосуванням багатосортної мови прикладної логіки. У статті побудована модель бази знань, яка описує нештатні і штатні ситуації, в розподілених інформаційних системах. Даний опис сформульовано в рамках багатосортної мови прикладної логічної теорії. В рамках розробленої моделі ставиться задача діагностики, яка полягає у визначенні всіх можливих альтернативних діагнозів на основі знань предметної області та даних спостережень, значень особливостей і виборі найбільш значущих з них в даній конкретній ситуації.

Ключові слова: бази знань, багатосортна мова прикладної логічної теорії, розподілені інформаційні системи.

Tishin P.M., Makovetskiy A.S. Описание закономерностей при диагностике распределенных информационных систем, с применением многосортного языка прикладной логики. В статье построена модель базы знаний, которая описывает нештатные и штатные ситуации, в распределенных информационных системах. Данное описание сформулировано в рамках многосортного языка прикладной логической теории. В рамках разработанной модели ставится задача диагностики, которая состоит в определении всех возможных альтернативных диагнозов на основе знаний предметной области и данных наблюдений, значений особенностей и выборе наиболее значимых из них в данной конкретной ситуации.

Ключевые слова: базы знаний, многосортный язык прикладной логической теории, распределенные информационные системы

Tishin P.M., Makovetskiy A.S. Description of regularities in the diagnosis of distributed information systems using polysorted language of applied logic. In this paper, a model of knowledge base that describes abnormal and staffing situations in distributed information systems was built. This description is formulated in the framework of polysorted language of applied logic theory. Within the developed model there is the problem of diagnostics which is to determine all the possible alternative diagnoses based on subject domain knowledge and observational data, the values of features and selecting the most important ones in this particular situation.

Keywords: knowledge bases, polysorted language of applied logic theory, distributed information systems.

Постановка науцнои проблеми. Современнне распределенные информационные системы (РИС) являются сложными системами сбора, обработки, передачи и хранения информации и определяют основную тенденцию развития информационных систем - переход от локальных к распределенным («облачным») вычислениям. Одним из необходимых условий эффективного функционирования РИС является штатное функционирование составляющих ее элементов. Они потенциально подвержены множеству внешних и внутренних воздействий, нарушающих штатное функционирование: перегрузки коммуникационных и обрабатывающих подсистем РИС; отказы оборудования; сбои программного обеспечения; компьютерные вирусы; атаки злоумышленников. Ввиду этого особую важность приобретают разработки, связанные с оперативным контролем функционирования всех подсистем РИС. Стремительный рост данных систем вызывает необходимость резкого увеличения количества администраторов, которые их обслуживают. Увеличения количества серверов, администрируемых одним специалистом, позволит не только сэкономить средства на обслуживание новых мощностей, но и обеспечить высокий темп их прироста. Такое увеличение может быть достигнуто только за счет большей автоматизации, что в свою очередь, потребует средств для более эффективного контроля ситуаций, за счет оперативной диагностики РИС.

Анализ исследований. С одной стороны, в настоящее время перспективным направлением построения систем диагностики является применение интеллектуальных подходов, основанных на знаниях, которые учитывают закономерности в работе программных и аппаратных компонентов РИС [1, 2]. Такие системы производят интеллектуальный анализ данных, поступающих из подсистем мониторинга компонентов РИС, основываясь на знаниях, накопленных в базе знаний. База знаний, при этом, является ключевым компонентом диагностической системы.

С другой стороны, при построении баз знаний для диагностических систем, применялся подход с использованием многосортного языка прикладной логики [3, 4]. Он предполагает использование многосортного языка прикладной логики [5, 6] для описания основных

закономірностей зустрічаються в роботі програмних і апаратних компонентів РИС. По порівнянню з іншими мовами подібного роду він має ряд переваг.

В перше, всі традиційні формальні мови проектуються як мови жорсткі, в які закладаються певний синтаксис і певна семантика. В відміння від них мова прикладної логіки з однієї сторони, має свої властивості, а з іншої ця мова може поповнюватися новими синтаксическими і семантическими конструкціями.

Друге переваження заключається в тому, що на мові прикладної логіки вже описано декілька моделей баз знань, суттєво більш складних, ніж на інших мовах.

Цілью данної роботи є розробка нової методики, для описання закономірностей, зустрічаються в роботі програмних і апаратних компонентів РИС [3, 4], заснованого на багатосортній мові прикладної логіки, результати якого можна використовувати при побудові інтелектуальних систем діагностики, заснованих на знаннях.

Основна частина. Задача діагностики, в межах данної роботи, складається в визначенні всіх можливих альтернативних діагнозів на основі знань предметної області і даних спостережень, значень особливостей (постійні в часі) і виборі найбільш значимих з них в даній конкретній ситуації. Припускається, що в кожному діагнозі можуть входити декілька причин відхилень, протікають одночасно або послідовно. Вважається, що всі спостереження протікають на певній часовій ділянці, називається періодом спостереження, на якому моменти часу вимірюються цілими неотрицательними числами, що минули з моменту початку спостережень.

При розв'язанні задачі діагностики припускається що визначені наступні бази знань:

- знання, що задовольняють обмеженням цілостності знань;
- знання про спостереження і їх можливі значення;
- знання описувальні нештатні і штатні ситуації, а також значення параметрів в цих ситуаціях;

- знання про причинно-наслідкові зв'язки між причинами відхилень і спостереженнями.

Крім того, припускається, що в системі знаходяться наступні результати спостережень:

- спостережені особливості і результати їх спостереження (значення);
- спостережені параметри, моменти їх спостереження і їх значення в ці моменти;

В термінах використовуваної моделі математична постановка задачі діагностики може бути сформульована наступним чином — потрібно знайти причини відхилень, які приводять РИС в ту або іншу нештатну ситуацію.

В якості прикладу наведемо модель бази знань, описувальну нештатні і штатні ситуації, а також значення параметрів в цих ситуаціях. Данне описання сформульовано в межах багатосортної мови прикладної логічної теорії.

Описувальна модель містить наступні аксіоми:

(1) Термін знання про штатні ситуації (ЗШС) це множина структурних значень з атрибутами наслідок, варіанти, фактори, кожне з яких описує знання про конкретну штатну ситуацію. Значенням першого атрибута є ім'я параметра, другого – множина варіантів штатних ситуацій для цього параметра, третього – множина особливостей.

ЗШС □ (наслідок □ собст_параметри_служби,
варіанти □ { } варіанти штатних ситуацій,
фактори □ { } особливості)

(2) Термін знань варіанти штатних ситуацій це множина структурних значень з атрибутами область значень наслідка, умови на фактори, кожне з яких описує знання про конкретний варіант штатної ситуації. Значенням першого атрибута є множина значень параметра в цьому варіанті, а другого – умова.

варіанти штатних ситуацій □ (
область значень наслідка □ множина значень,
умова на фактори □ умови)

(3) В знаннях про штатні ситуації в будь-якому варіанті норми область значень наслідка – це власне підмножина можливих значень параметра, являючогося наслідком цієї штатної ситуації.

(x: ЗШС)

(v: варианты(x))

область значений следствия(v) □ возможные значения(следствие(x))

(4) Термин действительности штатные ситуации (ШС) это множество структурных значений с атрибутами следствие и вариант, каждое из которых описывает наблюдающую штатную ситуацию. Значением первого атрибута является имя параметра, а второго – варианты штатных ситуаций.

ШС □ (следствие □ собст_параметры_службы,
вариант □ варианты штатных ситуаций)

(5) Если присутствует некоторая штатная ситуация, то в множестве знания о штатных ситуациях присутствует такой элемент, у которого следствие совпадает со следствием штатной ситуации, в множестве вариантов штатных ситуаций этого элемента содержится вариант штатных ситуаций из штатной ситуации и для него выполнено условие на факторы.

(x: ШС)(□ (z: ЗШС) следствие(z) = следствие(x)&& вариант(x) □ варианты(z)&& выполнено(условие на факторы(вариант(x))))

(6) Термин знания о нештатных ситуациях (ЗНШС) это множество структурных значений с атрибутами причина, период развития, следствие, варианты, факторы, необходимое условие(ну), модальность, каждое из которых описывает знания о конкретной нештатной ситуации. Значением причины является имя причины отклонения, значением периода развития – номер периода развития причины, значением следствия – имя параметра, значением вариантов – множество вариантов ншс, значением факторов множество особенностей, значением необходимого условия – множество условий, значением модальности – необходимость или возможность.

ЗНШС □ (причина □ причины отклонения,
период развития □ I[1, чпр(причина)],
следствие □ собст_параметры_службы,
варианты □ { } варианты ншс,
факторы □ { } особенности,
ну □ условия,
модальность □ приоритеты)

(7) Термин варианты ншс это множество структурных значений с атрибутами число периодов динамики(чпд), описание динамики, условие на факторы, каждое из которых описывает знания о конкретном варианте нештатной ситуации. Значением первого атрибута является положительное целое число. Второй атрибут является функцией, которая номеру периода динамики сопоставляет период динамики. Значение третьего атрибута – множество условий.

варианты ншс □ (чпд □ I[1, □],
описание динамики □ (I[1, чпд] □ периоды динамики),
условие на факторы □ условия)

(8) Термин действительности нештатные ситуации (НШС) это множество структурных значений с атрибутами причина, период развития, следствие, вариант, динамика значений, модальность, каждое из которых описывает некоторую нештатную ситуацию. Значением причины является имя причины отклонения, значением периода развития - номер периода развития, значением следствия – имя параметра, значением варианта – вариант ншс, значением динамики значений – разбиение, значением модальности значение приоритета.

НШС □ (причина □ причины отклонения,
период развития □ I[1, чпр(причина)],
следствие □ собст_параметры_службы,
вариант □ варианты ншс,
динамика значений □ разбиения,
модальность □ приоритеты)

(9) Если присутствует нештатная ситуация, на некотором периоде ее развития, то начало его динамики значений совпадает с моментом начала периода развития, а конец – с концом этого периода.

$$\begin{aligned} & (x: \text{НШС}) \text{el}(\text{динамика значений}(x), 0) = \\ & = \text{el}(\text{развитие}(\text{причина}(x)), \text{период развития}(x) - 1) \& \\ & \& \text{el}(\text{динамика значений}(x), \text{length}(\text{динамика значений}(x))) = \\ & = \text{el}(\text{развитие}(\text{причина}(x)), \text{период развития}(x)) \end{aligned}$$

(10) Если в ситуации присутствует нештатная ситуация, на некотором периоде ее развития, то в модель знаний входит элемент множества знания о нештатных ситуациях, у которого:- причиной является эта причина отклонения с тем же периодом развития,- следствие совпадает со следствием нештатной ситуации,- выполнено необходимое условие,- модальность совпадает с модальностью нештатной ситуации,- множеству вариантов ншс принадлежит вариант ншс нештатной ситуации, причём такой, что количество интервалов в динамике значений нештатной ситуации равно чпд этого варианта и выполнено условие на факторы.

$$\begin{aligned} & (x: \text{НШС})(\square (z: \text{ЗНШС}) \\ & \text{причина}(z) = \text{причина}(x) \& \\ & \& \text{период развития}(z) = \text{период развития}(x) \& \\ & \& \text{следствие}(z) = \text{следствие}(x) \& \\ & \& \text{выполнено}(\text{необходимое условие}(z))\& \\ & \& \text{модальность}(z) == \text{модальность}(x) \& \\ & \& \text{вариант}(x) \square \text{варианты}(z) \& \\ & \& \text{length}(\text{динамика значений}(x)) - 1 = \text{чпд}(\text{вариант}(x))\& \\ & \& \text{выполнено}(\text{условие на факторы}(\text{вариант}(x))). \end{aligned}$$

Основной алгоритм решения сформулированной задачи диагностики можно представить следующим образом:

1. Выбрать все наблюдаемые параметры с момента начала наблюдений до текущего момента времени.
 2. Выполнить проверку необходимых условий для данных параметров.
 3. Если для некоторого параметра не выполнено необходимое условие то принять решение, что нахождение причины отклонения невозможно, и выйти из задачи.
 4. Если для всех параметров выполнены необходимые условия то перейти к шагу 5.
 5. Проверить гипотезу о том, что наблюдается ШС (см. функцию Проверка_ШС).
 6. Если гипотеза о том, что наблюдается ШС, подтвердилась, то считать:
 - что в результате наблюдается ШС и отклонения отсутствуют;
 - перейти к шагу 8.
 7. Если гипотеза о том, что наблюдается ШС, не подтвердилась, то:
 - перебрать все причины отклонений из базы знаний, причем для каждой проверить выполнение необходимого условия для этой причины отклонений;
 - если необходимое условие выполнено, то проверить гипотезу о том, что наблюдается НШС вызванная этим отклонением (см. функцию Проверка_НШС).
 - в случае ее подтверждения добавить это отклонение к множеству решений.
 - перейти к шагу 8.
 8. Завершить работу, выдав полученные результаты.
- Приведем алгоритмы основных функций которые используются для решения задачи.

Функция Проверка_ШС предназначена для проверки гипотезы о том, что наблюдается ШС. Гипотеза о том, что наблюдается ШС, считается подтвержденной, если для каждого наблюдаемого параметра подтверждена гипотеза о том, что наблюдаемые значения этого параметра могут иметь место в некоторой ШС. Если хотя бы для одного наблюдаемого параметра такая гипотеза опровергается, то считается, что имеет место НШС.

Ниже приведено описание алгоритма функции Проверка_ШС

гипотеза ШС = ЛОЖЬ

ДЛЯ ВСЕХ параметр наблюдавшиеся параметры ВЫПОЛНЯТЬ

гіпотеза = ЛОЖЬ
гіпотеза = Проверка_ШС_по_параметру(параметр)
ЕСЛИ *гіпотеза == ЛОЖЬ*
ТО
гіпотеза ШС = ЛОЖЬ
ВЫХОД ИЗ ЦИКЛА
ИНАЧЕ
гіпотеза ШС = ИСТИНА
ЗАКОНЧИТЬ ТО
ЗАКОНЧИТЬ ДЛЯ ВСЕХ

Функция Проверка_ШС_по_параметру(параметр) предназначена для проверки гипотезы о том, что все наблюдаемые значения некоторого параметра могут иметь место в ШС. Эта гипотеза считается подтвержденной, если для каждого наблюдавшегося значения параметра найден некоторый вариант ШС, протекавший на протяжении всего периода наблюдений. Для осуществления этого поиска строится множество гипотез о развитии причинно-следственных связей (далее – МГ), куда входят варианты ШС, которые могут определять значения этого параметра. Наблюдаемые значения каждого параметра проверяются последовательно, начиная с самого раннего.

Ниже приведено описание алгоритма функции Проверка_ШС_по_параметру (параметр)

ДЛЯ ВСЕХ *x* ШС ВЫПОЛНЯТЬ
ЕСЛИ *следствие(x) == параметр*
ТО
ДЛЯ ВСЕХ *v* вариант(*x*) ВЫПОЛНЯТЬ
ЕСЛИ *условия на факторы (v) == ИСТИНА*
ДОБАВИТЬ В МГ(*x, v, период динамики = 1*)
ЗАКОНЧИТЬ ДЛЯ ВСЕХ
ЗАКОНЧИТЬ ТО
ЗАКОНЧИТЬ ДЛЯ ВСЕХ
гіпотеза = ЛОЖЬ
1: ДЛЯ ВСЕХ *t* МГ ВЫПОЛНЯТЬ
p = текущий период динамики
ЕСЛИ (*поиск_ШС(p, t, параметр) == ИСТИНА*)
2: ТО
запомнить, что все наблюдаемые значения рассматриваемого признака могут наблюдаться в ШС
гіпотеза = ИСТИНА
2: ИНАЧЕ
запомнить, что все наблюдаемые значения рассматриваемого признака не могут наблюдаться в ШС
гіпотеза = ЛОЖЬ
2: ЗАКОНЧИТЬ ТО
1: ЗАКОНЧИТЬ ДЛЯ ВСЕХ

Функция Проверка_НШС предназначена для проверки гипотезы о том, что наблюдается НШС, начавшаяся в момент начала наблюдений. Гипотеза о том, что наблюдается НШС, считается подтвержденной в том случае, если для каждого наблюдавшегося параметра не удастся опровергнуть гипотезу о том, что его наблюдавшиеся значения соответствуют знаниям о возможном развитии параметра в этой НШС.

Ниже приведено описание алгоритма функции Проверка_НШС(причина отклонения)

1: ДЛЯ ВСЕХ *параметр* наблюдавшиеся *параметры*
2: ДЛЯ ВСЕХ *x* НШС ВЫПОЛНЯТЬ
ЕСЛИ (*НЕ (следствие(x) == параметр)*) И (*причина(x) == причина отклонений*) ВЫПОЛНЯТЬ

result = Проверка_НШС_по_параметру (параметр, x, причина отклонения)
ЕСЛИ (result == ИСТИНА)
3: ТО
гипотеза НШС = ИСТИНА
3: ИНАЧЕ
гипотеза НШС= ЛОЖЬ
4: ВЫХОД ИЗ ЗАДАЧИ
3: ЗАКОНЧИТЬ ТО
2: ЗАКОНЧИТЬ ДЛЯ ВСЕХ
1: ЗАКОНЧИТЬ ДЛЯ ВСЕХ

Функция Проверка_НШС_по_параметру предназначена для проверки гипотезы о том, что все наблюдавшиеся значения некоторого параметра могут иметь место в НШС. Эта гипотеза считается подтвержденной, если для каждого наблюдаемого значения параметра найдена его причина:

- в случае, если данный параметр не входит в НШС – вариант ШС;
- в случае, если данный параметр входит в НШС – вариант НШС.

Для осуществления этого поиска строится множество гипотез о развитии причинных связей, куда входят указанные выше причинно-следственные связи (либо из первой группы, либо из второй), которые могут определять значения этого параметра. Наблюдаемые значения каждого параметра необходимо проверять последовательно (по времени наблюдения), начиная с самого раннего.

Ниже приведено описание алгоритма функции Проверка_НШС_по_параметру

ЕСЛИ (НУ(причина отклонения) == ИСТИНА) И (следствие(x) ==
параметр) И (причина(x) == причина отклонений)
1: ТО
2: ДЛЯ ВСЕХ v вариант(x) ВЫПОЛНЯТЬ
ЕСЛИ условия на факторы (v) == ИСТИНА
3: ТО
ДОБАВИТЬ В МГ(x, v, период динамики =1)
3: ЗАКОНЧИТЬ ТО
2: ЗАКОНЧИТЬ ДЛЯ ВСЕХ
1: ЗАКОНЧИТЬ ТО
ЕСЛИ МГ == пусто
1: ТО
Проверка_ШС
1: ИНАЧЕ
t = first (параметр)
ЕСЛИ поиск_НШС_приоритет (МГ, параметр, t) == ИСТИНА
2: ТО
гипотеза = ИСТИНА
2: ИНАЧЕ
гипотеза = ЛОЖЬ
2: ЗАКОНЧИТЬ ТО
1: ЗАКОНЧИТЬ ТО

Здесь НУ(причина отклонения)- функция принимающая значения истина или ложь, в зависимости от выполнения необходимых условий для наличия данной причины отклонения, а функция first (параметр) - определяет начальный момент наблюдения параметра.

Функция поиск_НШС_приоритет(МГ, параметр, t)

Функция поиск_НШС_приоритет предназначена для проверка гипотезы о том, что некоторое значение параметра возможно в некотором варианте НШС. Стратегия решения данной подзадачи учитывает следующие основные моменты:

- в случае отсутствия на момент наблюдения значения параметра в МГ вариантов НШС считается, что гипотеза опровергнута;

- выбор из МГ причинно-следственных связей, которые имеют наивысший приоритет (а значит, только среди них может быть причина рассматриваемого параметра) ведется с учетом их приоритетов.

Ниже приведено описание алгоритма функции поиск_НШС_приоритет

1: ДЛЯ ВСЕХ m МГ ВЫПОЛНЯТЬ

p = выбрать текущий период динамики в зависимости от m , t

1: ЗАКОНЧИТЬ ДЛЯ ВСЕХ

1: ДЛЯ ВСЕХ pr приоритеты ВЫПОЛНЯТЬ

2: ДЛЯ ВСЕХ m МГ ВЫПОЛНЯТЬ

ЕСЛИ (приоритет(m) = pr)

ТО

vr = поиск_НШС (p , m , параметр);

ЕСЛИ (vr = положительный)

ТО

временный результат текущей подзадачи = vr

выход из циклов 2 и 1

ИНАЧЕ

временный результат текущей подзадачи = vr

ЗАКОНЧИТЬ ЕСЛИ

ЗАКОНЧИТЬ ТО

2: ЗАКОНЧИТЬ ДЛЯ ВСЕХ

1: ЗАКОНЧИТЬ ДЛЯ ВСЕХ

Выводы и перспективы дальнейших исследований. Таким образом, был разработан новый подход к описанию закономерностей, возникающих в процессе функционирования аппаратных и программных компонентов РИС, основанный на многосортном языке прикладной логики. Созданные на его основе базы знаний могут быть использованы для построения интеллектуальных систем диагностики РИС с гибким синтаксисом описания знаний. Гибкость в описании знаний и расширенные возможности выразительных средств позволят упростить процесс модернизации базы знаний диагностической системы уже в процессе её функционирования.

Разработанный подход позволит в дальнейшем разрабатывать базы знаний, обладающие более сложной и гибкой структурой, а так же решать более сложные задачи.

1. Hanemann. A. A Hybrid Rule-Based/Case-Based Reasoning Approach for Service Fault Diagnosis [Текст] / A. Hanemann // in Proceedings of 20th International Conference on Advanced Information Networking and Application (AINA2006), includes proceedings of International Symposium on Frontiers in Networking with Applications (FINA 2006). – Vienna: 2006. – P. 734–738.
2. Lewis L. Managing Computer Networks: A Case-Based Reasoning Approach [Текст] / L. Lewis. – Norwood: Artech House Inc., 1995. – 205 p.
3. Нестеренко С. А., Тишин П. М., Маковецкий А. С. Разработка модели онтологии диагностики сервис-ориентированных сетевых структур на основе многосортного языка прикладной логики [Текст] / С.А. Нестеренко, П.М. Тишин, А.С. Маковецкий // Электротехнические и компьютерные системы. – Одесса: 2012. – № 07(83). – С 102-108.
4. Нестеренко С. А., Тишин П. М., Маковецкий А. С. Модель онтологии априорного подхода прогнозирования проблемных ситуаций в сложных вычислительных системах [Текст] / С.А. Нестеренко, П.М. Тишин, А.С. Маковецкий // Электротехнические и компьютерные системы. – Одесса: 2013. – № 10(86). – С 111-119.
5. Клещев А. С., Артемьева И. Л. Необогатенные системы логических соотношений [Текст] / А. С. Клещев, И. Л. Артемьева // Научно-техническая информация. Серия 2 –2000. – Ч. 1.– № 7.– С. 18–28. – Ч. 2. – № 8. – С. 8–18.
6. Клещёв А. С., Москаленко Ф. М., Черняховская М. Ю. Модель онтологии предметной области «Медицинская диагностика». Часть 2. Формальное описание причинно - следственных связей, причин значений признаков и причин заболеваний. [Текст] / А. С. Клещёв, Ф. М. Москаленко, М. Ю. Черняховская // НТИ. Серия 2. – 2006. – No 2. – С.19-30.