

Кретов В.С.
ИГУПИТ (Москва)

Поддержка принятия решений по выбору программных средств для аналитических подразделений и ситуационных центров

Проблема принятия решений по выбору программных средств, закупаемых для аналитических подразделений и ситуационных центров различного назначения, состоит в том, что, с одной стороны, рынок предлагает большое количество самых разнообразных программных средств, а с другой стороны, заказчику порою трудно определиться с приоритетностью выдвигаемых им требований (иногда достаточно противоречивых).

Предлагаемый нами подход к компьютеризации процесса выбора закупаемых программных средств рассмотрим на примере выбора информационно-аналитической системы (ИАС) для анализа структурированной информации. Актуальность этой задачи определяется не только важностью выполняемых ИАС функций, но и высокой ценой ошибки выбора, поскольку ИАС являются весьма дорогостоящим продуктом.

1. Принцип компьютерного выбора ИАС

Для выбора ИАС предлагается использовать разработанную нами специальную экспертную систему, основанную на идеях Дж. Ханта [1] и аппарате нечеткой математики [2].

Э. Хантом была разработана система формирования понятий (СФП), использующая для описания предъявляемых ей объектов схему "признак - значение". Описываемые по этой схеме объекты представляются матрицей A , строки которой соответствуют объектам, а столбцы - характеризующим эти объекты признакам. Элемент a_{ij} этой матрицы соответствует i -му значению j -го признака.

СФП сначала обучается путем предъявления ей выборки заранее отклассифицированных экспертом объектов, что позволяет автоматически

сформировать правило классификации. Это правило затем используется для классификации новых объектов.

Приведем формальную постановку задачи выбора ИАС для анализа структурированной информации

Дано:

1. Множество D^k отклассифицированных заказчиком описаний требований, предъявляемых к программным средствам k - го типа.

$$D^k = \{D_i^k\}; D_{ij}^k = \{Q^k\}; Q^k = \{Q_t^k\};$$

$$i = 1, I^k; j = 1, J^k; t = 1, Q_{\max}^k; k = 1, K;$$

I^k - число различных описаний требований к программным средствам k - го типа;

J^k - число признаков описаний требований к программным средствам k - го типа;

Q_{\max}^k - число различных значений признаков описаний требований к программным средствам k - го типа.

2. Множество Z^k присвоенных заказчиком "весовых" коэффициентов, определяющих "веса" всех значений признаков.

$$Z^k = \{Z_t^k\};$$

3. Множество W^k описаний программных средств-претендентов каждого из k типов.

$$W^k = \{W_m^k\}; W_{mj}^k = \{Q^k\}; Q^k = \{Q_t^k\}; m = 1, M^k$$

Требуется:

Выбрать из множества W^k описаний программных средств-претендентов по одному описанию программного средства каждого типа, удовлетворяющего предъявленным требованиям.

Совокупность R выбранных программных средств образует программный инструментарий ситуационного центра или аналитического подразделения организации

$$R = \{R^k\}; k = 1, K$$

Выбор программных средств каждого типа проводится в два этапа.

1 этап - формирование правила классификации. На практике заказчик обычно затрудняется сформулировать конкретные требования к выбираемому программному продукту (особенно когда число требований достаточно велико). Поэтому предлагается формировать правило классификации программных средств k -ого типа в процессе обучения - предъявляя системе имеющую $(J^k + 1)$ столбцов обучающую матрицу D^k , i -тая строка которой представляет собой некоторое отличное от других описание требований по схеме "признак-значение".

При этом заказчик на основании своего опыта должен, во-первых, высказать свои предпочтения в отношении рейтинга значений признаков описания требований, присваивая каждому из них "весовой" коэффициент, значение которого выбирается им в интервале от 0 до 1. Во-вторых, заказчик должен определить, какие из описаний требований, т.е. сочетаний характеризующих программные средства значений признаков, кажутся ему предпочтительными, а какие - нет. Первые он относит к положительным объектам, относя их к классу №1, а вторые - к отрицательным, относя их к классу №2.

На 2-ом этапе проводится автоматическая классификация программного продукта-претендента, который в соответствии с выработанным на 1-ом этапе классификационным правилом будет либо отнесен к положительным объектам (класс №1) и тогда данный программный продукт-претендент будет включен в состав закупаемого программного инструментария для ситуационного центра или аналитического подразделения организации, либо нет.

В качестве программной основы предлагаемого подхода к компьютеризации процесса выбора закупаемых программных средств предлагается использовать разработанную нами Экспертную систему поддержки принятия решений в кризисных ситуациях (ЭС ПРКС), которая предназначена для автоматической классификации текущей ситуации в условиях неполной и нечёткой входной информации и формирования управленческого решения, соответствующего распознанному классу ситуации [3].

1. Пример применения предлагаемого подхода

Сначала сформируем пространство описания объектов обучающей матрицы, используя показатели качества, выбранные для сравнительной оценки ИАС [4].

1-ая группа признаков «Переносимость программного обеспечения»:
подгруппа 1.1 «Переносимость серверных компонентов» (наличие версии для win32, наличие версии для Linux, наличие версии для FreeBSD); подгруппа 1.2 «Возможность использования различных СУБД» (возможность использования Oracle Database, возможность использования Microsoft SQL Server, возможность использования других СУБД); подгруппа 1.3 «Переносимость клиентских компонентов» (наличие полнофункционального тонкого клиента, наличие полнофункционального Win32-клиента, наличие полнофункционального Java-клиента).

2-ая группа признаков «Производительность и масштабируемость»:
подгруппа 2.1 «Зависимость производительности от объемов» (наличие средств фонового вычисления, наличие оценки трудоемкости вычислений); подгруппа 2.2 «Средства распараллеливания вычислений» (наличие средств многопоточной обработки, наличие средств многопроцессорной обработки, наличие средств распараллеливания по СВТ).

3-ья группа признаков «Средства защиты информации»: подгруппа 3.1 «Средства разграничения доступа» (системное разграничение доступа ОС и СУБД), средства прикладного разграничения доступа, наличие средств интеграции с LDAP (Microsoft Active Directory); подгруппа 3.2 «Средства защиты информации от сбоев» (использование системных средств защиты от сбоев, наличие прикладных механизмов защиты от сбоев).

4-тая группа признаков «Функции управления моделями данных»:
подгруппа 4.1 «Системная модель данных» (описание системной модели данных на специальном языке, визуальное описание системной модели данных, возможность модификации модели администратором, возможность работы с несколькими моделями данных); подгруппа 4.2 «Модели данных источников»

(описание моделей данных источников на специальном языке, визуальное описание моделей данных источников, возможность модификации моделей администратором); подгруппа 4.3 «Правила преобразования данных» (описание правил мэппинга данных на специальном языке, визуальное описание мэппинга данных, модификация мэппинга администратором, вычисления в процессе преобразования).

5-тая группа признаков «Функция сбора и накопления информации»:

подгруппа 5.1 «Регламенты ввода информации» (использование регламента ввода данных по времени, использование сценариев ввода данных, ввод информации из реляционных баз данных, ввод данных из структурированных файлов); подгруппа 5.2 «Автоматический ввод информации» (наличие средств мониторинга источников, возможность поиска данных в источниках (online), наличие средств преобразования форматов).

6-тая группа признаков «Функции логико-семантической обработки»:

подгруппа 6.1 «Идентификация объектов» (идентификация объектов по синонимам, идентификация объектов по ключевым атрибутам, идентификация связей, язык правил идентификации объектов, визуальное редактирование правил идентификации, редактирование правил администратором); подгруппа 6.2 «Логический вывод» (логический вывод в интерфейсе, логический вывод в базе знаний, язык правил логического вывода, визуальное редактирование правил вывода, редактирование правил администратором).

7-ая группа признаков «Функции поиска информации»: подгруппа 7.1

«Базовые режимы поиска объектов» (атрибутивный поиск объектов, использование комбинации условий в запросе, ограничение области поиска (тип объекта, класс и др.), сквозной полнотекстовый поиск объектов); подгруппа 7.2 «Расширенные режимы поиска объектов» (задание запроса на естественном языке, поиск объектов с учетом опечаток (нечеткий), визуальное формирование поискового запроса, визуальное формирование поискового запроса).

8-ая группа признаков «Аналитические функции»: подгруппа 8.1

«Аналитические запросы» (поиск группы связанных объектов (фактов, ситуаций), поиск связей объекта, поиск цепочек связей (траекторий), формирование досье объекта, мониторинг поведения объектов (фактов), прогнозирование поведения объектов); подгруппа 8.2 *«Работа с результатами запросов»* (теоретико-множественные операции, редактирование объектов, связей (сеть, таблица и др.), расширенные функции визуального редактирования, сценарии выполнения запросов, последовательное сужение запросов, отображение результатов запросов средствами ГИС).

9-ая группа признаков *«Функции использования результатов»:* подгруппа 9.1 *«Функции совместной работы»* (возможность использования общих папок, возможность открытия личных папок для чтения, возможность открытия личных папок для записи, возможность экспорта/импорта рабочих данных); подгруппа 9.2 *«Функции экспорта результатов»* (формирование статистических отчетов, экспорт в виде структурированных файлов, экспорт результатов в форме отчетов Microsoft Office, HTML), возврат результатов посредством API (Web Services и др.).

10-тая группа признаков "Ценовой диапазон" с признаками: "до 100 тысяч рублей", "от 101 тысячи рублей до 500 тысяч рублей", "от 501 тысячи рублей до 1 миллиона рублей", "свыше 1 миллиона рублей";

Для обучения ЭС ПРКС было выбрано 20 информационно-аналитических систем, предназначенные для анализа структурированной информации, в том числе такие известные системы, как:

- Система X-Files (разработчик - ЗАО «Ай-Теко») включает в свой состав: полнотекстовую поисковую подсистему на основе Hummingbird Search Server, набор автоматов для структуризации полнотекстовой информации, набор приложений на платформе .Net, реализующих прикладные функции системы, АРМ пользователей на основе web-интерфейса.
- Система TAIS/Ontos (разработчик – ЗАО «Авиакомп Сервисе») включает в свой состав: шлюзы для интеграции с источниками

структурированной информации; сервер транзакций, обеспечивающий интероперабельность (выполнение единого бизнес-процесса при работе с несколькими источниками и потребителями информации); сервер приложений, обеспечивающий работу пользователей с БД ТАIS;

- АРМ пользователей на основе тонкого клиента.
- Система АРИОН (разработчик – ООО «САЙТЭК») включает в свой состав: адаптеры для интеграции с источниками документальной информации; лингвистический процессор «АРИОН-Лингво»; сервер приложений; АРМ пользователей.

Кортеж признаков продукта-претендента: ценовой диапазон – свыше 1 млн. руб.; функции управления моделями данных – описание системной модели данных на специальном языке; функции поиска информации - атрибутивный поиск объектов; переносимость программного обеспечения – наличие версии для Linux; производительность и масштабируемость – наличие средств фонового вычисления, функции сбора и накопления информации – возможность поиска данных в источниках.

Рекомендованное системой управленческое решение – «закупать» с достоверностью в пределах 0,67-1,0. По экспертным оценкам результаты апробации признаны удовлетворительными.

Список литературы:

1. Хант Э., Мартин Дж., Стоун Ф. Моделирование процесса формирования понятий на вычислительной машине.- М.: Мир, 1970.
2. Заде Л. Понятие лингвистической переменной и ее применение к принятию приближенных решений.- М.: Мир, 1976.
3. Кретов В.С., Лебедев И.С. Экспертная система поддержки принятия решений в кризисных ситуациях (в настоящем сборнике).
4. Беляев К.В., Босов А.В., Краюшкин Д.В. Обзор и сравнительный анализ информационно - аналитических систем. -М.:ИПИ РАН, 2008.