

ОБРАБОТКА ДАННЫХ ИСКУССТВЕННЫМИ ИММУННЫМИ СЕТЯМИ В СИСТЕМАХ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

Г. А. Самигулина, к.т.н.

Институт проблем информатики и управления

Кеп елшемд1 мэл1меттерд1 ездеудщ жасанды иммундык жүйелерше непзделген интеллектуалдык технологиясы жасалган.

Туйшд| сездер: Жасанды иммундык жүйелер, кепелшегад мэл1меттер, интеллектуальдык технология.

The intellectual technology of multidimensional data processing on the basis of the Artificial Immune Systems is developed.

Key words: Artificial Immune Systems, multidimensional data, intellectual technology.

Стремительное развитие современных информационных технологий в области интеллектуальных систем нового поколения приводит к необходимости разработки нетрадиционных биологических подходов. В данном направлении наиболее развитыми в настоящий момент являются искусственные нейронные сети (нейрокомпьютеры), клеточные автоматы, генетические алгоритмы и искусственные иммунные сети (иммуннокомпьютеры). Особенно пристальное внимание во всем мире уделяется искусственным иммунным системам (Artificial Immune System, AIS), основанным на принципах обработки информации молекулами белка.

Прототипом AIS выступает иммунная система человека, высокопараллельные механизмы функционирования которой поражают быстродействием, экономичностью и эффективностью. Особенно привлекает способность естественных биологических систем решать многомерные задачи огромной вычислительной сложности в реальном масштабе времени. Пространственная конфигурация, принципы самосборки белков, их комплексов и сетей наиболее важны при обработке информации молекулами белка.

Разработаны интеллектуальные экспертные системы поддержки принятия решений на основе уникальной информационной техно-

логии обработки информации искусственными иммунными сетями в реальном масштабе времени.

Иммунные сети (ИС) - это адаптивные системы для обработки и анализа данных, которые представляют собой математическую структуру, имитирующую некоторые функции иммунной системы человека, обладающие способностью к обучению и прогнозированию на основе уже имеющихся временных рядов, принятию решения в незнакомой ситуации. ИС в принципе не нуждаются в заранее известной модели, а строят ее сами на основе полученной информации в виде временных рядов. Входными характеристиками AIS являются временные ряды, составленные из различных параметров, описывающих рассматриваемую систему. Вся необходимая информация заносится в базы данных (БД) и базы знаний (БЗ). Для построения иммунной сети в качестве признаков берутся поля таблиц БД.

Принципы действия механизмов репарации (т. е. исправления ошибок в процессе функционирования иммунной системы) позволяют строить принципиально новые алгоритмы распознавания образов на основе искусственных иммунных сетей.

Разработанная интеллектуальная технология обработки многомерной информации искусственными иммунными сетями решает две основные задачи.

Первая задача заключается в быстром обучении иммунной сети. Успешное развитие AIS напрямую зависит от создания алгоритмов, которые обеспечили бы наиболее эффективное обучение иммунной сети за минимально короткое время. В связи с данной проблемой используются различные методы предварительной обработки исходных данных для выделения основополагающих факторов, уменьшения анализируемого пространства признаков и снижения времени на обучение. Методы предварительной обработки многомерных данных предназначены для перевода данных в наиболее информативное для исследуемой задачи представление.

Для решения этой проблемы разработана процедура создания оптимальной структуры искусственной иммунной сети по весовым коэффициентам выделенных информативных признаков с целью эффективного обучения сети за минимально короткое время.

Процедура состоит из следующих шагов:

- нормировка исходных признаков, так как признаки измеряются в разных единицах;
- выделение информативных признаков с помощью метода главных компонент;

- выделение и анализ основополагающих факторов на основе развития варимаксного подхода;
- редукция малоинформативных признаков;
- визуальное представление многомерных данных на дисплее;
- ранжирование информативных признаков в зависимости от весовых коэффициентов;
- построение оптимальной структуры иммунной сети в зависимости от требований конкретного приложения.

Вторая проблема заключается в оценке энергетических погрешностей, которые неизбежно возникают при обработке многомерной совокупности данных иммунными сетями и решении задачи распознавания образов на основе AIS. Проблема особенно актуальна для схожих по структуре пептидов, которые имеют примерно одинаковые параметры и находятся на границах классов. Из-за энергетических ошибок они могут быть ошибочно отнесены не в свой класс, что может привести к аварийной ситуации.

Основной чертой белковых аминокислотных последовательностей, определяющей все их физические свойства, является повышенная стабильность нативной (функциональной) структуры, т. е. существование большой щели между энергией нативной структуры и минимальной энергией неверно свернутых структур. Определение нативной укладки цепи затруднено в силу различных погрешностей, которые возникают из-за неполноты и корреляции данных, ошибок измерения. Для решения данной проблемы разработан эффективный алгоритм, основанный на свойствах гомологичных белков.

Структура интеллектуальной технологии обработки многомерной информации на основе иммунных сетей основана на модульном подходе. Каждый из этапов данной технологии представляет собой отдельный модуль.

Разработанное программное обеспечение «РЕПТЮ» предназначено для реализации поставленных задач с помощью искусственных иммунных сетей. Данный пакет программ реализован на языке программирования DELPHI 7.0.

В состав предложенного программного продукта входят 9 взаимосвязанных подпрограмм, выполняющих определенные задачи: NORM (нормирование данных), INF_PRIZ (выделение информативных признаков), OPT_NET (построение оптимальной структуры иммунной сети), CLAS_EXP (классификация решений), LEARN (обучение иммунной сети с учителем), MIN_ENERG (определение минимума энергии связи между антителами и антигенами, решение задачи распознавания образов), ENERG_ERRORS (определение энергетических

погрешностей при решении задачи распознавания образов), K_PRIDIC (расчет коэффициентов риска прогнозирования), GRAFIC (прогноз и графическое представление данных).

С использованием полученных теоретических результатов разработаны интеллектуальные экспертные системы прогнозирования и анализа динамических процессов сложных нелинейных систем при априорной неопределенности различного типа (стохастической, интервально-заданной) в реальном масштабе времени.

Синтезированные интеллектуальные экспертные системы поддержки принятия решения и управления на основе биологического подхода AIS, работая в реальном масштабе времени, обеспечивают сложным современным нелинейным динамическим объектам высокое качество управления в условиях неопределенности и существенной переменности их параметров.

На основе описанной информационной технологии обработки информации искусственными иммунными сетями разработана интеллектуальная экспертная система дистанционного образования, достоинством которой является способность обрабатывать многомерные данные в реальном масштабе времени, что позволяет адаптировать учебный процесс на конкретного обучаемого и существенно повысить качество получаемого образования.

Представлен подход к разработке интеллектуальной системы дистанционного образования на основе AIS, включающий реализацию информационной, обучающей и контролирующей подсистем. «Информационная подсистема» осуществляет разработку методов и средств хранения информации, БД и БЗ и включает поисковые системы, электронные учебники по различным дисциплинам, электронные библиотеки, справочки, каталоги и т. д. «Обучающая подсистема» осуществляет разработку методов, средств и форм подачи обучающей информации, адаптированной на конкретного пользователя с учетом его индивидуальных характеристик. Составляется график выполнения объема требуемых работ и определяются сроки реализации. «Контролирующая подсистема» предназначена для комплексной оценки знаний обучающегося с целью оперативной корректировки программы и процесса обучения.

Достоинством разработанного подхода является возможность оперативной оценки достоверности прогноза на основе гомологичных белков, что существенно повышает эффективность функционирования интеллектуальных экспертных систем поддержки принятия решений.