

МНОГОКРИТЕРИАЛЬНАЯ ОПТИМИЗАЦИЯ КОРПОРАТИВНЫХ ИНФОРМАЦИОННЫХ СИСТЕМ

Лаздынь С.В., Землянская С.Ю.

Донецкий национальный технический университет,
кафедра автоматизированных систем управления
E-mail: slazd@ukr.net, zsaa@yandex.ru

Abstract

Lazdyn S.V., Zemlianskaya S.Y. Multicriteria optimization of the corporate information systems. It is proposed in the given article to use the comprehensive approach to the optimization of the corporate information systems at the stage of their designing or reconfiguration. The approach proposed consists in the use of the CIS object model and genetic algorithm taking into account several criterias of the optimization

Характеристика проблемы

Современные корпоративные информационные системы (КИС) – это сложные комплексы взаимосвязанных и согласованно функционирующих аппаратных и программных компонентов: компьютеров, коммуникационного оборудования, операционных систем, сетевых приложений [1,2], создаваемые для крупных предприятий и корпораций. КИС, как правило, имеют территориально распределённую архитектуру.

Главным требованием, предъявляемым к корпоративным информационным системам, является выполнение КИС её основной функции – обеспечение пользователям возможности быстрого доступа к необходимым им данным, расположенным зачастую на различных компьютерах [2]. При проектировании и эксплуатации корпоративных информационных систем необходимо обеспечить согласованную и эффективную работу всех их составляющих. Для того, чтобы система работала наиболее эффективным образом, необходимо решить ряд задач, связанных с определением конфигурации (топологии) компьютерной сети, параметров компьютерного и коммуникационного оборудования, с распределением приложений и данных по узлам системы. При этом, с одной стороны, параметры проектируемой системы должны быть подобраны таким образом, чтобы максимально удовлетворять условиям работы в реальном времени, т.е. обеспечивать высокую скорость реакции системы, а с другой стороны, стоимость системы не должна быть слишком высокой, кроме того необходимо обеспечить отказоустойчивость и безопасность данных в системе, что приводит к некоторой избыточности в аппаратуре и данных системы.

Для решения указанных выше задач необходимо разработать подход, позволяющий оптимизировать параметры КИС на этапе проектирования или реконфигурации при наличии нескольких критериев оптимизации.

Краткий анализ проведенных исследований по оптимизации КИС

Традиционные методы многокритериальной оптимизации можно условно разделить на 3 ключевых подхода. Первый связан с идеей ранжирования критериев по важности и дальнейшей последовательной оптимизации каждого критерия по отдельности с назначением допустимой величины изменения значения критерия, полученного на предыдущем шаге. Суть второго подхода состоит в выделении из всех критериев, а затем и оптимизации главного критерия и переводе остальных в ограничения. Наиболее часто используемым на практике является третий подход – скаляризация векторного критерия в один обобщенный критерий.

Основной проблемой, касающейся большинства традиционных методов, является необходимость выполнять алгоритм несколько раз для получения репрезентативной

аппроксимации множества эффективных точек (число итераций равно мощности предполагаемой аппроксимации множества Парето). Это не очень удобно при решении практических задач, однако имеется и еще более существенный недостаток – получаемая аппроксимация множества недоминируемых решений может оказаться нерепрезентативной, т.к. генерируемые эффективные точки будут неравномерно распределены как в пространстве альтернатив, так и в пространстве критериев. Зачастую можно получить набор очень близких друг к другу решений при потенциально бесконечном множестве Парето. Основной причиной такой ситуации является то, что в каждом из них задача многокритериальной оптимизации сводится к одной или нескольким задачам однокритериальной оптимизации. Таким образом, теряется суть решаемой задачи – одновременный учет многих критериев.

В работах [3,4,5] описаны результаты исследований различных методов численной оптимизации и приведены основные свойства рассмотренных методов. Результаты анализа методов оптимизации по нескольким критериям приведены в таблице 1.

Таблица 1. Результаты анализа методов оптимизации

Методы оптимизации	Большое кол-во парам.	Нечеткость	Многокритериальность	Сложная функция
Градиентные	Удовл.	Неуд.	Неуд.	Неуд.
Квазиньютоновские	Удовл.	Удовл.	Неуд.	Неуд.
Прямого поиска	Неуд.	Неуд.	Удовл.	Удовл.
Многоэкстремальные	Неуд.	Неуд.	Удовл.	Неуд.
Нелинейное программирование	Неуд.	Неуд.	Удовл.	Неуд.
Безусловная оптимизация	Неуд.	Неуд.	Удовл.	Удовл.

Из результатов анализа следует, что ни одна группа методов не может быть эффективно применена для решения задачи оптимизации параметров КИС, так как не обеспечивает достаточно высокой вероятности получения удовлетворительного решения.

В отличие от классических подходов к многокритериальной оптимизации, генетические алгоритмы (ГА) принадлежат к разряду многоточечных поисковых методов. Задача оптимизации с их помощью может быть решена даже в случае полимодального характера целевых функций. Более того, они также применимы к задачам с дискретным поисковым пространством. При решении многокритериальных задач генетические алгоритмы способны находить Парето-оптимальное множество за один прогон, благодаря заложенному в них полимодальному поиску.

Поэтому для решения поставленной задачи предлагается использовать генетические алгоритмы (ГА), широко применяющиеся в последнее время для различных задач оптимизации. Генетические алгоритмы - алгоритмы поиска оптимума, основанные на естественном отборе и природной генетике. Они комбинируют моделирование борьбы за выживание наиболее приспособленных особей с моделированием случайных явлений, подобных природным явлениям.

Указанные выше особенности задачи оптимизации КИС делают проблематичным применение традиционных и эвристических методов, использующих некоторые априорно известные свойства для организации поиска оптимального решения. При больших размерах дискретной или смешанной задачи, такой, как оптимизация КИС, методы полного перебора допустимых решений также оказываются неприемлемыми по сложности. Кроме того, традиционные методы вычислений не обладают необходимой гибкостью за пределами ограниченных условий. Генетические алгоритмы (ГА) позволяют решать прикладные задачи в условиях значительной неопределенности информации о сложных техногенных и природных системах. Это объясняется способностью ГА к адаптации и обучению в ходе работы.

Постановка задачі оптимізації КИС

При проектуванні інформаційних систем необхідно забезпечити досягнення наступальних результатів [6]:

- скорочення часу доступу до актуальної інформації;
- збільшення точності визначення вихідних значень;
- зменшення об'єму використовуваної оперативної пам'яті, зайнятої програмами і даними;
- зменшення часу виконання програм, зменшення завантаження каналів між центральною базою даних і сопряженими з нею автоматизованими робочими місцями і т. д.

На основі вищеизложеного можна запропонувати наступальний набір критеріїв оптимальності функціонування КИС:

1. Зменшення часу відклику системи
2. Зменшення черед к ресурсам
3. Зменшення завантаження ресурсів (коефіцієнт завантаження процесорів, каналів передачі даних)
4. Зменшення вартості проекту
5. Збільшення безпеки і надійності функціонування системи.

Таким чином, задача оптимізації КИС зводиться до задачі багатокритеріальної оптимізації і має наступальні особливості:

– кількість параметрів достатньо велика і не може бути однозначно визначено для будь-якої КИС в зв'язі з різноманітністю структури: КИС завжди складається з наступальних компонентів: апаратури, операційної системи, бази даних, додатків. Щоб оцінити стан системи, необхідно враховувати параметри її компонентів: мережі, пам'яті, процесорів, дискових підсистем. Після розрахунку часу відклику можна перейти до оцінки впливу цих або інших очікувань на загальний час відклику системи.

– експертний характер визначення оцінок безпеки і надійності обумовлює нечіткість значень параметрів. Крім того можлива неоднорідність оцінок, т.е. різні групи параметрів можуть оцінюватися різними експертами і відрізнятися шкалою значень.

– задачу можна охарактеризувати як багатокритеріальну, оскільки оптимальність рішення визначається кількома критеріями – якістю, вартістю, часом відклику і завантаженням ресурсів.

– оптимізуєму функція дискретна, недиференційована, має складний характер змін і т.д.

Задача оптимізації архітектури КИС формується наступальним чином.

Для заданих:

- НР – моделі поведінки додатку (частота і локалізація запуску, набір формуваних запитів);
- НВ – моделі архітектури КИС (територіальне розташування і розподіл даних).

Требують визначити:

- технічні параметри серверів (об'єм дискового простору, продуктивність, об'єм ОП, швидкість друку);
- технічні параметри робочих станцій (об'єм дискового простору, об'єм оперативної пам'яті);
- середовище передачі;
- типи операційних систем;
- топологію.

При вибраних параметрах досягається наступальне:

1. $TW \leq T_{гр}$ - час виконання додатку не повинен перевищувати граничний термін,

2. $T_{обслл} \rightarrow \min$ – среднее и предельное время обслуживания должно быть минимальным,
3. $\max\{T_{ож} i\} \rightarrow \min$ – среднее и предельное время ожидания должно быть минимальным,
4. $\max\{LQ i\} \rightarrow \min$ – предельная длина очереди в системе должна быть минимальной
5. $C_{syst} \rightarrow \min$ стоимость системы должна быть минимальной
6. $S_{syst} \rightarrow \max$ безопасность системы должна быть максимальной
7. $R_{syst} \rightarrow \max$ надёжность системы должна быть максимальной

Задача выбора оптимальной архитектуры КИС в данной постановке относится к классу смешанных задач комбинаторной и целочисленной оптимизации. При увеличении числа оптимизируемых параметров (принадлежащих множеству целых чисел или комбинаторных структур), ограничений, критериев оценки качества решений и уровня детализации моделей, задача синтеза архитектур не приобретает никаких новых особенностей, которые могли бы вывести ее из данного класса задач оптимизации.

Предлагаемая модель описания функционирования КИС позволяет определить задачу выбора оптимальной архитектуры КИС как смешанную многопараметрическую и многокритериальную экстремальную задачу с ограничениями:

$$\begin{aligned} f_i(HP, HW, \varphi) &\rightarrow \min/\max, i \in I_{f1}, I_{f2}; \\ g_i(HP, HW, \varphi) &\leq 0, i \in I_{g3}, I_{g4}; \\ g_i(HP, HW, \varphi) &= 0, i \in I_{g5}, I_{g6}; \\ HP &\in HP^*, HW \in HW^*; \end{aligned}$$

где HW - модель архитектуры КИС; HP - модель работы приложений; φ - оператор взаимодействия (задает взаимодействие между компонентами из HP, HW); f_i - критерии оценки качества решения; g_i - ограничения на допустимые решения; I_{f1}, I_{g3}, I_{g5} - множества динамических критериев и ограничений; I_{f2}, I_{g4}, I_{g6} - множества статических критериев и ограничений; (HP^*, HW^*) - пространство решений.

Оценки значений динамических критериев и ограничений из множеств I_{f1}, I_{g3}, I_{g5} для конкретного варианта решения задачи (HP_i, HW_i) могут быть получены лишь после выполнения оператора интерпретации φ ; оценки значений статических критериев и ограничений из множеств I_{f2}, I_{g4}, I_{g6} могут быть вычислены непосредственно по моделям HP и HW на основании экспертных оценок.

Динамические критерии и ограничения при выполнении программы HP на архитектуре HW определяются по результатам взаимодействия структур HP и структур HW в результате имитационного моделирования с использованием объектной модели.

Данная задача независимо от варианта постановки обладает следующими основными особенностями:

1. Пространство возможных решений по типу объектов является трехкомпонентным: структура КИС, логическая среда (системное программное обеспечение), прикладные программы; независимая оптимизация параметров отдельных компонентов невозможна.

2. Значения динамических критериев и ограничений для конкретного варианта решения задачи могут быть определены лишь в результате анализа взаимодействия φ . Взаимодействие реализуется путём эксперимента на объектной модели КИС.

3. f_i, g_i, φ - операторы, заданные правилами/алгоритмами их вычисления, т.е. их аналитическая структура не может быть использована для организации поиска оптимального решения (HP', HW') .

4. Критерии оценки качества решения f_i являются противоречивыми. Улучшение одного критерия, как правило, приводит к обязательному ухудшению другого.

5. Для любого из динамических критериев пространство решений может иметь большое число локальных минимумов и максимумов, точки разрыва и изолированные точки; глобальный оптимум может достигаться в нескольких точках.

Оптимизация КИС с использованием генетического алгоритма и объектной модели КИС

Для оптимизации параметров КИС предлагается использовать подход, основанный на совместном использовании объектной модели КИС и аппарата генетических алгоритмов (ГА), суть которого состоит в следующем.

Параметры отдельных компонентов КИС кодируются в виде хромосомы ГА. Популяция ГА представляет собой набор некоторых точек пространства поиска. На основе проектной информации пользователь строит начальный вариант проектируемой КИС, используя базу данных готовых образцов.

В процессе оптимизации с помощью операторов ГА генерируются хромосомы, то есть схемы распределения данных. Полученные схемы являются исходной информацией для объектной модели. Имитационное моделирование работы КИС на основе объектной модели позволяет получить оценки критериев эффективности КИС. Эти оценки, в свою очередь, являются значениями функции приспособленности ГА (фитнесс-функции) для данного варианта решения. То есть предлагается использовать разработанную объектную модель КИС для вычисления функции приспособленности ГА. Этот процесс повторяется до тех пор, пока не будет достигнут критерий останова генетического алгоритма.

Схема взаимодействия объектной модели КИС с ГА представлена на рис. 1.



Рисунок 1 – Оптимизация КИС с использованием ГА и объектной модели

При разработке ГА для поиска субоптимальных решений поставленной задачи необходимо учитывать специфику этой задачи. Прежде всего, это касается структуры хромосомы, принципов ее кодирования и декодирования.

Объектная модель КИС

Компьютерная информационная система как объект моделирования представляет собой совокупность соединенных друг с другом и взаимодействующих компонентов. Весь комплекс программно-аппаратных средств можно разделить на несколько уровней. На первом уровне системы находится аппаратный слой стандартизованных компьютерных платформ. Второй уровень – это коммуникационное оборудование. К третьему уровню, образующему программную платформу системы, относятся операционные системы (ОС) и приложения пользователей. На четвертом уровне располагаются различные сетевые приложения.

Проведенный структурный анализ показал, что в составе КИС можно выделить следующие типовые компоненты, параметры которых влияют на производительность системы: компьютерное оборудование, сетевое оборудование, программное обеспечение, данные. Для моделирования типовых компонентов КИС выбран объектно-ориентированный метод. Его преимущества по сравнению с традиционным имитационным моделированием – простота и наглядность отображения элементов и функций исследуемого объекта и возможность построения моделей для КИС с произвольной структурой. Для описания функционирования указанных компонентов разработаны соответствующие классы объектов [7].

Все компоненты КИС по своей сущности можно разделить на 2 больших класса:

- компоненты аппаратного обеспечения: серверы, рабочие станции, терминалы, сетевые кабели различных технологий, сетевое оборудование;
- компоненты программного обеспечения: операционные системы, СУБД, приложения, и т.п.

Все компоненты, составляющие конкретную КИС, относятся к этим классам и инкапсулируют свойства конкретного типа серверов, терминалов, операционных систем и сетевого оборудования.

Модель КИС представляет собой совокупность взаимодействующих объектов разработанных классов. Иерархия классов объектной модели КИС представлена на рис. 2.

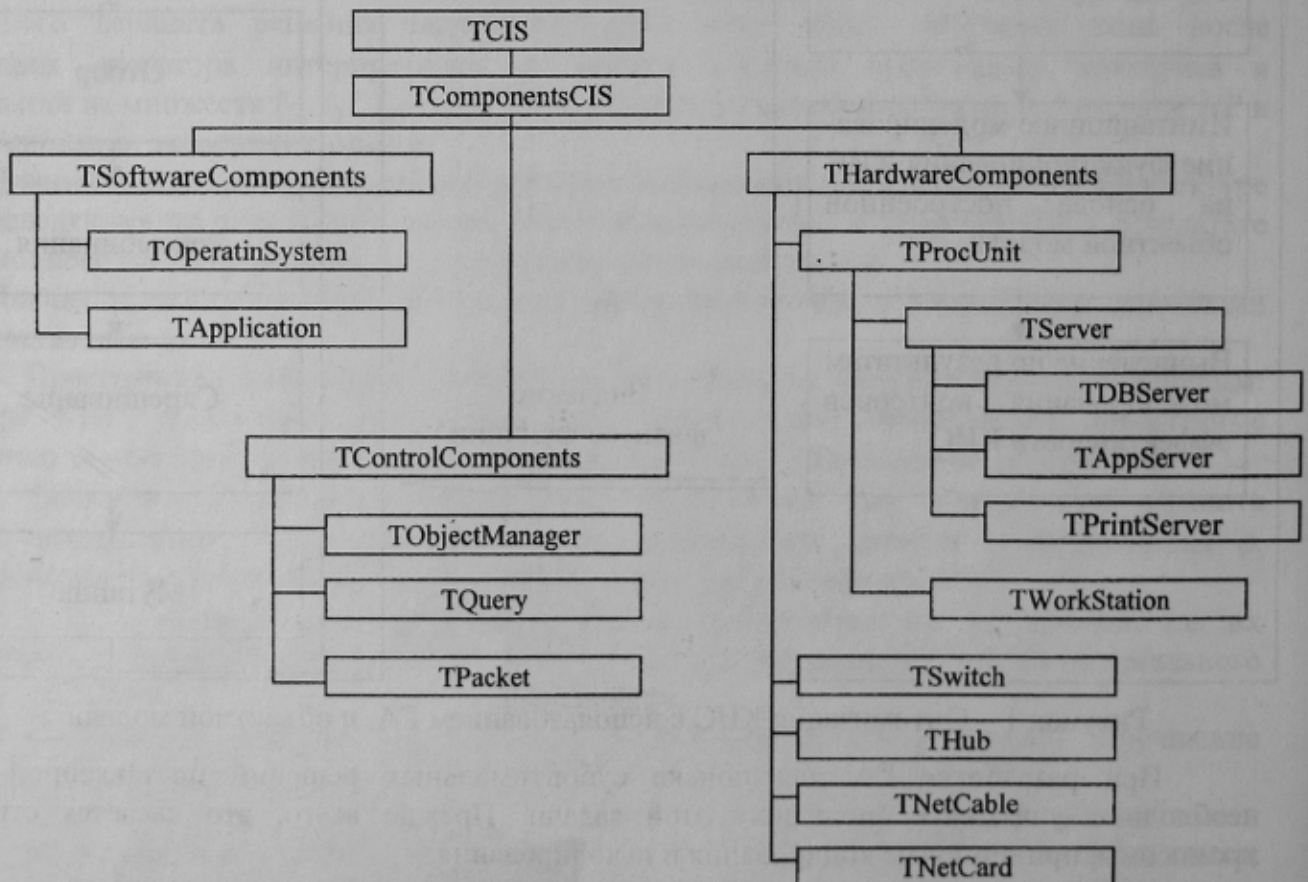


Рисунок 2 – Иерархия классов объектной модели КИС

Для организации взаимодействия между объектами в модель добавлен новый класс – менеджер объектов (TObjectManager), содержащий информацию о структуре системы и процессах, которые в ней происходят. Менеджер объектов содержит сведения обо всех объектах, входящих в рассматриваемую КИС. Кроме того, для управления работой модели необходимо обрабатывать события, происходящие в системе, поэтому менеджер объектов содержит информацию о состоянии объектов в виде списка происходящих событий: запуск приложения; появление запроса; переход пакета между узлами, освобождение узла системы. После обработки события оно удаляется из списка событий.

Менеджер объектов осуществляет координирование функционирования отдельных компонентов модели путем отслеживания событий, происходящих в системе: просматривает очередь событий, выбирает событие с минимальным временем наступления и активизирует соответствующий метод обработки события того объекта, состояние которого связано с этим событием.

В результате взаимодействия объектов через посредство менеджера событий происходит моделирование реальных процессов, протекающих в КИС: запуск приложений, формирование запросов, передача пакетов, обработка запросов, передача ответов и т.д.

По окончании моделирования для каждого объекта системы определяются такие параметры как коэффициент загрузки узла, длина очереди, количество необработанных запросов и др. Моделирование может осуществляться в течение заданного времени, или же может быть принудительно прервано пользователем.

Обоснование параметров генетического алгоритма

Рассмотрим способ кодирования хромосомы, используемый в ГА для решения задачи выбора оптимальной архитектуры КИС.

Схематично выбранный способ кодирования может быть представлен следующим образом. Предлагается представить множество P оптимизируемых параметров системы посредством не одной, а нескольких хромосом, предварительно разбив множество P на подмножества P_1, \dots, P_k , и поставив в соответствие каждому подмножеству P_i отдельную хромосому.

Оценка качества функционирования КИС на макроуровне состоит из нескольких групп: время обслуживания, время ожидания, длина очередей в системе, стоимость, надежность, безопасность. Поэтому очевидным является разбиение множества показателей качества на шесть подмножеств, соответствующих шести вышечисленным группам показателей качества. Таким образом формируется несколько хромосом различной длины:

1. $\bigcup_{i=1}^{K_s} \langle \text{параметры сервера} \rangle_i$,
2. $\left(\bigcup_{i=1}^{K_{rs}} \langle \text{параметры РС} \rangle_i \right)$,
3. $\bigcup_{i=1}^{K_{sp}} \text{физическая среда передачи}$
4. $\bigcup_{i=1}^{K_{tt}} \text{топология}$
5. $\bigcup_{i=1}^{K_{os}} \text{ОС}$

где \bigcup - операция конкатенации (склейки) строк; K_s – количество серверов; K_{rs} – количество рабочих станций; K_{sp} – количество каналов передачи; K_{tt} – количество типов топологий; K_{os} – количество ОС, установленных на рабочих станциях и серверах. Выбранный способ кодирования позволяет исключить появление недопустимых конфигураций в результате выполнения операций ГА.

Целевая функция позволяет оценивать решение комплексно по набору критериев. В качестве таких критериев в рассматриваемой задаче используются критерии 1-7. Введем целевую функцию в следующем виде:

$$F(k_i) = \sum_{i=1}^7 C_i \cdot f_i$$

где C_i – весовые коэффициенты, $C_i \in [-1; 1]$; $f_1 - f_4$ – значения критериев, которые вычисляются на основании взаимодействия компонентов объектной модели КИС по результатам имитационного моделирования; $f_5 - f_7$ – критерии, значения которых вычисляются статически, на основании экспертных данных.

С точки зрения ресурсоемкости многохромосомная схема представления информации также оказывается более предпочтительной. Обработка нескольких коротких хромосом требует меньших вычислительных ресурсов, чем обработка одной длинной хромосомы.

При исследовании и оптимизации КИС многохромосомная модель представления информации также обеспечивает больше возможностей, чем монохромосомная.

Если возникает необходимость произвести частичную оптимизацию – не по всем критериям, а только по некоторым, выборочно, то нет необходимости осуществлять различные сложные операции – маскирование отдельных участков длинной хромосомы, введение компенсирующих коэффициентов при расчете интегрального критерия и т. д. Достаточно только выбрать хромосомы, соответствующие тем группам, которые подлежат оптимизации.

При многохромосомном представлении значительно проще получить необходимую аналитическую информацию об оценке качества, а именно: вклад каждой группы параметров КИС в интегральный критерий; каким образом, количественно и качественно, влияет тот или иной параметр на каждую группу критериев.

При монохромосомном представлении информации вклад каждой группы параметров в интегральную оценку качества и влияние различных факторов на каждую группу показателей качества в отдельности и на интегральную оценку в целом сложно оценить, поскольку влияние этих факторов сглаживается при формировании интегральной оценки, отражающей общую картину, и для определения этого влияния на группы необходимо осуществлять дополнительные исследования. Кроме того, значения критериев качества функционирования КИС по каждой группе могут иметь неоднородный характер. В этом случае многохромосомное представление более устойчиво к влиянию неоднородностей, чем монохромосомное, поскольку обработка одной длинной хромосомы с неоднородным характером генов более сложна с точки зрения программирования и менее эффективна с точки зрения использования вычислительных ресурсов, чем обработка нескольких более коротких и однородных хромосом.

Таким образом, многохромосомная модель представления информации имеет следующие преимущества перед традиционным монохромосомным представлением:

- обработка информации требует меньших вычислительных затрат;
- упрощается реализация различных видов частной оптимизации – по одной или нескольким группам параметров;
- упрощается определение вклада каждой группы параметров КИС в процессе формирования интегральной целевой функции и влияния различных факторов на каждую группу критериев в отдельности;
- модель устойчива к возможной неоднородности значений показателей качества, в частности, при различных способах кодирования, диапазонах значений;
- снижается алгоритмическая сложность программы, реализующей процесс оптимизации.

Существенное влияние на работу ГА оказывает способ отбора особей в новое поколение [8].

Отбор производится среди особей предыдущего поколения и особей, образовавшихся в результате применения генетического отбора (ГО). Среди существующих подходов можно выделить различные варианты пропорционального отбора и элитного отбора. При пропорциональном отборе вероятность перехода особи в следующее поколение прямо пропорциональна ее приспособленности. Элитный метод заключается в формировании новой популяции из наиболее приспособленных особей, выбираемых из множества родительских особей и особей, полученных в результате действия ГО.

Недостатком такого подхода является риск преждевременной сходимости к локальному минимуму [8]. Можно использовать комбинацию этих методов, идея которой заключается в следующем.

Выбор метода отбора в каждом поколении происходит случайно с определенной вероятностью. Периодическое добавление элитного отбора будет приводить к более быстрому схождению алгоритма к найденным экстремумам. Использование пропорционального метода отбора ведет к долгому поиску решения, и решения находятся достаточно неэффективные. Применение элитного метода отбора позволяет более быстро находить решения, но они сильно отклоняются от оптимальных. При использовании смешанного метода решение ищется дольше, чем при элитном методе, но в результате находятся более оптимальные решения.

Критерий останова зададим следующим образом. Алгоритм останавливается, если за заданное число итераций, максимальное значение целевой функции улучшено не было.

Заклучение

КИС является сложной многоуровневой системой. Для оптимизации параметров КИС на этапе проектирования или реконфигурации при наличии нескольких критериев оптимизации предложено использовать комплексный подход, заключающийся в совместном использовании объектной модели КИС и генетического алгоритма. При этом генетический алгоритм используется для генерации новых вариантов КИС, а объектная модель – для определения значения фитнес-функции. Фитнес-функция получена путём свёртки критериев оптимальности КИС в один суперкритерий с использованием коэффициентов значимости. Предлагаемый подход позволяет получить субоптимальные параметры КИС для заданного распределения данных по узлам системы.

Литература

1. Марков А.А. Моделирование информационно-вычислительных процессов: Уч. пособие для вузов.- М: Изд-во МГТУ им. Н.Э. Баумана, 1999.- 360 с.
2. Олифер Н.А., Олифер В.Г. Средства анализа и оптимизации локальных сетей. Интернет-ресурс <http://www.citforum.ru/nets/optimize>
3. Батищев Д.И. Методы оптимального проектирования. М.: Радио и связь, 1984
4. Гилл Ф., Мюррей У., Райт М. Практическая оптимизация. М.: Мир, 1985
5. Исаев С. А. Решение многокритериальных задач. Интернет-ресурс <http://bspu.ab.ru/Docs/~saisa/ga/ideal.html>.
6. Макаров И.М., Соколов В.Б и др. "Теория выбора и принятия решений"
7. Лаздынь С.В., Землянская С.Ю. Разработка динамической модели корпоративной информационной системы // Наукові праці ДонНТУ. Серія "Обчислювальна техніка та автоматизація" - Випуск 74. - Донецьк: ТОВ Лебідь, 2004. - С. 152 - 159.
8. Батищев Д.И. Генетические алгоритмы решения экстремальных задач. Воронеж, 1995.