

3. Карпук, А. А. Математические модели и алгоритмы вычисления чистых позиций участников в клиринговых системах расчетов / А. А. Карпук, Е. С. Шейнкман // Современные информационные компьютерные технологии: сб. науч. статей: в 2 ч. / ГрГУ им. Я. Купалы; редкол.: Е. А. Ровба, А. М. Кадан (отв. редактор) [и др.]. – Гродно : ГрГУ, 2008. – Ч. 2. – С. 62–65.

4. Матальцкий, М. А. Применение НМ-сети с двумя типами заявок при моделировании межбанковских платежей / М. А. Матальцкий, А. А. Карпук, Е. В. Колузаева // Современные математические методы анализа и оптимизации информационно-телекоммуникационных сетей. материалы междунар. конф., 26–29 янв. 2009 г. – Минск, 2009. – С. 151–154.

Карпук Анатолий Алексеевич, начальник управления развития и сопровождения клиринговой и взаимодействующих систем Расчетного центра Национального банка Республики Беларусь, кандидат технических наук, доцент, Anatoly_Karpuk@bisc.by

Кищенко Марина Георгиевна, главный специалист управления качества и системного анализа Расчетного центра Национального банка Республики Беларусь, Marina_Kischenko@bisc.by

УДК 004

А. Т. Кобзарь

ПЛАНИРОВАНИЕ ЗАГРУЗКИ СИСТЕМЫ РЕЗЕРВНОГО КОПИРОВАНИЯ НА ОСНОВЕ МЕТОДА ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ

Разработана методика оценки интенсивности потоков файлов для резервного копирования, позволяющая оценивать затраты на процедуру резервного копирования и планировать загрузку систем резервного копирования

Введение

Для увеличения производительности информационных систем и увеличения скорости доступа к информации данные стараются расположить локально на том узле сети, где они будут обрабатываться. Размещение данных как можно ближе к пользователям не только уменьшает время отклика и высвобождает пропускную способность, но и сокращает трафик через глобальные сети. Кроме того, распределение данных между несколькими системами повышает скорость выполнения запросов [1]. Необходимо помнить, что в распределенных системах важным показателем является показатель пропускной способности, отражающий количество данных, которое можно скопировать в данный период времени.

На современных предприятиях объем данных, обрабатываемых информационными системами, растет в геометрической прогрессии. При этом ценность информации в современном мире становится все ощутимее. Наличие больших объемов данных позволяет повысить эффективность работы и сохранить конкурентоспособность компании. Но возникает необходимость обеспечения сохранности столь ценной информации от случайного или намеренного повреждения. Для этих целей используют системы резервного копирования.

Резервное копирование – это процесс создания резервных копий документов, данных или программ для оперативного их восстановления или для длительного хранения на случай их разрушения или порчи [2].

Система резервного копирования является служебной подсистемой системы хранения данных. Система резервного копирования предназначена для создания резервных копий и восстановления данных. Она позволяет защитить данные от разрушения не только в случае сбоя или выхода из строя аппаратуры, но и в результате ошибок программных средств и пользователей [3].

Резервное копирование необходимо для защиты:

- от физической порчи носителей информации;
- уничтожения информации компьютерными вирусами;
- случайного или преднамеренного удаления информации самими пользователями;
- удаления информации из-за сбоя в системе и т. п.

Процесс резервного копирования требует времени. И при некоторых типах резервного копирования процедура снятия резервной копии может привести к простоем системы и отсутствию у пользователей возможности работать в информационной системе. Резервное копирование большого количества данных за небольшой период времени требует от системы высокой производительности. Время загрузки системы резервного копирования зависит, в том числе и от выбора стратегии резервного копирования.

Перечень предлагаемого на рынке программного обеспечения для резервного копирования ежегодно увеличивается. Программное обеспечение способно осуществлять резервное копирование данных и зачастую включает в себя или совместимо с дополнительными программами, отвечающими за управления резервным копированием. Несмотря на это задачи организации резервного копирования в современных информационных системах постоянно усложняются. Это обусловлено, с одной стороны, сложностью выбора нужного программного продукта, а с другой – повышением требований к составу и качеству исполнения функций резервного копирования. Многие известные программные продукты имеют удобный пользовательский интерфейс и разнообразный набор сервисов, но все равно выбор стратегии резервного копирования перекладывается на плечи администратора. От выбора стратегии резервного копирования зависит время, которое необходимо потратить на осуществление резервного копирования и для полного восстановления потерянных данных, т. е. время загрузки системы резервного копирования. Скорость и качество процесса восстановления данных можно улучшить за счет уменьшения объема восстанавливаемой информации. Описываемый в работе метод позволяет не только осуществлять отбор файлов для резервного копирования, тем самым уменьшая объем резервируемой информации и увеличивая производительность системы резервного копирования, но и осуществлять планирование загрузки системы резервного копирования.

Модель системы резервного копирования на основе генетических алгоритмов

Предлагается подход к выбору стратегии резервного копирования, основанный на минимизации затрат на резервное копирование и в том числе минимизации загрузки системы резервного копирования путем отбора из массива файлов, хранящихся в информационной системе, только тех файлов, к которым целесообразно применить процедуру резервного копирования.

При проектировании подсистем резервного копирования необходимо выбрать стратегию резервного копирования, наилучшим образом соответствующую нуждам системы. Стратегия резервного копирования определяет объем необходимой памяти для хранения данных, частоту резервного копирования, скорость восстановления данных после сбоя, время простоя системы при снятии резервной копии или восстановлении данных и т. п., а соответственно стоимость всего процесса создания резервной копии.

Для минимизации затрат на резервное копирование рассматривается задача отбора из массива файлов, хранящихся в информационной системе, только тех файлов, к которым целесообразно применить процедуру резервного копирования. Существует большое количество факторов, которые следует учитывать при отборе файлов для резервного копирования, например, ценность данных, частота изменений и т. п. Для решения задачи отбора с учетом многофакторности модели предлагается использовать процедуры поиска решения, основанные на механизме естественного отбора и наследования – генетическом алгоритме [4].

Предполагается, что каждый файл информационной системы – это особь, а конечное множество особей или конечный набор файлов – это популяция. Каждая особь характеризуется набором определенных генов (свойств), соответствующих определенным параметрам файла:

- стоимость хранимой в файле информации;
- стоимость потерь, возникших в случае утраты информации, хранимой в данном файле;
- размер файла;
- частота вносимых изменений;
- дата снятия последней резервной копии;
- дата внесения последнего изменения и т. п.

Для каждого файла упорядоченную последовательность отобранных параметров $X = (x_1, x_2, x_3, \dots, x_n)$ переведем в закодированную форму, составив для каждого файла последовательность двоичных цифр (единиц и нулей). Последовательность (вектор) из нулей и единиц называется хромосомой, а каждая позиция в этой последовательности называется геномом.

Набор и содержание файлов в системе постоянно меняется, так же как и популяция особей в живой природе, которые скрещиваются и изменяются (мутируют). Процессы изменений в файловой системе можно рассматривать как эволюционные процессы в поколениях хромосом.

Механизм селекции в нашем случае будет заключаться в выборе наиболее приспособленных хромосом для следующей популяции. На каждой итерации генетического алгоритма приспособленность каждой особи данной популяции оценивается при помощи функции приспособленности.

Репродукция следующей популяции заключается в создании новых хромосом в результате рекомбинации генов родительских хромосом. Рекомбинация – это процесс, в результате которого возникают новые комбинации генов. Для этого используются две операции:

Скрещивание (кроссовер) – операция, при которой создаются две совершенно новые хромосомы потомков путем комбинирования генетического материала пары родителей.

Мутация – случайное изменение одной или нескольких позиций в хромосоме. Примером операции мутации файла может служить изменение размера файла.

При скрещивании и мутации свойства файлов (особей) наследуются в соответствии с механизмом наследования признаков.

Схема функционирования генетического алгоритма в классическом варианте выглядит следующим образом [3]:

- случайным образом формируется начальная популяция, состоящая из N особей (файлов);
- вычисляется приспособленность каждой особи $F_j(x)$, $j=1 \dots N$ и популяции в целом. Значение этой функции определяет, насколько хорошо подходит особь, описанная данной хромосомой, для решения задачи;
- отбирается особь из популяции;
- с определенной вероятностью отбирается вторая особь из популяции и применяется оператор скрещивания (кроссовера);
- с определенной вероятностью выполняется оператор мутации;
- помещаем полученную хромосому в новую популяцию.

Если $F(x_1, x_2, x_3, \dots, x_n) \geq H$, где H – экспертная оценка порога селекции, то особь отбирается для резервного копирования.

Планирование загрузки системы резервного копирования

Рассмотрим файловую систему, состоящую из множества файлов N . Каждый файл файловой системы имеет n различных свойств. Ценность каждого свойства определяется его стоимостью a_i , где $i=(1, n)$. Файлы из множества файлов N будем выбирать случайным образом. Пусть x_{ij} – одна из характеристик j -го файла, где $j=(1, N)$. Функцию приспособленности зададим следующим выражением:

$$F_j(x) = \sum_{i=1}^n x_{ij} a_i. \quad (1)$$

Механизм наследования представим в следующем виде:

$$F(x) = \sum_{i=1}^n x_i a_i \alpha_i, \quad (2)$$

где $\alpha_i = 0$ или 1 в зависимости от механизма наследования.

При скрещивании файлов количество файлов в файловой системе останется неизменным. При скрещивании двух файлов появятся 2 новых файла с одинаковыми характеристиками, которые они унаследуют от родительских файлов. Пусть один из родительских файлов имеет набор характеристик $x = (x_1, x_2, x_3, \dots, x_n)$ и вектором наследования α_1 , а второй – $y = (y_1, y_2, y_3, \dots, y_n)$ и вектором наследования α_2 .

Тогда вектор наследуемых характеристик для файлов, получившихся в результате скрещивания двух файлов $z = (z_1, z_2, z_3, \dots, z_n)$ будет рассчитываться по следующей формуле:

$$z_i = x_i \alpha_{1i} + y_i \alpha_{2i}. \quad (3)$$

Процедуру скрещивания файлов в общем виде можно записать следующей формулой:

$$z_i = f_i(x_i). \quad (4)$$

При мутации у файла изменяется любая характеристика x_{ij} . Тогда процедуру мутации файла можно записать формулой:

$$\varphi k(x_k) = (x_{k_1}, x_{k_2}, x_{k_3}, \dots, x_{k_k}), \quad (5)$$

$$z_i = \varphi k(x_k). \quad (6)$$

Иными словами, значение функции приспособленности для файла, полученного в результате скрещивания или мутации, можно представить следующей формулой:

$$F_j(x) = \sum_{i=1}^n z_i a_i. \quad (7)$$

Рассмотрена эволюция файловой системы за период времени T . Для каждого файла в файловой системе в конце периода эволюции можно определить с заданной вероятностью, какие характеристики он будет иметь, оценить стоимость данного файла и в соответствии с этим сделать вывод о необходимости резервного копирования данного файла.

В результате имеется возможность определять суммарную интенсивность потока файлов, подлежащих резервному копированию, т. е. загрузку системы резервного копирования. Интенсивность входящего потока данных оценим как сумму значений функций приспособленности всех файлов, отобранных для резервного копирования. Далее через отношение всех входящих файлов к числу отобранных для резервного копирования можно посчитать нагрузку на систему резервного копирования.

Литература

1. Сайт издательства «Открытые системы». Распределенное тиражирование данных [Электронный ресурс] / Джерри Голик. – Россия, 2000. – Режим доступа: <http://www.osp.ru/lap/2000/04/131050/> – Дата доступа: 15.03.2009.
2. Воройский, Ф. С. Информатика. Энциклопедический систематизированный словарь-справочник. (Введение в современные информационные и телекоммуникационные технологии в терминах и фактах) / Ф. С. Воройский. – М., 2007.
3. IBS Системы резервного копирования [Электронный ресурс] / Алексей Лобанов. – Россия, 2000. – Режим доступа: <http://si.ibs.ru/content/si/70/708-article.asp> – Дата доступа: 16.03.2009.
4. Генетические алгоритмы – математический аппарат [Электронный ресурс] / Алексей Стариков. – Россия, 2000. – Режим доступа: http://www.basegroup.ru/library/optimization/ga_math/ – Дата доступа: 14.03.2009.
5. Рутковская Д., Нейронные сети, генетические алгоритмы и нечеткие системы / Д. Рутковская, М. Пилиньский, Л. Рутковский; под ред. А. С. Попова. – М. : Горячая линия – Телеком, 2006. – 452 с.

Кобзарь Анастасия Тимуровна, аспирантка кафедры теории вероятностей и математической статистики факультета физико-математических и естественных наук Российского университета дружбы народов, jasya@mail.ru

УДК 519.872

Е. В. Колузаева, М. А. Маталыцкий., С. Э. Статкевич

ПРИМЕНЕНИЕ НМ-СЕТЕЙ С ОГРАНИЧЕННЫМ ВРЕМЕНЕМ ОЖИДАНИЯ ЗАЯВОК В ОЧЕРЕДЯХ ПРИ МОДЕЛИРОВАНИИ МЕЖБАНКОВСКИХ ПЛАТЕЖЕЙ

Рассматривается применение НМ (Howard-Matalytski)-сетей массового обслуживания в качестве стохастической модели системы межбанковских расчетов в случае, когда электронные платежные документы могут быть досрочно отзываны банком-плательщиком из системы.

Введение

Сети массового обслуживания (МО) используются в качестве математических моделей различных объектов в компьютерной технике, экономике, производстве, страховании, медицине. С методами их исследования и применениями можно познакомиться, например, в [1], где, в частности, показано, что собой представляют заявки и системы сети (СМО) в различных ситуациях. Понятие НМ (Howard-Matalytski)-сетей впервые было введено в рассмотрение в работах [2–5]. В таких сетях заявка при переходе из одной СМО в другую приносит ей некоторый доход, а доход первой СМО уменьшается на эту же величину. Подробная библиография работ по НМ-сетям приведена в [6, 7].