ФАЗОВЫЙ ПЕРЕХОД ОТКАЗОУСТОЙЧИВОСТИ В РАСТУЩИХ СЕТЯХ ПРОИЗВОЛЬНОЙ НАЧАЛЬНОЙ КОНФИГУРАЦИИ

Коганов А. В., Сазонов А. Н.

Исследуется отказоустойчивость конечной вычислительной сети с произвольным графом, элементы которой имеют вероятность отказа и вероятность восстановления после отказа. Работа сети происходит по трехэтапным тактам (разрушение-восстановление-функционирование). Предлагается алгоритм наращивания сети в начале каждого такта ее работы. При этом граф увеличенной конфигурации сети формируется путем добавления новых экземпляров исходной сети и соединения их определенным образом с элементами старой конфигурации сети. Доказывается, что при достаточно быстром росте сеть имеет положительную вероятность неограниченной безотказной работы. Параметрическая оценка критической скорости роста сети имеет логарифмический порядок по числу тактов

1. Введение. В нескольких предыдущих работах [6-9] авторы исследовали отказоустойчивость вычислительных сетей на графе в модели отказа, которая будет описана ниже, для случаев стационарного или растущего графа сети. При этом было обнаружено свойство появления положительной вероятности неограниченно долгой безотказной работы сети при достаточно быстром наращивании. Этот результат был строго доказан [7] в частном случае корпоративных сетей (полный граф связей элементов) и подтвержден вычислительно [8] для случая планетарного графа (несколько звездообразных графов с центрами, связанными полным графом). Сходный эффект описан в [10]. Доказательство этого факта в общем случае оказалось возможным для специально выбранного алгоритма наращивания графа произвольной исходной сети. Доказательство опирается на результат, полученный в [7] для полного графа. При этом в некотором смысле сохраняется структура исходной сети, что важно для практического применения метода, где структура может диктоваться условиями задачи, решаемой на вычислительной сети.

2. Вычислительные среды с разрушением и восстановлением. Обобщенная модель вычислительной среды (ВС) рассматривается в виде направленного графа с помеченными вершинами, где метки (цвета) вершин соответствуют типам узлов, а дуги – односторонним каналам передачи данных.

Каждая из задач, выполняющихся на ВС, может занимать один или несколько узлов. Следовательно, можно рассматривать модель задачи (МЗ), основанную на модели вычислительной сети (МВС). Соответственно, МЗ также будет направленным графом с мечеными вершинами. Для возможности выполнения задачи на данной ВС необходимо, чтобы МЗ была изоморфна какому-либо подграфу МВС.

В случае выхода из строя некоторых компонентов ВС, оставшаяся работоспособной часть системы называется разрушенной ВС (РВС). Вследствие того, что задача может использовать не все узлы исходной ВС, а также из-за возможности дублирования элементов в некоторых случаях исходная задача сможет выполняться и на РВС.

Для выполнения задачи на ВС необходимо ассоциировать каждую вершину МЗ с вершиной МВС того же типа так, чтобы при этом в полученном образе задачи на графе МВС существовали все дуги, соответствующие дугам на графе МЗ. В данной работе рассматривается случай, когда несколько однотипных вершин графа МЗ могут быть ассоциированы с одной и той же вершиной МВС того же типа. При этом предполагается, что локальные пересылки данных с одного узла ВС на самого себя всегда возможны. Под разложимостью графа МЗ на граф МВС понимается возможность указанного отображения вершин этих графов.

Для исследования отказоустойчивости МЗ может быть выбрана так, чтобы ее граф совпадал с графом исходной ВС. Такая МЗ называется полной (ПМЗ). В этом случае работа задачи на РВС (при условии, что граф РВС не совпадает с графом ВС) будет возможна только если в ВС имело место дублирование узлов и существуют все необходимые задаче дуги графа РВС. Если на РВС выполнима модель полной задачи, то и модель любой задачи, разрешимой на исходной МВС, разрешима на этой РВС.

Модель разрушения и восстановления МВС на каждом шаге следующая. Такт работы ВС состоит из трех последовательных подтактов. В первом подтакте на множестве исправных элементов реализуется испытание Бернулли с вероятностью разрушения p. Вероятности разрушения элементов в совокупности независимы. Во втором подтакте на множестве всех разрушенных элементов (включая оставшиеся с предыдущих тактов) реализуется испытание Бернулли с вероятностью восстановления q. Вероятности восстановления элементов в совокупности независимы. В третьем подтакте проверяется разрешимость задачи на полученной РВС. Наработкой на отказ в реализации считается номер первого такта, на котором задача неразрешима. Средней наработкой T (или наработкой сети) называется математическое ожидание наработки на отказ для полной задачи.

В общем случае вычисление или оценка наработки ВС требует длительного численного эксперимента. Примеры такого имитационного исследования даны в [8-9]. Для простейших графов иногда удается провести аналитический расчет [5-7].

3. Растущие вычислительные сети. Изменяемой вычислительной сетью (ИВС) назовем сеть, у которой на каждом такте изменяется граф связи элементов. При этом новый граф включает предыдущий, т.е. сеть растет. Это соответствует введению *нулевого подтакта*, на котором вводится несколько новых вершин и дуг графа сети. Каждой новой вершине сопоставлен некоторый тип элемента, который первоначально находится в рабочем состоянии. Однако уже на первом подтакте он может разрушиться. Соответственно, текущее состояние элементов вершин графа (рабочее/разрушен) на каждом такте после второго подтакта (частичное восстановление) назовем разрушенной ИВС (РИВС).

Случай ИВС для исходного полного графа (корпоративные сети) с последующим наращиванием графа как полного рассмотрен в [7]. В дальнейшем нам потребуются из этой статьи две теоремы. Номера теорем в исходной работе совпадают с нижеприведенными. Рассматривается один режим (более эффективный) из двух, рассмотренных в [7]. Формулировки адаптированы к тексту данной статьи.

Теорема 3.1. Обозначим $m_i(t)$ — это число элементов типа i на такте номер t. Пусть $p < 1-q$. Тогда, если для всех типов элементов выполнено условие

$$\sum_{t=1}^{\infty}(1-q)^{m_i(t)} < \infty, \tag{3.1}$$

то имеется положительная вероятность H безотказной работы корпоративной сети неограниченное время.

Если хотя бы для одного типа элементов выполнено условие

$$\sum_{t=1}^{\infty}(p(1-q))^{m_i(t)} = \infty \quad , \tag{3.2}$$

то с вероятностью 1 отказ сети наступит за конечное число тактов.□

Теорема 3.2. Если принять параметрический закон роста

$$m_i(t) \sim \xi \cdot \ln(t), \tag{3.3}$$

то критическая точка параметра $\xi = \xi_k$ заключена в интервале

$$\ln^{-1}(p(1-q))^{-1} = \xi_0 \le \xi_k \le \xi_c = \ln^{-1}(1-q)^{-1}, \tag{3.4}$$

Фазовый переход отказоустойчивости корпоративной сети по скорости роста имеет вид:

$$\begin{aligned} &\xi > \xi_c \Rightarrow H > 0; \\ &\xi \le \xi_0 \Rightarrow H = 0; \end{aligned} \tag{3.5}$$

□

4. Алгоритмы роста сети. Изменяемые ВС моделируются алгоритмом нулевого подтакта. Рассмотрим два алгоритма, которые сохраняют РВС предыдущего такта в качестве подграфа новой сети. Оба алгоритма применимы к любой начальной ВС и не вводят связей между элементами двух разных типов, если их не было в исходной сети. Последнее требование с прикладной точки зрения соответствует учету технологических ограничений, которые действуют для данной ВС. При этом полная задача соответствует исходной сети.

4.1. Простое дублирование сети. На нулевом подтакте добавляется $n(t)$ новых экземпляров (копий) исходной сети. Новый граф ИВС состоит из $m(t) = m(t-1) + n(t)$ несвязанных компонент, каждая из которых изоморфна как граф с меченными вершинами исходной ВС; $m(1) = 1$.

Для такой сети отказа нет, если и только если хотя бы одна компонента не дает отказа. В теории надежности это соответствует горячему резерву. Вероятность отказа $H(t)$ для ИВС рекуррентно оценивается через вероятность $h(t)$ отказа для исходной ВС:

$$\begin{aligned} &H(1) = h(1)^{n(1)+1}; \\ &H(t) \le (1 - H(t-1))h(1)^{n(t)+1} < h(1)^{n(t)+1} \end{aligned} \tag{4.1}$$

4.2. Связанное дублирование сети. На нулевом подтакте добавляется $n(t)$ новых экземпляров (копий) исходной сети. Потом на элементах всех экземпляров, которые соответствуют одному элементу исходного графа, строится полный граф. Потом на объединении вершин всех экземпляров исходной ВС строятся дуги, которые соединяют те пары элементов (с учетом ориентации), у которых соединены дугой прообразы в исходном графе. Новый граф соответствует дублированию каждой вершины исходного графа с кратностью $m(t)$.

Такая сеть соответствует снижению вероятности отказа элемента исходной сети:

$$p' \le p^{n(t)}. \tag{4.2}$$

При этом вероятность восстановления q' после отказа дублирующей совокупности элементов оценивается как

$$1-(1-q)^{m(t)} \le q'. \tag{4.3}$$

4.3. Селективное дублирование элементов. Строится полный граф на элементах простого дублирования сети, а потом убираются все дуги между парами элементов тех типов, которые не имеют связи в исходном графе сети. Алгоритмы 4.1 и 4.2 дают подграфы (иногда собственные) этого графа.

Теорема 4.1. Если исходная ВС имеет дугу, соединяющую элементы разных типов, то наработка при простом дублировании ниже, чем при связном дублировании той же кратности $m(t)$.

Доказательство. Представим случайный процесс отказов и восстановлений элементов сетей двух указанных видов как параметрические функции на пространстве случайных событий с вероятностной мерой $(W \mid s)$ с параметром времени t — число тактов работы всех сетей. Предположим синхронную одновременную работу двух сетей. Обозначим эти функции соответственно для простого дублирования и связного дублирования

$$Fs(t,w), Fc(t,w); \quad w \in W, \quad t = 0,1,\ldots \tag{4.4}$$

Значением каждой такой функции будет конфигурация отказов на элементах соответствующей сети на указанном такте. Для наших целей удобно определить функции, соответствующие разным сетям как коррелированные, но без корреляции отказов элементов для каждой сети. Соответственно, сети обозначим *A*, *B*. Исходную сеть обозначим *C* (дублирование состоит из экземпляров *C** этой сети).

Назовем элементы из одной сети родственными, а из разных сетей — адекватными, если они соответствуют одному элементу исходной сети. В сетях *A* и *B* родственные элементы образуют наборы из $m(t)$ элементов. У каждой группы есть адекватная в другой сети. Будем считать, что элементы каждой сети пронумерованы двойными числами. Порядок нумерации определен нумерацией элементов на *C* (первая компонента номера) и нумерацией вводимых новых экземпляров *C** (вторая компонента). Поэтому адекватные элементы сетей *A* и *B* образуют одинаковые номера, и можно говорить об *адекватности по номеру* для элементов с одинаковыми номерами.

Функцию $Fs(t,w)$ зададим произвольно, с выполнением условий на значения вероятностей и независимости событий отказов и восстановлений. Функцию $Fc(t,w)$ определим для каждого значения аргументов как конфигурацию отказов адекватных по номерам отказам из $Fs(t,w)$. Очевидно, что при этом выполнены все условия на вероятности. Для каждой дуги из *C* отказ ее образа в дублирующей сети *A* наступает, когда на каждом экземпляре разрушена хотя бы один из ее концов (событие ***M***), а в сети *B* — когда на всех экземплярах разрушен один и тот же ее конец (событие ***K***); разрушения второго конца при этом произвольны, и порядок концов не важен. Очевидно, что событие ***K*** влечет событие ***M***, и потому имеет не большую вероятность. При не менее чем двукратном дублировании событие ***K*** как подмножество $(W \mid s)$ строго вложено в событие ***M*** и имеет строго меньшую вероятность. Поэтому при наличии в сети C связанных дугой элементов разных типов вероятность отказа для функции $Fs(t,w)$ строго выше, чем для функции $Fc(t,w)$, поскольку для полной задачи любая такая связь должна сохраниться хотя бы для одной пары вершин этих типов. □

Теорема 4.2. Отказ сети селективного дублирования наступает при тех и только тех конфигурациях отказов вершин графа, при которых наступает отказ соответствующей корпоративной системы. (Доказательство очевидно, поскольку изоморфное вложение полной задачи не может использовать отброшенных дуг полного графа: дуги между элементами этих типов в исходном графе отсутствуют.)

Следствие. Вероятность отказа при селективном дублировании на любом такте не выше, чем на корпоративной сети с тем же составом элементов.

5. Фазовый переход наработки по скорости роста в ИВС. При реализации ИВС любым из методов дублирования для произвольной начальной ВС возникает эффект, аналогичный теоремам 3.1, 3.2 для корпоративных ИВС. Далее такие ИВС будем называть дублирующими.

Замечание 5.1. Рассмотрим корпоративную сеть с одним типом элементов. Отказ такой сети соответствует разрушению всех элементов, поскольку полная задача в этом случае может быть отображена в одну вершину полного графа. По той же причине отказ произвольной ВС с одним типом элементов соответствует отказу всех элементов.

Определение 5.1. Макросетью для дублирующей ИВС назовем сеть, элементами которой являются экземпляры исходной ВС, составляющие ИВС, а граф связей элементов соответствует дугам между экземплярами сети.

Для простого дублирования макросеть состоит из $m(t)$ вершин без дуг.

Для связного дублирования макросеть будет полным графом на $m(t)$ вершинах.

Элементы макросети (макроэлементы) дублирующей ИВС все имеют один тип: это экземпляры исходной сети. Но у макроэлементов кроме пары состояний «исправен/отказ» имеются промежуточные состояния частичного разрушения РВС. На этих состояниях имеется вероятностная матрица переходов Марковского процесса. В нашей постановке все вероятности в ней положительны и матрица конечная. Поэтому можно определить минимальную P' и максимальную P'' вероятность перехода в «отказ» по всем предыдущим возможным состояниям. Аналогично, имеются минимальная Q' и максимальная Q'' вероятности перехода в одно из работоспособных состояний из состояния «отказ». Тогда по замечанию 5.1. макросеть имеет наработку не меньшую чем корпоративная ИВС с одним типом элементов и вероятностями отказа и восстановления P',P'' и Q',Q''. Число элементов при этом равно $m(t)$. При достаточно малом исходном значении p выполняется $P''<1-Q''$. Поэтому применимы теоремы 3.1 и 3.2.

Замечание 5.2. Для не простого дублирования оценка наработки, полученная по замечанию 5.1, заведомо занижена, так как полная задача строится относительно исходной ВС, а не относительно макросети. Поэтому возможно сохранение работоспособности ИВС при формальном отказе всех экземпляров исходной ВС. Для получения верхней оценки наработки требуется оценить сверху число возможных вложений исходной сети в расширенную сеть. Обозначим это число $S(t)$. Тогда для простого дублирования $S(t)=m(t)$. Для связного дублирования каждая связь продублирована не более чем $m(m+1)/2$ раз, и, следовательно, $S(t)\le m(t)(m(t)+1)/2$. Для селективного дублирования существующая в исходной сети связь между парой типов элементов i, j дублируется ровно $m^2(t)n_i n_j$ раз, где n_i — это число элементов указанного типа в исходной сети, и поэтому $S(t)\le (m(t)n_{\max})^2$. Для всех методов дублирования $S(t)\ge m(t)$.

Из замечания 5.2 и теорем 3.1, 3.2 следует:

Теорема 5.1. Пусть $P<1-Q$. Тогда, если для всех типов элементов выполнено условие

$$\sum_{t=1}^{\infty}(1-Q')^{m(t)}<\infty, \tag{5.1}$$

то имеется положительная вероятность H безотказной работы корпоративной сети неограниченное время.

Если хотя бы для одного типа элементов выполнено условие

$$\sum_{t=1}^{\infty}(P'(1-Q''))^{m(t)}=\infty \quad , \tag{5.2}$$

то с вероятностью 1 отказ сети наступит за конечное число тактов.□

Теорема 5.2. Если принять параметрический закон роста

$$\begin{aligned}&(1-Q')^{m(t)}\sim t^{-r} \quad ,\\ &(P'(1-Q''))^{S(t)}\sim t^{-d},\end{aligned} \tag{5.3}$$

то условия (5.1)(5.2) означают $r>1;\quad d\le 1$. Критическая скорость роста, выше которой обеспечена положительная вероятность неограниченно долгой безотказной работы

$$m(t)>-\log(t)/\log(1-Q'). \tag{5.4}$$

Критическая скорость роста, ниже которой гарантировано, что отказ произойдет за конечное время, соответствует неравенству

$$S(t)<-\log(t)/(\log(P')+\log(1-Q'')); \tag{5.5}$$

Исходя из оценок замечания 5.2 из (5.5) для разных типов дублирования получаем:

Простое дублирование

$$m(t)<-\log(t)/(\log(P')+\log(1-Q'')); \tag{5.6}$$

Связное дублирование:

$$m(t)<-\sqrt{\log(t)}\sqrt{2/\log(P'(1-Q''))}; \tag{5.7}$$

Селективное дублирование (N — число элементов исходной ВС):

$$m(t)<-\sqrt{\log(t)}\ /(N\sqrt{\log(p'(1-Q''))}). \tag{5.8}$$

Для разных типов дублирования величины $P', P'', Q', Q'', S(t)$ принимают разные значения. В интервале между двумя оценками роста сети (5.4) и (5.6-5.8) данный метод анализа не дает определенного вывода о наработке ВС. □

Из доказанного видно, что неограниченное время безотказной работы ИВС возможно при логарифмическом росте по времени числа дублирующих экземпляров сети. Разумеется, при этом есть положительная вероятность отказа за конечное время. Коэффициент при логарифме должен быть не ниже указанного в (5.4). Различие между способами дублирования хорошо проявляется в оценке (5.5-5.8) для нижней критической скорости роста. Для методов дублирования со связями между экземплярами исходной сети оценка роста снижается до корня из логарифма времени. Однако при этом увеличивается методический интервал неопределенности критерия, поскольку полученная оценка верхней критической скорости не зависит явно от способа дублирования.

Работа выполнена при финансовой поддержке РФФИ, проект № 07-01-00101-а

СПИСОК ЛИТЕРАТУРЫ

1. Автоматы.// Сб. ред. К. Шеннон, Дж. Маккарти. М.: ИЛ, 1956.
2. *С. К. Клини*. Представление событий в нервных сетях и конечных автоматах. // Сб. Автоматы, ред. К. Шеннон, Дж. Макарти. М.: ИЛ, 1956.
3. *Дж. Нейман*. Вероятностная логика и синтез надежных организмов из ненадежных компонент. // Сб. ред. К. Шеннон, Дж. Маккарти. М.: ИЛ, 1956.
4. *Н. Мур, К. Шеннон*. Надежные схемы из ненадежных реле.// Кибернетический сборник. М.: ИЛ, 1960.
5. *Коганов А. В*. Индукторные пространства как средство моделирования // Вопросы кибернетики (Алгебра, Гипергеометрия, Вероятность, Моделирование). 1999. С. 119–181.
6. *Коганов А. В., Сазонов А. Н.* Анализ отказоустойчивости вычислительной среды корпоративного типа // Математика. Компьютер. Оборазование: Сб. научных трудов. Том. 2 / Под ред. Г.Ю.Ризниченко. – М.-Ижевск: НИЦ "Регулярная и хаотическая динамика". 2006. С. 228–234.
7. *Коганов А. В., Сазонов А. Н.* Корпоративные вычислительные сети с разрушением и восстановлением. Фазовый переход в растущих сетях // Математика. Компьютер. Образование: Сб. научных трудов. Том 2 / Под ред. Г.Ю.Ризниченко. – М.-Ижевск: НИЦ "Регулярная и хаотическая динамика". 2007. С. 48–59.
8. *Коганов А. В., Сазонов А. Н.* Анализ критических точек отказоустойчивости вычислительной среды планетарного типа // Программные продукты и системы, приложение к международному журналу «Проблемы теории и практики управления». 2007. №3. С. 27–31.
9. *Коганов А. В., Сазонов А. Н.* Анализ отказоустойчивости вычислительной среды планетарного типа // Математика. Компьютер. Оборазование: Сб. научных трудов. Том. 2 / Под ред. Г.Ю.Ризниченко. – М.-Ижевск: НИЦ "Регулярная и хаотическая динамика". 2006. С. 235–248.
10. *Цициашвили Г. Ц., Маркова Н. В.* Переходные явления в объединённой системе резервирования с восстановлением // Дальневосточный математический журнал. 2001. Т. 2. № 2. С. 106–114.

PHASE CONVERSION OF FAULT-TOLERANCE OF INCREASING NETWORK WITH ARBITRARY INITIAL CONFIGURATION

Koganov A. V., Sazonov A. N.

Fault-tolerance of a finite computing net with arbitrary graph, containing elements with certain probability of fault and restore, is analyzed. Algorithm for net growth at each work cycle is suggested. It is shown that if the rate of net increase is sufficiently big then the probability of infinity faultless work is positive. Estimated critical net increase rate is logarithmic over the number of work cycles