

УДК 519.248

© Л. В. Уткин, В. В. Фомин

МИНИМАКСНЫЙ ПОДХОД ДЛЯ ПОСТРОЕНИЯ КЛАССА МОДЕЛЕЙ НАДЕЖНОСТИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ НА ОСНОВЕ КОМБИНАЦИИ ОБОБЩЕННОГО БАЙЕСОВСКОГО ВЫВОДА И ЭЛЕМЕНТОВ ТЕОРИИ ОБУЧЕНИЯ

В работе предлагается новый подход к построению моделей надежности программного обеспечения, использующий комбинацию обобщенного байесовского вывода и элементов теории обучения. Для построения моделей рассматривается множество распределений вероятностей, отражающее недостаток имеющихся статистических данных, и на этом множестве осуществляется поиск максимума функционала риска. Параметры модели вычисляются минимизацией полученного функционала. Этот подход является достаточно общим и позволяет создавать различные модели надежности.

Кл. сл.: надежность, программное обеспечение, статистическое моделирование

ВВЕДЕНИЕ

Проявление ошибок в программном обеспечении (ПО) интенсивно изучалось в литературе в целях улучшения характеристик программ. Огромное количество моделей надежности программного обеспечения (МНПО) было разработано в последние десятилетия, но ни одна модель не дает точного прогноза для различных программ. Это следствие нереалистичности ряда предположений, используемых в моделях. Детальный и всесторонний критический обзор вероятностных моделей надежности ПО был предложен в работах [1, 2].

Основной идеей моделирования надежности ПО является попытка статистически описать процесс отладки программного обеспечения (исправления ошибок и тестирования до момента проявления очередной ошибки). При этом обычно предполагается, что после исправления очередной ошибки их общее количество уменьшается или программа становится лучше в смысле надежности. В данном случае говорят о росте надежности программы. Таким образом, для анализа надежности ПО с использованием статистических моделей их параметры в общем случае оцениваются из набора имеющихся отладочных данных и данных об отказах. Необходимо отметить, что слово "рост" на самом деле используется условно, т. к. для многих статистических данных наблюдается падение надежности из-за внесенных ошибок в процессе исправления выявленных и вследствие ряда других причин.

В работе предлагается новый класс моделей надежности ПО. Основная идея построения моде-

лей заключается в комбинированном использовании обобщенного байесовского вывода [3], где вместо одного априорного распределения вероятностей рассматривается множество распределений, и метода минимизации функционала риска при выборе "наихудшего" в определенном смысле распределения вероятностей. Выбор этого распределения определяет минимаксную стратегию.

Общий алгоритм построения модели надежности программного обеспечения можно представить в виде трех основных этапов.

1) построение множества функций распределения вероятностей (ФРВ) по обучающей выборке на основе обобщенного байесовского вывода;

2) выбор "оптимального" распределения вероятностей, максимизирующего функционал риска в соответствии с минимаксной стратегией;

3) вычисление оптимальных параметров модели, минимизирующих максимальный функционал риска.

1. МОДЕЛИ НАДЕЖНОСТИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Пусть X_i — случайное время между $(i - 1)$ -м и i -м отказами ПО. Случайная величина X_i имеет некоторое распределение вероятностей с функцией плотности $p_i(x)$ и с вектором параметров α_i . Основная цель анализа надежности ПО после исправления n ошибок в процессе отладки заключается в определении функции распределения времени до отказа $F_{n+1}(x)$ после n -го отказа, т. е. в период эксплуатации.

Большинство моделей надежности ПО различается предположениями о виде функции плотности $p_i(x)$ и функции f изменения параметров α_i в процессе отладки, которая называется функцией роста надежности ПО, заданной в явном или неявном виде.

Одной из наиболее известных и простых моделей надежности ПО является модель Джелински—Моранда [2]. Эта модель строится на следующих предположениях.

1) Исходное число ошибок в программе равно N . Ошибки не вносятся в программу в процессе отладки.

2) Время между отказами имеет экспоненциальное распределение с параметром $\lambda(N - (i - 1))$, который пропорционален числу оставшихся ошибок в ПО. Нарботка на отказ равна $1 / \lambda(N - (i - 1))$.

3) Интенсивность отказов постоянна в пределах интервала между отказами.

4) Каждая ошибка в ПО имеет одинаковые шансы быть обнаруженной.

5) Отказы независимы.

Функция плотности вероятности времени до отказа имеет вид

$$p(t) = \lambda(N - (i - 1)) \exp(-\lambda(N - (i - 1))t).$$

Задача построения модели сводится к получению параметров λ и N по статистическим данным в виде множества $\mathbf{X} = (x_1, \dots, x_n)$ наблюдаемых значений случайных величин $X_1, \dots, X_n$ времен до отказа в процессе отладки ПО.

2. ОБОБЩЕННЫЕ БАЙЕСОВСКИЕ МОДЕЛИ ДЛЯ ПОЛУЧЕНИЯ МНОЖЕСТВА РАСПРЕДЕЛЕНИЙ ВРЕМЕНИ ДО ОТКАЗА ПО

Стандартный байесовский подход

Одним из эффективных подходов к оценке параметров модели является байесовский анализ [5]. В соответствии с ним анализируемые параметры рассматриваются как случайные величины, для которых назначается некоторое априорное распределение вероятностей перед получением наблюдений или статистических данных. Если предположить, что случайная величина имеет распределение вероятностей с вектором неизвестных параметров $\mathbf{b}$, то эти параметры следует рассматривать, как случайные величины с априорной плотностью вероятностей $\pi(\mathbf{b} | \mathbf{c})$, которая в свою очередь характеризуется параметрами $\mathbf{c}$. В этом случае байесовский подход применяется для вычисления ФРВ случайной величины

$$F(k | \mathbf{c}) = \int_{\Omega} F(k | \mathbf{b}) \cdot \pi(\mathbf{b} | \mathbf{c}) d\mathbf{b}.$$

Здесь Ω — множество значений параметров $\mathbf{b}$.

Центральным элементом байесовского подхода является вывод апостериорного распределения неизвестных параметров при получении статистических данных. Это осуществляется на основе теоремы Байеса. Предположим, что априорное распределение $\pi(\mathbf{b} | \mathbf{c})$ отражает неопределенность относительно $\mathbf{b}$ до получения статистической информации в виде множества $\mathbf{X}$. Пусть $p(k)$ — функция плотности времени до отказа при условии $\mathbf{b}$. Тогда апостериорная плотность $\pi(\mathbf{b} | \mathbf{X}, \mathbf{c})$ как условная плотность параметра $\mathbf{b}$ при полученных наблюдениях $\mathbf{X}$ и априорном параметре $\mathbf{c}$ вычисляется как

$$\pi(\mathbf{b} | \mathbf{X}, \mathbf{c}) \propto p(x_1) \cdots p(x_n) \cdot \pi(\mathbf{b} | \mathbf{c}).$$

Здесь $\pi(\mathbf{b} | \mathbf{X}, \mathbf{c})$ отражает модифицированное представление о параметрах $\mathbf{b}$ с учетом полученной информации $\mathbf{X}$.

Априорное распределение зачастую выбирается таким образом, чтобы упростить вычисления. Наиболее эффективный путь в этом направлении — это выбор *согласованных априорных распределений* [5]. Если апостериорное распределение $\pi(\mathbf{b} | \mathbf{X}, \mathbf{c})$ и априорное распределение $\pi(\mathbf{b} | \mathbf{c})$ принадлежат одному семейству распределений, то π и p называются согласованными распределениями, а π называется согласованным априорным распределением для p .

Интервальные априорные модели

Одним из факторов выбора априорного распределения, т. е. его параметров, является то, что при наличии информативных статистических данных даже плохое априорное распределение не повлияет существенно на апостериорное. С другой стороны, до появления каких-либо статистических данных или наблюдений о параметрах зачастую ничего неизвестно. Поэтому если отбросить требование простоты вычислений, то предпочтительное, или хорошее (подходящее), априорное распределение должно минимально влиять на вывод, т. е. на апостериорное распределение, а также учитывать отсутствие априорной информации о параметрах. Априорные распределения, моделирующие отсутствие априорной информации, называются *неинформативными*.

В литературе существует достаточно большое количество подходов для выбора того или иного неинформативного априорного распределения [3], имеющих свои достоинства и недостатки. Однако наиболее интересным является подход, полностью

отличающийся от большинства традиционных. Суть этого подхода заключается в следующем.

Определим не одно априорное распределение, а целый класс $\mathcal{R}$ распределений π , для которого можно найти нижнюю и верхнюю вероятности события A как

$$\underline{P}(A) = \inf\{P_{\pi}(A) : \pi \in \mathcal{R}\},$$

$$\overline{P}(A) = \sup\{P_{\pi}(A) : \pi \in \mathcal{R}\}.$$

При определенных условиях множество $\mathcal{R}$ полностью определяется нижней и верхней функциями распределения вероятностей. Следует отметить, что класс $\mathcal{R}$ следует рассматривать не как класс подходящих априорных распределений, а как подходящий класс априорных распределений. Это значит, что каждое отдельное распределение из класса не является подходящим, или хорошим априорным распределением, т. к. ни одно отдельное распределение не может удовлетворительно моделировать отсутствие информации. Но весь класс в целом, определяемый верхним и нижним распределениями вероятностей, является подходящей моделью отсутствия информации. Когда априорной информации почти нет, $\underline{P}(A)$ для этого класса должно быть близко к 0, а $\overline{P}(A)$ близко к 1. Это означает, что априори событие A может иметь любую вероятность.

В большинстве случаев класс распределений $\mathcal{R}$ определяется параметрически. К наиболее известным таким классам априорных распределений следует отнести обобщенную модель Дирихле [6], обобщенные модели экспоненциального семейства распределений [7]. В нашем подходе множество $\mathcal{R}$ используется для того, чтобы учесть статистические данные и сгенерировать множество $\mathcal{F}$ прогнозируемых распределений с нижней и верхней границами.

Обобщенная гамма-экспоненциальная модель

Одной из неинформативных моделей, основанных на определении класса $\mathcal{F}$, является гамма-экспоненциальная модель, которую можно рассматривать как частный случай моделей, использующих экспоненциальное семейство распределений вероятностей [7]. Эта модель также может рассматриваться как класс распределений Парето с определенным множеством параметров. Если наблюдаем событие отказа в момент времени t и предполагаем, что это время имеет экспоненциальное распределение с параметром λ , то согласованное априорное распределение для параметра $\lambda > 0$ является гамма-распределением, имеющим функцию плотности вероятности

$$\pi(\lambda) = \text{Gamma}(a, b) = \frac{1}{\Gamma(a)} b^a \lambda^{a-1} \exp(-b\lambda).$$

Здесь $a > 0$ и $b > 0$ — параметры сдвига и масштаба гамма-распределения; $\Gamma(a)$ — стандартная гамма-функция. Апостериорная плотность $\pi(\lambda | x)$ после наблюдения события в момент времени x имеет вид

$$\pi(\lambda | x) = \text{Gamma}(a + 1, b + x).$$

Отсюда апостериорная функция распределения равна

$$\begin{aligned} \Pr\{X < t\} &= F(t) = \\ &= \int_0^t (1 - e^{-\lambda x}) \cdot \text{Gamma}(a + 1, b + x) d\lambda = \\ &= 1 - \left(\frac{b + x}{b + t + x} \right)^{a+1}. \end{aligned}$$

Здесь байесовский вывод приводит к распределению Парето времени до отказа программы с параметрами a и b . Заменим параметры $a = s\gamma$ и $b = s - s\gamma$ и предположим, что новый параметр γ может быть любым в интервале от 0 до 1. Изменяя этот параметр, получаем множество всех распределений Парето с фиксированным параметром s и множеством параметров $0 \leq \gamma \leq 1$. Это множество ограничено некоторой верхней и нижней граничными функциями распределения вида

$$\underline{F}^{(s)}(t) = 1 - \left(\frac{s + x}{s + t + x} \right)^1, \quad \overline{F}^{(s)}(t) = 1 - \left(\frac{x}{t + x} \right)^{s+1}.$$

Здесь параметр $s > 0$ определяет влияние априорного распределения на апостериорное. В частности, если $s = 0$, то апостериорное распределение полностью определяется только информацией о числе отказов n и k . Используя множество распределений вероятностей вместо одного распределения, можно найти только нижнюю и верхнюю границы для вероятностей событий вместо точных значений. При этом нижняя и верхняя границы для вероятностей событий могут быть получены минимизацией и максимизацией вероятностей событий по всем значениям параметра γ в интервале $[0, 1]$. Параметр s также определяет, насколько быстро границы вероятностей сходятся при поступлении статистических данных. Меньшие значения s дают быструю сходимость и более рискованные заключения, в то время как большие значения s приводят к более осторожному статистическому выводу.

Перед получением каких-либо наблюдений имеем $x = 0$ и, следовательно, нижняя вероятность отказа равна 0, а верхняя вероятность равна 1.

Обобщенная байесовская модель роста надежности ПО

Для моделирования надежности ПО мы используем обобщенную гамма-экспоненциальную модель, предполагая, что i -е время X_i до отказа программы имеет экспоненциальное распределение с параметром $\lambda_i = \lambda(N - (i - 1))$, $i = 1, \dots, n$.

Заметим, что среднее значение параметра λ_i , имеющего гамма-распределение с априорными параметрами a_i и b_i , равно a_i / b_i , т. е. $E\lambda_i = a_i / b_i$.

Если зафиксировать параметр $a_i = a$, то запишем

$$\lambda(N - (i - 1)) = a / b_i.$$

Отсюда априорный параметр $b_1 = b$ равен $a / (\lambda N)$, и выполняется условие $\lambda = a / (bN)$. Другими словами, априорный параметр b_i , учитывающий рост надежности, равен

$$b_i = a / (\lambda(N - i + 1)) = b \cdot f(i),$$

$$\text{где } f(i) = \frac{N}{N - i + 1}.$$

После исправления соответствующей ошибки i -й параметр b_i получается равным $b_i g(i)$, где $g(i) = \frac{N - i}{N - i + 1}$.

Если априорный параметр гамма-распределения равен b_1 , то апостериорный параметр после отказа в момент времени x_1 равен $b_1 + x_1$. После исправления ошибки в программе мы получим априорный параметр $b_2 = (b_1 + x_1)g(1)$. Тогда апостериорный параметр после отказа в момент времени x_2 равен

$$b_2 + x_2 = (b_1 + x_1)g(1) + x_2.$$

После исправления соответствующей ошибки в программе мы получим априорный параметр $b_3 = (b_2 + x_2)g(2)$. Тогда апостериорный параметр после отказа в момент времени x_3 равен

$$\begin{aligned} b_3 + x_3 &= (b_2 + x_2)g(2) + x_3 = \\ &= ((b_1 + x_1)g(1) + x_2)g(2) + x_3. \end{aligned}$$

Последовательно продолжая рассуждения, мы получим выражение для априорного параметра

после исправления n ошибок:

$$b_n = b_1 \prod_{i=1}^n g(i) + \sum_{i=1}^n x_i \left(\prod_{j=i}^n g(j) \right).$$

Заметим, что верны следующие преобразования:

$$\begin{aligned} \prod_{i=1}^n g(i) &= \frac{N-1}{N} \frac{N-2}{N-1} \dots \frac{N-n}{N-n+1} = \frac{N-n}{N}, \\ \prod_{j=i}^n g(j) &= \frac{N-i}{N-i+1} \frac{N-i-1}{N-i} \dots \frac{N-n}{N-i-n+1} = \\ &= \frac{N-n}{N-i+1}. \end{aligned}$$

Отсюда

$$b_n = b_1 \frac{N-n}{N} + \sum_{i=1}^n x_i \frac{N-n}{N-i+1}.$$

Апостериорный параметр a_n после n -го отказа равен $a_n = a + n$. Обозначим

$$B = \sum_{i=1}^n x_i \frac{N-n}{N-i+1}, \quad C = \frac{N-n}{N}.$$

Тогда можно записать функцию распределения вероятностей $F_{n+1}(t)$ после n -го шага отладки программы следующим образом:

$$F_{n+1}(t) = 1 - \left(\frac{Cb + B}{Cb + B + t} \right)^{a+n}.$$

Для построения обобщенной модели роста необходимо выбрать ограниченное множество векторов (a, b) . Предлагается выбрать множество векторов (a, b) в треугольнике $(0, 0)$, $(s, 0)$, $(0, s)$. Это означает, что все возможные априорные интенсивности отказов λ учитываются в этом случае, т. к. математическое ожидание интенсивности отказов равно $E\lambda = a / b$. При (a, b) из приведенного выше треугольника ожидаемая интенсивность отказов изменяется в интервале $(0, \infty)$.

Нижняя граница функции распределения $F_{n+1}^{(s)}(t)$ определяется как

$$\underline{F}_{n+1}^{(s)}(t) = 1 - \left(\frac{Cs + B}{t + Cs + B} \right)^n.$$

Верхняя граница функции распределения $\overline{F}_{n+1}^{(s)}(t)$ определяется как

$$\overline{F}_{n+1}^{(s)}(t) = 1 - \left(\frac{B}{t + B} \right)^{s+n}.$$

Соответствующие функции плотности равны

$$\underline{p}_{n+1}^{(s)}(t) = \frac{n(B + Cs)^n}{(t + B + Cs)^{n+1}}, \quad \overline{p}_{n+1}^{(s)}(t) = \frac{(s + n)B^{s+n}}{(t + B)^{s+n+1}}.$$

3. ОЦЕНКА ПАРАМЕТРОВ МОДЕЛИ НАДЕЖНОСТИ ПО В РАМКАХ ТЕОРИИ ОБУЧЕНИЯ

Функционал риска

Сформулируем стандартную регрессионную модель в терминах теории обучения. Согласно работе Вапника [10], задача обучения может быть описана следующим образом. Выбирается наилучшая функция $f(x, \alpha_0)$ из параметрического множества функций $f(x, \alpha)$ на множестве параметров $\alpha \in \Lambda$. Выбор подходящей функции основан на "зашумленной" обучающей выборке, состоящей из n независимых и одинаково распределенных наблюдений (x_i, y_i) , $i = 1, \dots, n$. Качество аппроксимации определяется функцией потерь $L(y - f(x, \alpha)) = L(z)$, например $L(z) = z^2$. Основная цель обучения — минимизация функционала риска

$$R(\alpha) = \iint L(y - f(x, \alpha)) p(y - f(x, \alpha)) dx dy.$$

Минимизация функционала риска осуществляется в классе функций $f(x, \alpha)$, $\alpha \in \Lambda$. Другими словами, функция $f(x, \alpha_0)$ обеспечивает минимум $R(\alpha)$ так, что $R(\alpha_0) = \min_{\alpha \in \Lambda} R(\alpha)$.

Если совместная ФПВ $p(x, y)$ неизвестна, то функционал риска $R(\alpha)$ можно заменить так называемым эмпирическим функционалом риска

$$R_{\text{emp}}(\alpha) = \frac{1}{n} \sum_{i=1}^n L(y_i - f(x_i, \alpha)).$$

Когда распределения вероятностей известны, то общепринятый метод вычисления наилучшей функции $f(x, \alpha_0)$ — метод максимума функции правдоподобия [10].

Проблема обучения при условии наличия множества распределений

Предположим, что мы не знаем точное распределение шума, но имеется информация о том, что это распределение принадлежит множеству $\mathcal{F}$, ограниченному некоторой нижней $\underline{F}$ и верхней $\overline{F}$ функциями распределения вероятностей. Эти границы можно получить на основе наблюдений (x_i, y_i) , $i = 1, \dots, n$. Введем также множество $\mathcal{R}$ со-

ответствующих функций плотности вероятности, т. е. для каждой функции из $\mathcal{F}$ ставим в соответствие функцию p из $\mathcal{R}$. Рассмотрим минимаксную (пессимистическую) стратегию построения модели.

Обозначим $z = y - f(x, \alpha)$. Заметим, что значения x при анализе надежности есть номера обнаруженных ошибок. Тогда соответствующая величина (номер обнаруженной ошибки) не является случайной и функционал риска в соответствии с минимаксной стратегией имеет следующий вид:

$$\overline{R}(\alpha_0) = \min_{\alpha \in \Lambda} \max_{p(z) \in \mathcal{R}} \int L(z) p(z) dz.$$

Приведенное выражение можно объяснить достаточно просто. Мы не знаем точную функцию плотности p ошибки или шума, и каждое распределение из $\mathcal{R}$ может быть выбрано. Поэтому нам следует выбрать наихудшее распределение, приводящее к наибольшему значению функционала риска. Минимаксную стратегию можно рассматривать как некоторую страховку против наихудшей ситуации, т. к. эта стратегия минимизирует ожидаемые потери в наименее благоприятной ситуации.

Главное условие, которое будет использоваться ниже, заключается в том, что функция потерь $L(y - f(x, \alpha))$ имеет один минимум для каждого $\alpha \in \Lambda$. Это предположение верно для самых различных известных типов функций потерь. Например, квадратичная функция потерь имеет минимум 0 в точке $y - f(x, \alpha)$.

В работах [9, 10] доказано, что оптимальная функция распределения вероятностей из множества $\mathcal{F}$, обеспечивающая верхнюю границу математического ожидания функции L , имеет вид

$$F_U(x) = \begin{cases} \overline{F}(x) & \text{для } x \leq \overline{F}^{-1}(\tau), \\ \tau & \overline{F}^{-1}(\tau) < x < \underline{F}^{-1}(\tau), \\ \underline{F}(x) & x \geq \underline{F}^{-1}(\tau). \end{cases}$$

Соответствующая оптимальная функция плотности записывается как

$$p_U(x) = \begin{cases} d\overline{F}(x)/dx, & x \leq \overline{F}^{-1}(\tau); \\ 0, & \overline{F}^{-1}(\tau) < x < \underline{F}^{-1}(\tau); \\ d\underline{F}(x)/dx, & x \geq \underline{F}^{-1}(\tau). \end{cases}$$

Здесь значение величины τ определяется из условия $\underline{F}^{-1}(\tau) + \overline{F}^{-1}(\tau) = 0$.

Используя эти результаты можно получить

$$\bar{R}(\alpha) = \int L(z) p_U(z) dz.$$

Упростим выражение для $\bar{R}(\alpha)$

$$\begin{aligned} \bar{R}(\alpha) = & \int_{-\infty}^{\bar{F}^{-1}(\tau)} L(y - f(x, \alpha)) d\bar{F}(y - f(x, \alpha)) + \\ & + \int_{\bar{F}^{-1}(\tau)}^{\infty} L(y - f(x, \alpha)) d\bar{F}(y - f(x, \alpha)). \end{aligned}$$

Оптимальное значение параметров α находится минимизацией $\bar{R}(\alpha)$ на множестве $\alpha \in \Lambda$.

Функционал риска в модели надежности ПО

Качество прогнозирования надежности программного обеспечения во многих задачах оценивается некоторым показателем расстояния между прогнозируемым ожидаемым значением времени до отказа программы и фактическим значением времени до отказа. Так как ожидаемое время до отказа (называемое также наработкой до отказа) после n -го этапа отладки программы равно $M = 1 / \lambda(N - n + 1)$, то функцию потерь $L(y, f(x, \alpha))$ можно определить следующим образом:

$$L(y, f(x, \alpha)) = L(y - M).$$

Здесь $\alpha = (\lambda, N)$. Можно принять, для примера,

$$L(y, f(x, \alpha)) = (y - M)^2$$

$$\text{или } L(y, f(x, \alpha)) = |y - M|.$$

Заметим, что функции плотности вероятности, полученные в разделе 2, характеризуют время до отказа X , но не разность между временем до отказа и наработкой на отказ, т. е. $X - M$. Поэтому необходимо определить нижнюю и верхнюю "смещенные" функции плотности

$$\underline{p}^{(s)}(t) = \frac{n(B + Cs)^n}{(t + B + Cs + M)^{n+1}},$$

$$\bar{p}^{(s)}(t) = \frac{(s + n)B^{s+n}}{(t + B + M)^{s+n+1}}.$$

Здесь $t \geq -M$. Соответствующие "смещенные" границы функций распределения вероятностей равны

$$\underline{F}^{(s)}(t) = 1 - \left(\frac{Cs + B}{t + Cs + B + M} \right)^n,$$

$$\bar{F}^{(s)}(t) = 1 - \left(\frac{B}{t + B + M} \right)^{s+n}.$$

Для вычисления функционала риска необходимо решить уравнение $\underline{F}^{-1}(\tau) + \bar{F}^{-1}(\tau) = 0$, где

$$\underline{F}^{-1}(\tau) = (Cs + B)(1 - \tau)^{1/n} - (Cs + B + M),$$

$$\bar{F}^{-1}(\tau) = B(1 - \tau)^{1/(s+n)} - (B + M).$$

Так как мы предположили, что минимум функции потерь L достигается в точке $y - f(x, \alpha) = 0$, то верхняя граница функционала риска, соответствующая минимаксной стратегии, принимает вид

$$\begin{aligned} \bar{R}(N, \lambda) = & \int_{-M}^{\bar{F}^{-1}(\tau)} \frac{(s + n)B^{s+n} \cdot L(y - M)}{(t + B + M)^{s+n+1}} dy + \\ & + \int_{\bar{F}^{-1}(\tau)}^{\infty} \frac{n(B + Cs)^n \cdot L(y - M)}{(y + B + Cs + M)^{n+1}} dy. \end{aligned}$$

Оптимальные значения параметров N, λ могут быть найдены минимизацией функционала $\bar{R}(N, \lambda)$ по множеству положительных значений N, λ .

ЗАКЛЮЧЕНИЕ

В работе предложен новый подход к построению моделей надежности ПО, использующий комбинацию обобщенного байесовского вывода и элементов теории обучения. Этот подход является достаточно общим и позволяет создавать различные модели надежности, меняя вид функции потерь, вид функциональной зависимости параметров модели в процессе отладки ПО и способ построения множества распределений.

Необходимо отметить, что в работе не приводятся результаты сравнительных статистических экспериментов с разработанной моделью. Целью работы являлась не разработка конкретной модели, а разработка подхода к созданию различных моделей, адекватно отражающих недостаток статистической информации, который обычно имеет место при моделировании надежности ПО.

Необходимо также отметить, что рассматриваемое множество функций распределения вероятностей не является параметрическим множеством распределений одного типа, совпадающего с типом граничных функций. Множество образуется из всех возможных функций распределения с заданными границами. Это — важная особенность предлагаемого подхода, которая отличает его от существующих методов, использующих только параметрические классы функций распределения.

Другой особенностью предлагаемого подхода является возможность сужать множество распределений вероятностей по мере поступления статистической информации об отказах и исправлениях соответствующих ошибок. Очевидно, что величина

на множества распределений влияет на ширину итоговых границ распределений времени до отказа после завершения отладочного процесса. Это позволяет создавать правило остановки отладочного процесса не только по анализу среднего времени до отказа, но и с учетом неточности результатов.

Направлением дальнейшего исследования является поиск методов и эффективных алгоритмов решения задач оптимизации на основе функционала риска для тех случаев, когда в явном виде сложно записать функционал риска и найти соответствующий интеграл.

СПИСОК ЛИТЕРАТУРЫ

1. Уткин Л.В., Шубинский И.Б. Нетрадиционные методы оценки надежности информационных систем. СПб.: Любавич, 2000. 173 с.
2. Cai K.-Y., Wen C.-Y., Zhang M.-L. A Critical Review on Software Reliability Modeling // Reliability Engineering and System Safety. 1991. V. 32, N 3. P. 357–371.
3. Уткин Л.В. Анализ риска и принятие решений при неполной информации. СПб.: Наука, 2007. 404 с.
4. Jelinski Z., Moranda P.B. Software Reliability Research // Statistical Computer Performance Evaluation. N.Y.: Academic Press, 1972. P. 464–484.
5. Bernardo J.M., Smith A.F.M. Bayesian Theory // Statistical Methods and Applications. 1994. V. 3, N 1. P. 155–160.
6. Walley P. Inferences from Multinomial Data: Learning about a Bag of Marbles // Journal of the Royal Statis-

- tical Society: Series B. 1996. V. 58. P. 3–57.
7. Quaeghebeur E., de Cooman G. Imprecise Probability Models for Inference in Exponential Families // Proc. of the 4th Int. Symposium on Imprecise Probabilities and Their Applications, ISIPTA'05. Pennsylvania, Pittsburgh: Carnegie Mellon University, July 2005. P. 287–296.
8. Vapnik V. The Nature of Statistical Learning Theory. N.Y.: Springer, 1995. 188 p.
9. Utkin L.V. Risk Analysis under Partial Prior Information and Non-Monotone Utility Functions // International Journal of Information Technology and Decision Making. 2007. V. 6, N 4. P. 625–647.
10. Utkin L.V., Destercke S. Computing Expectations with Continuous P-Boxes: Univariate Case // International Journal of Approximate Reasoning. 2009. V. 50, N 5. P. 778–798.

Санкт-Петербургская государственная лесотехническая академия (Уткин Л.В.)

Российский государственный педагогический университет им. А.И. Герцена, Санкт-Петербург (Фомин В.В.)

Контакты: Уткин Лев Владимирович,
lev.utkin@mail.ru

Материал поступил в редакцию 19.04.2010.

THE MINIMAX APPROACH TO THE CONSTRUCTION OF A MODEL CLASS SOFTWARE RELIABILITY BASED ON A COMBINATION OF GENERALIZED BAYESIAN INFERENCE, AND ELEMENTS OF LEARNING THEORY

L. V. Utkin¹, V. V. Fomin²

¹Saint-Petersburg State Forest Technical Academy

²Herzen State Pedagogical University of Russia, Saint-Petersburg

The paper proposes a new approach to constructing models of software reliability, using a combination of generalized Bayesian inference, and elements of learning theory. To construct the models we consider a set of probability distributions, reflecting the lack of available statistical data, and on this set, search for the maximum of functional risk is done. The model parameters are computed by minimizing the obtained functional. This approach is quite general and allows to develop various models of reliability.

Keywords: reliability, software, statistical modeling