УДК 004.415.5+53.081
Б.М. Конорев, В.Л. Петрик

Обоснование выбора системы единиц физических величин для независимой верификации при сертификации программного обеспечения

Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ»

Введение

Современные информационно-управляющие системы (ИУС) характеризуются все возрастающей частью функций, реализованных программно, что позволяет отнести их к системам с интенсивным использованием программного обеспечения (ПО).

Наличие остаточных программных дефектов (ПД) в программных реализациях критически важных функций определяет риски возникновения нештатных ситуаций, связанных с нарушением функциональной безопасности (ФБ) ИУС [1]. Снижение рисков до предельно допустимого уровня обеспечивается независимой верификацией (НВ) в процессе сертификации ПО, основная цель которой заключается в определении вероятности наличия остаточных ПД.

В основу НВ положен метод семантического контроля (СК). СК позволяет посредством анализа семантического программного инварианта – физических размерностей программных переменных, выполнять семантическую независимую верификацию (СНВ) программного кода в режиме статического анализа и обнаружить семантические программные дефекты (СПД) [2, 3].

Постановка задачи

СНВ ПО выполняется в семантическом пространстве, базис которого определяется выбранной системой единиц (СЕ). Чувствительность методов СНФ, определяющая долю обнаруживаемых ПД, а следовательно, и вероятность наличия остаточных программных дефектов при условии их необнаружения в программном коде, зависит от СЕ. Поэтому возникает необходимость решения проблемы выбора оптимальной СЕ с точки зрения обеспечения максимальной эффективности СК и минимальной потребности в вычислительных ресурсах.

Цель статьи — обоснование выбора оптимальной СЕ с точки зрения достоверности, определяющей эффективность обнаружения СПД, и ресурсоемкости СНВ ПО.

Результаты исследования

В процессе исследований под ресурсоемкостью в соответствии со стандартом ISO/IEC 9126-4 [4] будем понимать комплексный показатель, оценивающий дополнительные накладные расходы: объем оперативной памяти, требуемой для хранения семантической информации, и ресурсы центрального процессора, необходимые для реализации.

При СНВ возможны три исхода событий:

1) СПД отсутствуют, обозначим вероятность такого события $P_{норма}$;
2) СПД присутствуют, и они обнаружены, обозначим вероятность события $P_{обн}$;
3) СПД присутствуют, но не обнаружены, обозначим вероятность события $P_{необн}$.

Под эффективностью будем понимать вероятность обнаружения СПД при условии их существования

$$\eta = \frac{P_{обн}}{P_{обн} + P_{необн}}. \qquad (1)$$

Ввиду того, что СНВ основана на совпадении размерностей операндов аддитивных операций (отношения, сложения, вычитания, присваивания), достоверность метода зависит от количества физических величин, имеющих в выбранной СЕ совпадающие размерности, что влияет на $P_{необн}$.

Предположим, что СЕ определяет N физических величин, часть которых имеет совпадающие размерности. Например, в СИ размерность поверхностной плотности теплового потока [q]=MT^{-3}, ее единица – Вт/м^2. Ту же размерность и единицу измерения имеют интенсивность излучения, интенсивность звука, излучательность. В СГС одинаковые размерности, например, имеют сила тока [I]=$L^{3/2}M^{1/2}T^{-2}$ и магнитный поток [Ф]=$L^{3/2}M^{1/2}T^{-2}$ [5].

Далее будем считать, что использование физических величин, определяющих физический тип программных переменных, подчинено равномерному закону распределения внутри выбранной СЕ, и предполагать, что СПД могут быть вызваны только совпадением физических размерностей, обусловленных несовершенством СЕ.

Объединим физические величины, имеющие совпадающие размерности, во множества, количество которых обозначим M. Каждое из множеств M_i содержит n_i (i=1,2,…,M) физических величин совпадающей размерности, причем $\sum_{i=1}^{M} n_i = N$.

События отсутствия, обнаружения и необнаружения СПД в пределах i-го множества имеют вероятности $P_{норма_i}, P_{обн_i}$ и $P_{необн_i}$ и образуют полную группу событий

$$P_{норма_i} + P_{обн_i} + P_{необн_i} = 1.$$

Сумма данных вероятностей для объединения всех M множеств, т.е. для всей СЕ:

$$\sum_{i=1}^{M} (P_{норма_i} + P_{обн_i} + P_{необн_i}) = \sum_{i=1}^{M} 1 = M \ .$$

Считаем, что СПД отсутствует, если мы выбираем физическую величину, принадлежащую множеству необходимой размерности и совпадающую с требуемой. Поэтому вероятность отсутствия СПД в пределах i-го множества $P_{норма_i}$ является произведением вероятности выбора i-го множества P_{Mi} и вероятности P_i выбора величины, совпадающей с требуемой:

$$P_{норма\ i} = P_{Mi} P_i = \frac{n_i}{N} \cdot \frac{1}{n_i} = \frac{1}{N}.$$

Вероятность отсутствия СПД для объединения множеств

$$P_{норма} = \sum_{i=1}^{M} P_{норма\ i} = \sum_{i=1}^{M} \frac{1}{N} = \frac{M}{N}.$$

СПД не обнаруживается, если его причиной является выбор физической величины, принадлежащей множеству необходимой размерности, но не являющейся требуемой. Отсюда вероятность необнаружения СПД в пределах i-го

множества $P_{необн_i}$ является произведением вероятности выбора i-го множества P_{Mi} и вероятности $\overline{P_i}$ невыбора требуемой величины внутри i-го множества

$$P_{необн_i} = P_{Mi}\overline{P_i} = \frac{n_i}{N}\cdot\frac{n_i - 1}{n_i} = \frac{n_i - 1}{N}.$$

Вероятность необнаружения СПД для объединения множеств

$$P_{необн} = \sum_{i=1}^{M} P_{необн_i} = \sum_{i=1}^{M} \frac{n_i - 1}{N} = 1 - \frac{M}{N}.$$

СПД обнаруживается, если мы выбираем физическую величину, принадлежащую i-му множеству с отличной от требуемой размерностью. Вероятность такого события

$$P_{обн_i} = \frac{N - n_i}{N}.$$

Определим вероятность обнаружения СПД для СЕ

$$P_{обн} = \sum_{i=1}^{M} P_{обн_i} = \sum_{i=1}^{M} \frac{N - n_i}{N} = M - 1.$$

Найденные вероятности $P_{норма}$, $P_{обн}$ и $P_{необн}$ подставим в формулу (1):

$$\eta = \frac{P_{обн}}{P_{обн} + P_{необн}} = \frac{M - 1}{M - 1 + 1 - \frac{M}{N}} = \frac{N(M-1)}{NM - M} = \frac{N(M-1)}{M(N-1)}.$$

Таким образом, достоверность СНВ, обусловленная статистическими характеристиками выбранной СЕ:

$$\eta = \frac{N(M-1)}{NM - M} = \frac{N(M-1)}{M(N-1)}. \qquad (2)$$

Данные результаты были подтверждены экспериментально путем моделирования совокупности неких условных СЕ, для которых количество элементов N и множеств M генерировалось случайным образом. Подсчитывались статистические вероятности $\tilde{P}_{норма_i}$, $\tilde{P}_{необн_i}$ и $\tilde{P}_{обн_i}$, которые совпали с теоретическими до четвертого знака после запятой. На рис. 1 показана зависимость достоверности СНВ в СЕ, обладающей заданным количеством множеств и общим количеством физических величин N>>M.

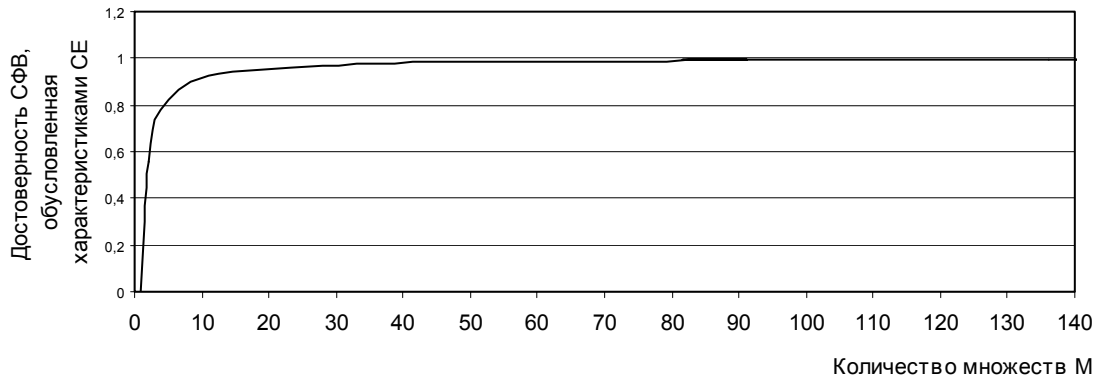

Рис 1. Достоверность СНВ, обусловленная статистическими характеристиками СЕ

Проведем анализ трех реальных систем единиц – СИ, СГС, МКГСС и оценим их влияние на достоверность СНВ.

Так как МКГСС включает в себя только механические и геометрические единицы, то сначала оценим достоверность выше указанных СЕ в механическом базисе (длина, время, масса). Подсчитаем количество геометрических и механических величин с совпадающими размерностями для каждой СЕ и объединим их во множества. Результаты показаны на рис.2.

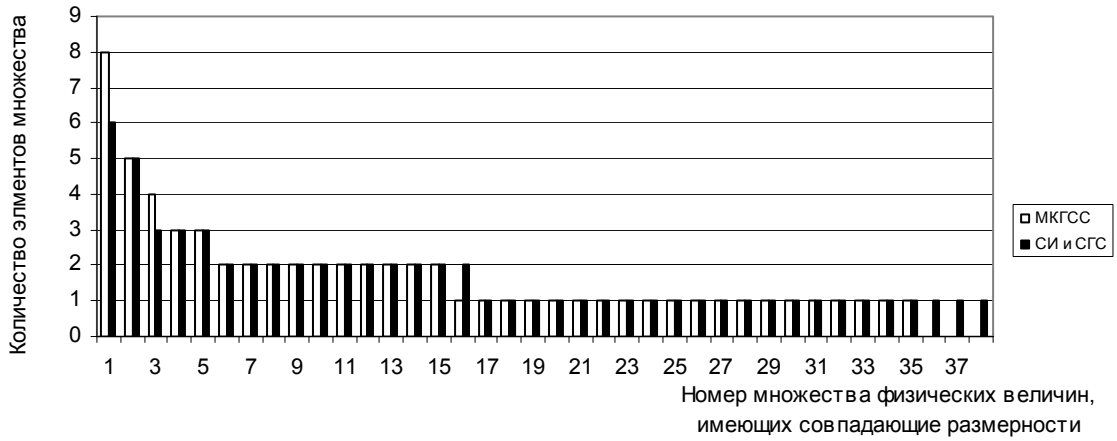

Рис. 2. Статистические характеристики систем МКГСС, СИ и СГС в механическом базисе

Тогда достоверность для СИ, СГС и МКГСС в механическом базисе, вычисленная по формуле (2):

$$\eta_{МКГСС} = \frac{63(35-1)}{35(63-1)} \approx 0{,}987,$$

$$\eta_{СИ} = \frac{63(38-1)}{38(63-1)} \approx 0{,}989, \quad \eta_{СГС} = \frac{63(38-1)}{38(63-1)} \approx 0{,}989.$$

Таким образом, предпочтительнее системы СИ и СГС.

На рис. 3 изображены распределения множеств физических величин с совпадающими размерностями.

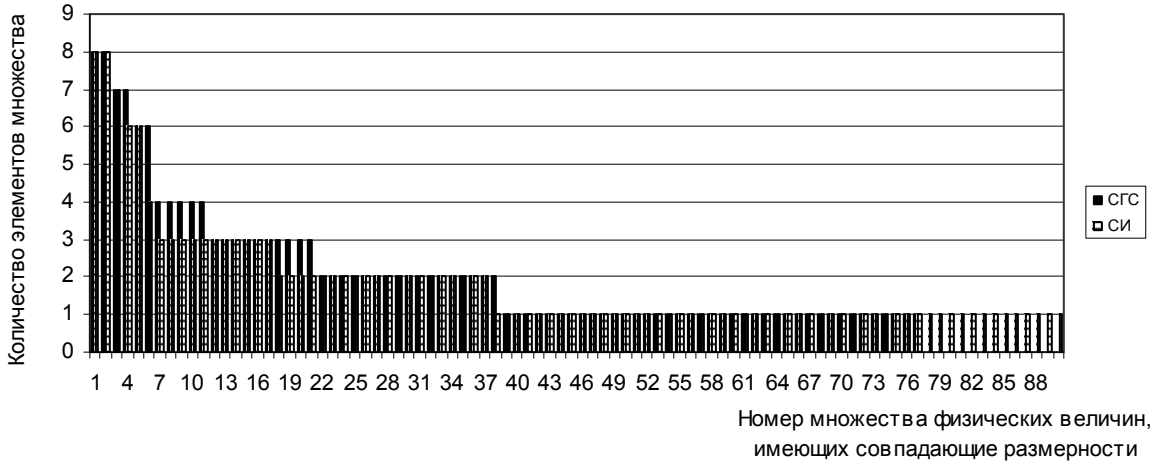

Рис. 3. Статистические характеристики систем СИ и СГС

Достоверность, вычисленная по формуле (2) для СИ и СГС:

$$\eta_{СИ} = \frac{165(90-1)}{90(165-1)} \approx 0{,}995,$$

$$\eta_{СГС} = \frac{165(77-1)}{77(165-1)} \approx 0{,}993.$$

Как уже отмечалось, современные системы критического применения ПО должны удовлетворять повышенным требованиям к ФБ. Поэтому для окончательного выбора воспользуемся величиной, дополнительной к достоверности

$$\overline{\eta} = 1 - \eta \,. \qquad (3)$$

Тогда

$$\overline{\eta}_{СИ} = 1 - 0{,}995 = 0{,}005,$$

$$\overline{\eta}_{СГС} = 1 - 0{,}993 = 0{,}007.$$

Поэтому эффективность СФВ, основанная на использовании СИ, выше на 30%.

Выбор СЕ влияет также на ресурсоемкость СФВ, так как в основе метода – решение нескольких систем линейных алгебраических уравнений, количество которых определяется количеством основных единиц, например: для СГС – 3, СИ – 9, МКГСС – 3. Наличие в СГС физических типов с дробными значениями основных единиц требует при СНВ операций над числами с плавающей точкой, что увеличивает требуемые ресурсы.

Выводы и направления дальнейших исследований

В статье была проведена оценка влияния статистических характеристик СЕ СИ, СГС, МКГСС на ресурсоемкость СНВ и эффективность, определяемую достоверностью СНВ. При этом с точки зрения эффективности оптимальной является СИ, а с точки зрения ресурсоемкости – СГС.

Дальнейшие исследования целесообразно направить на снижение ресурсоемкости семантической независимой верификации с одновременным сохранением ее эффективности, а также на преодоление неполноты документирования ПО, обусловленной отсутствием и искажением данных в документации, что потребует реализации семантической верификации в условиях неопределенных проектных спецификаций.

Список литературы

1. Липаев В.В. Надежность программных средств. Сер. «Информатизация России на пороге XXI века».— М.: СИНТЕГ, 1998. – 232 с.
2. Харченко В.С., Манжос Ю.С., Петрик В.Л. Статистический анализ ПО системы управления КА и оценка проверяющей способности семантического контроля. //«Технологии машиностроения».– Харьков. – 2002. –Вып. 2. – С. 32-43.
3. Калибровка чувствительности методов статического анализа, используемых для оценки качества и безопасности по ИУС АЭС. /Конорев Б.М., Манжос Ю.С., Петрик В.Л. и др.// Междунар. симпозиум «Измерения, важные для безопасности в реакторах». – М. –23-25 ноября 2004. –С. 15-1–15-12.
4. ISO/IEC PTDR 9126-4: Software Engineering – Product Quality – Part 4: Quality In Use Metrics. 2000.
5. Сена Л.А. Единицы физических величин и их размерности. – М.: Наука, 1988. – 431 с.