

УДК 001.891.573; 519.872

ПОГРЕШНОСТИ МОДЕЛИРОВАНИЯ ВЫСОКОНАГРУЖЕННЫХ
СИСТЕМ В GPSS WORLD

Т.И. Алиев

Излагаются результаты исследования влияния программных генераторов псевдослучайных величин на точность моделирования высоконагруженных систем массового обслуживания в системе имитационного моделирования GPSS World. Показано, что при некоторых сочетаниях генераторов, называемых проблемными, погрешности определения среднего времени ожидания заявок в высоконагруженной системе могут достигать десятков процентов. Выполненный анализ проблемных сочетаний позволил выявить факторы, обуславливающие большие погрешности результатов моделирования.

Ключевые слова: имитационное моделирование, высоконагруженная система, генераторы псевдослучайных величин, погрешность моделирования, GPSS World.

Введение

Одной из наиболее апробированных и широко распространенных систем имитационного моделирования (СИМ) является GPSS (General Purpose Simulation System) и, в частности, реализованная в среде Microsoft Windows версия GPSS World [1], в которой используются программные генераторы равномерно распределенных в интервале (0;1) псевдослучайных величин. Последовательность формируемых в процессе моделирования значений псевдослучайных величин зависит от задаваемого в качестве параметра целочисленного начального значения, что позволяет говорить о множестве используемых в GPSS World генераторов случайных величин, которые будем обозначать как 1, 2, 3, Кроме того, в GPSS World имеется более 20 библиотечных процедур для формирования случайных величин с заданными законами распределений, базирующихся на встроенных генераторах равномерно распределенных величин.

Полувекковая история существования СИМ GPSS и ее широкое применение в различных областях сформировали мнение, что при достаточно большой и обоснованно выбранной длительности моделирования погрешности результатов имитационного моделирования обычно не превышают 3–5%. Однако при моделировании высоконагруженных систем массового обслуживания (СМО), загрузка которых $\rho \geq 0,8$, было выявлено, что при некоторых сочетаниях генераторов случайных величин, задаваемых в виде $\{N1; N2\}$ (здесь $N1$ и $N2$ – номера генераторов, используемых для формирования значений интервалов между поступающими в систему заявками и длительностей обслуживания заявок соответственно), погрешность результатов моделирования может составлять 30% и более. При этом для других сочетаний генераторов эта погрешность при той же длительности моделирования, задаваемой числом прошедших через систему заявок, оказывается менее 3%.

Для выяснения обстоятельств, связанных с появлением такого эффекта, было проведено исследование генераторов псевдослучайных величин с целью оценки качества используемых в СИМ GPSS World генераторов и выявления причин, обуславливающих большую погрешность результатов моделирования при использовании некоторых сочетаний. Исследование проводилось в несколько этапов.

На первом этапе были выявлены так называемые проблемные сочетания генераторов, при использовании которых погрешность результатов моделирования оказывается значительной.

Для объяснения возникающей погрешности были протестированы проблемные встроенные генераторы равномерно распределенных псевдослучайных величин и библиотечные генераторы на соответствие декларируемым законам распределения [2].

Поскольку большая погрешность моделирования возникает только при использовании некоторых конкретных сочетаний генераторов, необходимо было выполнить корреляционный анализ соответствующих последовательностей случайных значений, рассчитать коэффициенты корреляции и сравнить эти коэффициенты для проблемных и непоблемных сочетаний генераторов случайных величин.

На последнем этапе была проанализирована зависимость результатов моделирования, а именно, среднего времени ожидания заявок, от длительности моделирования, задаваемого количеством заявок, прошедших через модель.

Описание экспериментов

Для исследования погрешностей результатов имитационного моделирования и выявления проблемных сочетаний генераторов использовались GPSS-модели одно- и многоканальных СМО типа M/M/1, M/G/1 и M/M/2, поддающиеся точному аналитическому расчету [2]. Варьируя загрузку систем от 0,1 до 0,9 и пропуская через каждую модель по 1 млн. заявок, было выявлено, что погрешность результа-

тов моделирования лежит в приемлемых пределах (1–5%) для значений загрузки примерно до 0,7. Начиная со значения $\rho \geq 0,8$, при некоторых сочетаниях генераторов случайных величин погрешность начинала расти и составляла 10% и более, а при $\rho=0,95$ достигала 40%. При исследовании реальных систем (маршрутизаторов, коммутаторов, серверов в компьютерных сетях) наибольший интерес обычно представляют сведения о качестве функционирования системы при большой загрузке, что делает задачу обеспечения приемлемой точности в процессе имитационного моделирования весьма актуальной. В связи с этим все последующие эксперименты выполнялись для систем, загрузка которых составляла 0,9 и 0,95. Для указанных значений загрузки системы были выполнены более 10 тыс. экспериментов при различных сочетаниях генераторов с номерами от 1 до 100 и выявлены сочетания генераторов, при которых достигается наибольшая погрешность результатов моделирования. В процессе экспериментов было установлено, что погрешности моделирования зависят только от загрузки системы $\rho=\lambda b$ и длительности моделирования, задаваемого количеством заявок, прошедших через систему, и не зависят от значений параметров загрузки – интенсивности поступления заявок λ и средней длительности обслуживания заявок b . Кроме того, было установлено, что использование проблемных сочетаний генераторов случайных величин приводит к значительным погрешностям результатов моделирования для других, более сложных моделей, в частности, для многоканальных СМО и двухузловых сетей массового обслуживания (СeМО), при других законах распределений, использующих эти генераторы.

Проблемные сочетания генераторов

На модели М/М/1 с однородным потоком заявок и загрузкой $\rho=0,9$ было проведено 10 000 экспериментов с целью выявления всех проблемных сочетаний генераторов с номерами от 1 до 100. В каждом эксперименте через модель пропускался 1 млн. заявок. Относительная погрешность моделирования рассчитывалась для среднего времени ожидания заявок как $\delta w = \frac{w - w^*}{w^*} \cdot 100\%$, где w – значение, полученное в процессе имитационного эксперимента; w^* – точное значение, рассчитанное для СМО М/М/1. Для большинства сочетаний погрешность не превышала 5%. В то же время, как видно из табл. 1, для 40 сочетаний погрешность превысила 20% и для 128 сочетаний – 10%.

Погрешность	Проблемные сочетания генераторов										
10–20%	{1;2}	{3;6}	{3;9}	{4;6}	{4;12}	{5;15}	{6;3}	{6;18}	{7;14}		
	{7;21}	{8;16}	{8;32}	{8;48}	{9;3}	{10;30}	{11;77}	{12;4}	{12;6}		
	{13;26}	{13;65}	{14;7}	{14;42}	{15;3}	{15;5}	{15;10}	{15;45}	{16;8}		
	{16;32}	{16;48}	{16;64}	{17;51}	{17;68}	{18;6}	{19;57}	{20;5}	{20;30}		
	{20;60}	{20;100}	{21;7}	{21;14}	{21;63}	{22;44}	{22;66}	{24;48}	{24;72}		
	{24;96}	{25;50}	{25;75}	{26;13}	{27;9}	{27;54}	{28;56}	{29;87}	{30;6}		
	{30;10}	{30;20}	{30;60}	{32;8}	{32;48}	{33;22}	{33;66}	{33;99}	{36;12}		
	{38;19}	{38;76}	{39;78}	{40;8}	{40;10}	{40;80}	{41;82}	{42;6}	{42;14}		
	{42;28}	{44;22}	{44;88}	{46;69}	{46;92}	{47;94}	{48;8}	{48;16}	{48;24}		
	{48;96}	{50;25}	{51;17}	{60;15}	{60;20}	{60;30}	{63;9}	{63;21}	{64;16}		
	{65;13}	{66;22}	{68;17}	{69;46}	{70;35}	{72;24}	{75;15}	{76;38}	{80;40}		
	{82;41}	{84;12}	{84;56}	{87;29}	{88;44}	{92;46}	{93;31}	{94;47}	{96;24}		
	{96;48}	{99;33}	{100;20}	{100;50}							
	>20%	{2;1}	{4;8}	{5;10}	{8;4}	{9;18}	{10;5}	{10;20}	{12;24}	{14;28}	
		{15;30}	{18;9}	{20;10}	{20;40}	{21;42}	{24;12}	{26;52}	{28;14}	{30;15}	
		{31;62}	{32;16}	{32;64}	{36;72}	{37;74}	{40;20}	{42;21}	{43;86}	{45;90}	
{49;98}		{52;26}	{54;27}	{57;19}	{62;31}	{64;32}	{66;33}	{72;36}	{74;37}		
{78;39}		{86;43}	{90;45}	{98;49}							

Таблица 1. Проблемные сочетания генераторов

Анализ представленных результатов показывает, что наиболее проблемными являются сочетания генераторов, номера которых соотносятся как 1:2, 1:3, 2:1 или 3:1.

Большинство проблемных сочетаний генераторов были проверены также для СМО общего вида G/G/1, в частности, U/U/1 и многоканальных СМО. Для сравнения в табл. 2 представлены результаты имитационного моделирования при использовании различных сочетаний генераторов в СМО М/М/1, загрузка которой $\rho=0,95$, а длительность обслуживания $b=9,5$ единиц времени. В этом случае точное значение среднего времени ожидания $w=180,5$ единиц.

Генераторы	Время ожидания, ед.	Погрешность, %
{2; 1}	114,4	-36,6
{2; 2}	173,6	-3,8
{2; 4}	155,8	-13,7
{20; 40}	111,8	-38,1
{40; 20}	113,2	-37,3
{110; 40}	179,9	-0,3
{200; 400}	136,5	-24,4
{999; 40}	177,1	-1,9
{2000; 4000}	119,3	-33,9

Таблица 2. Погрешности моделирования СМО М/М/1

Как видно из табл. 2, для проблемных сочетаний генераторов случайных величин погрешность результатов моделирования достигала 40%, в то время как у непроблемных сочетаний {2; 2}, {110; 40} и {999; 40} при той же длительности моделирования погрешность не превышала 5%. Более того, использование одного и того же генератора с номером 2 {2; 2} для формирования значений интервалов между поступающими заявками и длительности обслуживания позволяет получить существенно более точный результат, чем использование этого же генератора в сочетании с другим генератором. Перестановка генераторов в сочетании с {10; 20} на {20; 10} незначительно изменяет погрешность моделирования.

Аналогичные эксперименты были выполнены для двухузловой разомкнутой экспоненциальной СеМО, в которой заявки последовательно проходят через узлы, нагрузка которых $\rho_1=\rho_2=0,95$. Расчетные значения среднего времени ожидания составили в узлах $w_1=w_2=180,5$ и в сети $W=w_1+w_2=361$ единиц. В табл. 3 представлены результаты имитационного моделирования при различных сочетаниях генераторов, задаваемых в виде {N1; N2; N3}, где N1, N2 и N3 – номера генераторов, используемых для формирования значений интервалов между поступающими в сеть заявками и длительностей обслуживания заявок в узлах 1 и 2 соответственно.

Сочетание генераторов	Узел 1		Узел 2		Сеть	
	w_1 , ед.	δw_1 , %	w_2 , ед.	δw_2 , %	W , ед.	δW , %
{10; 20; 40}	121,1	-32,9	134,7	-25,39	255,72	-29,16
{20; 10; 40}	106,6	-41,0	122,0	-32,39	228,61	-36,67
{40; 20; 10}	113,2	-37,3	128,0	-29,11	241,13	-33,20
{21; 15; 719}	169,0	-6,4	170,7	-5,43	339,68	-5,91
{20; 10; 20}	172,7	-4,3	185,4	2,69	358,04	-0,82
{14; 28; 56}	110,4	-38,9	157,2	-12,90	267,58	-25,88
{28; 56; 14}	138,4	-23,3	134,2	-25,65	272,62	-24,48
{140; 280; 560}	124,4	-31,1	133,5	-26,06	257,88	-28,57
{280; 560; 140}	116,1	-35,7	128,7	-28,71	244,80	-32,19

Таблица 3. Погрешности моделирования двухузловой СеМО

Как видно из табл. 3, относительная погрешность (разброс) результатов моделирования для проблемных сочетаний генераторов составляет более 25% как для узловых, так и для сетевых характеристик. Большой разброс наблюдается также для такой важной характеристики, как максимальная длина очереди заявок. Так, например, максимальная длина очереди заявок в узлах 1 и 2 для сочетания {20; 10; 20} составляет 153 и 186 соответственно, а для проблемного сочетания {20; 10; 40} – 79 и 95, т.е. имеем почти двукратное отличие.

Тестирование генераторов

Генератор равномерно распределенных случайных величин вырабатывает базовую последовательность значений в интервале (0; 1), которая используется для формирования случайных величин, распределенных по другим законам и в других интервалах. Поскольку погрешность может появиться как на этапе генерации базовой последовательности, так и на этапе формирования экспоненциально распределенной величины, было проведено тестирование базовых последовательностей и последовательностей экспоненциально распределенных случайных величин.

Относительная погрешность математического ожидания случайных величин для генераторов, входящих в состав проблемных сочетаний, в большинстве случаев составляла более 0,1%, в то время как

для других генераторов – менее 0,1%. Этот результат справедлив как для базовых последовательностей, так и для последовательностей экспоненциально распределенных случайных величин.

Для базовой последовательности относительная погрешность среднеквадратического отклонения для всех исследованных генераторов, включая генераторы, входящие в состав проблемных сочетаний, составила менее 0,01%. Для последовательностей экспоненциально распределенных случайных величин эта погрешность оказалась больше 0,01%, а для генераторов, входящих в состав проблемных сочетаний, достигала значений 0,2% и более.

Результаты корреляционного анализа

Корреляционный анализ базовых последовательностей и последовательностей, экспоненциально распределенных случайных величин на выборке по 1 млн. значений каждого генератора, показал, что коэффициенты корреляций проблемных сочетаний генераторов случайных величин в 5–10 раз превышают значения коэффициентов корреляции непроблемных сочетаний, причем они принимают положительные значения, т.е. большему (меньшему) значению первого генератора соответствуют большие (меньшие) значения второго генератора. Это означает, что с большой вероятностью заявке, поступившей в систему с большим интервалом (сформированном первым генератором) от момента поступления предыдущей заявки, будет соответствовать (сформировано вторым генератором) большее время обслуживания, и, наоборот, заявке, поступившей в систему с небольшим интервалом от момента поступления предыдущей заявки, будет соответствовать маленькое время обслуживания. Вероятность же ситуации, когда заявке, поступившей в систему с маленьким интервалом, будет соответствовать большое время обслуживания, оказывается незначительной. Все это приводит к тому, что результаты имитационного моделирования, в частности, среднее время ожидания, оказываются существенно заниженными даже при достаточно длительном моделировании.

Анализ переходных процессов в модели

Для проблемных сочетаний генераторов с целью выяснения характера изменения среднего времени ожидания от длительности имитационного моделирования, задаваемого числом заявок, проходящих через моделируемую систему, были выполнены эксперименты, при которых через систему пропусклось до 100 млн. заявок, а в некоторых моделях до 200 млн. заявок. Изучение переходного процесса в СМО позволило установить, что даже при таком большом количестве заявок, пропущенных через систему, среднее время ожидания не достигало расчетного значения, а погрешность могла составлять до 10%. На рис. 1 представлены зависимости среднего времени ожидания заявок и относительной погрешности результатов имитационного моделирования от числа заявок, прошедших через имитационную модель СМО М/М/1 с загрузкой 0,95, для сочетания генераторов {20;10}.

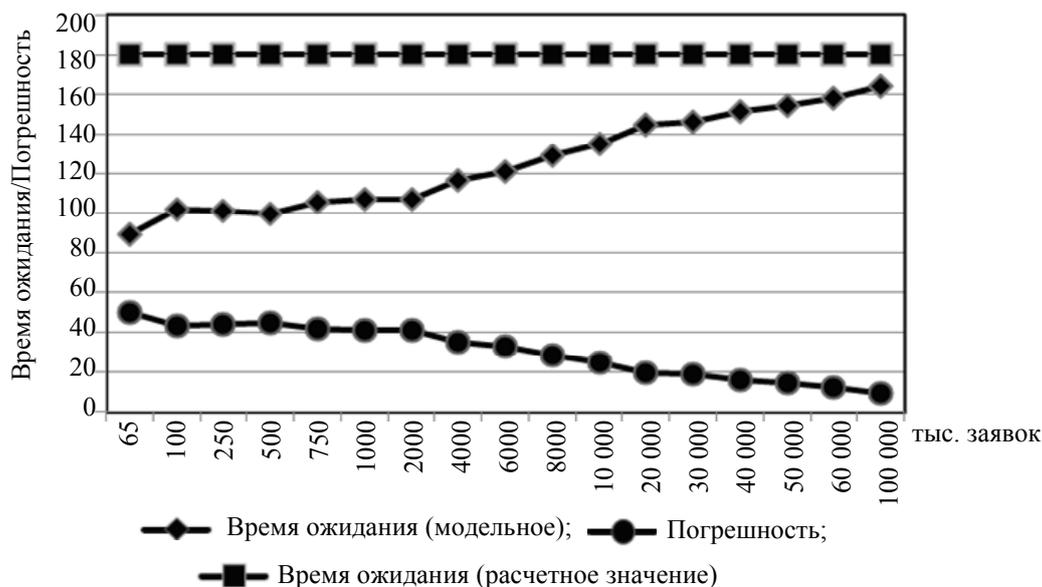


Рис. 1. Проблемное сочетание генераторов {20; 10}

Из графика видно, что даже после прохождения через модель 100 млн. заявок относительная погрешность составляет около 10%. Отметим, что коэффициент корреляции генераторов 20 и 10 положителен и равен 0,00247.

Заметим также, что использование широко применяемого на практике метода «срезов» для определения вхождения системы в стационарный режим может привести к ошибочным результатам. Например, как видно из графика, среднее время ожидания при пропускании 100 тыс., 250 тыс. и 500 тыс. заявок остается практически неизменным (отличие менее чем на 1%), откуда можно сделать вывод, что система вошла в стационарный (установившийся) режим и, следовательно, процесс моделирования может быть завершен. При этом погрешность результата моделирования оказывается более 40%.

Для сравнения на рис. 2 показана аналогичная зависимость для той же модели при использовании беспроблемного сочетания генераторов {110; 40}, откуда видно, что установившийся режим в системе существует уже после пропускания 500 тыс. заявок, и погрешность результатов моделирования в дальнейшем не превышает 4%. Следует обратить внимание, что такой характер зависимости модельного времени ожидания по отношению к расчетному может быть объяснен тем, что коэффициент корреляции генераторов 110 и 40 – отрицательный и равен $-0,00176$.



Рис. 2. Непроблемное сочетание генераторов {110;40}

Для двухузловой разомкнутой СеМО на рис. 3 представлены зависимости относительных погрешностей среднего времени ожидания заявок в узлах и в сети в целом от количества заявок, прошедших через имитационную модель при использовании проблемного сочетания генераторов {20; 10; 40}. Из графика видно, что даже после пропускания через модель 200 млн заявок относительная погрешность результатов моделирования составляет около 10%.

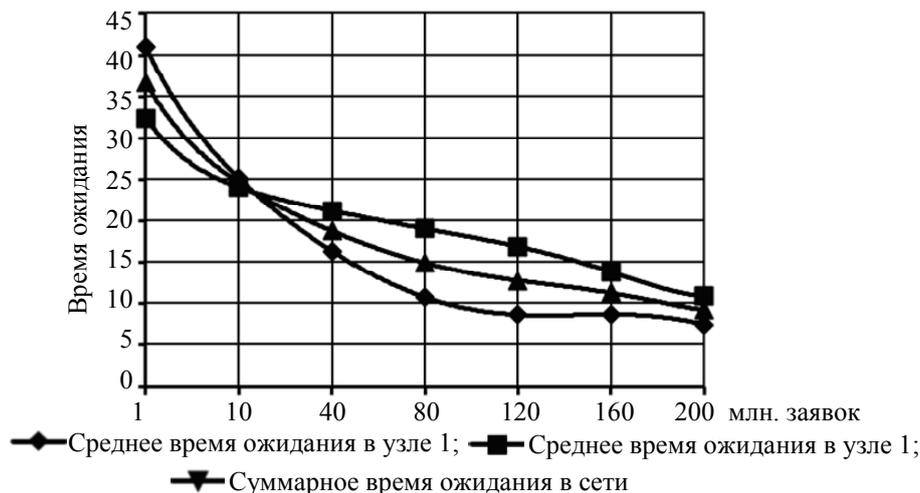


Рис. 3. Относительные погрешности для сочетания генераторов {20; 10; 40}

Заключение

Погрешность результатов моделирования высоконагруженных систем ($\rho \geq 0,8$) в системе имитационного моделирования на GPSS World при некоторых сочетаниях генераторов случайных величин, используемых для формирования значений интервалов между поступающими в систему заявками и длительностей обслуживания заявок, может оказаться значительной и достигать десятков процентов.

Результаты многочисленных экспериментов позволили выявить, что наибольшая погрешность достигается для сочетаний генераторов, номера которых соотносятся как 1:2, 1:3, 2:1 и 3:1.

Выполненный корреляционный анализ показал, что коэффициенты корреляций проблемных сочетаний генераторов принимают положительные значения и по абсолютной величине в 5–10 раз превышают значения коэффициентов корреляции непроблемных сочетаний, что обуславливает значительные погрешности результатов моделирования высоконагруженных систем.

Анализ переходных процессов в моделях высоконагруженных систем при использовании проблемных сочетаний генераторов случайных величин показывает, что среднее время ожидания заявок медленно увеличивается с увеличением длительности моделирования (числа заявок, пропускаемых через модель), но так и не достигает точного значения, полученного аналитически, в некоторых случаях даже после прохождения через моделируемую систему более 100 млн. заявок.

Литература

1. Бражник А.Н. Имитационное моделирование: возможности GPSS World. – СПб: Реноме, 2006. – 439 с.
2. Алиев Т.И. Основы моделирования дискретных систем. – СПб: СПбГУ ИТМО, 2009. – 363 с.

Алиев Тауфик Измаилович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, профессор, зав. кафедрой, aliev@d1.ifmo.ru