НЕКОТОРЫЕ БАЙЕСОВСКИЕ МОДЕЛИ В ТЕОРИИ МАССОВОГО ОБСЛУЖИВАНИЯ И ТЕОРИИ НАДЕЖНОСТИ[8]

Чиро д'Апиче
Отдел информатики и прикладной математики Университета Салерно, Италия

Росанна Манцо
Отдел информатики и прикладной математики Университета Салерно, Италия

Сергей Шоргин
Институт проблем информатики Российской академии наук, Москва, Россия

исследования

Аннотация

Вводится байесовский подход для определнных задач теории массового обслуживания и теории надежности. Соответствующий метод предусматривает рандомизацию характеристик систем относительно некоторых априорных распределений. Данный подход может использоваться, в частности, для вычисления средних значений и построения доверительных интервалов для вероятностно-временных и надежностных характеристик больших групп систем или устройств.

1. Введение и основные предположения. Теория массового обслуживания является весьма развитой математической дисциплиной. В рамках этой теории получено огромное количество глубоких с точки зрения математики и важных с прикладной точки зрения результатов, относящихся к исследованию систем и сетей массового обслуживания, представляющих собой модели широкого класса реальных систем, прежде всего информационно-телекоммуникационных систем и сетей. В настоящее время развитие теории массового обслуживания ведется в основном в направлении рассмотрения все более сложных по вероятностным характеристикам входящих потоков и распределений времени обслуживания, все более сложных дисциплин обслуживания, в интересах более адекватного отражения реальных процессов.

Одним из направлений обобщения и усложнения постановок является усложнение вероятностной структуры тех или иных входных параметров СМО; в частности, рассмотрение вместо традиционных входящих потоков – потоков Кокса, самоподобных потоков, марковских и полумарковских потоков и т.п.; аналогичные обобщения осуществляются и по отношению к распределениям времен обслуживания. В определенной степени эти обобщения могут интерпретироваться как результат рандомизации тех или параметров более «простых» потоков и распределений обслуживания. Так, процесс Кокса получается в результате специальной рандомизации интенсивности пуассоновского потока, и т.п.

Все эти обобщенные современные постановки предполагают, что стохастический механизм рандомизации «влияет» на параметры системы непосредственно в период ее

[8] Работа поддержана РФФИ, проект 05-07-90103

функционирования, то есть мы изначально знаем, с какой системой «имеем дело», пусть даже эта система достаточно сложна, и исследуем характеристики функционирования именно этой «изначально фиксированной» системы. Однако в реальной практике часты ситуации, при которых сама исследуемая система задана в определенном смысле «неточно»; скажем, если даже говорить о простейших системах типа M|G|1, исследователю могут быть априори неизвестны параметр входящего потока λ и параметры обслуживания μ и σ^2. Такие ситуации возникают, скажем, в случае, когда рассматривается целый класс однотипных СМО, относительно которых известны только типы входящего потока и распределения обслуживания, а также дисциплина обслуживания, но конкретные параметры этих потоков и распределений, вообще говоря, различны для различных СМО данного класса. Исследователь априори не знает, с какой СМО из данного класса он имеет дело (это может иметь место, например, при исследовании серии однотипных устройств коммутации или передачи, выпускаемых одним и тем же производителем, для которых разброс значений тех или иных показателей обуславливается естественными технологическими вариациями при производстве; возможны и другие примеры). В этом случае, поскольку неизвестными являются именно «исходные» параметры потоков и времени обслуживания, естественным является рандомизационный подход, при котором элементами вероятностного пространства становятся (если рассматривать приведенный выше пример) значения λ, μ и σ^2 (а в общем случае можно говорить о вероятностном пространстве, элементами которого являются сами однотипные СМО). При этом подлежащие вычислению характеристики такой «рандомизированной» СМО, естественно, являются рандомизацией аналогичных характеристик «обычной» СМО аналогичного типа – с учетом того априорного распределения входных параметров СМО, которое взято исследователем за основу.

Таким образом, в том же примере с СМО типа M|G|1 возникают задачи рандомизации «обычных» характеристик таких систем с учетом априорных распределений входных параметров; скажем, может приниматься предположение о показательном, равномерном или каком-то другом распределении одной или нескольких из величин λ, μ и σ^2 (которые при таком подходе становятся случайными величинами), об их независимости или зависимости, и т.п. Полученные результаты могут применяться, например, для вычисления средних значения, построения доверительных интервалов для тех или иных характеристик рассматриваемого класса СМО «в целом». Такой подход к построению моделей массового обслуживания естественно назвать байесовским. Впервые этот подход сформулирован в [1].

Другим направлением применения байесовского подхода является оценка надежности. Как известно [2], коэффициент готовности восстанавливаемого устройства в стационарном режиме может быть вычислен по формуле

$$k = \frac{\lambda^{-1}}{\lambda^{-1} + \mu^{-1}} = \frac{\mu}{\lambda + \mu},$$

где λ^{-1} – среднее время безотказной работы, μ^{-1} – среднее время восстановления.

Если мы примем сформулированное выше предположение, в соответствии с которым любое изучаемое устройство выбирается случайным образом из некоторого множества сходных устройств, различающихся средними величинами показателей надежности, то, в соответствии с приведенными выше рассуждениями, значения λ и μ могут рассматриваться в качестве случайных. Следовательно, при таких предположениях коэффициент готовности k также является случайной величиной, и его распределение зависит от распределений величин λ, μ. Результаты, получаемые в рамках этой постановки, могут использоваться, в частности, для вычисления средних значений и построения доверительных интервалов для надежностных характеристик всей изучаемой группы устройств.

2. Вводный пример. Для того, чтобы лучше объяснить суть постановки, в качестве вводного пример рассмотрим ситуацию, когда некий наблюдатель (исследователь) имеет дело с достаточно большой серией систем обслуживания M|M|1|0, различающихся только параметром распределения обслуживания. В частности, это могут быть некие станки, коммутаторы, маршрутизаторы или другие обслуживающие средства, о которых заранее известно, что их функционирование описывается системой вышеуказанного типа. То есть эти системы идентичны по дисциплине обслуживания и по типам входящего потока и распределения времени обслуживания. В данном примере также и характеристики входящего потока предполагаются идентичными для всех систем данной серии; различаются только численные характеристики **обслуживания** (то есть параметры показательного распределения).

Разброс характеристик обслуживания обуславливается технологическими (конструктивными) причинами. И главным аспектом рассматриваемой постановки является то, что исследователю не известно, каково конкретное значение параметра обслуживания рассматриваемой им «взятой наугад» системы из данной серии. Известно лишь «априорное» распределение этого параметра (поскольку серия систем мыслится большой, то можно в рамках этой серии рассматривать стохастические явления и вводить вероятностные распределения). Исследователя интересуют характеристики обслуживания для серии в целом (или характеристики «наугад взятой» системы). Очевидно, что, помимо традиционных факторов стохастичности, имеющих место в СМО (стохастичность поступающего потока и процессов обслуживания), появляется еще один фактор, связанный со **случайностью выбора рассматриваемой системы.**

Пусть, скажем, параметр обслуживания μ у рассматриваемых систем может принимать только два значения: μ_1 и μ_2 с вероятностями соответственно p_1 и p_2. «Физически» это означает, что среди рассматриваемой серии систем (маршрутизаторов, станков и т.п.) встречаются только две «разновидности» обслуживающих приборов: приборы первой разновидности осуществляют обслуживание с параметром μ_1, приборы второй разновидности - с параметром μ_2. Тогда у «взятой наугад» системы коэффициент загрузки становится случайной величиной, принимающей значения λ/μ_1 с вероятностью p_1 и λ/μ_2 с вероятностью p_2; стационарная вероятность блокировки «выбранной» системы в связи с вмешательством случайного фактора выбора конкретной системы сама становится «случайной» и принимает значения $\lambda/(\lambda+\mu_1)$ с вероятностью p_1 (это вероятность того, что исследователю «попалась» система первой «разновидности») и $\lambda/(\lambda+\mu_2)$ с вероятностью p_2 («попалась» система второй «разновидности»). Естественно, «усредненная» вероятность блокировки в такой «байесовской» СМО равна $p_1\lambda/(\lambda+\mu_1)+p_2\lambda/(\lambda+\mu_2)$.

Как видим, для исследования байесовских СМО нет необходимости проводить анализ системы методами теории массового обслуживания. Байесовская система является «рандомизацией» некоторой «обычной» СМО и, соответственно, характеристики байесовской СМО вычисляются путем рандомизации последующего усреднения (по априорному распределению параметра или параметров) уже вычисленных ранее методами теории массового обслуживания характеристик «обычной» СМО. То есть математическая часть работы сводится именно к этой рандомизации и усреднению. При этом как с технической, так и с концептуальной точек зрения целесообразно осуществлять рандомизацию **стационарных** характеристик «обычных» СМО, получая стационарные характеристики байесовских СМО.

Отметим ещё одну содержательную модель, математическим описанием которых может служить байесовская СМО. Предположим, что исследователь рассматривает не серию систем, а некоторую одну систему, количественные параметры которой меняются со временем. Например, имеется обслуживающее устройство, один из элементов которого в заранее не известные моменты заменяется на другой, потом на третий, и т.п. Скажем, такой системой является пограничный пост в аэропорту, пограничник на котором иногда сменяется – в моменты времени, не известные наблюдателю (пассажиру); наблюдатель знает лишь вероятность того, что он «наткнётся» на некоторого конкретного пограничника и среднее время проверки паспорта каждым из возможных пограничников.

При таком подходе структура системы и дисциплина обслуживания не изменяется со временем, а изменяется только количественный параметр распределения обслуживания (например, интенсивность). Аналогичным образом может меняться и параметр поступающего потока. О том, когда происходят изменения, информации нет. Исследователю известно только **распределение** значений «изменчивых», случайных параметров, с которыми он «сталкивается», рассматривая систему в «случайный» момент времени.

Поскольку предполагается, что информации о моментах «перестройки» системы или хотя бы о распределении этих моментов у исследователя нет, описание переходных процессов в системе такого рода невозможно. Следовательно, возможен анализ (и последующая рандомизация) только **стационарных** распределений рассматриваемой СМО. Для того, чтобы такая постановка имела смысл, необходимо ввести предположение, что система изменяется достаточно «редко», - так, чтобы на каждом интервале постоянства параметров СМО «успевала» достичь стационарного состояния. Естественно, результаты такого анализа будут приближенными, поскольку в реальной жизни стационарное состояние, строго говоря, не достигается.

3. Простые модели «байесовских» систем

Ниже в настоящей статье рассмотрены еще две простейшие модели «байесовских» СМО, предназначенные для дальнейшего разъяснения специфики задач, возникающих при этом подходе, и получаемых при этом результатов.

Равномерное распределение λ и μ: коэффициент загрузки. Рассмотрим произвольную систему массового обслуживания с интенсивностью входящего потока λ и интенсивностью обслуживания μ. Загрузка такой системы $\rho = \lambda / \mu$. Как известно, от значения ρ зависит наличие стационарного режима у рассматриваемой системы; величина ρ входит во многие формулы, описывающие характеристики разнообразных СМО. В связи с этим рассмотрение величины ρ представляется одной из первоочередных задач, которые следует рассмотреть в рамках байесовской теории СМО.

Множество возможных и интересных для приложений совместных распределений величин λ и μ весьма обширно. Мы рассмотрим один из простейших, и в то же время достаточно распространённых на практике случаев, когда величины λ и μ независимы и распределены равномерно на некоторых заданных отрезках. Такая модель приемлема для описания ситуации, когда для обоих значений λ и μ (или для любого из них) задан допустимый интервал значений, но реальная величина λ и/или μ может варьироваться в этих пределах.

Пусть носителем случайной величины λ является некоторый отрезок $[a_\lambda, b_\lambda]$, носителем случайной величины μ - некоторый отрезок $[a_\mu, b_\mu]$, причем $0 \le a_\lambda \le b_\lambda$, $0 \le a_\mu \le b_\mu$.

При этом функция распределения случайной величины $\rho = \lambda / \mu$ может быть записана таким образом:

$$P\{\rho < x\} = \iint\limits_{\lambda/\mu < x} \frac{1}{b_\lambda - a_\lambda} \frac{1}{b_\mu - a_\mu} d\lambda d\mu$$

Дальнейшие выкладки существенно зависят от соотношения между величинами a_λ / a_μ и b_λ / b_μ. Предположим для определенности, что $a_\lambda / a_\mu \le b_\lambda / b_\mu$. Тогда:

при $x \le a_\lambda / b_\mu$

$$P\{\rho < x\} = 0,$$

при $a_\lambda / b_\mu \le x \le a_\lambda / a_\mu$

$$P\{\rho < x\} = K \frac{(b_\mu x - a_\lambda)^2}{2x},$$

при $a_\lambda / a_\mu \le x \le b_\lambda / b_\mu$

$$P\{\rho < x\} = K\left(\frac{a_\mu + b_\mu}{2} x - a_\lambda\right)(b_\mu - a_\mu),$$

при $b_\lambda / b_\mu \le x \le b_\lambda / a_\mu$

$$P\{\rho < x\} = 1 - K \frac{(b_\lambda - a_\mu x)^2}{2x},$$

при $x \ge b_\lambda / a_\mu$

$$P\{\rho < x\} = 1,$$

где

$$K = \frac{1}{(b_\mu - a_\mu)(b_\lambda - a_\lambda)}.$$

Выпишем плотность случайной величины ρ:

при $x \le a_\lambda / b_\mu$

$$f_\rho(x) = 0,$$

при $a_\lambda / b_\mu \le x \le a_\lambda / a_\mu$

$$f_\rho(x) = K\left(\frac{b_\mu^2}{2} - \frac{a_\lambda^2}{2x^2}\right),$$

при $a_\lambda / a_\mu \le x \le b_\lambda / b_\mu$

$$f_\rho(x) = K\left(\frac{b_\mu^2 - a_\mu^2}{2}\right),$$

при $b_\lambda / b_\mu \le x \le b_\lambda / a_\mu$

$$f_\rho(x) = K\left(\frac{b_\lambda^2}{2x^2} - \frac{a_\mu^2}{2}\right),$$

при $x \geq b_\lambda / a_\mu$

$$f_\rho(x) = 0.$$

Проведя элементарные выкладки, получаем, что среднее значение и второй момент случайной величины ρ, соответственно, равны:

$$\mathbf{E}\,\rho = \frac{b_\lambda + a_\lambda}{2(b_\mu - a_\mu)} \ln \frac{b_\mu}{a_\mu},$$

$$\mathbf{E}\,\rho^2 = \frac{a_\lambda^2 + a_\lambda b_\lambda + b_\lambda^2}{3 a_\mu b_\mu}.$$

Очевидно, что при $b_\lambda - a_\lambda \to 0$ и при $b_\mu - a_\mu \to 0$, то есть при стягивании носителя случайной величины λ к некоторой фиксированной точке λ_0 и при стягивании носителя случайной величины μ к некоторой фиксированной точке μ_0, значение $\mathbf{E}\,\rho$, как это и должно быть, стремится к λ_0 / μ_0, а значение $\mathbf{E}\,\rho^2$ к λ_0^2 / μ_0^2.

Заметим также, что зависимость среднего значения величины ρ от распределения λ сводится к зависимости от математического ожидания λ. В то же время зависимость $\mathbf{E}\,\rho$ от параметров распределения μ носит более сложный вид.

В случае $a_\lambda / a_\mu \geq b_\lambda / b_\mu$ формулы для функции распределения и плотности случайной величины ρ аналогичны; здесь мы их опускаем для сокращения объема изложения. Математическое ожидание и второй момент случайной величины ρ в указанном случае совпадают с выписанными выше значениями.

На основании полученных результатов нетрудно вычислить другие необходимые характеристики величины ρ.

Заметим, что рассмотренная модель позволяет изучать важную ситуацию, в которой $\lambda < \mu$ с вероятностью 1. При этом $\rho < 1$, что является условием эргодичности систем с одним обслуживающим прибором. В силу постулируемой независимости случайных величин λ и μ обеспечить выполнение условия $\lambda < \mu$ может только соответствующее взаимное расположение отрезков $[a_\lambda, b_\lambda]$ и $[a_\mu, b_\mu]$, то есть выполнение условия $0 \leq a_\lambda \leq b_\lambda \leq a_\mu \leq b_\mu$.

Показательное распределение λ и μ: коэффициент загрузки и вероятность потерь в системе M|M|1|0; коэффициент готовности

Рассмотрим другую вероятностную модель для величин λ и μ. В ситуации, когда никакой априорной информации о значениях λ и μ нет, за исключением данных об их средних значениях, в качестве «первого приближения» целесообразно рассмотреть модель, в которой и λ, и μ распределены по показательному закону с известными средними (пусть они равны соответственно *1/l* и *1/m*). Предположение о независимости λ и μ сохраняется.

Итак, функция распределения случайной величины λ равна $1-\exp(-lu)$, функция распределения случайной величины μ равна $1-\exp(-mu)$. Как и в предыдущем разделе, прежде всего рассмотрим $\rho = \lambda/\mu$. Очевидно, что при $x \geq 0$

$$P\{\rho < x\} = P\{\lambda < \mu x\} = \int_0^\infty P\{\lambda < xy\} dP\{\mu < y\} = \int_0^\infty [1-\exp(-lxy)] m \exp(-my) dy =$$

$$= \frac{lx}{m+lx}$$

Отсюда, в частности, следует, что случайная величина $\rho = \lambda/\mu$ в данном случае, в отличие от ситуации, рассмотренной в предыдущем разделе, не имеет моментов первого и более высоких порядков. Однако некоторые другие характеристики байесовских систем массового обслуживания, зависящие от случайного $\rho = \lambda/\mu$, могут иметь конечные моменты. Рассмотрим, например, СМО M|M|1|0. Вероятность того, что поступивший в эту систему вызов не будет потерян, в стационарном режиме равна, в соответствии с формулами Эрланга, $\pi = 1/(1+\rho)$. В «байесовской» постановке эта вероятность, как было отмечено выше, сама становится «случайной». Рассмотрим распределение случайной величины π в условиях рассматриваемой модели.

При $0 \leq y \leq 1$

$$P\{\pi < y\} = P\{\rho > (1-y)/y\} = \frac{my}{my + l(1-y)}$$

Соответственно, плотность распределения случайной величины π равна $\frac{ml}{[my + l(1-y)]^2}$, а усредненная вероятность непотери вызова

$$\mathbf{E}\pi = \int_0^1 \frac{mly}{[my + l(1-y)]^2} dy = \frac{ml}{(m-l)^2}\left(\ln\frac{m}{l} + \frac{l}{m} - 1\right).$$

Нетрудно вычислить также второй момент случайной величины π и другие ее характеристики. Заметим, что при $m=l$

$\mathbf{E}\pi = 1/2$.

Значение

$$\pi = 1/(1+\rho) = \frac{\mu}{\lambda+\mu}$$

совпадает с величиной коэффициента готовности k (см. выше). Следовательно, распределение случайного коэффициента готовности в случае величин λ и μ, имеющих показательное распределение, представлен выше в качестве распределения случайной величины π.

4. Выводы. Представленные в статье результаты являются самыми начальными, «пробными» в рамках проблематики байесовских СМО и байесовских моделей надежности. Очевидно, что дальнейшее продвижение в рамках данной проблематики требует рассмотрения тех априорных распределений величин λ, μ и других традиционных входных параметров для СМО и восстанавливаемых устройств, которые могут представлять интерес для практики, и вычисление соответствующих распределений показателей функционирования и надежности различных типов систем после их рандомизации с учетом указанных априорных распределений.

ЛИТЕРАТУРА

1. Шоргин С.Я. О байесовских моделях массового обслуживания. II Научная сессия Института проблем информатики РАН: Тезисы докладов. – М.:ИПИ РАН, 2005, с.120-121.
2. B.A. Kozlov, I.A. Ushakov, Reliability Handbook, Holt, Rinehart & Winston (1970).