Схеми за симулация на редки събития при метода на случайната значима извадка Importance Sampling

Екатерина Оцетова-Дудин, Светла Радева

Importance Sampling Simulation Schemes: *Rare event simulation for estimation of Quality of Service parameters at broadband convergence networks is under consideration. Different kinds of Importance Sampling simulation schemes are considered as basic schemes, schemes using topological information, and schemes trying to approach the zero-variance change of measure. The robustness properties of simulation schemes are discussed. Implementation of bounded relative error and bounded normal approximation are analyzed.*

Key words: *Rare Event Simulation, Importance Sampling, Queuing systems.*

ВЪВЕДЕНИЕ

От особенно значение за трафика в широколентовите безжични мрежи е гарантирането на качеството на обслужване на потребителите. Вероятностните параметри, които характеризират степента и качеството на обслужване в съвременните широколентови безжични мрежи се появяват с много малка вероятност, която е от порядъка на 10-8 [3, 5]. Поради тази причина те не могат да бъдат изследвани с помощта на стандартни методи на симулация, като метода Монте Карло, който е приложим ако вероятността за появата на определено събитие е не по-малка от 10-5 [1, 2]. Това налага прилагането на методи, които биха могли да ускорят стохастичната симулация.

При ускоряване на стохастичната симулация разглеждания модел се трансформира в нов модел, който е свързан с изходния модел. Новият модел дава възможност да бъдат получени по-точни оценки на параметрите, които ни интересуват – в случая на параметрите на качеството на обслужване [3]. Един методите за ускоряване на стохастичната симулация е симулацията на редки събития [2, 5].

СИМУЛАЦИЯ НА РЕДКИ СЪБИТИЯ

Съществуват два основни подхода, които се използват при симулацията на редки събития - метода на случайната значима извадка и метода на разклонение на траекторията на образците. При симулацията на редки събития по метода на случайната значима извадка се извършва промяна на вероятностния закон за разпределение, за да се увеличи честотата на поява на „значими” за симулацията събития. Основната цел на тази симулационна техника е да се намали дисперсията или друга оценъчна функция, която е получена в резултат на компютърна симулация. В хода на такава симулация получаването на изходните образци е пропорционално на тяхната относителна важност спрямо очаквания резултат.

Оценките при метода на случайната значима извадка могат да приемат предварително зададена точност, за да се съкрати симулационното време. За генериране на значимата извадка се използват краен брой от независими променливи с нормален закон на разпределение [2]. Тогава условната вероятност за появата на рядко събитие със зададена функция на разпределение се заменя с условната вероятност за поява на по-малко рядко събитие с подобно разпределение [4, 5].

Симулацията на редки събития по метода на разклонение на траекторията на образците при достигане на определени прагови величини дава възможност без да бъде променян закона за разпределение на случайните величини да бъде избрана такава траектория на симулационната пътечка, която с най-голяма вероятност ще доведе до появата на рядко събитие [5]. За целта при достигането на определени прагови величини перспективните симулационни пътечки се разклоняват и симулацията продължава, а по неперспективните симулационни пътечки симулацията се прекратява.

ВИДОВЕ СХЕМИ И ПРЕСМЯТАНЕ НА ОТНОСИТЕЛНАТА ГРЕШКА

Схемите за симулация при метода на случайната значима извадка Importance Sampling се подразделят на три категории:

- Основни схеми;
- Схеми, които се базират на използване на топологична информация;
- Схеми, при които директно се опитва да се достигне нулева вариация при смяна на мярката.

Освен избора на симулационна схема е необходимо да бъде избиран и начина за пресмятане на относителната грешка при изчисленията. Относителната грешка може да бъде определена като се извършва [4] :

- Пресмятане на ограничената относителна грешка BRE (bounded relative error), при която относителната грешка остава ограничена при $\varepsilon \to 0$, така че относителната точност на доверителния интервал не е чувствителна спрямо редкостта на събитието;
- Пресмятане на ограничената нормална апроксимация BNA (bounded normal approximation), която представлява едно достатъчно условие за това, че обхвата на доверителния интервал остава валиден за редки събития независимо от степента на вероятността на тяхната поява.

Разглежда се прехвърлянето на сигнала (хендовъра) от една базова станция към друга при преместване на мобилната станция. Основната цел на хендовъра е да разпредели текущото повикване, когато силата на сигнала към базовата станция спадне под определен праг при преместване на мобилната станция. Нека F е множеството от състояния, при които всички канали на базовата станция са заети (или опашките на хендовъра са пълни) и текущото повикване не може да бъде обслужено и R е множеството от състояния, при които има свободни канали в базовата станция (места в опашките на хендовъра за очакване на обслужване. Ако x ≠ 0 може да се обозначи $F_x = \{y : (x, y) \in F\}$ и $R_x = \{y : (x, y) \in R\}$. Нека $f_x = \sum_{y \in F_x} \mathbf{P}(x, y)$ да бъдат вероятностите за това, че текущото повикване се губи, защото не може да бъде обслужено, защото всички канали на базовата станция са заети (или опашките на хендовъра са пълни) и $r_x = \sum_{y \in R_x} \mathbf{P}(x, y)$ вероятностите за това, че текущото повикване може да бъде обслужено, тъй като има свободни канали и/или места в съответните опашки на хендовъра.

ОСНОВНИ СХЕМИ

Първата основна схема е свързана с увеличаване на вероятностите за неуспешни попадения, в случая за това, че текущото повикване се губи, защото не може да бъде обслужено и се нарича FB (failure biasing). При нея се увеличава вероятността за загуба на повикване до една фиксирана стойност α ∈ (0, 1), като обикновено 0,5≤ α ≤ 0,9. Вероятностите за преход от едно състояние към друго се променят както следва:

$$\forall x \in U, x \neq \mathbf{0}, \quad (x, y) \in \mathrm{F}: \quad \widetilde{\mathrm{P}}(x, y) = \alpha \frac{\mathrm{P}(x, y)}{f_x} \tag{1}$$

$$\forall x \in U, x \neq \mathbf{0}, \quad (x, y) \in \mathrm{R}: \quad \widetilde{\mathrm{P}}(x, y) = (1 - \alpha) \frac{\mathrm{P}(x, y)}{r_x} \tag{2}$$

Следва да се отбележи, че общата вероятност от загуба на текущо повикване е равна на α. След получаване на загуба на повикване се извършва превключване обратно към P. Получава се една балансирана система, при която от всяко състояние x прехода към загуба има вероятност от същия порядък с отклонение в рамките на ε. Недостатък на този метод е, че някои пътечки продължават да остават твърде редки, тъй като един от нейните преходи към загуба за дадено състояние все още има вероятност о(1) в резултат на по-малко редките загуби при началния закон за разпределение, което не води до получаването на „интересни" състояния.

Втората основна схема представлява едно разширение на първата и се нарича балансирано увеличение на неуспешните попадения BFB (balanced failure biasing). При нея се разглежда подмножество от преходи към неуспешни попадения, за да се пресметнат условните индивидуални вероятности, които са взети пропорционално на първоначалната вероятност, пресметната в рамките на първата основна схема FB. Така първоначалната вероятност се замества с равномерно разпределени вероятности.

$$\forall x \in U, x \neq \mathbf{0}, \quad (x, y) \in \mathrm{F}: \quad \tilde{\mathrm{P}}(x, y) = \alpha \frac{1}{Card\,(F_x)} \tag{3}$$

$$\forall x \in U, x \neq \mathbf{0}, \quad (x, y) \in \mathrm{F}: \quad \tilde{\mathrm{P}}(x, y) = \alpha \frac{1}{Card\,(F_x)} \tag{4}$$

От **0** се използва следствието, че α = 1. Тази схема удовлетворява и двата начина за пресмятане на относителната грешка.

Третата основна схема инверсно (обратно) увеличение на неуспешните попадения IFB (inverse failure biasing) се основава на ефективната симулация на M/M/1 опашка, включваща превключване на заемане и обслужване на заявките. За нея е в сила:

$$\text{If } x = \mathbf{0}, \forall y: \quad \mathrm{P}(\mathbf{0}, y) > 0, \quad \tilde{\mathrm{P}}(\mathbf{0}, y) = \alpha \frac{1}{Card\,(F_{\mathbf{0}})} \tag{5}$$

$$\forall x \in U, x \neq \mathbf{0}, \quad \forall y: (x, y) \in \mathrm{F}: \quad \tilde{\mathrm{P}}(x, y) = \frac{r_x}{Card\,(F_x)} \tag{6}$$

$$\text{and if} \quad (x, y) \in \mathrm{R}: \quad \tilde{\mathrm{P}}(x, y) = \frac{f_x}{Card\,(R_x)} \tag{7}$$

Вероятността за обслужване на повикванията е *o*(1). Схемата е много ефективна, когато повечето пътечки към загуби на повиквания включват само загуби.

Четвъртата основна схема се нарича проста балансирана степен на подобие (simple balanced likelihood ratio). При нея се увеличава честотата на поява на загуби на повиквания, като същевременно се запазва ограничена степента на подобие, свързана с циклите за възстановяване. Основната идея в случая е да се дефинира множество (стек), инициализирано като празно множество, съответстващо на загубите с даден размер в рамките на ε.. По време на циклите за симулация степента на подобие за загуба на повикване се поставя на върха на съответния стек и тя се отстранява от стека, ако съществува обслужено повикване с магнитуд (отклонение) от същия порядък, за да се отмени текущата стойност на степента на подобие. По такъв начин е удовлетворена ограничената относителна грешка BRE.

СХЕМИ ОСНОВАНИ НА ИЗПОЛЗВАНЕ НА ТОПОЛОГИЧНА ИНФОРМАЦИЯ

При основните схеми се използва само локална информация, с оглед на математическия модел, който отчита само директните преходи. От друга страна, възможно е да се използва по-обща топологична информация, която да покаже колко далече сме от множеството от състояния при загуба на повикване. Първата техника е ***избирателното увеличаване на вероятностите за неуспешни попадения***, в случая за това, че текущото повикване се губи, защото не може да бъде обслужено SFB (selective failure biasing).

Разликата между FB и SFB е че за всяко x $\in$ U, множеството от преходи към неуспешни попадения се разделя на две подмножества: множество от преходи към неуспешни попадения при които поне един компонент (всички канали са заети или опашките са пълни) вече има поне едно неуспешно попадение от този тип, за което условните вероятности образуват множество α_1; и множеството от преходи, при които винаги се наблюдават успешни попадения (с условна вероятност $1 - \alpha_1$). Ако първото множество е празно, то $\alpha_1 = 0$; ако второто множество е празно, тогава $\alpha_1=1$. Във всяко подмножество индивидуалните вероятности се задават пропорционално на първоначалните. Колкото повече се доближаваме до множеството от преходи към неуспешни попадения, толкова повече увеличаваме вероятността за поява на рядко събитие.

Подобен подход е ***избирателното увеличаване на вероятностите за неуспешни попадения за системи от подобни редове*** SFBSS (selective failure biasing for ‘series-like’ systems). Предполага се, че структурата на системата е близка до ситуацията, при която системата функционира тогава и само тогава, когато за всеки вид на компонента k, броят на функциониращите компоненти е по-голям или равен на някаква прагова величина l_k. За да се подобри избирателното увеличаване на вероятностите за неуспешни попадения, трябва да се направи по-вероятно неуспешното попадение за компонентите от клас k, при фиксиран брой от успешни попадения $n_k - x_k$, ако състоянието x е по-близо до праговата величина l_k. Тогава се търси множеството от преходи, при което за състояние x се задава вероятност α1, включващо компонентите от клас k, такива че съотношението $(n_k - x_k) - l_k$ е минимално.

Друг подход е ***избирателното увеличаване на вероятностите за неуспешни попадения за паралелни системи*** SFBPS (selective failure biasing for ‘parallel-like’ systems), който действа по подобен начин, но е създаден за системи, работещи с множества от l_k извън N_k паралелни модули, където $1 \le k \le K$. От $x \in U$, един компонент от тип k се казва, че е критичен, ако броя на преходите с успешни попадения е по-голям от l_k. Един преход (x, y) се счита за критичен, ако $y_k > x_k$ за някакво k. Методът използва два параметъра – α и α_1, като най-напред се ускоряват критичните преходи, след което не-критичните съответно с тегла αβ и α(1 - β). По същия начин за всяко множество индивидуалните вероятности се вземат пропорционално на оригиналните.

Следващ подход е ***дистанционното избирателно увеличаване на вероятностите за неуспешни попадения*** DSFB (distance-based selected failure biasing). Прилага се за системи с обобщена структура, предаване на неуспешни попадения и по-малко общи свойства. Всичко това изисква наличието на повече информация за системата, за нейната топология и нейния модел. От всяко текущо състояние x се пресмята y такова, че $(x, y) \in \Gamma$, разстоянието d(y) спрямо D се определя като $d(y) = \min_{z \in D} \sum_k (y_k - z_k)$, което може изчислително да изисква много голямо множество от неуспешни попадения, но за някои специфични модели може да е много ефективно. Множеството от преходи към неуспешни попадения, на които е присвоена вероятността α се декомпозира до множество от преходи към неуспешни попадения при l-тото най-малко разстояние до D, което получава условна вероятност $\alpha_1(1 - \alpha_1)^{l-1}$, освен последната, която е условната вероятност $(1 - \alpha_1)^l$. Отново за всяко подмножество индивидуалните вероятности се вземат пропорционално на оригиналните.

СХЕМИ ПРИ КОИТО ДИРЕКТНО СЕ ДОСТИГА НУЛЕВА ВАРИАЦИЯ ПРИ СМЯНА НА МЯРКАТА

До сега разгледаните схеми дават възможност да се достигне по-бързо до множеството от преходи за неуспешни попадения, в сравнение със стандартните методи за симулация, но не третират въпросите, свързани с достигането на нулева вариация при смяна на мярката. За целта трябва да се оцени вероятността за преход $\gamma(x) = P[T_D < T_0]$.

Нека $\gamma(x) = P[T_D < T_0 \mid X_0 = x]$. Тогава Марковската верига с нулева вариация при смяна на мярката може да се представи като

$$\tilde{\mathbf{P}}(x,y) = \frac{\mathbf{P}(x,y)\gamma(y)}{\sum_z \mathbf{P}(x,z)\gamma(z)} \qquad (8)$$

Прилагането на такава смяна на мярката предполага, че за всяко z е известна стойността на $\gamma(z)$, но в случая е интересна стойността на $\gamma(\mathbf{0})$. Заменяме $\gamma(z)$ с една апроксимация на ( ) за да се доближим до оценителя за нулева вариация. Подходът дава добри резултати при малки модели, но не е подходящ при модели с много състояния. Тъй като състоянията трябва да се съхраняват, в процеса на симулация възниква проблем с компютърното пространство при съхранение на всички възможни преходи.

СИМУЛАЦИОНЕН АЛГОРИТЪМ ЗА ЗАГУБИ НА ПОВИКВАНИЯ ПРИ ПРЕХВЪРЛЯНЕ НА СИГНАЛА

Разработен е симулационен алгоритъм по схемата FB (failure biasing), който увеличава вероятностите за неуспешни попадения, в случая увеличава вероятността за загуба на текущо повикване, защото то не може да бъде обслужено, тъй като всички канали на базовата станция са заети или опашките на хендовъра са препълнени. Разглежда се модел на хендовър с две опашки: за гласови повиквания и за повиквания от данни, представен в [6]. Тук се предлага следния ***Алгоритъм*** за ускоряване на стохастичната симулация:

Стъпка 1: Инициализация на системните параметри, генериране на стойности за входящите повиквания, и изпълнение на алгоритъма, представен в [6].

Стъпка 2: Задава се стойност за общата вероятност от загуба на текущо повикване равна на $\alpha \in (0, 1)$, като обикновено $0{,}5 \le \alpha \le 0{,}9$, с която се увеличава вероятността за загуба на повикване.

Стъпка 3: Изгражда се Марковска верига за преход към загуба от състояние x.

Стъпка 4: Вероятностите за преход от едно състояние към друго се променят съгласно (1) и (2).

Стъпка 5: След получаване на загуба на повикване в каналите на базовата станция или опашките на хендовъра се извършва превключване обратно към P.

Стъпка 6: Изчислява се вероятността за преход към загуба от състояние x.

Стъпка 7: Прави се проверка дали вероятността за преход към загуба от състояние x има отклонение в рамките на ε. Ако е по-малка от ε, се преминава към стъпка 4, ако не към Стъпка 3.

Предложеният алгоритъм дава възможност да бъдат оценени вероятностите за загуба на текущо повикване и се прилага съвместно с алгоритъма, представен в [6].

ИЗВОДИ

Разгледани са симулационни схеми, които може да бъдат приложени за ускоряване на стохастичната симулация с помощта на редки събития. Анализирани са различни схеми за симулация при метода на случайната значима извадка Importance Sampling. За оценка за загубите на повиквания при прехвърлянето на сигнала от една базова станция към друга е предложен симулационен алгоритъм по схемата FB (failure biasing), който увеличава вероятностите за неуспешни попадения, в случая увеличава вероятността за загуба на текущо повикване, защото то не може да бъде обслужено, тъй като всички канали на базовата станция са заети или опашките на хендовъра са препълнени.

ЛИТЕРАТУРА

[1] Blanchet, J., Rudoy, D. Rare Event Simulation and Counting Problems. Book chapter in Rare Event Simulation using Monte Carlo Methods. John Wiley & Sons, 2009, pp. 171 – 192.

[2] Bucklew, J. An Introduction to Rare Event Simulation. Springer Series in Statistics, XI. Springer-Verlag, Berlin, 2004.

[3] Ivanova, E., Radeva, S., Radev, D. Rare Events and Quality of Services for IPv6 networks, Journal of Electrotechniques and Electronics E+E, Vol. 46 No 7/8, 2011, pp.12-17.

[4] L'Ecuyer, P., Mandjes, M., Tuffin, B. Importance Sampling in Rare Event Simulation. Book chapter in Rare Event Simulation using Monte Carlo Methods. John Wiley & Sons, 2009, pp. 17 – 38.

[5] Радев, Д., Симулация на редки събития в широколентови цифрови мрежи, Vol. 6, Колбис, 2006.

[6] Оцетова – Дудин, Е. Радева, С. Моделиране и симулация на прехвърляне на сигнала в клетъчни радио мрежи, Сборник доклади: ТЕЛЕКОМ'2012, София, (под печат).

За контакти:

ас. инж. Екатерина Оцетова-Дудин - Катедра "Безжични комуникации и разпръскване", Висше Училище Колеж по Телекомуникации и Пощи - София,
тел.: 02/ 86 22 244, e-mail: eotsetova@abv.bg

проф. дтн Светла Радева - Катедра "Безжични комуникации и разпръскване", Висше Училище Колеж по Телекомуникации и Пощи - София, тел.: 02/ 86 22 244,
e-mail: svetla_ktp@abv.bg

Докладът е рецензиран.