

DOI:10.22144/ctu.jsi.2020.092

SỬ DỤNG THUẬT TOÁN ENTROPY CHÉO VÀ CHỌN MẪU GIBBS ĐỂ ƯỚC LƯỢNG XÁC SUẤT SỰ KIỆN HIẾM

Trần Văn Lý^{1*}, Nguyễn Tử Thịnh¹, Nguyễn Dương Thanh Phú², Trà Đức Phô³ và Trần Văn Trọng⁴

¹Khoa Khoa học Tự nhiên, Trường Đại học Cần Thơ

²Trường THPT Nguyễn Việt Dũng, Thành phố Cần Thơ

³Nhà xuất bản Giáo dục - Chi nhánh Cần Thơ

⁴Trường Cao đẳng Kinh tế - Kỹ thuật Cần Thơ

*Người chịu trách nhiệm về bài viết: Trần Văn Lý (email: tvly@ctu.edu.vn)

Thông tin chung:

Ngày nhận bài: 04/03/2020

Ngày nhận bài sửa: 15/04/2020

Ngày duyệt đăng: 29/06/2020

Title:

Using Cross-Entropy algorithm for estimating rare event probability

Từ khóa:

Chọn mẫu Gibbs, đối độ đo xác suất, Entropy chéo, mô phỏng Monte Carlo, sự kiện hiếm

Keywords:

Cross-entropy, Gibbs sampler, Monte-Carlo simulation, probability measure change, rare events

ABSTRACT

In this study, Monte-Carlo simulation samples are generated by the periodic Gibbs sampling algorithm. The probability of a rare event will be estimated from these simulated samples. By using the Naïve Monte Carlo method to estimate the very small probability of rare events, it is necessary to create very large simulation samples that take a long time to initialize. This limitation was significantly improved by combining the cross-entropy method with the Gibbs sampling algorithm to create Monte-Carlo simulation samples. Using the technique of probability measure change in the cross-entropy method, rare events will occur in the simulation sample at a higher frequency according to the new probability measure. The probability of these rare events can be well estimated by returning the results for the initial probability measure.

TÓM TẮT

Trong nghiên cứu này, các mẫu mô phỏng Monte Carlo được khởi tạo bởi thuật toán chọn mẫu Gibbs quét tuần tự. Xác suất của sự kiện hiếm sẽ được ước lượng từ các mẫu mô phỏng này. Khi sử dụng phương pháp Monte Carlo đơn giản, để ước lượng được các xác suất rất bé của sự kiện hiếm thì cần phải tạo các mẫu mô phỏng có kích thước rất lớn, mất nhiều thời gian khởi tạo. Hạn chế này được cải thiện đáng kể khi phương pháp Entropy chéo được sử dụng kết hợp với thuật toán Gibbs để tạo các mẫu mô phỏng Monte Carlo. Với kỹ thuật đối độ đo xác suất trong phương pháp Entropy chéo, các sự kiện hiếm sẽ xuất hiện trong mẫu mô phỏng với tần số cao hơn theo độ đo xác suất mới, nhờ đó không cần khởi tạo mẫu có kích thước quá lớn cũng có thể ước lượng tốt được xác suất của các sự kiện hiếm này khi trả ngược các kết quả tính toán về độ đo xác suất ban đầu.

Trích dẫn: Trần Văn Lý, Nguyễn Tử Thịnh, Nguyễn Dương Thanh Phú, Trà Đức Phô và Trần Văn Trọng, 2020. Sử dụng thuật toán Entropy chéo và chọn mẫu Gibbs để ước lượng xác suất sự kiện hiếm. Tạp chí Khoa học Trường Đại học Cần Thơ. 56(Số chuyên đề: Khoa học tự nhiên)(1): 46-53.

1 GIỚI THIỆU

Trong nghiên cứu này phương pháp Entropy chéo được giới thiệu qua một thuật toán phù hợp để ước lượng xác suất của các sự kiện hiếm trong một trường ngẫu nhiên phức hợp (De Boer *et al.*, 2005). Phương pháp được mô tả là một quá trình lặp trong đó mỗi bước lặp có hai xử lý cơ bản sau:

(i) Khởi tạo một mẫu dữ liệu ngẫu nhiên (các quỹ đạo, các vector ngẫu nhiên, ...) từ phân phối của một trường ngẫu nhiên phức hợp được quan tâm;

(ii) Cập nhật các tham số của mô hình (từ dữ liệu mô phỏng) rồi sử dụng để khởi tạo mẫu mô phỏng khác thực hiện cho bước lặp tiếp theo.

Trong nghiên cứu này các mẫu mô phỏng cho một vector ngẫu nhiên đa biến $X = (X_1, X_2, \dots, X_V)$ được khởi tạo từ một phân phối mục tiêu $\pi(X)$ xác định. Nghiên cứu được xem xét với giả thiết là có thể dễ dàng thực hiện việc lấy mẫu từ các phân phối điều kiện $\pi(X_i|X_{[-i]})$, $i = 1, 2, \dots, V$, trong đó $X_{[-i]} = (X_j; j \neq i)$. Chọn mẫu Gibbs tuần tự là phương pháp được chọn lựa để thực hiện việc cập nhật các thành phần X_i của X với mẫu lấy từ các phân phối điều kiện $\pi(X_i|X_{[-i]})$ trên trạng thái hiện tại của các thành phần khác. Thứ tự của thành phần được cập nhật có thể thay đổi. Với điều kiện cân bằng tổng quát và giới hạn về số lần lặp, mẫu của các biến ngẫu nhiên sẽ tiệm cận về phân phối mục tiêu $\pi(X)$. Các khái niệm cũng như các cơ sở lý thuyết liên quan đến phép chọn mẫu Gibbs được tham khảo ở Brémaud (1999); sự hội tụ của thuật toán được khẳng định bởi các kết quả trong Liu *et al.* (1995) và Levine and Casella (2006).

2 THUẬT TOÁN CHỌN MẪU GIBBS TUẦN TỰ

Giả sử ta muốn mô phỏng vector ngẫu nhiên V chiều $X = (X_1, X_2, \dots, X_V)$ có phân phối $\pi(X)$. Phép chọn mẫu Gibbs quét tuần tự cập nhật thay đổi các thành phần X_i của X theo một trình tự qui định trước. Thuật toán của phép chọn mẫu này được trình bày như dưới đây.

Thuật toán 2.1. Lấy mẫu Gibbs quét tuần tự.

1) Chọn $X^{(0)} = (X_1^{(0)}, X_2^{(0)}, \dots, X_V^{(0)})$.

2) Ở lần lặp thứ $t (\geq 1)$:

– Lấy ngẫu nhiên $X_1^{(t)} \sim \pi(X_1|X_2^{(t-1)}, X_3^{(t-1)}, \dots, X_V^{(t-1)})$;

– Lấy ngẫu nhiên $X_2^{(t)} \sim \pi(X_2|X_1^{(t)}, X_3^{(t-1)}, \dots, X_V^{(t-1)})$; ;

– Lấy ngẫu nhiên $X_V^{(t)} \sim \pi(X_V|X_1^{(t)}, X_2^{(t)}, \dots, X_{V-1}^{(t)})$.

3) Lặp lại bước 2) cho tới khi đạt được yêu cầu dừng (xấp xỉ với phân phối mục tiêu ở một mức độ nào đó).

Sự hội tụ của thuật toán được đảm bảo bởi **Định lý 2.1** bên dưới nếu các điều kiện sau đây được thỏa mãn:

(d1) Sau một số bước hữu hạn n_0 bước lặp, khoảng cách Pearson $d_\pi(p, \pi)$ từ hàm mật độ $p_{n_0}(X)$ đến hàm mật độ mục tiêu π là hữu hạn, trong đó $d_\pi(p_{n_0}, \pi)$ được xác định bởi

$$d_\pi^2(p_{n_0}, \pi) = Var \left\{ \frac{p_{n_0}(X)}{\pi(X)} \right\} = \int \frac{p_{n_0}^2(X)}{\pi(X)} dX - 1;$$

(d2) $\int \left\{ \frac{K(Y|X)}{\pi(Y)} \right\}^2 \pi(X)\pi(Y) dXdY < \infty$, trong đó K là hàm xác suất chuyển;

(d3) Mẫu tạo bởi **Thuật toán 2.1** là một xích Markov tối giản.

Định lý 2.1. Cho mật độ ban đầu của phép lấy mẫu Gibbs tuần tự là $p_0(X)$ và $p_n(X)$ mật độ nhận được ở bước lặp thứ n . Khi các điều kiện (d1), (d2) và (d3) thỏa mãn thì khoảng cách Pearson từ P_n đến hàm mục tiêu π sẽ hội tụ đơn điệu giảm về 0 theo một tốc độ hình học.

Chứng minh.

Ký hiệu K^n là hàm chuyển sau n bước lặp, sử dụng quy nạp ta được

$$K^n(Y|X) = \int K(Y|X')K^{n-1}(X'|X)dX'$$

Không mất tính tổng quát, ta giả sử $n_0 = 0$ trong điều kiện (d1).

Đặt $g(X) = \frac{p_0(X)}{\pi(X)} - 1$, khi đó với bất kỳ hàm $h \in L_0^2(\pi)$ ta có

$$\begin{aligned} & |E_{p_n}\{h(X)\} - E\{h(X)\}| \\ &= \left| \int (Y)K^n(Y|X) \left\{ \frac{P_0(X)}{\pi(X) - 1} \right\} \pi(X) dXdY \right| \\ &= |\langle F_s^n, g \rangle| \leq \|F_s^n\| \|h\| \|g\|, \end{aligned}$$

Trong đó, $\|h\| = \sqrt{E(h^2)}$, $\|F_s^n\| = \sup_{h \in L_0^2(\pi), \|h\|=1} \|F_s^n(h)\|$ và

$$L_0^2(\pi) = \{h(X): E[h(X)] = 0, E[h^2(X)] < \infty\}.$$

Lấy $t(X) = \frac{p_n(X)}{\pi(X)-1}$ thì biểu thức vế bên trái sẽ là $d_\pi^2(p_n, \pi)$. Và vì $\|g\| = d_\pi(p_0, \pi)$ nên $d_\pi(p_n, \pi) \leq \|F_S^n\| d_\pi(p_0, \pi)$.

Tính đơn điệu dễ dàng được nhận ra với lưu ý rằng $\|F_S\| \leq 1$ khi ta thay thế p_0 bởi p_{n-1} và tốc độ hình học nhận được là

$$\lim_{n \rightarrow \infty} \|F_S^n\|^{\frac{1}{n}} = r.$$

Kết luận của Định lý nhận được nhờ kết quả của Định lý 3 trong (Liu *et al.*, 1995) đã chứng minh được rằng $r < 1$ khi các điều kiện (d2) và (d3) thỏa mãn. \square

3 PHƯƠNG PHÁP ENTROPY CHÉO

Xét một họ các hàm mật độ $\pi_p = \pi(X, p)$ với $X = (X_1, X_2, \dots, X_V)$ là một vector ngẫu nhiên và $p = (p_1, p_2, \dots, p_V)$ là một vector tham số. Entropy chéo giữa $\pi_q = \pi(X, q)$ và $\pi_p = \pi(X, p)$, còn được gọi là khoảng cách Kullback-Leibler, được xác định bởi

$$\mathcal{D}(\pi_q, \pi_p) = E_q \left(\log \frac{\pi_q}{\pi_p} \right) = \sum_X \pi(X, q) \log(\pi(X, q)) - \sum_X \pi(X, q) \log(\pi(X, p)).$$

Xét tổng $S(X) = X_1 + X_2 + \dots + X_V$ và giả sử rằng ta quan tâm đến xác suất l của sự kiện hiếm dạng $\{S(X) \geq \gamma\}$ theo mật độ $\pi(X, \bar{p})$ với γ là một mức giá trị đủ lớn nào đó, nghĩa là:

$$l = P_{\bar{p}}(S(X) \geq \gamma) = E_{\bar{p}}(1_{\{S(X) \geq \gamma\}}). \quad (1)$$

Sử dụng một mẫu ngẫu nhiên $(X^{(1)}, X^{(2)}, \dots, X^{(N)})$ được lấy từ hàm mật độ $\pi(X, \bar{p})$ ta có thể ước lượng Monte Carlo đơn giản cho xác suất l :

$$\hat{l} = \frac{1}{N} \sum_{i=1}^N 1_{\{S(X^{(i)}) \geq \gamma\}}. \quad (2)$$

Tuy nhiên để ước lượng được các xác suất l khá bé theo cách này, cần phải tạo một mẫu có kích thước rất lớn, dẫn đến “chi phí dung lượng” và thời gian tạo mẫu sẽ lớn.

Một thuật toán được giới thiệu theo phương pháp Entropy chéo dựa trên phép đổi độ đo xác suất, qua đó có thể ước lượng được các xác suất khá bé với một mẫu không cần quá lớn, và do đó thời gian tạo mẫu sẽ ít hơn.

Thay vì sử dụng (2), xác suất l sẽ được ước lượng dưới dạng

$$\hat{l} = \frac{1}{N} \sum_{i=1}^N 1_{\{S(X^{(i)}) \geq \gamma\}} \frac{\pi(X, \bar{p})}{\pi^*(X)}. \quad (3)$$

Trong đó, độ đo mới

$$\pi^*(X) = \frac{1_{\{S(X) \geq \gamma\}} \pi(X, \bar{p})}{l}$$

được chọn từ họ $\pi(X, p)$ sao cho khoảng cách Kullback-Leibler giữa $\pi^*(\cdot)$ và $\pi(\cdot, p)$ sau đây

$$\mathcal{D}(\pi^*, \pi_p) = \sum_X \pi^*(X) \log(\pi^*(X)) - \sum_X \pi^*(X) \log \pi(X, p)$$

càng nhỏ càng tốt. Điều này dẫn đến việc phải chọn p sao cho

$$\begin{aligned} & \sum_X \pi^*(X) \log \pi(X, p) \\ &= \sum_X \frac{1_{\{S(X) \geq \gamma\}} \pi(X, \bar{p})}{l} \log \pi(X, p) \end{aligned}$$

là lớn nhất, tương đương chuỗi thủ tục tính toán sau đây

$$\begin{aligned} & \max_p \sum_X \frac{1_{\{S(X) \geq \gamma\}} \pi(X, \bar{p})}{l} \log \pi(X, p) \Leftrightarrow \\ & \max_p E_{\bar{p}} [1_{\{S(X) \geq \gamma\}} \log \pi(X, p)]. \end{aligned}$$

Tiếp tục thêm bước làm tương tự, khi đặt trong một thuật toán lặp, giả sử rằng $p^{(t-1)}$ đã được xác định ở bước lặp thứ t , thực hiện phép đổi độ đo với mẫu trọng tâm lấy từ phân phối $p^{(t-1)}$, độ đo mới $\pi(X, p^{(t)})$ được xác định bởi

$$\begin{aligned} & p^{(t)} = \\ & \arg \max_p [1_{\{S(X) \geq \gamma\}} W(X, \bar{p}, p^{(t-1)}) \log \pi(X, p)], \end{aligned}$$

trong đó $x_0 = \arg \max_x [y(x)]$ là hàm xác định tham số x_0 để đạt y_{max} và

$$W(X, \bar{p}, p^{(t-1)}) = \frac{\pi(X^{(t)}, \bar{p})}{\pi^*(X, p^{(t-1)})}.$$

Cập nhật của $p^{(t)}$ có thể tìm được từ lời giải của hệ sau

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N 1_{\{S(X^{(i)}) \geq \gamma\}} W(X^{(i)}, \bar{p}, p^{(t-1)}) \nabla \log \pi(X^{(i)}, p) = \\ & 0, \quad (4) \end{aligned}$$

trong đó $(X^{(1)}, X^{(2)}, \dots, X^{(N)})$ được lấy từ mật độ $\pi(X, p^{(t-1)})$ và toán tử gradient được tính tương ứng theo p .

Để ước lượng cho xác suất l trong (1) theo (3), chúng ta sẽ sử dụng độ đo xác suất $\pi^*(\cdot) = p^{(t)}$ được cập nhật bởi (4) ở một số bước lặp phù hợp theo thuật toán Entropy chéo (CEA, Cross-Entropy Algorithm) dưới đây (De Boer *et al.*, 2005).

Thuật toán 3.1. CEA dùng để ước lượng các xác suất sự kiện hiếm.

1) Chọn mật độ ban đầu $p^{(1)}$ và bắt đầu đếm số bước lặp $t = 1$.

2) Ở lần lặp thứ t ($t \geq 2$):

Khởi tạo mẫu ngẫu nhiên $(X^{(1)}, X^{(2)}, \dots, X^{(N)})$ được lấy từ hàm mật độ $\pi(X, p^{(t-1)})$, tính các tổng $S(X^{(i)}), i = \overline{1, N}$, rồi sắp thứ tự từ nhỏ đến lớn $S_{(1)}, S_{(2)}, \dots, S_{(N)}$. Với $\rho \in (0; 1)$, đặt $\hat{p} = S_{(\lceil(1-\rho)N\rceil)}$ là số lượng $(1 - \rho)$ phần sau của mẫu đã sắp thứ tự này;

Nếu $\hat{p} < \gamma$: sử dụng mẫu khởi tạo để tính cập nhật các xác suất $p^{(t)}$ từ (4) rồi tiếp tục lặp lại bước 2).

Nếu $\hat{p} \geq \gamma$, dừng bước 2), lấy $\hat{p} = \gamma$ và chuyển sang bước 3).

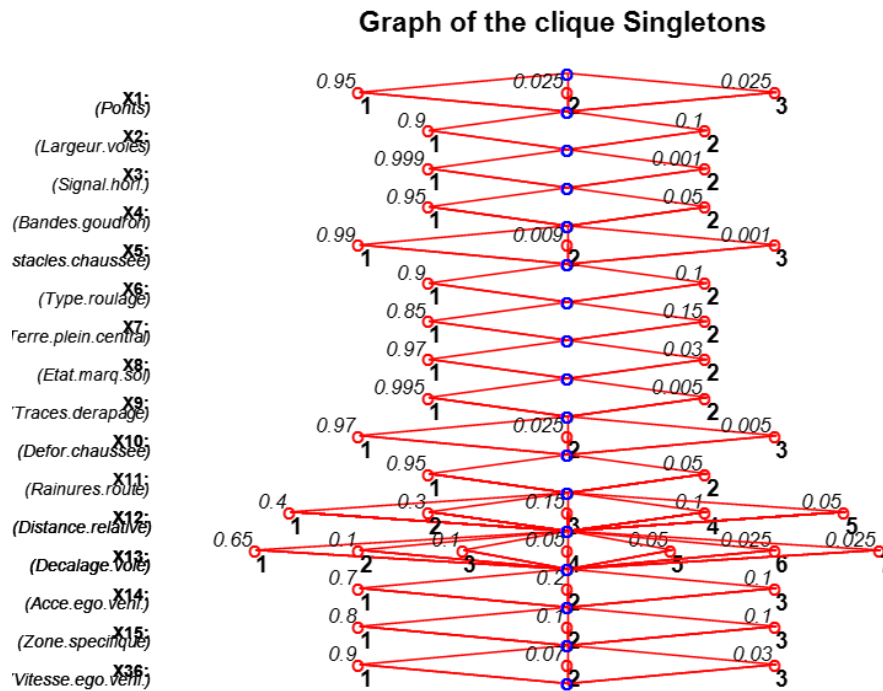
3) Đặt T là bước lặp sau cùng, khởi tạo mẫu $(X^{(1)}, X^{(2)}, \dots, X^{(N_1)})$ từ phân phối $\pi(X, p^{(T)})$ và ước lượng xác suất l bởi

$$\hat{l} = \frac{1}{N_1} \sum_{i=1}^{N_1} 1_{\{S(X^{(i)}) \geq \gamma\}} W(X^{(i)}, \bar{p}, p^{(T)}).$$

4 ÁP DỤNG

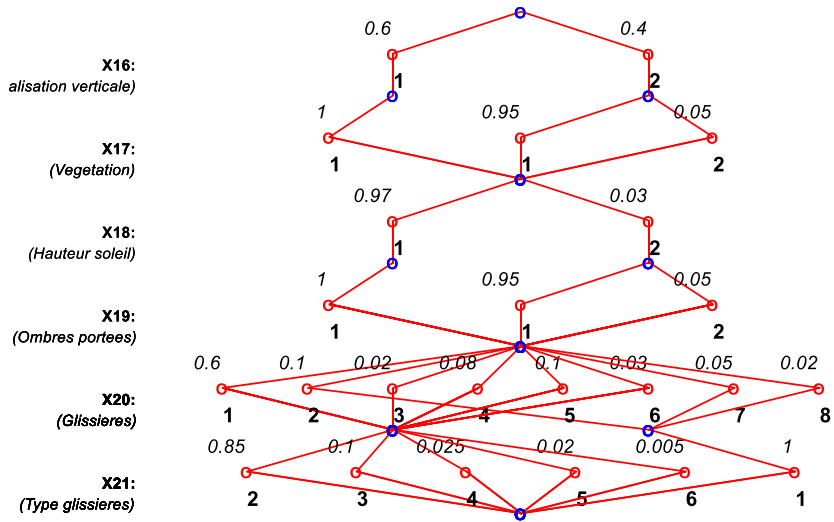
4.1 Mô hình áp dụng

Chúng tôi xét một mô hình gồm 61 tham số ngẫu nhiên dưới dạng một vector ngẫu nhiên $X = (X_1, X_2, \dots, X_{61})$ theo 4 nhóm các biến tham số độc lập cực đại. Nhóm Singletons có 16 tham số, nhóm Binaires có 6 tham số, nhóm Meteo có 9 tham số và nhóm Route có 30 tham số có phân bố xác suất lần lượt theo cấu trúc các Hình 1, 2, 3 và 4. Lưu ý là trong phân bố này các trạng thái có xác suất nhỏ sẽ được mã hóa bởi các mức giá trị lớn và chúng tôi gọi phân phối ban đầu này là phân phối mục tiêu π .



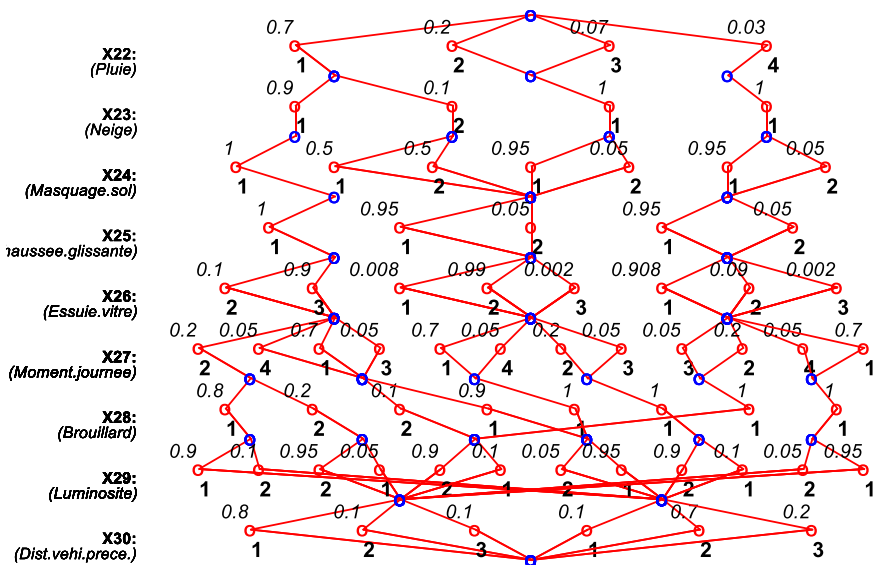
Hình 1: Mã hóa trạng thái và phân bố xác suất của nhóm Singletons

Graph of the clique Binaires

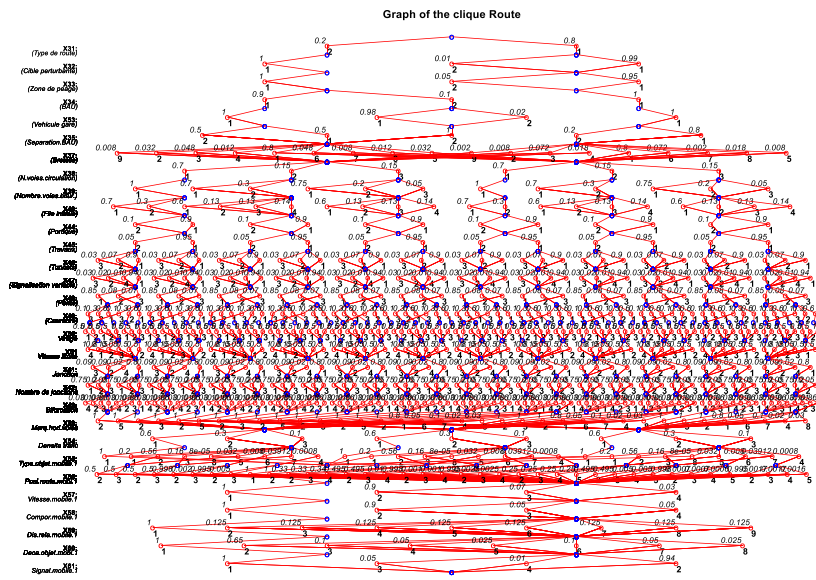


Hình 2: Mã hóa trạng thái và phân bố xác suất của nhóm Binaires

Graph of the clique Meteo



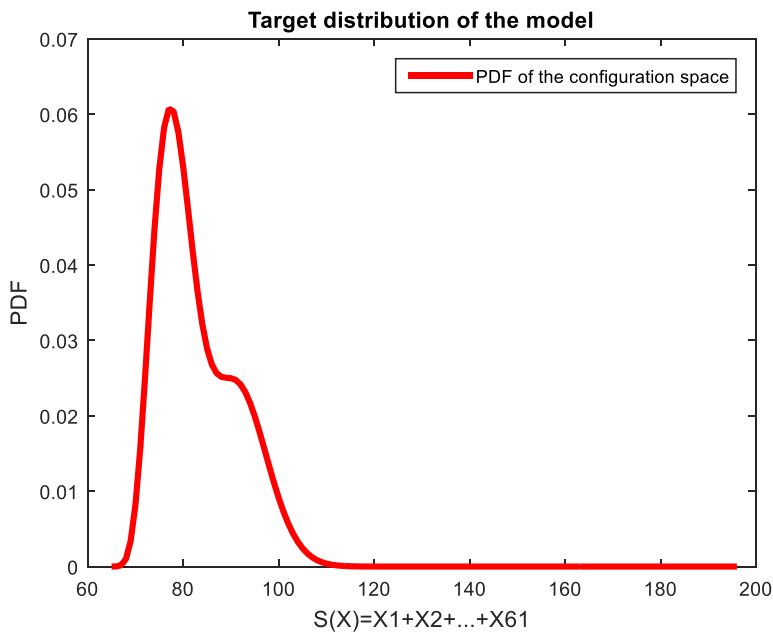
Hình 3: Mã hóa trạng thái và phân bố xác suất của nhóm Meteo



Hình 4: Mã hóa trạng thái và phân bố xác suất của nhóm Route

Với mỗi cấu hình cụ thể của vector ngẫu nhiên $X = (X_1, X_2, \dots, X_{61})$, thiết lập tổng quan sát $S(X) = X_1 + X_2 + \dots + X_{61}$. Theo sự mã hóa trạng thái

trong phân phối mục tiêu của mô hình, tổng quan sát $S(X)$ có giá trị nhỏ nhất là 65 và giá trị lớn nhất là 196. Hình 5 biểu diễn mật độ xác suất của mô hình theo tổng quan sát $S(X)$.



Hình 5: Phân phối mục tiêu của mô hình áp dụng

Trong mô hình này, những cấu hình có tổng quan sát càng lớn, xác suất xảy ra càng bé. Theo đó chúng ta có thể xem xét các sự kiện hiếm (tập các cấu hình có xác suất xảy ra khá bé hay rất bé) trong mô hình dưới dạng $\{S(X) \geq \gamma\}$ với γ là một mức giá trị đủ lớn.

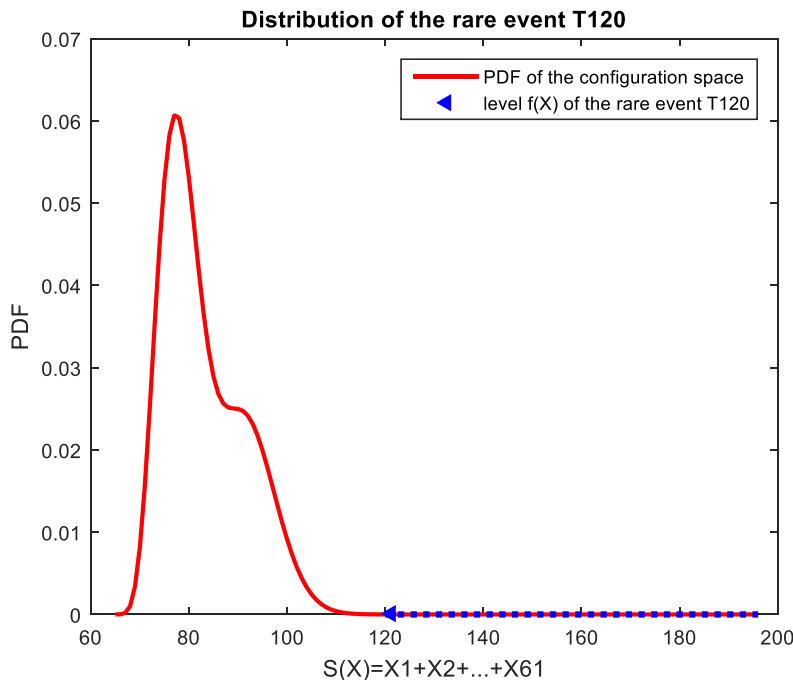
Trong thực nghiệm áp dụng tiếp theo dưới đây chúng tôi xem xét ước lượng xác suất của sự kiện hiếm $120 = \{S(X) \geq 120\}$, Hình 6, với xác suất kiểm chứng rất bé $53 \cdot 10^{-7}$.

4.2 Kết quả thực nghiệm

Các thực nghiệm này được thực hiện bởi phần mềm Matlab 2015a trên CPU 2.5 GHz, processor Intel Core i5, RAM 4GB. Các dữ liệu đầu vào gồm các trạng thái tham số, các điều kiện phụ thuộc, các giá trị mã hóa và các xác suất trong phân phối mô hình ban đầu được biên tập ở dạng các bảng dữ liệu Excel, khi đưa vào môi trường Matlab chúng tôi chuyển thành các siêu ma trận (super-matrix) và các

mảng tế bào. Bằng các chương trình thuật toán Matlab trong một toolbox được thiết lập, chúng tôi xuất ra các dữ liệu mô phỏng ở dạng các mảng tế bào. Từ đó chuyển sang các định dạng khác để xử lý tính toán, mô phỏng hình ảnh,

Để ước lượng cho sự xác suất kiện hiếm R120, chúng tôi tạo các mẫu mô phỏng có kích thước $N=100000$ và $N=1500000$ theo phương pháp Monte Carlo đơn giản. Đối với phương pháp Entropy chéo chúng tôi tạo mẫu có kích thước $N=100000$. Thời gian tạo mẫu, tần suất xuất hiện sự kiện hiếm R120 và các khoảng tin cậy 95% thu được từ các mẫu thực nghiệm này được trình bày ở Bảng 1. Kết quả thu được cho thấy với cỡ mẫu $N=100000$ tạo ra từ phương pháp Entropy chéo, có thể ước lượng được xác suất của R120. Nhưng với mẫu tạo ra từ phương pháp Monte Carlo đơn giản có cùng kích thước này thì sự kiện R120 chưa thấy xuất hiện, mà phải tạo mẫu đến kích thước $N=1500000$ chúng ta mới ước lượng được xác suất của R120.



Hình 6: Phân phối xác suất của sự kiện hiếm R120

Bảng 1: Ước lượng xác suất của sự kiện hiế R120 (xác suất đúng là $53 \cdot 10^{-7}$)

Phương pháp, Cỡ mẫu	Thời gian (s)	Tần suất của R120	Khoảng tin cậy 95%
Naïve Monte Carlo, $N=100000$	2319	0	NaN
Naïve Monte Carlo, $N=1500000$	71791	$82 \cdot 10^{-7}$	$(82 \pm 45) \cdot 10^{-7}$
Entropy chéo, $N = 100000$	2322	$59 \cdot 10^{-7}$	$(59 \pm 39) \cdot 10^{-7}$

5 KẾT LUẬN

Khảo sát các sự kiện hiếm với xác suất xảy ra nhỏ thường gặp khó khăn do không có nhiều dữ liệu thực. Mô phỏng Monte Carlo thường được dùng để tạo các dữ liệu mô phỏng nghiên cứu. Nhưng với những sự kiện có xác suất rất bé, đòi hỏi phải khảo sát trên các mẫu mô phỏng có kích thước rất lớn, mất nhiều “dung lượng” và thời gian khởi tạo. Điều này là một hạn chế không nhỏ trong những trường hợp ứng dụng xử lý trực tiếp. Bằng cách kết hợp với phương pháp Entropy chéo, trong đó kỹ thuật đổi độ đo được vận dụng, một cách tiếp cận hiệu quả đã được giới thiệu để khắc phục hạn chế này. Cụ thể, để ước lượng được xác rất bé của những sự kiện hiếm, chúng tôi không cần tạo mẫu quá lớn. Điều này thể hiện qua kết quả thực nghiệm được trình bày trong bài báo. Trên thực tế phương pháp này cũng đã được áp dụng trong nhiều dự án công nghiệp. Trong đó nhiều thuật toán tạo mẫu Gibbs cải tiến khác đã được sử dụng và ứng dụng trong các vấn đề khác nữa, không chỉ đơn giản là để ước lượng xác

suất sự kiện hiếm. Tất nhiên là tính hội tụ và yêu cầu xấp xỉ phân phối (của dữ liệu mô phỏng với không gian cấu hình) của các thuật toán cải tiến cũng rất được chú trọng trong các ứng dụng đó. Lưu ý rằng đây là những yêu cầu rất quan trọng mà các ứng dụng công nghiệp thường đòi hỏi rất nghiêm ngặt.

TÀI LIỆU THAM KHẢO

- Brémaud, P., 1999. Markov chains – Gibbs Fields, Monte Carlo Simulation and Queues. Springer – New York, 444 pages.
- De Boer, P.T., Kroese, D.P., Mannor, S. and Rubinstein, R.Y, 2005. A Tutorial on the Cross-Entropy Method. *Annals of Operations Research*, 134: 19-67.
- Levine, R.A. and Casella, G., 2006. Optimizing random scan Gibbs samplers. *Journal of Multivariate Analysis*, 97(10):2071-2100.
- Liu, J.S., Wong, W.H. and Kong, A., 1995. Covariance Structure and Convergence Rate of the Gibbs Sampler with Various Scans. *J. R. Statist. Soc. B*, 57(1): 157-169.