

APLICACION DEL MUESTREO ENFATIZADO A LA EVALUACION DE TRANSMISIONES DIGITALES

Aníbal R. Figueiras Vidal*, José B. Mariño Acebal**, Miguel A. Lagunas Hernández** y Ramón García Gómez**

* ETSI Telecomunicación, Cdad. Universitaria, Madrid 3

** ETSI Telecomunicación, Jorge Girona Salgado s/n, Barcelona 34

ABSTRACT.- The Importance Sampling Technique is a method for reducing the computational effort in Montecarlo simulations for obtaining the relative frequency of an event having a very low probability. While this method is well known in general Operations Research /1//2//3/ and Radar /4//5//6//7//8//9//10/, it has not been considered in digital communication problems. This paper aims to introduce the Importance Sampling concept in this communication context.

INTRODUCCION

El método de Montecarlo es una alternativa para la evaluación de muchos procesos de detección cuyo análisis, sea por la estructura del receptor o por las estadísticas de la señal recibida, resulta complejo o inviable.

En sistemas de comunicaciones digitales (CCDD) de prestaciones elevadas (los habituales), la característica básica (probabilidad de error, P_e) toma valores muy reducidos; por lo que se requeriría, en principio, un gran esfuerzo computacional para su obtención. La técnica de muestreo enfatizado ("Importance Sampling", IS), bien conocida en Investigación Operativa general /1//2//3/ y Radar /4//5//6//7//8//9//10/, constituye una opción que se ha mostrado eficaz para reducir dicho esfuerzo; pero, curiosamente, su aplicación a CCDD es prácticamente inexistente.

En el presente trabajo, se introduce el concepto del IS con la nomenclatura y en el contexto habitual de las CCDD, insistiendo en sus posibilidades como herramienta de evaluación de sus prestaciones.

EL PROBLEMA DE LA DETECCION

Nos ocuparemos de casos discretos, ya que se pueden reducir a éstos los continuos /11/. Dado que las ideas básicas se generalizan con facilidad, también nos limitaremos inicialmente al caso binario. El correspondiente esquema genérico del detector se muestra en la Fig. 1.

Los observables \bar{r} (representativos de la señal y el ruido recibidos en cada intervalo de símbolo) son procesados por un circuito en general no lineal, $g(\bar{r})$; a cuya salida se sustrae un cierto valor numérico T (umbral), tras de lo que se aplica

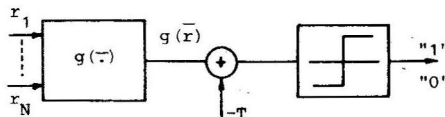


Fig. 1. Detector binario discreto

una decisión por signo, adoptando como cierta la hipótesis "i" ("0", "1"). $g(\bar{r})$ puede establecerse con arreglo a algún criterio de optimización (caso de receptores MAP y ML), siguiendo un esquema intuitivamente razonable o que se ha mostrado eficaz en la práctica, o buscando un comportamiento estable (y poco degradado respecto a óptimos) frente a variaciones en las estadísticas de la señal recibida (detectores robustos, libres en distribución o no paramétricos) /12//13//14//15/. En todo caso, $g(\bar{r})$ puede resultar tan complicada para ciertos observables que la evaluación de P_e puede hacerse muy dificultosa o imposible.

En general,

$$P_e = \Pr("0") \int_{-T}^{\infty} f_0(g) dg + \Pr("1") \int_{-\infty}^T f_1(g) dg \quad (1)$$

siendo $f_i(\cdot)$ la densidad de probabilidad (ddp) de g cuando se ha transmitido "i". Dada la analogía existente, nos ocuparemos de la primera integral, P_{e0} (que, en muchas situaciones "simétricas", coincide con P_e).

LA LIMITACION DEL METODO DE MONTECARLO

El método de Montecarlo recurre a generar repetidamente (en ordenador) la variable aleatoria (va) \bar{r} según $p_0(\bar{r})$ (siendo $p_0(\bar{r})$ la ddp de dicha va bajo la hipótesis "i"); obteniendo tras ello $g(\bar{r})$ y estimando P_{e0} por la frecuencia relativa del suceso $g(\bar{r}) > T$:

$$\hat{P}_{eN} = \frac{n^\circ \text{ veces que } f(\bar{r}_j) > T}{N} = \frac{1}{N} \sum_{j=1}^N u[g(\bar{r}_j) - T] \quad (2)$$

Claramente, el estimador es insesgado: su media tiende al valor a estimar

$$P_{e0} = \int_T^\infty f_0(g) dg = 1 - F_0(T) = E[u(g - T)] \quad (3)$$

con $N \rightarrow \infty$; pero su varianza vale

$$\text{Var}[\hat{P}_{eN}] = [1 - F_0(T)] F_0(T) / N \quad (4)$$

de modo que, para que el resultado sea confiable si $P_{e0} = F_0(T)$ es muy reducido, se necesita elevar grandemente N (p. ej., $N \approx 10/P_{e0}$), y el esfuerzo computacional implícito en ello puede no ser soportable.

EL MUESTREO ENFATIZADO

El principio de aplicación del IS se puede comprender a través de la siguiente discusión: si $h_0(\cdot)$ es una ddp

$$P_{e0} = \int_T^\infty f_0(g) dg = \int_T^\infty \frac{f_0(g)}{h_0(g)} h_0(g) dg \quad (5)$$

brindándose la posibilidad de utilizar un estimador

$$\hat{P}_{eN} = \frac{\{f_0[g(\bar{r}_j)] / h_0[g(\bar{r}_j)]\} n^\circ \text{ veces que } g(\bar{r}_j) > T}{N} = \frac{1}{N} \sum_{j=1}^N \frac{f_0 g(\bar{r}_j)}{h_0 g(\bar{r}_j)} u[g(\bar{r}_j) - T] \quad (6)$$

que también es insesgado, y ofrece la posibilidad de reducir grandemente la varianza de la estimación si $h_0(\cdot)$ se escoge adecuadamente.

¿Qué quiere decir, en lenguaje corriente, $h_0(\cdot)$ adecuado?. Está claro que el problema del primer estimador radica en que, si P_{e0} es muy reducido, resulta muy probable una simulación en la que no se produzca error alguno (salvo si N es suficientemente elevado). Esto es debido a que la va $g(\bar{r}_j)$ generada tiene una ddp $f_0(\cdot)$ casi completamente a la izquierda de T . (Fig. 2).

Si se genera, por el contrario, una va con ddp $h_0(\cdot)$ cuya área cubra apreciablemente el intervalo de valores a la derecha de T , se producirán muchos más "errores" en una simulación de igual N ;

por lo que (previa la ponderación de (6)) se conseguirá una estimación mucho más confiable con N moderado.

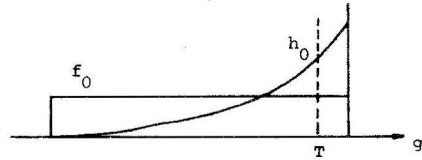


Fig. 2. Enfatizado de la ddp de g.

Realmente, en el ordenador se generan las \bar{r}_j ; de modo que lo que debe buscarse es una ddp $q_0(\bar{r})$ (sustituyendo a $p_0(\bar{r})$) cuyo efecto sobre g sea el indicado; entonces, y como también

$$P_{e0} = \int_{Z_1} P_0(\bar{x}) d\bar{x} \quad (7)$$

siendo Z_1 la región del espacio de observación que corresponde a decidir "1", puede utilizarse

$$\hat{P}_{eON} = \frac{1}{N} \sum_{j=1}^N \frac{P_0(\bar{r}_j)}{q_0(\bar{r}_j)} u[g(\bar{r}_j) - T] \quad (8)$$

PRECAUCIONES ADICIONALES: UN EJEMPLO

Pese a la sencillez del principio del IS, no han de pasarse por alto unas mínimas atenciones en su aplicación para evitar obtener resultados no fiables pese a su "apariencia". Recurriremos para señalar las precauciones fundamentales a un ejemplo que permite además una inmediata extensión de la discusión para binario recien realizada: considerando el receptor (coherente) por "correlación máxima" para cuatro señales ortogonales de igual energía en ruido gaussiano y blanco (Fig. 3), que es un ML e implica $P_{e0} = P_e$ (se admite independencia entre símbolos consecutivos).

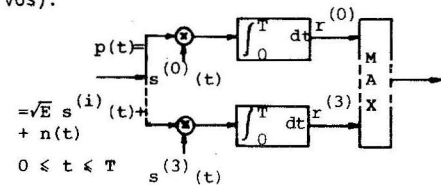


Fig. 3. Detector del ejemplo

Puede probarse //1/

$$P_e = 1 - \int_{-\infty}^{\infty} \phi^3(x) \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(x - \sqrt{E})^2}{2}\right] dx \quad (9)$$

habiendo tomado la densidad espectral de ruido unitaria, y siendo $\phi(\cdot)$ la integral gaussiana. (9) no admite expresión más

simple.

Los observables tienen, bajo la hipótesis "0", las ddp:

$$p_0[r^{(0)}] = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{[r^{(0)} - \sqrt{E}]^2}{2}\right\} \quad (10a)$$

$$p_0[r^{(i)}] = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{[r^{(i)}]^2}{2}\right\}, \quad i \neq 0 \quad (10b)$$

y son independientes (la ddp conjunta es producto de las individuales). Se produce error (bajo la hipótesis "0") cuando

$$g_j = \max\{r^{(i)} - r^{(0)}\} > 0, \quad i \neq 0 \quad (11)$$

y puede enfatizarse tomando:

$$q_0[r^{(0)}] = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{[r^{(0)}]^2}{2}\right\} \quad (12)$$

con lo que se constituirá el estimador:

$$\hat{p}_{eN} = \frac{1}{N} \sum_{j=1}^N \exp\left\{-\frac{[r_j^{(0)} - \sqrt{E}]^2 - [x_j^{(0)}]^2}{2}\right\} u(g_j) \quad (13)$$

La Tabla 1 muestra los resultados de 3 pruebas (P1, P2, P3) en función de E y N, precedidos de los valores derivados de una integración numérica (IN). Para los efectos ilustrativos y la discusión aquí pretendida, es suficiente con los decimales mostrados.

	E = 10	E = 20	E = 40
IN	$3.0 \cdot 10^{-2}$	$2.5 \cdot 10^{-3}$	$1.2 \cdot 10^{-5}$
P1	10	$1.22 \cdot 10^{-2}$	$1.55 \cdot 10^{-4}$
N=	10^2	$1.81 \cdot 10^{-2}$	$4.37 \cdot 10^{-4}$
	10^3	$3.15 \cdot 10^{-2}$	$1.54 \cdot 10^{-3}$
P2	10	$1.82 \cdot 10^{-3}$	$1.17 \cdot 10^{-5}$
N=	10^2	$2.72 \cdot 10^{-2}$	$9.95 \cdot 10^{-4}$
	10^3	$5.40 \cdot 10^{-2}$	$3.98 \cdot 10^{-3}$
P3	10	$3.67 \cdot 10^{-3}$	$3.60 \cdot 10^{-5}$
N=	10^2	$2.44 \cdot 10^{-2}$	$8.28 \cdot 10^{-4}$
	10^3	$3.41 \cdot 10^{-2}$	$2.01 \cdot 10^{-3}$

Tabla 1. Resultados de las pruebas "enfanzadas".

La Tabla 2 muestra resultados de 3 pruebas de simulación "directa" (Q1, Q2, Q3) con iguales parámetros

	E = 10	E = 20	E = 40
Q1	10	-	$1.00 \cdot 10^{-1}$
N=	10^2	$1.00 \cdot 10^{-2}$	$1.00 \cdot 10^{-2}$
	10^3	$4.00 \cdot 10^{-2}$	$2.00 \cdot 10^{-3}$
Q2	10	$1.00 \cdot 10^{-1}$	$1.00 \cdot 10^{-1}$
N=	10^2	$3.00 \cdot 10^{-2}$	$2.00 \cdot 10^{-2}$
	10^3	$3.10 \cdot 10^{-2}$	$5.00 \cdot 10^{-3}$
Q3	10	-	-
N=	10^2	$2.00 \cdot 10^{-2}$	-
	10^3	$3.50 \cdot 10^{-2}$	$2.00 \cdot 10^{-3}$

Tabla 2. Resultados de pruebas "directas"

Es inmediato observar la ventaja del empleo del IS a medida que se viaja hacia la derecha y arriba: es decir, que la probabilidad buscada es mucho menor que $1/N$; tal como era de esperar.

Pero deben observarse los siguientes efectos:

- Quando nos desplazamos hacia la izquierda y abajo, la simulación "directa" es preferible al IS: "distorcionar" las ddp se hace innecesario y perturbador.
- Si N es muy bajo respecto a las probabilidades involucradas en el IS ($N=10$, p. ej., en este caso), hay una "oscilación" de valores debido a la "brevedad" de la experiencia (interviene el punto de "arranque" de las muestras seudoraleatorias). Es preciso comprobar la estabilización de los resultados a medida que se hace avanzar N.
- Si E es alto (la probabilidad a estimar es pequeña), puede ocurrir que, incluso para valores de N aparentemente "razonables", los resultados de las experiencias del IS sean dispersos (que sean sesgados se debe al efecto anterior). Ello se debe a una no muy afortunada elección de la enfatización: en nuestro caso, cuando E es elevado, el factor exponencial de (13) es tan pequeño que deteriora la varianza del estimador. Un enfatizado menos violento (p.ej., tomando E' entre 0 y E como medio para $r^{(0)}$), y como su varianza una "moderadamente" inferior a 1) conduce a los resultados adecuados.

De ello se deducen como precauciones a tener en cuenta:

- Recurrir al IS cuando es efectivamente necesario;
- Comprobar la estabilización de resultados con el crecimiento de N;

- 3) Elegir cuidadosamente $q_0(\bar{r}_j)$, y controlar en todo caso la dispersión en varias experiencias.

BIBLIOGRAFIA

- /1/ J.J. Hammersley, D.C. Handscomb, Monte Carlo Methods. London: Methuen, 1964.
- /2/ F.S. Hillier, G.J. Lieberman, Introduction to Operations Research. New York: Holden-Day; 1967.
- /3/ J.J. Moder, S.E. Elmaghraby (Eds.), Handbook of Operations Research. New York: Van Nostrand; 1978.
- /4/ V.G. Hansen, IEEE IT-16, 309-318, (1970).
- /5/ V.G. Hansen, H.R. Ward, IEEE AES-8, 648-652, (1972).
- /6/ V.G. Hansen, IEEE IT-18, 664-667, (1972).
- /7/ B.H. Cantrell, G.V. Trunk, IEEE AES-9, 649-653, (1973).
- /8/ V.G. Hansen, Comp. & Elect. Engng., 1, 545-550, (1974).
- /9/ G.V. Trunk, B.H. Cantrell, F.D. Queen, IEEE AES-10, 574-582, (1974).
- /10/ B.H. Cantrell, G.V. Trunk, IEEE AES-10, 878-880, (1974).
- /11/ H.L. Van Trees, Detection, Estimation, and Modulation Theory (I). New York: Wiley; 1968.
- /12/ V.G. Hansen, B.A. Olsen, IEEE AES-7, 942-950, (1971).
- /13/ G.W. Zeoli, T.S. Fong, IEEE AES-7, 951-959, (1971).
- /14/ G.M. Dillard, C.E. Antoniak, IEEE AES-6, 629-635, (1970)
- /15/ E.K. Al-Hussaini, IEEE AES-15, 573-575, (1979).