

GRÁFICOS DE CONTROL MULTIVARIABLES APLICADOS A LA ESTIMACIÓN DE LA CALIDAD PERCIBIDA DE LA VOZ SOBRE IP.

APPLIED MULTIVARIATE CONTROL CHART TO THE PERCEIVED VOICE QUALITY PREDICTION TO VOICE OVER IP

LEONARDO FRANCISCO PANES MARTÍNEZ ¹

Universidad de Concepción, Concepción - Chile

RESUMEN

Con el desarrollo del protocolo de red IP se ha generado un crecimiento explosivo de Internet en los últimos diez o quince años, que ha llevado a la proliferación de múltiples servicios de comunicación, dada la naturaleza de este protocolo se hace necesario implementar mecanismos que permitan asegurar la calidad de servicio a los usuarios finales. El presente artículo presenta un sistema de evaluación en línea efectiva en la vigilancia y medición de la calidad percibida de la voz sobre IP consistente en el tiempo. El modelo presentado combina el control estadístico de procesos y los métodos objetivos de estimación de la calidad percibida de voz. Para esto se ha implementado el gráfico de control multivariable T_2 de Hotteling y el método objetivo no intrusivo de estimación de la calidad percibida de voz "Modelo-E".

Palabras Claves: Control estadístico de procesos, Protocolo de red IP, Modelo E, Voz sobre IP

ABSTRACT

The development of the Internet Protocol (IP) it's generating an explosive growth of Internet and that have taken the proliferation of multiples services of communication. But this protocol needs to implement mechanisms that allow to assure the quality of service to the end users. The present article presents a system of effective on-line evaluation in the surveillance and measurements of the perceived quality of the voice on IP consistent in the time. The presented model combines the statistical control of processes and the objective methods of estimate of the perceived quality of voice. For this it has been implemented Multivariable Control Chart T_2 de Hotteling and the perceived method for voice quality prediction to voice over IP.

Key words: Control Chart, Internet Protocol IP, E-Model, Voice on IP.

¹ Programa de Magíster en Ingeniería Industrial, Departamento Ingeniería Civil Industrial, Universidad de Concepción, Concepción, Chile.
E-mail: lpanes@udec.cl

1. INTRODUCCIÓN

Estimar la calidad de servicio percibida de la voz es un requisito fundamental en los sistemas de comunicación modernos; ya sea por razones técnicas, legales o comerciales. Para medir la calidad de voz percibida por el usuario, existen dos tipos de métodos utilizados: los métodos subjetivos y objetivos. Los métodos subjetivos representan la medida más fiable y respetada para medir la calidad percibida de la voz, Keagy (2001), pero resultan ser costosos y difíciles de implementar de forma permanente, en este sentido los métodos objetivos resultan ser los más apropiados. En la actualidad existen distintas aplicaciones que han implementado métodos objetivos, por ejemplo PSQM (1996), PESQ (2001) o el modelo E (2005).

Dada la naturaleza de las redes de datos, esta nos obliga a compartir el ancho de banda con todo el conjunto de aplicaciones que se ejecutan en la red, los paquetes de voz pueden perderse, no hay una garantía sobre el tiempo que tardarán en llegar de un extremo al otro, y menos si llegaran en forma ordenada. Lograr transportar voz de alta calidad sobre redes IP no es una tarea fácil de alcanzar, esto requiere el manejo de las capacidades de la red para el control del tráfico, protocolos de tiempo real y anchos de banda suficientes durante el tiempo que tome la realización de la llamada, por tanto se hace necesario entonces implementar mecanismos que permitan asegurar la calidad de la voz a los usuarios finales, estos mecanismos deben incluir métodos de medición y vigilancia de los parámetros funcionales de red.

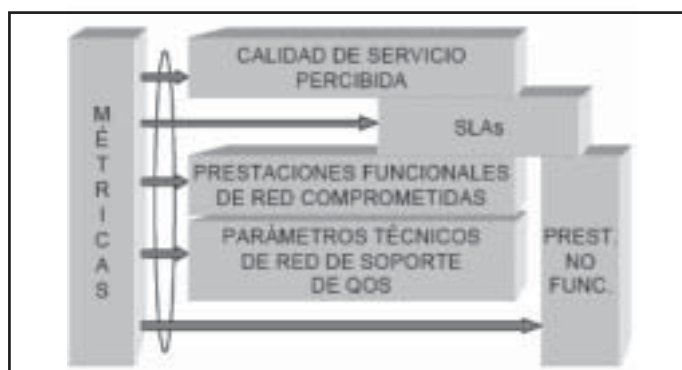
Para esto último, una de las herramientas más utilizada en la administración y control de calidad de procesos son los gráficos de control, esto dado su efectividad en la vigilancia, reducción de la variabilidad y estimación de los parámetros que definen la calidad de los procesos.

De acuerdo a lo expuesto anteriormente, en las siguientes secciones se presenta un sistema de evaluación en línea, efectiva en la vigilancia y medición de la calidad percibida de la voz sobre IP, el modelo presentado combina el control estadístico de procesos y métodos objetivos de estimación de la calidad percibida de voz.

• Trabajo Relacionado

La calidad de servicio en telecomunicaciones es un concepto que comprende muchas áreas, incluye no sólo el comportamiento de la red, sino también el grado de satisfacción del usuario al utilizar el servicio, en el que influyen aspectos como atención al cliente, comportamiento de la red, tasa de fallas, contratos, etc. Ante la necesidad de enmarcar la investigación, se ha utilizado el modelo conceptual de calidad de servicio en telecomunicaciones presentado por Bellido (2004), ver Figura 1.

Figura 1: Modelo conceptual de QoS



El modelo identifica distintos niveles y perspectivas dentro de la calidad de servicios en telecomunicaciones: calidad de servicio percibida, SLAs o contratos, prestaciones funcionales de la red, prestaciones no funcionales de la red, parámetros de red de soporte y métricas para todos los niveles. En lo referente a la investigación presentada se abordan dos aspectos del modelo, la calidad de servicio percibida y el nivel de prestaciones funcionales de la red.

Los parámetros que definen las prestaciones funcionales de red se definen en la recomendación Y.1540 (2002) de la ITU-T, los rangos de valores y la clasificación de estos parámetros de acuerdo a la clase de servicio se definen en la recomendación Y.1541 (2002) de la ITU-, obteniendo así que para la voz sobre IP es Clase 0, y los rangos de valores son los siguientes:

- IPTD: Retardo de transferencia de paquetes IP, no mayor a 150ms .
- IPDV: Variación del retardo en la transferencia de paquetes IP, no mayor a 50ms².
- IPER: Tasa de errores en los paquetes IP, 1×10^{-4}
- IPLR: Tasa de pérdida de paquetes IP, 1×10^{-3} .

• Métodos de estimación de la calidad de Voz percibida.

Para estimar la calidad percibida de la voz sobre IP pueden usarse dos tipos de métodos, los métodos objetivos y los métodos subjetivos.

Los métodos subjetivos consisten en evaluar la opinión media de un grupo de personas a las que se le hacen pruebas de escucha, los resultados de estas pruebas se pueden clasificar en tres categorías dependiendo del esquema usado: *Absolute Category Rating (ACR)*, que tiene como resultado el *Mean Opinion Score (MOS)*; *Degradation Category Rating (DCR)*, que tiene como resultado el *Degradation Mean Opinion Score (DMOS)* y *Comparison Category Rating (CCR)*, que tiene como resultado el *Comparison Mean Opinion Score (CMOS)*.

Estos métodos están normalizados por la recomendación P.800 (1996) de la ITU-T. El más utilizado es el *Mean Opinion Score*. En la tabla 1 se muestra la categorización en terminos absolutos de la calidad de escucha (MOS). El principal inconveniente de estos métodos es el costo y tiempo involucrados en la realización de las pruebas lo que hace que este tipo de método no sea viable de implementar de forma permanente para medir en línea la calidad percibida.

Tabla 1: Escala de Calidad De Escucha, MOS

MOS	Escala de Calidad De Escucha
5	Excelente
4	Buena
3	Media
2	Pobre
1	Mala

Los métodos objetivos se pueden clasificar en dos tipos: los métodos intrusivos, que basan su estimación en la comparación de señales; y los métodos no intrusivos, que basan su estimación en los parámetros funcionales de red.

Los métodos intrusivos requieren inyectar señales al medio en estudio para luego compararla con señales patrones, identificando las degradaciones y distorsiones de la señal

² Este valor va a depender directamente a la capacidad de *buffer Jitter* de los dispositivos de red

transmitida con respecto a la señal patrón y en base a esto estimar la calidad percibida. Entre los métodos intrusivos mas utilizados se encuentran: Perceptual Speech Quality Measure, PSQM; Measuring Normalizing Blocks, MNB (1998); Enhanced Modified Bark Spectral Distortion, EMBSD (2000); y Perceptual Evaluation of Speech Quality, PESQ.

Los métodos no intrusivos estiman la calidad percibida utilizando directamente los parámetros funcionales de la red. Entre los métodos objetivos no intrusivos de estimación de la calidad percibida se encuentran: el Modelo E, Método Factor de Alteración de Equipamiento definido en la recomendación G.113 (2002) de la ITU-T y la Aplicación de Redes Neuronales presentada por Mohamed (2003).

Es importante mencionar en este punto que esto concuerda con lo presentado por Bellido (2004) en su modelo, en donde los parámetros funcionales de la red y la calidad percibida por el usuario están directamente relacionados y forman junto a otras medidas, la calidad de servicio entregada al usuario.

El principal inconveniente de los métodos objetivos es que estos no dan como resultado una representación directa de la calidad percibida por el usuario, por tanto se hace necesaria su calibración en base a los resultados obtenidos de los métodos subjetivos.

Si se comparan los métodos objetivos, los métodos intrusivos presentan un mejor desempeño en la correlación con los métodos subjetivos, PESQ por ejemplo presenta un grado de correlación con el método DMOS cercano a 0,95 comparado con un 0,7 del Modelo E, resultados obtenidos por Casas (2005). El inconveniente de los métodos intrusivos es que para estimar la calidad percibida hacen uso completo del canal de transmisión haciendo inviable su implementación como sistema de medición en línea, además que su desempeño está sujeto a la correcta alineación de la señal comparada con la señal patrón, en esto último el método PESQ es el más robusto de los métodos mencionados, los métodos como EMBSD y MNB que no tienen en cuenta este efecto comparan partes distintas de la señal y consideran como distorsión segmentos de la señal que simplemente están desfasados, dando un resultado poco confiable.

Los métodos no intrusivos, no requieren de señales extra para estimar la calidad percibida, lo que los hace candidatos potentes para monitorear la calidad percibida de la voz. De estos métodos, el que presentan un mejor desempeño en su correlación con los métodos subjetivos, es la implementación de una red de neuronas “*feed-forward*” que implemento Mohamed (2003) presentando un grado de correlación de 0.93; su principal inconveniente es la fuerte dependencia de su desempeño con la etapa de entrenamiento, por ejemplo de los resultados obtenidos por Mohamed, la correlación entre con el MOS fue en un rango entre [0,71 – 0,93], el valor más bajo fue para un entrenamiento con 56 configuraciones. y el más alto, con 102 configuraciones, esto hace necesario crear en la fase de entrenamiento la mayor cantidad de escenarios posibles, lo que dificulta su implementación. EL Modelo E siendo un modelo matemático que fue diseñado originalmente para ser utilizado en la etapa de diseño y planificación telefónica tiene un desempeño no menor, cercanos a 0,7 Casas (2005); su fácil implementación y facilidad de configuración se ha demostrado en De Lima (2005) con su implementación en el software METRONET.

Hoene (2005) y Sun (2004) presentaron modelos híbridos “PESQ-Modelo E” buscando mejorar el comportamiento del Modelo E. Estos métodos calculan las degradaciones de equipo del sistema a través de PESQ obteniendo medidas más fiables que sirven de entrada al Modelo

E, mejorando por supuesto el desempeño de este último, pero a la vez introduce mayores dificultades a la hora de implementar el método.

Por último, podemos decir que si bien es cierto existe un extenso trabajo en el desarrollo de métodos de estimación de la calidad percibida de la voz, estos están enfocados al diseño, planificación de sistemas de voz o *troubleshooting*. Una excepción es la aplicación de De Lima (2005) que diseñó un sistema de medición en línea que obtiene los parámetros funcionales desde los ticket de terminación de llamadas; sin embargo, el trabajo realizado hasta ahora no considera la etapa posterior a la implementación de un método de estimación, es aquí donde creemos que la aplicación de gráficos de control presentará un real aporte en la vigilancia, reducción de la variabilidad y estimación de los parámetros funcionales de la red que definen la calidad percibida de la voz sobre IP.

• Gráficos de control Multivariantes

Los gráficos de control son una importante herramienta utilizada en la administración y control estadístico³ de la calidad de procesos, cual se representan los valores de algún tipo de medición sobre las características de calidad a partir de una muestra durante el funcionamiento de un proceso continuo, los gráficos de control tienen tres usos fundamentales:

- Rastreo y vigilancia del proceso.
- Reducción de la variabilidad del proceso.
- Estimación de los parámetros y capacidad del proceso

Uno de los objetivos más importantes del control estadístico de procesos es detectar rápidamente la ocurrencia de cambios en el proceso, a fin de que se puedan investigar y tomar acciones sobre éste. En cuanto al funcionamiento de los gráficos de control, existe una estrecha relación entre estos y la prueba de hipótesis estadística. Esencialmente, los gráficos de control son pruebas de hipótesis para cada muestra del proceso a lo largo del tiempo, comprobando que éste se encuentra bajo control estadístico.

En los procesos de producción actuales y en particular para controlar la calidad percibida de voz es necesario vigilar más de una característica de calidad simultáneamente. Una primera posibilidad sería utilizar gráficos de control univariantes para cada una de las características de calidad; sin embargo, sería imposible controlar el efecto de las interacciones entre ellas, así como su variación a lo largo del tiempo; en términos estadísticos, se estaría obviando la información contenida en las covarianzas y su estabilidad (Alfaro, 2004). Dentro los gráficos multivariantes tenemos: el Gráfico de Control T^2 de Hotelling y los llamados gráficos de control con memoria, MCUSUM (*multivariable cumulative sum*) y el MEWMA (*multivariable exponentially weighted moving average*).

El Gráfico T^2 de Hotelling es una extensión multivariable del gráfico de control de Shewhart (1931) y fue introducido por Hotelling (1947), de la misma forma, el gráfico MCUSUM fue propuesto por Crosier (1988) y MEWMA por Lowry (1992), y son una extensión multivariable de los gráficos de control CUSUM, Page (1961) y EWMA, Roberts (1959) respectivamente.

Si se comparan los gráficos antes mencionados, los gráficos de T^2 de Hotelling al igual que el gráfico de control de Shewhart calcula su estadístico de la información obtenida en la última muestra del proceso, el principal inconveniente de esto es que lo hace incapaz de detectar

³ Desde el punto de vista de la variabilidad de un proceso, cuando éste sólo funciona con causas comunes de variabilidad, se considera bajo control estadístico.

pequeños cambios en el proceso, para esto ultimo los gráficos de control con memoria utilizan en el calculo de su estadístico información proveniente de periodos pasados, haciéndolos más sensibles para detectar pequeños cambios en el vector de medias del proceso. A su vez los gráficos con memoria son mas sensibles al error de tipo II y menos apropiados para detectar cambios de gran magnitud que los gráficos T^2 de Hotelling, por tanto la elección del tipo de grafico va a estar dado por las necesidades del proceso, es decir, dependiendo del proceso y magnitud de cambio que se quiera detectar es el gráfico de control que se debe usar, existirá un compromiso entre la magnitud de los cambios que se desea detectar y la robustez del gráfico.

En el uso de los gráficos de control multivariantes, el mayor inconveniente es la interpretabilidad, el gráfico sólo indica que se ha producido un cambio en el proceso, pero no hace explícito cuál o cuáles de las variables son las que originan este cambio. Para esto último Avendaño (2003) presentan una combinación de gráficos de control univariantes y multivariantes aplicados a un proceso, además de la utilización de una Red Neuronal que permite detectar que variables han cambiado en el caso en que no se hace explícito en el gráfico de control univariable correspondiente.

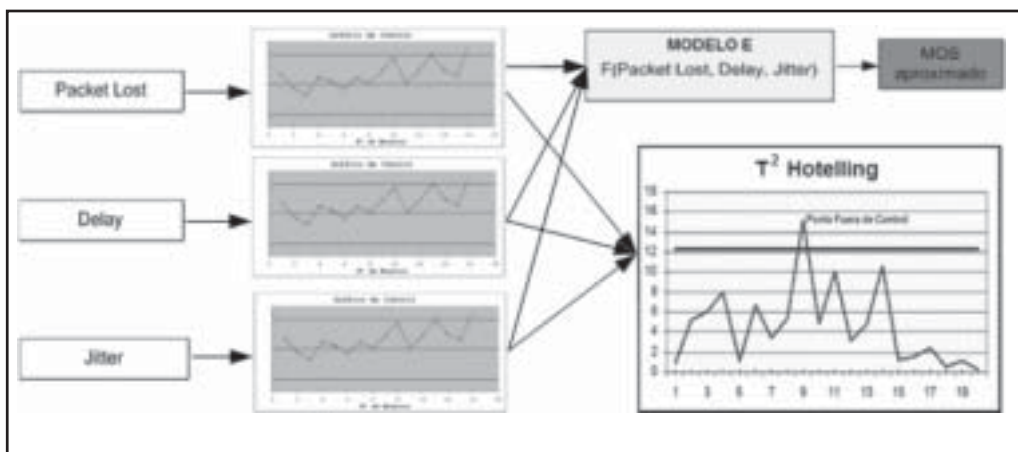
2. TRABAJO REALIZADO

El modelo propuesto en esta investigación consiste en la implementación conjunta de un sistema de control estadístico en línea utilizando gráficos de control aplicados a los parámetros funcionales de red, los que a su vez sirven de entrada al método de estimación de la calidad percibida. Esta estrategia conjunta nos permitirá conseguir un sistema de evaluación en línea de la calidad percibida de voz para usuario VoIP.

Para estimar la calidad percibida de voz se utiliza el Modelo E, este método resulta ser fácil de implementar en forma permanente y de acuerdo a los resultados obtenidos por De Lima (2005) y Casas (2004) presenta buenos resultados en la estimación de la calidad percibida.

Dada las características de proceso, se requiere detectar cambios mayores a dos desviaciones estándar (Oliveira, 2005), por tanto como mecanismo de de estimación y vigilancia de los parámetros funcionales de red, se utilizan en conjunto p gráficos univariantes y un gráfico multivariante, el gráfico T^2 de Hotelling, siendo estos últimos mas robustos que los gráficos con memoria CUSUM o EWMA.

Figura 2: Modelo de medición propuesto



Modelo E

El Modelo E, definido en la Recomendación G.107 de la ITU-T, se basa en el método del factor de degradación, introducido por J. Allnatt (1983) y tiene como resultado un factor llamado índice de transmisión R , que combina todas las degradaciones pertinentes para la conexión considerada. Este factor R está constituido por:

$$R = R_0 - I_S - I_d - I_{e\text{-eff}} + A$$

- R_0 Representa los efectos del ruido del ruido de la conexión.
- I_S es una combinación de todas las degradaciones que aparecen de forma más o menos simultánea con la señal vocal.
- I_d es la suma de las degradaciones producidas por el retardo.
- $I_{e\text{-eff}}$ es el factor de degradación efectiva del equipo, tiene en cuenta la distorsión de la voz producida por el CODEC de velocidad binaria baja y el efecto de la pérdida de paquetes.
- El factor de mejora A permite compensar los factores de degradación cuando existen otras ventajas de acceso para el usuario.

El factor de determinación de índices de transmisión R puede variar entre 0 y 100, donde $R = 0$ representa una calidad extremadamente mala y $R = 100$ representa una calidad muy alta, la siguiente ecuación muestra la relación de R con la medida subjetiva MOS.

Para $R < 0$: $MOS = 1$

Para $0 \leq R \leq 100$: $MOS = 1 + 0,035 \cdot R + R \cdot (R - 60) \cdot (100 - R) \times 7 \cdot 10^{-6}$

Para $R > 100$: $MOS = 4,5$

Implementación del Modelo.

En el Modelo E se definen alrededor de 15 ecuaciones que representan las distintas degradaciones consideradas en el modelo de conexión de referencia que se utiliza para estimar la calidad percibida. A continuación se presentaran las relaciones y consideraciones que se llevaron a cabo en la implementación del método, los valores para los parámetros del modelo que no tienen relación directa con los parámetros funcionales de la red se fijaron de acuerdo a la recomendaciones G107 y G.108 de la ITU-T.

- **Calculo de R_0**

El parámetro R_0 representa los efectos del ruido del ruido de la conexión, para el caso de voz sobre IP el efecto del ruido solo esta presente en los equipos terminales.

$$R_0 = 94,7297$$

- **Calculo de I_S**

El factor de degradación I_S es la suma de todas las degradaciones que pueden producirse, de forma más o menos simultánea con la transmisión de señales vocales, este se divide en tres factores de degradación específicos.

$$Is(T, qdu) = Iolr + Ist(T) + Iq(qdu)$$

El factor $Iolr$ representa la disminución de calidad producida por valores demasiado bajos de OLR . El factor Iq representa la degradación producida por R_0 y la distorsión de cuantificación qdu , El factor Ist representa la degradación producida por efectos locales no óptimos y el retardo de la señal.

- **Factor de degradación por retardo, Id**

El factor de degradación Id , que representa todas las degradaciones debidas al retardo de las señales vocales, este se divide en tres factores de degradación específicos.

$$Id(T) = Idte + Idle + Idd(T)$$

El factor $Idte$ e $Idle$ expresan una estimación para las degradaciones debidas al eco para el hablante y oyente respectivamente, para el caso de voz sobre IP no son relevantes y se asocian a equipos terminales analógicos. El factor Idd representa la degradación producida por retardos absolutos demasiado largos $T > 100ms$, para $T \leq 100ms$ $Idd = 0$.

- **Factor de degradación del equipo, I_{e-eff}**

Este factor es el mas importante a la hora de utilizar el modelo E, el valor de este factor de degradación efectiva de equipo depende directamente del CODEC utilizado y del porcentaje de paquetes perdidos, los valores especificos han sido tabulados de acuerdo a pruebas subjetivas y se presentan en la recomendación G.113 de la ITU-T.

El factor de degradación efectiva del equipo I_{e-eff} , se obtiene utilizando el valor del factor degradación I_e específico para el CODEC con pérdida de paquetes cero, la probabilidad de pérdida de paquetes Ppl el factor de robustez contra pérdida de paquetes Bpl que depende directamente del CODEC utilizado. En este caso, $BurstR = 1$, ya que el sistema es en condiciones de pérdida de paquetes aleatoria.

$$I_{e-eff}(I_e, Ppl, Bpl) = I_e + (95 - I_e) \cdot \frac{Ppl}{\frac{Ppl}{BurstR} + Bpl}$$

- **Factor de mejora, A**

El factor de mejora A permite compensar los factores de degradación cuando existen otras ventajas de acceso para el usuario, para nuestra investigación se consideró una conexión alámbrica, por lo tanto se considero $A = 0$ de acuerdo a la referencia G.107 de la ITU-T.

Gráfico T^2 de Hotteling

Considerando un proceso con p variables a controlar, un vector de medias $\mu_0 = [\mu_{01} \mu_{02} \dots \mu_{0p}]$ y matriz de varianzas y covarianzas Σ_0 de orden p . El estadístico T^2 de Hotteling en el instante i se calcula de acuerdo a la siguiente ecuación:

$$T_i^2 = n \cdot (\bar{X}_i - \mu_0) \cdot \Sigma_0^{-1} \cdot (\bar{X}_i - \mu_0) \quad \text{con} \quad \bar{X} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{bmatrix} \quad \bar{x}_j = \frac{1}{n} \sum_{k=1}^n x_{k,j} \quad j = 1, 2, 3, \dots, p$$

Donde \bar{X}_i es el vector de medias muestrales de dimensión $p \times 1$ y tamaño de muestra n para las p -características, en la i -ésima muestra. Si μ_0 y Σ_0 son conocidos y el proceso está bajo control, el estadístico T_i^2 se distribuye como una chi-cuadrado central con p grados de libertad ($T_i^2 \approx \chi_p^2$), este estadístico corresponde al ensayo de hipótesis: $H_0: \mu_i = \mu_0$ vs. $H_1: \mu_i \neq \mu_0$. Cuando el proceso está bajo control ($\mu_i = \mu_0$), existe una probabilidad α de que el estadístico T_i^2 exceda al valor crítico $\chi_{p,\alpha}^2$, probabilidad de error de tipo I. Cuando el proceso no está bajo control, el estadístico T_i^2 se distribuye como una variable chi-cuadrado no central con p grados de libertad y con parámetro de no centralidad λ . Por ello se toma como indicación o señal de “fuera de control” un valor del estadístico T_i^2 que supera al valor $\chi_{p,\alpha}^2$. Por tanto, el límite de control superior en el gráfico T^2 será: $LSC = \chi_{p,\alpha}^2$ y el límite inferior igual a 0.

- **Número de variables a controlar**

La interpretabilidad y del gráfico está inversamente relacionada con la dimensionalidad del proceso, es decir, mientras menos variables se pretendan controlar, más fácil será interpretar el gráfico y éste será a su vez más estable. Es por esto, que se hace necesario un esfuerzo adicional en la determinación del número de variables o en la reducción de éste mediante técnicas estadísticas.

Una técnica para lograr esto, es la que se conoce como “componentes principales”. Ésta técnica de reducción de la dimensión nos va a permitir transformar las variables originales en una combinación lineal de las mismas, obviamente de menor dimensión, sin que esto suponga una excesiva pérdida de información.

- **Elección del tamaño de la muestra n**

El tamaño de muestra n está directamente relacionado con el parámetro de no centralidad λ , lo que se traduce que en la medida en que si aumentamos el tamaño de la muestra, se disminuye el error de Tipo II y adicionalmente se asegura la normalidad de los datos muestrales⁴.

$$\lambda = n \cdot d^2 \quad \text{con} \quad d = \sqrt{(\mu_1 - \mu_0)' \cdot \Sigma_X^{-1} \cdot (\mu_1 - \mu_0)}$$

⁴ Teorema central del limite

d es la distancia de Mahalanobis, que representa la magnitud del cambio entre el vector de medias original μ_0 y el nuevo vector de medias μ_1 .

- **Cálculo de μ_0 y Σ_0**

Si los valores μ_0 y Σ_0 son desconocidos, es necesario estimarlos a partir de m muestras de tamaño n tomadas desde el proceso bajo control.

- **Estimación de μ_0**

$$\bar{X} = [\bar{X}_1 \quad \bar{X}_2 \quad \dots \quad \bar{X}_p] , \text{ siendo } \bar{X}_j = \frac{1}{m} \sum_{k=1}^m \bar{X}_{jk}, \quad j = 1, 2, \dots, p$$

\bar{X}_k es el vector de medias muestrales, calculadas de la Ecuación 3.20 para m muestras.

- **Estimación de Σ_0**

La varianza muestral se calcula de acuerdo a la siguiente ecuación

$$S^2_{jk} = \frac{1}{n-1} \sum_{i=1}^n (x_{ijk} - \bar{x}_{jk}), \quad \text{para } \begin{matrix} j = 1, 2, \dots, p \\ k = 1, 2, \dots, m \end{matrix}$$

Donde X_{ijk} es la i -ésima observación de la j -ésima característica de calidad en la k -ésima muestra. La covarianza, entre la característica de calidad j y la característica de calidad h , en la k -ésima muestra es:

$$S_{jhk} = \frac{1}{n-1} \sum_{i=1}^n (x_{ijk} - \bar{x}_{jk})(x_{ihk} - \bar{x}_{hk}) \quad \text{para } \begin{matrix} j = 1, 2, \dots, p \\ k = 1, 2, \dots, m \\ j \neq h \end{matrix}$$

Luego se calculan los promedios de las estadísticas S^2_{jk} y S_{jhk} para las m muestras, con el fin de obtener el estimador insesgado de Σ_0 , con lo que finalmente obtenemos la matriz S de dimensiones $p \times p$:

$$\begin{aligned} S^2_j &= \frac{1}{m} \sum_{k=1}^m S^2_{jk}, \quad \text{para } j = 1, 2, \dots, p \\ S_{jh} &= \frac{1}{m} \sum_{k=1}^m S_{jhk}, \quad \text{para } j \neq h \end{aligned} \quad S = \begin{bmatrix} S^2_1 & S_{12} & S_{13} & \dots & S_{1p} \\ \cdot & S^2_2 & S_{23} & \dots & S_{2p} \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & S^2_p \end{bmatrix}$$

Finalmente sustituyendo los resultados obtenidos el cálculo del estadístico T^2 en el instante i se calcula de acuerdo a la siguiente ecuación

$$T^2_i = n \cdot (\bar{X}_i - \bar{\bar{X}}) \cdot S^{-1} \cdot (\bar{X}_i - \bar{\bar{X}})$$

- **Cálculo de los Límites de Control del Gráfico.**

Para datos agrupados, el estadístico T^2 de la ecuación anterior sigue una distribución F de Snedecor, por lo que los límites de control van a ser calculados de la siguiente manera:

$$LSC = \frac{p(m-1)(n-1)}{(mn - m - p + 1)} \cdot F_{\alpha, p, mn-m-p+1} \quad LIC = 0$$

- **Consideraciones adicionales**

En la práctica hay que tener especial cuidado con el cumplimiento de las hipótesis de partida, ya que no hay que olvidar que de partida se supone que las observaciones van a seguir una distribución normal, o aproximadamente normal e independientes. La falta de cumplimiento de estas hipótesis constituye uno de los principales problemas que nos podemos encontrar cuando se lleva a cabo la aplicación práctica de estos métodos estadísticos de control de la calidad.

La correlación cruzada existente entre las variables, va a ser recogida mediante la matriz de varianzas y covarianzas, utilizada por el gráfico; sin embargo para la autocorrelación habrá que recurrir a un análisis complementario que nos permita detectarla y, en caso afirmativo, intentar eliminarla. La principal consecuencia del incumplimiento de esta hipótesis consiste en un aumento del número de falsas alarmas dadas por el gráfico, situación bastante común en la práctica.

Escenario de Pruebas del Modelo

Para realizar las pruebas se dispuso de 3000 llamadas con sus respectivos valores de Jitter, porcentaje de paquetes perdidos y retardo y CODEC G.729.

Los valores para los parámetros asociados al CODEC se obtienen de acuerdo a la tabla I.3 de la recomendación G.113 de la ITU-T

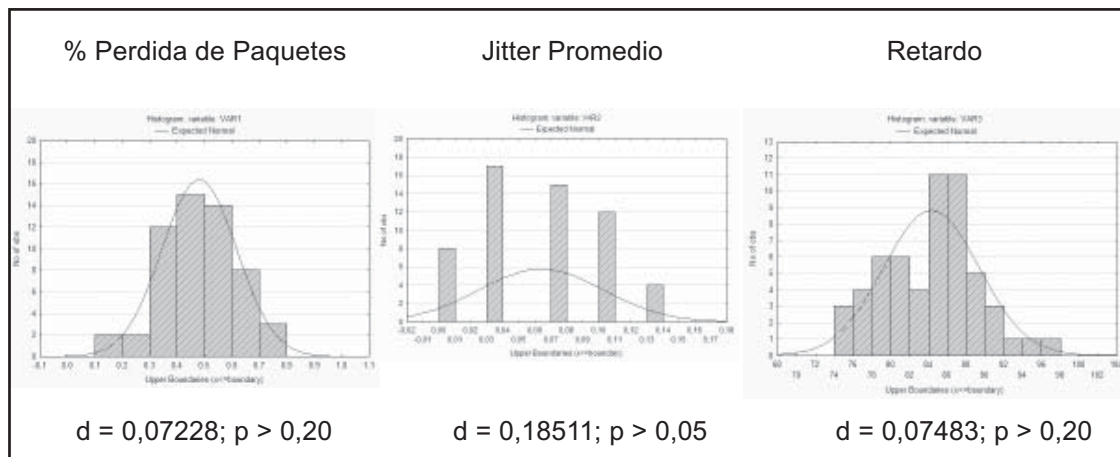
Nombre	Descripción	Valor
qdu	Número de unidades de distorsión de cuantificación	1.7
Ie	Factor de degradación de equipo	15
Bpl	Grado de robustez del CODEC ante la pérdida de paquetes	16,1

Para estimar los parámetros necesarios para implementar los gráficos de control se utilizaron los datos de 1700 llamadas del sistema bajo control con un tamaño de muestra de 30.

Se define X_1 = Porcentaje de Pérdida de Paquetes, X_2 = Jitter en ms y X_3 = Retardo en ms; los parámetros calculados son los siguientes:

Para verificar la normalidad de los datos se realizó el test de normalidad de Kolmogorov-Smirnov a las características de calidad seleccionadas, ver figura 3

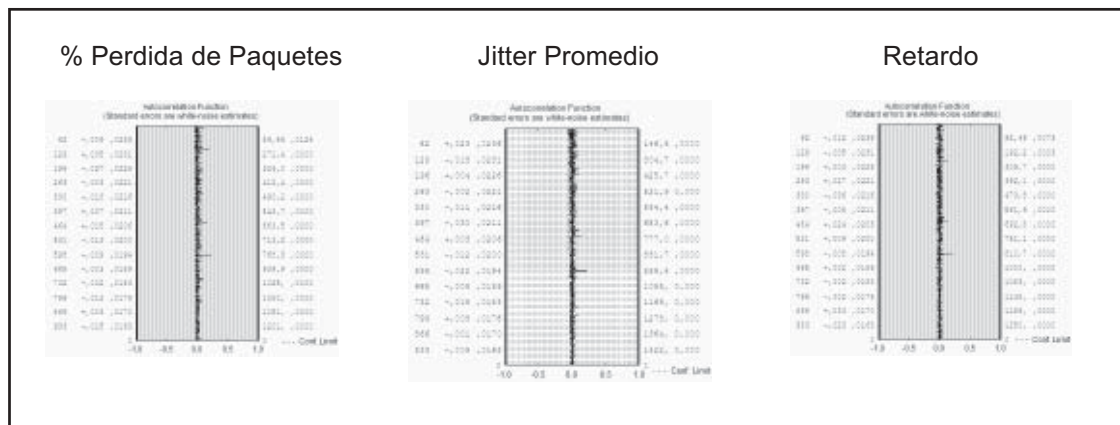
Figura 3: test de Normalidad de Kolmogorov-Smirnov



Podemos observar que la variable Jitter no pasa el test de normalidad, habrá que tener cuidado en el comportamiento del Gráfico de Control respectivo, ya que pueden existir falsas alarmas o bien que los datos utilizados para calcular los parámetros no son suficientes.

Otro aspecto importante es el estudio de la autocorrelación de los datos, para esto se utilizó el programa estadístico STATISTICA 6.0, obteniéndose los siguientes resultados.

Figura 4: Análisis de la función de auto correlación para las características de calidad



Como se puede ver en la Figura 4, para las características de calidad de nuestro sistema no existen elementos que muestren algún grado de dependencia o estacionalidad de los datos en la función de autocorrelación.

Para el cálculo del estadístico tenemos lo siguiente:

$$\bar{X} = \begin{bmatrix} 0.481 \\ 0.065 \\ 84.25 \end{bmatrix} S^{-1} = \begin{bmatrix} 1.821 & 0.071 & 0.002 \\ 0.071 & 17.874 & -0.001 \\ 0.002 & -0.001 & 0.001 \end{bmatrix}$$

Con lo que se obtiene la ecuación que calcula el estadístico T_i^2 en el instante i

$$T_i^2 = 30 \cdot \begin{bmatrix} \bar{X}_1 - 0.481 & \bar{X}_2 - 0.065 & \bar{X}_3 - 84.25 \end{bmatrix} \begin{bmatrix} 1.821 & 0.071 & 0.002 \\ 0.071 & 17.874 & -0.001 \\ 0.002 & -0.001 & 0.001 \end{bmatrix} \cdot \begin{bmatrix} \bar{X}_1 - 0.481 \\ \bar{X}_2 - 0.065 \\ \bar{X}_3 - 84.25 \end{bmatrix}$$

Para datos agrupados, el estadístico T^2 de la ecuación anterior sigue una distribución F de Snedecor, con nivel de significancia $\alpha = 0.01$:

$$p = 3$$

$$n = 30$$

$$m = 1700 / 30 = 57$$

$$\alpha = 0.01$$

$$LSC = \frac{p(m-1)(n-1)}{(mn - m - p + 1)} \cdot F_{\alpha, p, mn - m - p + 1} = \frac{3 \cdot (57 - 1) \cdot (30 - 1)}{(30 \cdot 57 - 57 - 3 + 1)} \cdot F_{0.01, 3, 30 \cdot 57 - 57 - 3 + 1} = 5.58$$

$$LIC = 0$$

El Algoritmo del Gráfico T^2 de Hotelling y el Modelo E, se implementó en MATLAB 7.0.

3. RESULTADOS

De acuerdo a la hipótesis nula de esta investigación, nuestro objetivo es probar que se puede realizar una medición en línea de la calidad percibida de la voz, para probar la hipótesis de trabajo se realizaron las siguientes pruebas con los 1300 datos restantes :

- Corrida con un cambio en el Retardo
- Corrida con un cambio en el porcentaje de paquetes perdidos
- Pruebas adicionales.

Corrida con un cambio en el Retardo

Esta prueba corresponde a verificar el comportamiento de la estructura de control cuando existe un cambio en la media del retardo y las demás variables se encuentran bajo control. Para realizar esta prueba se introdujo un retardo adicional de 20 ms a esta característica de calidad, desde la muestra 20 hasta la 25 a la media del sistema bajo control, en la figura 5 se muestran los gráficos de control univariados para el porcentaje de paquetes perdidos y el retardo, además del gráfico de MOS y el gráfico multivariado T^2 de Hotelling

Corrida con un cambio en el porcentaje de paquetes perdidos

Esta prueba corresponde a verificar el comportamiento de la estructura de control cuando existe un cambio en la media del retardo y las demás variables se encuentran bajo control. Para

realizar esta prueba se introdujo una variación aleatoria de dos desviaciones estándar desde la muestra 32 hasta la 45 sobre la media de la pérdida de paquetes del sistema bajo control, en la Figura 6 se muestra los gráficos de control univariables para el porcentaje de paquetes perdidos y el retardo, además del gráfico de MOS y el grafico miultivariable T^2 de Hotelling.

Figura 5: cambio en el Retardo

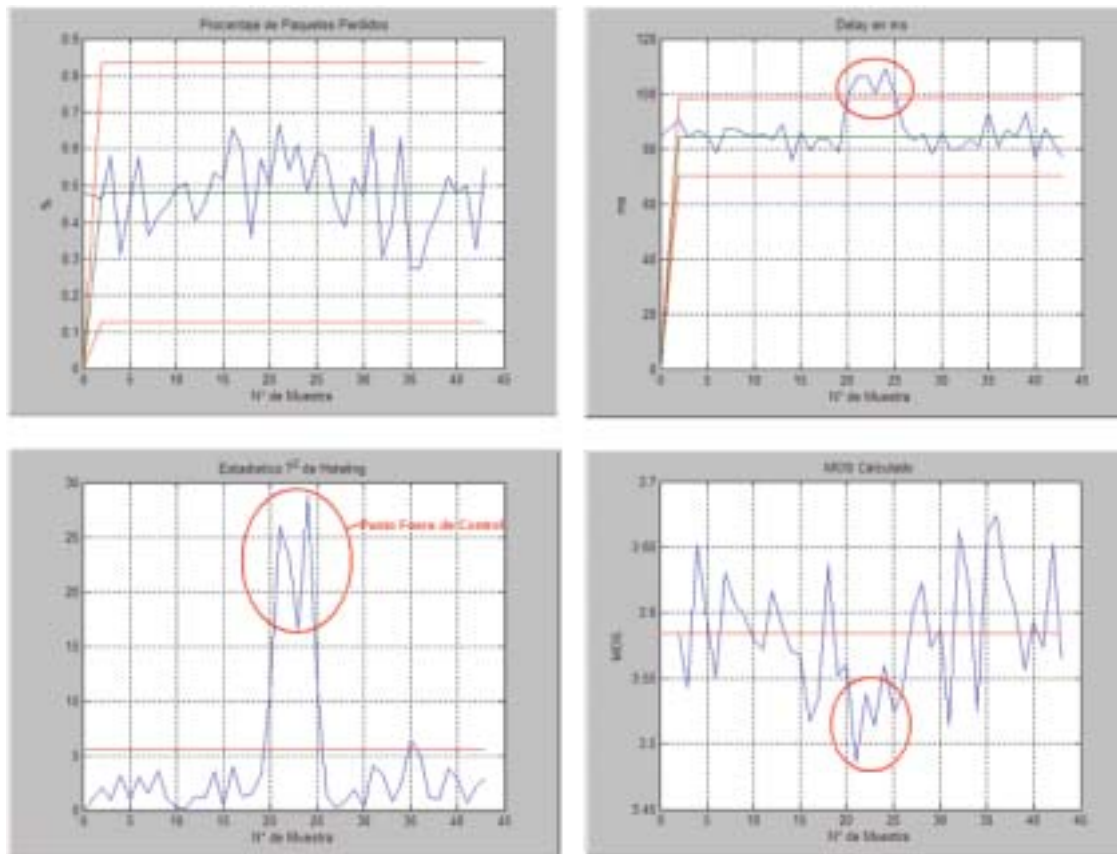
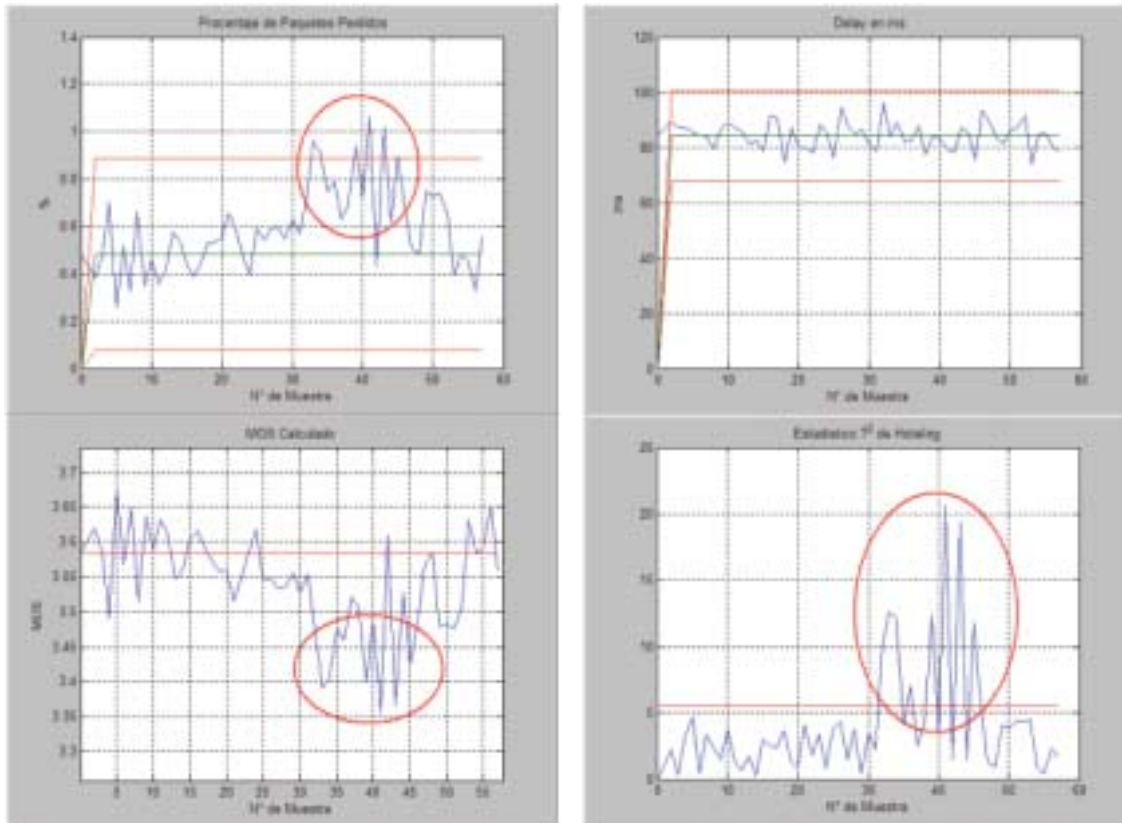
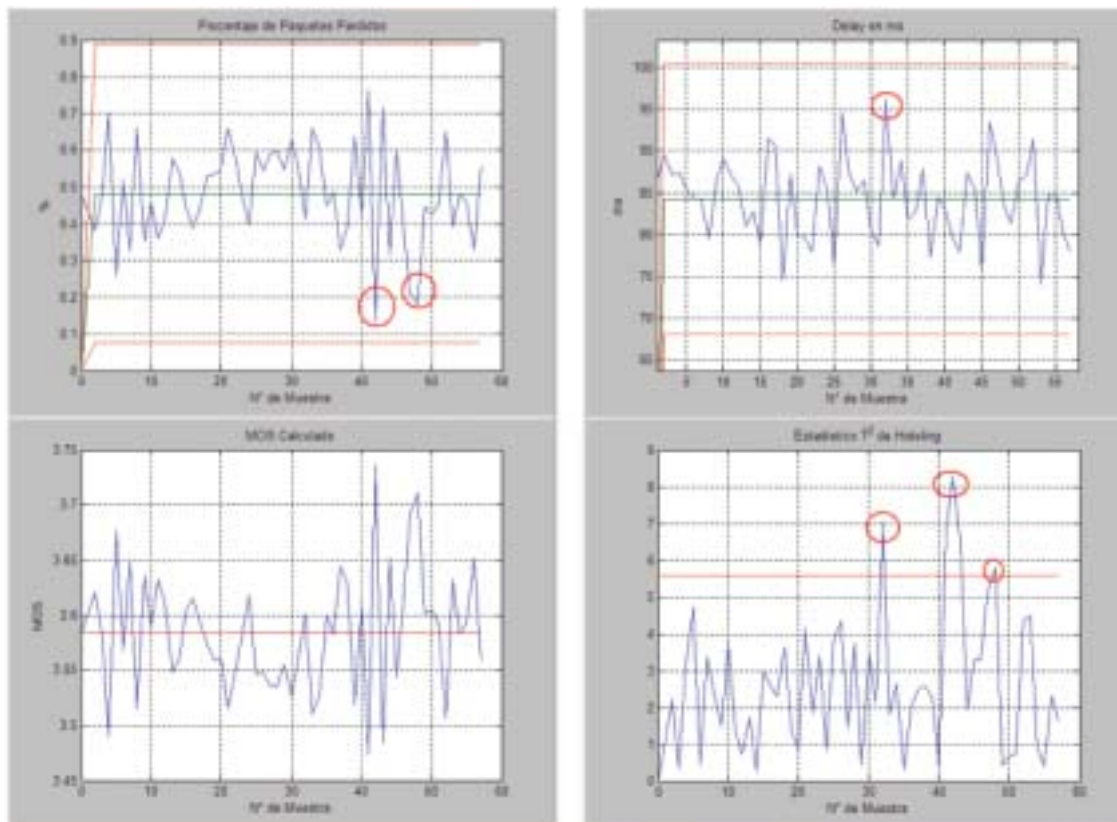


Figura 6: cambio en el Retardo**Pruebas adicionales.**

Como parte de esta investigación se analiza la siguiente corrida de datos que representa modificaciones de la media cercana a las dos desviaciones estándar, de manera de verificar el comportamiento del Gráfico de Control multivariable. La Figura 7 muestra el porcentaje de paquetes perdidos para la corrida de datos de para esta prueba, donde se presentan los puntos 42 y 48 mas halla de dos desviaciones estándar, pero no sobrepasan el límite de control inferior.

Figura 7: cambio en el Retardo



4. CONCLUSIONES

Aquí se exponen las conclusiones más importantes de la investigación, destacando sus principales contribuciones y posibles líneas de continuación del trabajo realizado.

- La investigación realizada nos permite concluir que la combinación de control estadístico de procesos y métodos objetivos para la estimación de la calidad percibida es, sin lugar a dudas, una estructura efectiva en la vigilancia y medición de la calidad percibida de voz.
- Las pruebas realizadas presentaron los siguientes resultados:
 - Es posible obtener una estimación de la calidad percibida de la voz sobre IP a través del Modelo E, utilizando como parámetros de entrada al modelo las medias de las variables principales: Porcentaje de paquetes perdidos, Jitter y Retardo; obtenidas de los Gráficos de Control, que sirven como un estimador insesgado de las medias de las variables.
 - Es posible implementar Gráficos de Control a las variables de red, que permite un esquema de monitoreo en línea de la calidad percibida de la voz sobre IP, lo que permite detectar e investigar causas asignables que afecten al servicio de voz sobre IP.
- El modelo E implementado presenta buenos resultados de estimación de la calidad percibida de la voz; sin embargo, no hay que olvidar que para implementar este algoritmo se hicieron

una serie de supuestos sobre las variables de punto final que deben estar en constante validación.

- El gráfico T^2 de Hotelling, es más sensible que los Gráficos de Control univariables ante cambios en la media del proceso. Pueden generarse puntos fuera de control debido a pequeños cambios en conjunto de más de una variable, debido a la correlación, o cambios de más de dos desviaciones de alguna variable, pero sin sobrepasar los límites de control de su gráfico respectivo.
- Para la implementación de una estrategia de control estadístico, es necesario implementar en conjunto los gráficos de control univariables con los multivariables, ya que los últimos no entregan visualmente información de las medias ni de la evolución de las variables de proceso, si no mas bien las desviaciones y correlaciones existentes entre las variables medidas y las medias de los procesos.

Trabajo Futuro

- Investigar variables adicionales que se puedan medir e incluir a la estructura propuesta; por ejemplo, los parámetros no funcionales de la prestación del servicio, como lo son: porcentaje de reclamos, pérdida total de servicio, SLA's , etc. Con esto generar algún tipo de medida que incluya estos últimos en adición a la calidad percibida de la voz.
- Existen investigaciones que buscan mejorar el comportamiento del Modelo E, desarrollando modelos híbridos entre métodos subjetivos y objetivos de estimación de la calidad, mejorando la correlación entre MOS real y MOS estimado para valores extremos en los parámetros de transmisión, lo que implica profundizar aún más en distintas relaciones existentes entre las variables de punto final y las de transmisión.

5. BIBLIOGRAFÍA

- ALFARO, José Luís (2004); *Métodos Multivariantes en Control Estadístico de La Calidad*. Universidad de Castilla-La Mancha, España.
- Avendaño G., Aparisi F. y Sanz J. (2003) “Una red neuronal para la interpretación de la señal fuera de control del gráfico multivariante T^2 de Hotelling”. Universidad Politécnica de Valencia, España.
- BELLIDO, Luis (2004). *Contribución a las metodologías para la evaluación de la calidad de servicio en redes homogéneas*. Universidad Politécnica de Madrid, Madrid.
- CASAS, Pedro; et.al (2005). *Calidad de Servicio Percibida en Servicios de Voz y Video sobre IP*. Universidad de la República, Montevideo, Uruguay.
- Crosier, R (1988). “Multivariate Generalizations of Cumulative Sum Quality- Control Schemes”. *Technometrics*,30 (3), 291-303.
- DE LIMA, Ana Flávia M; et al (2005). “Monitoring based on statistical analysis for evaluating quality of calls in VoIP environment”. University of Ceará, Brazil.

- HOENE, Christian, Holger Karl, Adam Wolisz (2005), *A Perceptual Quality Model Intended for Adaptive VoIP Applications*, International Journal of Communication Systems, p 1-20.
- Hotelling, H. (1947), "*Multivariate Quality Control, in Techniques of Statistical Analysis*". Eds. C. Eisenhart, M. Hastay y W. A. Wallis, McGraw-Hill. 111-184.
- ITU-T Recommendation P.861 (1996), *Objective quality measurement of telephone-band speech codecs*. Genève.
- ITU-T Recommendation P.862 (2001), *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*. Genève.
- ITU-T, Recommendation G.107 (2005), "*The E-model, A Computational Model for Use in Transmission Planning*", Genève.
- ITU-T Recommendation G.108 (1999), "Application of the E-model: A planning guide." Genève, 1999
- ITU-T, Recommendation G.113 (2002), "*Transmission impairments due to speech processing*," Genève.
- ITU-T Recomendación Y.1540 (2002), "*Servicio de comunicación de datos con protocolo Internet. Parámetros de calidad de funcionamiento relativos a la disponibilidad y la transferencia de paquetes de protocolo Internet.* ", Geneve.
- ITU-T Recomendación Y.1541 (2002), "*Objetivos de calidad de funcionamiento de red para servicios basados en el protocolo Internet*".
- ITU-T Recommendation P.800 (1996), *Methods for subjective determination of transmission quality*. Genève.
- KEAGY, Scott. (2001) *Integración de Redes de Voz y Datos*, Person Educación, Madrid, ISBN: 84-205-3187-1.
- Lowry and Woodall (1992), "*A multivariate exponentially weighted moving average control chart*", Technometrics, 34.
- MOHAMED, Samir; et al. (2003) "*Performance evaluation of real-time speech through a packet network: a random neural networks-based approach*", INRIA/IRISA.
- OLIVEIRA, Adriana,; et al. (2005) *Using Control Charts for Traffic Change's Detection in MPLS Networks*, IEEE Computer Society's 12th Annual International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunications Systems.
- Page, E.S (1961). "*Cumulative Sum Charts*". Technometrics 3, (1), 1-9.
- Roberts, S. (1959). "*Control Chart Tests Based on Geometric Moving Average*," Technometrics, 1, pp. 239–250.

- Shewhart, W (1931). "*The Economic Control of Quality of Manufactured Products*, D. Van Nostrand", New York.
- Sun, Lingfen; Ifeakor, Emmanuel (2004) "*New Models for Perceived Voice Quality Prediction and their Applications in Playout Buffer Optimization for VoIP Networks*". University of Plymouth.
- Voran. S (1998) *Objective Estimation of Perceived Speech Quality Using Measuring Normalizing Blocks*, U.S. Department of Commerce
- Yang, Wonho (2000); *Enhanced Modified Bark Spectral Distortion (EMBSD): An objective speech quality measure based on audible distortion and cognition model*, Thesis (PhD). Temple University

