

CONTROL BORROSO APLICADO A LAS BASES DE DATOS DEL SECTOR TERCIARIO. PRIMER MODELO SIMPLE CAPA PARA EL SUBSECTOR SALUD

Carlos A. Discoli[&], Fernando Romero^{**}

Instituto de Estudios del Hábitat, IDEHAB, U.I. N°2, FAU, UNLP.

Callc 47 N°162. CC 478 (1900) La Plata.

Tel-fax: 021-214705. E-mail: erosenfe@isis.unlp.edu.ar

RESUMEN

El control y diagnóstico temprano de las redes complejas del sector terciario, requiere del desarrollo de una metodología amplia que interprete a los sistemas con exigencias crecientes. Por tal motivo se han desarrollado, adaptado e implementado teorías y técnicas borrosas compatibles a la determinística, que permitan evaluar e inferir escenarios alternativos en función de variables estructurales y críticas. En éste trabajo se desarrollan los avances y dificultades encontradas.

INTRODUCCIÓN

El sector terciario y en este caso el subsector salud, analizado como un sistema complejo de redes productoras de servicios de gran heterogeneidad, abarca un universo importante de variables críticas y estructurales, con una dinámica de múltiples entradas y salidas lineales, no lineales, invariantes o variables en el tiempo.

La amplitud del campo experimental requirió la elaboración de una metodología abarcativa, que sistematice la información y permita generar diagnósticos tempranos de los sectores en estudio, utilizando herramientas de análisis determinista y de lógica borrosa. Inicialmente se elaboró un sistema de control y diagnóstico temprano, publicado en encuentros anteriores (C.Discoli et al 1993a¹ y 1993b²), el que ha permitido sistematizar la información, correr rutinas estadísticas descriptivas y determinísticas, generando múltiples salidas de información básica y calificada. Se han desarrollado los primeros índices y perfiles de comportamiento de una red productora de servicios mostrando las características antes detalladas (C.Discoli, 1995a³ y 1995b⁴). Con el objeto de considerar dentro del mismo sistema la información subjetiva de informantes calificados, se ha incorporado como alternativa complementaria la teoría de conjuntos borrosos adoptando una generación de expresiones de valor o calificaciones de cualidades de los entes estudiados (redes edilicias de servicios) expuestos en (A.Bignoli, 1993⁵, C.Discoli, 1995c⁶ y C.Discoli y F.Romero, 1997⁷), flexibilizando y enriqueciendo al sistema. Este trabajo muestra los cambios metodológicos y los avances realizados en la integración de las disciplinas y del modelo piloto planteado en C.Discoli y F.Romero, 1997.

METODOLOGÍA

Se mejora la metodología de análisis, trabajando en serie-paralelo y en forma iterativa, incorporando la lógica determinista y borrosa en el diseño de un sistema de diagnóstico integrado. La figura 1 muestra el nuevo esquema síntesis de la propuesta.

La información sistematizada puede ser analizada por ambas vertientes, además de utilizar las salidas de una de las áreas de análisis como entradas de la otra. Las vertientes determinística y borrosa pueden analizar las variables objetivas y subjetivas significativas del sistema y provenientes de informantes calificados de los sectores en estudio. En el caso del análisis borroso, a cada variable se considera en particular, se le asigna una valoración lingüística y un grado de pertenencia o grado de aceptación para rangos definidos discretamente. El estado de relación entre las mismas se define por medio de pautas (reglas) que se elaboran en función del conocimiento y/o comportamiento previo entre variables y la experiencia acumulada. La construcción de las reglas no contempla una lógica y una metodología definida, requiriendo solo consistencia, coherencia y evitar la duplicidad.

El tratamiento de la información, comprensión de la situación, obtención de diagnósticos y generación de escenarios dependerá de la cantidad de variables, de las reglas y sus interacciones y del tamaño del controlador difuso. En el modelo planteado en (⁷), las reglas incorporadas han generado distorsiones en las salidas del modelo. Los problemas se detectaron en la definición, en la probabilidad del disparo y corrida de las mismas. Para su verificación y corrección se han utilizado los perfiles de caracterización de las variables desarrollados en C.Discoli, 1995b y precisando con mayor detalle las frecuencias de distribución de los rangos discretos de las variables analizadas, con el objeto de corregir el proceso en la etapa de fuzzificación (Esto es adoptar grados de aceptación o de pertenencia precisos para las variables lingüísticas planteadas).

Al controlador difuso del modelo simple capa presentado en C.Discoli y F.Romero, 1997. se le incorporaron las modificaciones mencionadas, y se está ensayando con las variables camas, superficie, tipología y producción. Las salidas que se están obteniendo, en este caso de energía requerida, son equiparables a las del sistema real (nodos de la red de salud).

[&] Investigador CONICET.

^{**} Prof. Apoyo. FAU-UNLP

Nota Técnica

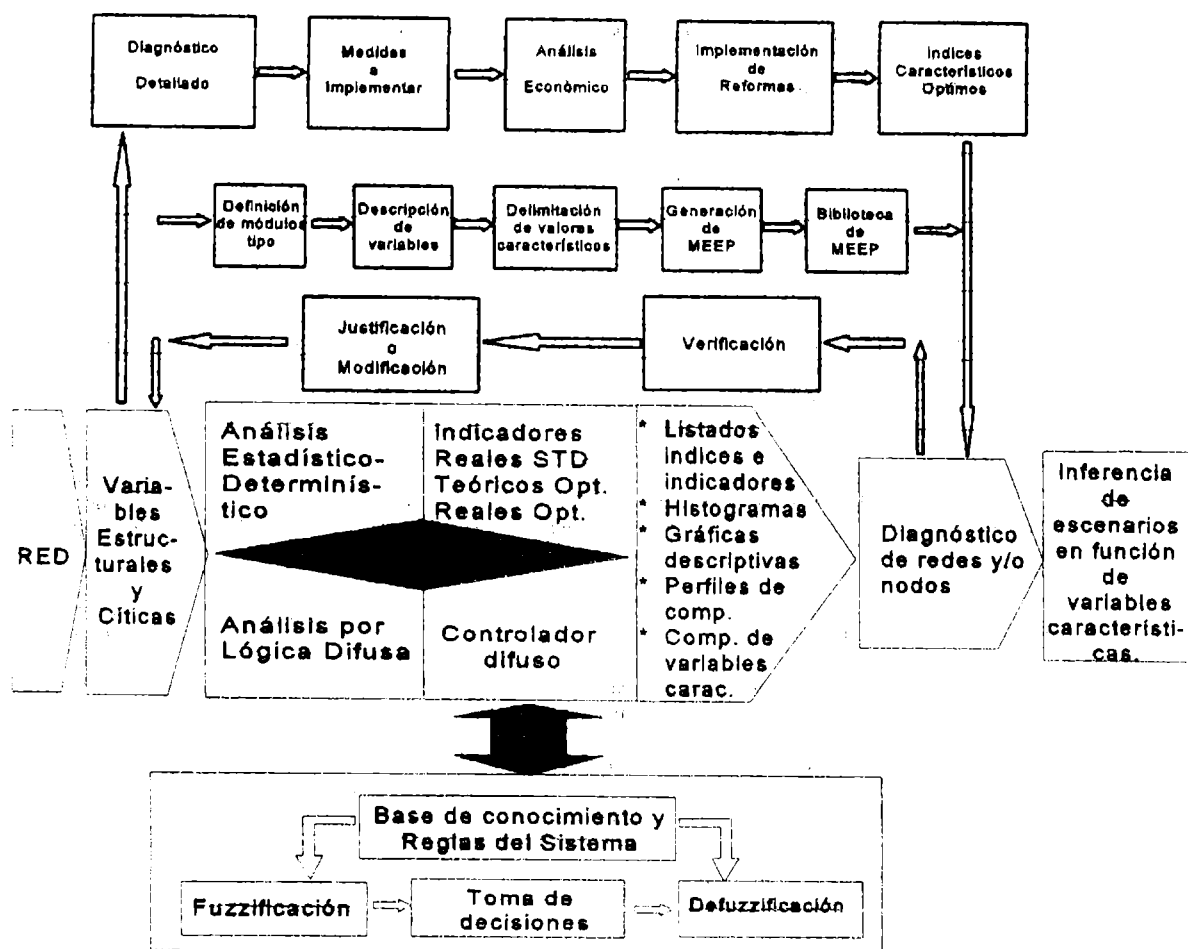


Figura 1: Esquema simplificado de la metodología propuesta

El modelo simple capa apunta a dar respuestas a escala global, acorde al sistema de control y diagnóstico propuesto para el sector en estudio. Considera macrovariables particionadas como en la metodología de diagnóstico presentada en C. Discoli et al. 1993a. (esto es localización, sanitarias, edicias, etc.), involucrándolas con la infraestructura funcional del sector de servicio tratado (administración, Internación, Consultorios Externos, Laboratorio y diagnósticos y CAP) y orientando las salidas según los requerimientos del caso (energía fig.2). Incorporar la diversidad de variables mencionadas y abordarlas detalladamente Y. Rosenfeld et al. 1997⁸, implica generar modelos multicapa, en el cual estas salidas pasarían a ser entradas de otras etapas posteriores. Para ello se emplearían técnicas de redes neuronales. Estas son técnicas que permiten enlazar módulos de programa entre sí, de tal manera que las distintas salidas sean entradas de otros programas, con un determinado peso, lográndose una arquitectura similar a la de las neuronas en el cerebro humano, en la cual la salida del axón está determinada por las distintas excitaciones que llegan a las dendritas provenientes de otros axones pertenecientes a otras neuronas.

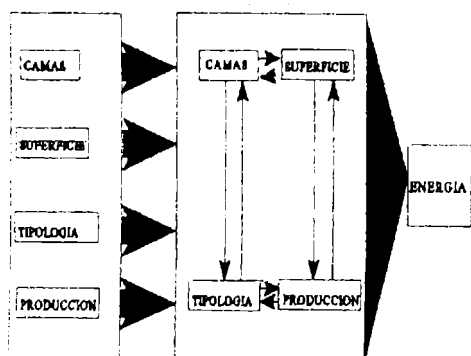


Figura 2: Modelo simple capa a escala global

En un modelo multicapa, inserto dentro del sistema integral, la modificación de una variable de entrada, (ej: número de camas), para un análisis detallado, repercute prácticamente en todas las áreas planteadas en la partición funcional, debiendo considerar el nivel de capacidad ociosa de cada una de ellas y el grado de cambios que soporta manteniendo sus indicadores de eficiencia en rangos aceptables en la prestación de servicios. Si quisiéramos representar gráficamente esta situación (la instancia camas) partiríamos de un diagrama combinado serie- paralelo, semejante a las interrelaciones de las neuronas. La fig.3 muestra una aproximación del diagrama combinado.

El modelo piloto planteado cuenta con:

- i. **Variables de entrada:** es el conjunto mínimo de variables que describen el estado del sistema. Este conjunto de variables define el espacio de entrada y el de salida.
- ii. **Fuzzificación:** es el proceso mediante el cual los valores que definen un estado, originalmente discretos, son convertidos en conjuntos difusos. Para ello, seguimos los siguientes pasos: a) Transformamos sus rangos (obtenidos en estudios estadísticos) en universos de discurso. b) En el caso de ser un universo continuo, lo discretizamos. c) Particionamos en conjuntos difusos (asignamos funciones de pertenencia). d) Asignamos etiquetas lingüísticas a los distintos conjuntos difusos tales como bajo, mediano y alto (ver figuras 6 y 7 del modelo). Los valores discretos que entren a este proceso serán convertidos en conjuntos difusos.
- iii. **Base de conocimientos:** Está compuesto por las variables lingüísticas y las reglas. Esta base debe cumplir con ciertas condiciones, a saber: a) **Completitud:** La unión de las funciones de pertenencia de los conjuntos difusos debe presentar un valor mínimo mayor o igual que 0.5, de tal manera que siempre habrá una regla con un grado significativo de pertenencia (peso). El incumplimiento de esta condición, como también las instancias fuera de rango, implica una falta de nivel mínimo de información, condición bajo la cual no se podrán tomar decisiones. Esta situación implicaría modificar la base de conocimiento para tener este nivel. b) **Consistencia:** Se debe evitar que dos reglas se contradigan. O sea que para una misma entrada, se disparen varias reglas con igual peso pero con salidas distintas. c) **Interacción entre reglas:** definida por el tipo de implicancia usada (Zadeh, Sugeno, Mandani).
- iv. **Base de reglas:** Contiene reglas del tipo "if A1 is a1 and A2 is a2...then V1 is v1", donde "Ai" son las entradas, "ai" las etiquetas lingüísticas, "Vi" la salida, y "vi" la etiqueta de la salida. Ellas permiten introducir en forma simple conocimiento proveniente de estudios anteriores o de experiencia de personas expertas.
- v. **Toma de decisiones:** Para la toma de decisiones, se utiliza la lógica de inferencias y la regla composicional de inferencias. Se basa en inferir el estado de una variable a partir del estado de otra. Se trata de emular el funcionamiento de la inteligencia humana (inteligencia artificial), que tiene como virtud, según Zadeh, atribuir grados de pertenencia a un conjunto de un elemento dado en forma rápida y sencilla (A. Bignoli, 1993). En este módulo se define que reglas se activan y con que peso.
- vi. **Defuzzificación:** Es el proceso mediante el cual se obtiene a partir de una variable difusa una salida discreta. Esto se hace por distintos métodos: a) Criterio del máximo: Es el valor del universo de salida con máximo valor de pertenencia. Si hay más de un valor no queda determinado. b) Criterio del medio de máximos: Toma el promedio de los máximos. Se desperdicia información, pues no tiene en cuenta la forma de las funciones de pertenencia de salida, la cual se pierde. c) Criterio del centro del área: Se usa el centro de gravedad de la superficie bajo la curva de la unión de las funciones de pertenencia de salida.

CORRECCIONES A REALIZAR EN EL MODELO PILOTO

Se ha trabajado con variables estructurales y algunas críticas, según el escenario de coyuntura, generando un sistema de una capa, con el objeto de definir los estados energéticos en función de diversos requerimientos como el clima, la tipología, la producción, la complejidad, el número de camas y los presupuestos (variables de entrada del sistema). Se detalla en forma sintética, a título de ejemplo, los componentes principales del modelo.

El proceso de fuzzificación de las variables definidas generan los conjuntos difusos con sus grados de pertenencia, contemplando las restricciones mencionadas en el punto iii (completitud, consistencia e interacción). Para definir las funciones de pertenencia debemos especificar que entendemos por las etiquetas que asignamos a cada variable lingüística, y la cantidad de ellas. Para esta tarea se tomaron datos estadísticos, sobre los que se hizo un histograma. Este histograma nos suministra información para:

- * Definir los rangos de los universos: Será el mismo del histograma, ya que no tiene sentido definir etiquetas para casos inexistentes, al menos en esta etapa del proyecto.
- * Definir la cantidad de etiquetas: Debemos localizar la moda y los puntos sobresalientes del histograma. Luego definir un valor de corte. En general, se trató de no definir muchas etiquetas porque ello hace más pesada la tarea de procesamiento al tener necesidad de iterar el sistema en una cantidad mayor de matrices. Esto es debido a que cada etiqueta es un vector definido para todos los puntos del universo.
- * Dar forma a la función de pertenencia: Estará determinada por la forma del histograma a los lados de los puntos sobresalientes del histograma. Dichos puntos representarán el valor 1 para la función de pertenencia, y luego irá bajando dicho valor de acuerdo a los valores de frecuencia. Se debe tener en cuenta que si después de este proceso hay lugares en el universo en los cuales ninguna etiqueta presenta un valor, para dichos valores no suministrará salida el sistema, debido a que no se disparará ninguna regla. Debido a ello es que conviene expandir las funciones de pertenencia para asegurar un valor mínimo de función de pertenencia en el universo. Dicho valor se tomó como 0.5.

Las figuras 4 y 5 muestran los histogramas de algunas de las variables de entrada y las figuras 6 y 7 los diagramas de función de pertenencia de cada variable con el número definido de etiquetas.

Para la definición de las reglas del sistema: Se tomaron valores reales de las variables de entrada y salida de un hospital para cada regla, se ubicó dentro de las respectivas variables lingüísticas para asignarlo a una etiqueta, y se le dio el formato de regla. Para mayor claridad damos un ejemplo: **Hospital:** Interzonal San Martín. **Superficie:** 31.950 m². **Etiqueta a la que pertenece:** alto. **Camas:** 522. **Etiqueta a la que pertenece:** alto. **Producción:** 14.299. **Etiqueta a la que pertenece:** alto. **Tipología (Valor del G):** 2,55. **Etiqueta a la que pertenece:** alto. **Energía:** 7.384.778 Kwh. **Etiqueta a la que pertenece:** alto. En función de las etiquetas la regla quedaría: if Sup. is alto and camas is alto and producción is alto and tipología is alto then energía is alto. Con respecto a la reducida información de alguna de las variables en juego (consumo de energía en los establecimientos chicos), parte de las reglas se construyeron con los perfiles de caracterización, los cuales permiten inferir valores aproximados. En la medida que se amplíen las bases de datos y en consecuencia la base de conocimiento sobre el tema, estas estimaciones serán corregidas.

La redefinición de las reglas no necesariamente contemplan todas las combinaciones posibles entre las variables ya que el grado de dependencia entre ellas no es simétrico. La base de reglas requiere de una construcción cuidadosa, dado que ante una determinada

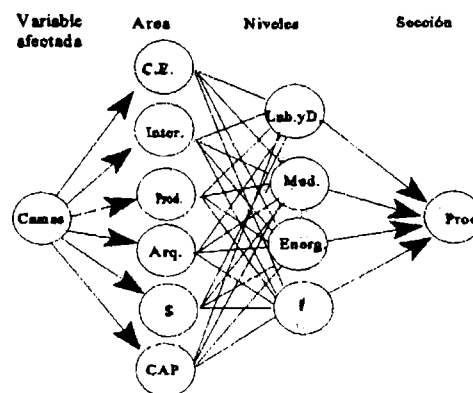


Figura 3: Diagrama combinado serie-paralelo.

Nota Técnica

Histograma de cantidad de camas
Agudos de la Pcia. Bs.As.

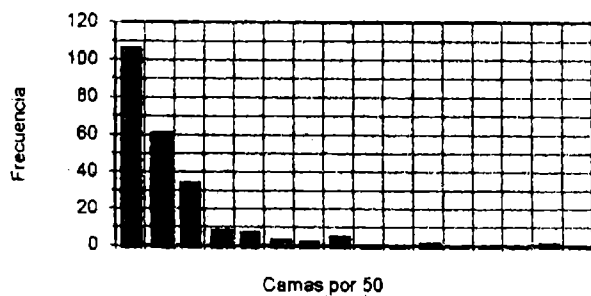


Figura 4: Histograma de la variable camas.

Histograma de producción
Agudos en la Pcia. Bs.As.

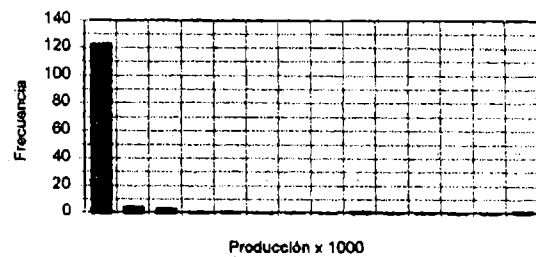


Figura 5: Histograma de variable producción

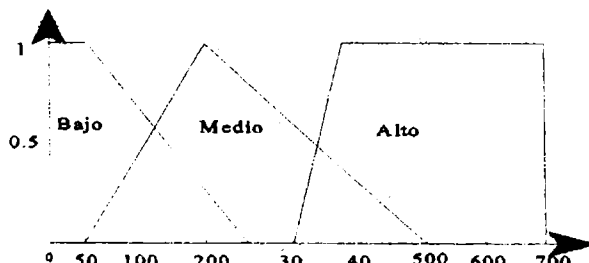


Figura 6: Diagrama función de pertenencia, CAMAS.

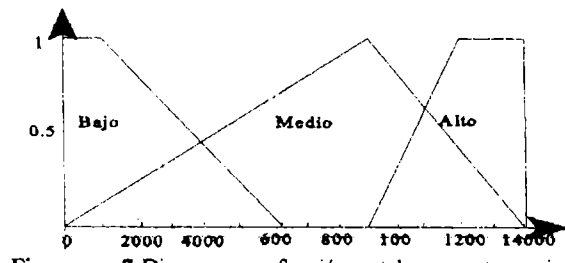


Figura 7: Diagrama función de pertenencia, PRODUCCION.

entrada, las reglas deberán ser activadas correctamente para el análisis de la situación y la toma de decisiones. Un detalle de la configuración se publicó en C. Discoli y F. Romero, 1997. En la etapa de desfuzzificación, las salidas difusas del sistema (variable requerida) se discretiza por los métodos mencionados en el punto vi del modelo piloto. Para la simulación del modelo se utilizó el software Matlab con su extensión para lógica borrosa, Fuzzy Inference Systems toolbox (FISMAT).

CONCLUSIONES

- * Metodológicamente se plantea un análisis en paralelo e iterativo, integrando la lógica determinística y borrosa en un mismo sistema de diagnóstico. De ésta manera existiría un mayor aprovechamiento en el manejo de ambos instrumentos, mejorando la convergencia de las salidas requeridas. Ambas vertientes parten de una base de datos común, con tratamientos diferenciados, adoptando en algunos casos la salida de una de ellas como entrada de la otra, mejorando el proceso de fuzificación, de obtención y de justificación de las reglas, ajustando la inferencia en la toma de decisiones y en la salida de requerimientos (nuevos inputs para futuros modelos multicapa de "Redes Neuronales" y realimentar a ambos tratamientos.
- * El modelo piloto planteado en C. Discoli y F. Romero, 1997 ha permitido mejorar algunas fases del proceso (Fuzificación y Base del Conocimiento, ver fig.1), ajustando los grados de pertenencia de las variables lingüísticas consideradas y la construcción y justificación de las reglas.
- * Se ha mejorado, sistematizado y unificado la fuzificación y desfuzificación para los distintos algoritmos utilizados (Zadeh, Sugeno, Mandani).
- * Reestructurar el proceso de diagnóstico ha permitido mejorar la adaptación de teoría y técnicas relacionadas con el control moderno, aproximándonos a los sistemas complejos con exigencias crecientes.

REFERENCIAS

- Discoli, C. y Rosenfeld, E. (1993a) Diagnóstico temprano y control de la gestión energético-productiva del hábitat en el sector terciario. 16ª Reunión de Trabajo de ASADES. Tomo I, pp.67-71. La Plata.
- Discoli, C. Rosenfeld, Y y Tesler, J. (1993b) Sistema informatizado para el control y diagnóstico en redes edilicias de salud. Versión Beta. 16ª Reunión de Trabajo de ASADES. Tomo I, pp.53-58. La Plata.
- Discoli, C. (1995a) Perfil energético en las redes complejas de servicios del sector terciario. El caso salud. 18ª Reunión de Trabajo de ASADES. San Luis, 1995.
- Discoli, C. (1995b) Diagnóstico de redes complejas de servicios. perfiles energéticos-productivos para el sector salud. Seminario Internacional de la Gestión del territorio: Problemas Ambientales y Urbanos, CEI-UNQ. En prensa, 8p.
- A. Bignoli (1993). Teoría de los conjuntos borrosos, Academia Nacional de Ingeniería. Bs. As.
- Discoli, C. (1995c) Aplicación de los conjuntos borrosos a las bases de información del sector terciario. Prueba piloto para el subsector salud. 18ª Reunión de Trabajo de ASADES, San Luis.
7. Discoli, C y Romero, F. (1997) Desarrollo metodológico aplicando control borroso a las bases de datos del sector terciario, subsector salud. 19ª Reunión de Trabajo de ASADES. Mar del Plata.
8. Rosenfeld, Y, Martini, I. Discoli, C., Tesler, J. Rosenfeld, E. (1997) Sistematización y biblioteca de módulos energéticos-productivos (MEEP) del subsector salud. 17ª Reunión de Trabajo de ASADES. Mar del Plata 1997.
- 9 Matlab for Windows. Manual de uso.