

ANÁLISIS DEL ALGORITMO DE HSUAN-SHIH EN LA IDENTIFICACIÓN DE SUBSISTEMAS DE INFORMACIÓN

*D.E. Jara Díaz¹, B.Sc, Docente Fundación Universitaria Compensar – Panamericana,
F. A. Cortés Aldana, M.Sc; Ph.D, Profesor Asociado de Ingeniería de Sistemas y Computación Universidad Nacional
Recibido octubre 18, 2012 – Aceptado, Febrero 4, 2013
<http://dx.doi.org/10.18566/puente.v7n1.a01>*

Resumen— Uno de los resultados del proceso de Planeación Estratégica de Sistemas de Información, en el marco de la aplicación de la metodología Business Systems Planning/Strategic Alignment (BSP/SA), es la matriz Procesos – Clases de Datos mediante la cual se define la Arquitectura de Información y se establecen cuáles son los subsistemas de información. El procedimiento para la identificación de los subsistemas de información se apoya principalmente en un análisis de afinidad entre los procesos que conforman la Matriz Procesos – Clase de Datos. Los especialistas generalmente establecen dicho análisis con base en su experiencia, entremezclando las relaciones administrativas y operativas identificadas. El algoritmo de Hsuan-Shih genera una aproximación a la identificación de subsistemas de información a partir de una relación de equivalencia difusa definida a partir de la información consignada en la matriz Procesos Clases de Datos. En este trabajo se presentan los resultados obtenidos después de desarrollar una implementación de este algoritmo y validarlo frente a estudios previos de Planeación Estratégica de Sistemas de Información.

Palabras claves: *Algoritmo de Hsuan - Shih, Identificación de subsistemas de información, Matriz Procesos Clases de Datos, Planeación de Sistemas de Información, Systems Planning/Strategic Arrangement (BSP/SA).*

Abstract— **One of the results of the Information System Planning based on the application of the Business Systems Planning/Strategic Arrangement**

Este trabajo es postulado dentro del marco del IV Congreso Iberoamericano Soporte al Conocimiento con la Tecnología.

E. Jara es Matemática de la Universidad Nacional de Colombia. Actualmente es estudiante de la Maestría en Ingeniería de Sistemas y Computación de la misma universidad y se desempeña como docente de la Fundación Universitaria Compensar - Panamericana. Teléfono 3212259158 e-mail: dejarad@unal.edu.co.

F. Cortés es Doctor en Proyectos de Ingeniería e Innovación de la Universidad Politécnica de Valencia, Magister en Ciencias Económicas, Especialista en Administración de la Universidad Santo Tomás e Ingeniero de Sistemas de la Universidad Nacional de Colombia. Actualmente se desempeña como profesor asociado al Departamento de Ingeniería de Sistemas y Computación de la Universidad Nacional de Colombia. Miembro del grupo de investigación ALGOS (<http://dis.unal.edu.co/grupos/algos/>) E-mail: facortesa@unal.edu.co. Teléfono 3165000 Ext. 14011.

(BSP/SA) methodology is the Process/Data Class matrix (CRUD Matrix). The CRUD Matrix allows defining the information architecture and identified the information subsystems. The procedure for identifying information subsystems is based primarily on an analysis of affinity between the processes in the CRUD Matrix. The specialists do the analysis based on their experience, mixing administrative and operational relationships identified. The Hsuan-Shih algorithm generates an approach to identifying information subsystems from a fuzzy equivalence relation defined on the basis of information contained in the CRUD matrix. This work shows the results after de implementation of the Hsuan-Shih algorithm and validation with Strategic Information System Planning Process, done before in different scenarios.

Keywords: *Hsuan - Shih algorithm, Information subsystem identification, CRUD Matrix, Information Systems Planning, Systems Planning/Strategic Arrangement (BSP/SA)*

I. INTRODUCCIÓN

LA identificación de la arquitectura de información de una empresa/entidad juega un papel importante en el diseño, desarrollo, implementación y mantenimiento de un sistema de información de una organización, toda vez que los subsistemas de información de la organización y sus interrelaciones determinan el éxito de la gestión de la información de la misma.

La identificación de subsistemas de información es un tema que se ha abordado desde diferentes áreas de investigación de la ingeniería informática como el diseño de las bases de datos, la ingeniería de software, el pensamiento sistémico y la planeación de sistemas de información. Por ejemplo, la metodología Business Systems Planning/Strategic Alignment (BSP/SA de aquí en adelante) propone un método manual para la identificación de los subsistemas de información [1].

BSP/SA es una de las técnicas más importantes para la Planeación de Sistemas de Información de

una empresa, su enfoque se centra en la importancia de los datos de modo que su objetivo es asegurar que tanto los datos necesarios para soportar los requerimientos de una organización como la arquitectura de información, estén disponibles y sean relativamente estables para la organización en el largo plazo. Las etapas críticas dentro de la metodología son la identificación de los procesos de negocio, la identificación de las clases de datos y la definición de la arquitectura de información de la organización en donde se establecen las relaciones entre los procesos y las clases de datos por medio de una matriz de Procesos/Clases de Datos. La metodología indica cómo manipular la matriz de modo que se puedan identificar los diferentes subsistemas de información y los flujos de información entre ellos.

En este punto surgen las siguientes preguntas: ¿Es posible automatizar la identificación de los subsistemas de información de una organización/dependencia? y concretamente: ¿Desde la metodología BSP/SA es posible automatizar dicha identificación?

Félix Cortés, profesor del Departamento de Ingeniería de Sistemas e Industrial de la Universidad Nacional e integrante del grupo de investigación ALGOS, ha estado trabajando con los estudiantes en el análisis, diseño e implementación de herramientas de software para facilitar el proceso de planeación de los sistemas de información sin dejar a un lado la estrategia organizacional [2]. Durante la primera etapa de la investigación, se han desarrollado varios estudios de casos en diversos contextos empresariales con el fin de analizar detalladamente el método BSP/SA de la IBM. Como producto de esta etapa, ha sido posible proponer la automatización de las matrices Procesos vs. Organización y Procesos vs. Clases de Datos (ver [3] y [4]). Dentro del curso de la investigación se están examinando los posibles caminos de automatización de la matriz Procesos vs. Clases de Datos Agrupados.

El estudio del estado del arte mostró que la identificación automatizada de subsistemas de información se ha abordado desde diferentes áreas. En [5] por ejemplo, se propone e implementa un algoritmo por medio del cual se identifican los subsistemas de información a partir del modelo entidad relación; para ello se mide la distancia que está inversamente relacionada con la semejanza de

interacciones de una pareja de entidades con relaciones en la matriz binaria entidad relación asociada al diagrama. En [6], el profesor Hsuan-Shih, basado en la metodología BSP/SA, propone un algoritmo que determina un número óptimo de subsistemas de información a partir de la matriz Procesos vs. Clases de Datos.

Los objetivos del trabajo son describir el algoritmo de Hsuan-Shih, implementarlo y validarlo frente a varias matrices de Procesos/Clases de Datos obtenidas manualmente en procesos previos de Planeación Estratégica de Sistemas de Información. Este documento registra el proceso y los resultados obtenidos, sin embargo, como el algoritmo emplea varias propiedades de la lógica difusa y asume algunos conocimientos de la metodología BSP/SA, entonces se presenta un marco teórico con los elementos más importantes.

II. MARCO TEÓRICO

A. Antecedentes Matemáticos: Lógica Difusa

De acuerdo con Sivanandam en [7], en el mundo existen adjetivos que son vagos o difusos, como por ejemplo ser alto, ser viejo, ser agradable, etc. que generan incertidumbre y falta de información sobre un problema. La lógica difusa proporciona una estructura o modelo que permite representar construcciones lingüísticas como las mencionadas anteriormente.

1) Conjuntos difusos

En la teoría clásica de conjuntos, éstos se visualizan como colecciones de elementos para las cuales se sabe con certeza si un objeto dado pertenece a él o no. En cambio, los conjuntos difusos representan conceptos ambiguos, de modo que el grado de pertenencia de un elemento particular expresa el grado de compatibilidad del elemento para dicho concepto. Para definir un conjunto difuso, es necesario representar la compatibilidad de los elementos a través de funciones de pertenencia o membresía que van desde un universo dado X al intervalo $[0,1]$. Dada $\chi: X \rightarrow [0,1]$ una función de pertenencia y a un número real entre cero y uno $-a \in (0,1)$, se puede formar un conjunto difuso \underline{A} de la siguiente manera: Dado x un elemento del universo ($x \in X$), diremos que x pertenece al conjunto difuso \underline{A} ($x \in \underline{A}$) si y solo si $\chi(x)$ es un valor entre a y uno $-\chi(x) \in [a, 1]$ - es importante tener en cuenta que los valores más grandes de $\chi(x)$ indican que x tiene un mayor

grado de pertenencia al conjunto A .

2) *Relaciones Difusas*

Las relaciones difusas son subconjuntos difusos de $X \times Y$. Por consiguiente, de acuerdo con lo expuesto anteriormente, es necesario definir una función de pertenencia especificada en la expresión matemática (1):

$$\mu_R: X \times Y \rightarrow [0,1] \quad (1)$$

Y establecer un conjunto difuso R de modo que sus elementos satisfagan la función de pertenencia μ_R para un valor α específico.

AFIRMACIÓN 1: Si R_1 es una relación difusa de tolerancia sobre el producto cartesiano $X \times X$, donde $|X| = n$ entonces la relación R_1 se puede llevar a una relación de equivalencia calculando a lo sumo como en la ecuación (2):

$$R_1^{n-1} = \underbrace{R_1 \circ R_1 \circ \dots \circ R_1}_{n-1 \text{ veces}} = R \quad (2)$$

Ahora bien, sea R una relación difusa de X en X con función de pertenencia μ_R . R es una relación de equivalencia difusa si se satisfacen las propiedades, especificadas en las ecuaciones (3), (4) y (5), a saber:

- Reflexividad: $\mu_R(x_i, x_i) = 1$ (3)

- Simetría: $\mu_R(x_i, x_j) = \mu_R(x_j, x_i)$ (4)

- Transitividad Si $\mu_R(x_i, x_j) = \lambda_1$ y $\mu_R(x_j, x_k) = \lambda_2$ entonces $\mu_R(x_i, x_k) = \min\{\lambda_1, \lambda_2\}$ (5)

Si R_1 es una relación difusa que es reflexiva y simétrica, entonces es una relación difusa de tolerancia.

B. Business Systems Planning/Strategic Alignment (BSP/SA)

De acuerdo con Rondón en [3], existen tres tipos de enfoques para la planeación estratégica de Sistemas de Información.

- El enfoque basado en métodos de las Tecnologías de la Información y la Comunicación (TIC): Canalizado hacia la agregación de valor a los productos y servicios de la empresa a través del empleo de las TIC, cambiando la forma como se hacen negocios.

- El enfoque híbrido: El cual está orientado hacia la combinación de los enfoques anteriores.

La metodología BSP/SA pertenece al enfoque híbrido y su objetivo es “determinar la estructura estable de la información que apoya todos los procesos de la organización mediante la generación de un plan priorizado de Sistemas de Información, tomando en cuenta las necesidades de información que están asociadas al desarrollo de todas y cada una de las funciones de la empresa identificadas” ([3], página 205).

La metodología BSP/SA se compone de trece etapas, las cuales se ilustran en la Fig. 1. Este trabajo está particularmente interesado en la matriz Procesos Clases de Datos Agrupados desarrollada en el marco de la etapa 8. Una vez se han identificado todos los procesos y clases de datos, se da inicio a esta fase mediante la definición de la matriz procesos clases de datos cuya estructura se muestra en la tabla 1. En la primera columna se listan todos los procesos. En BSP/SA se recomienda iniciar listando los procesos de planeación estratégica, seguir con los procesos asociados al ciclo de vida del producto/servicio y finalizar con los procesos para la gestión de los recursos de apoyo. En la primera fila se deben registrar todas las clases de datos identificadas. Se debe establecer cuáles son las clases de datos que son creadas por cada proceso, colocando la letra “C” en la intersección respectiva. Luego se determinan cuáles son las clases de datos que son empleadas por cada proceso, colocando la letra “U” en la intersección correspondiente. Es importante verificar que cada clase de datos es creada por un único proceso (ver Tabla I).

TABLA I.
ESQUEMA DE LA MATRIZ PROCESOS/CLASES DE DATOS

Procesos	Datos								
	D_1	D_2	D_3	D_4	D_5	D_6	D_7	...	D_n
P_1	U	U	U	U	U	U	U		C
P_2	C		U	U	U	U			
P_3	U	C	U	U	U	U			
P_4									
⋮									
P_m	U	U	U	U	U	U	C		

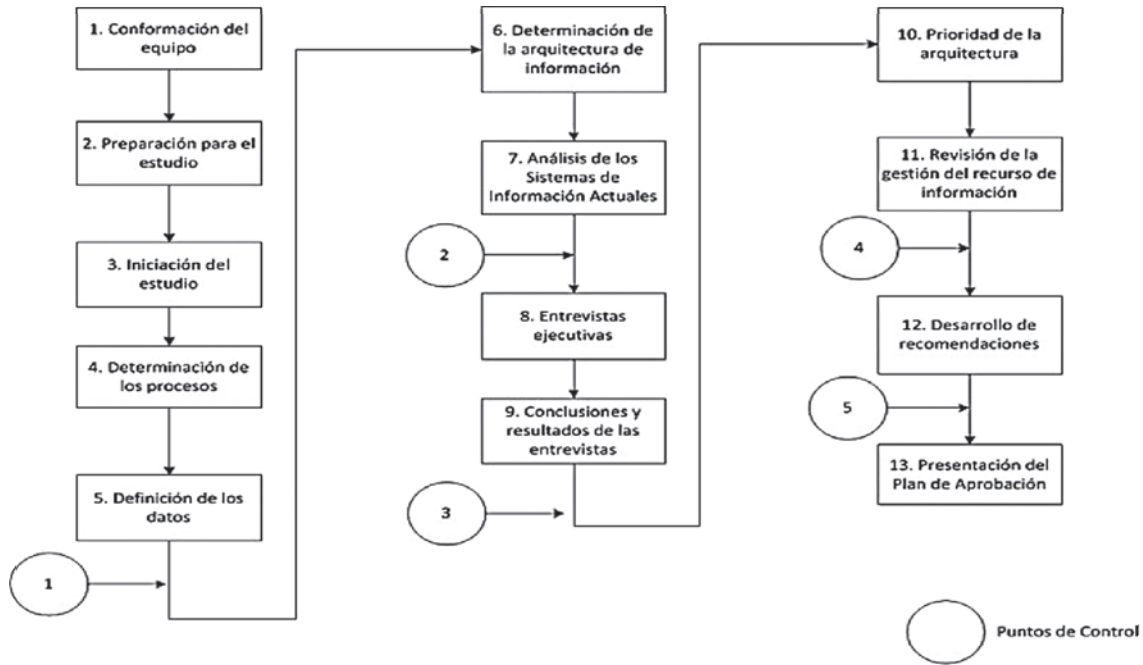


Fig. 1 Etapas BSP/SA

Antes de poder representar la arquitectura de información de la organización como un diagrama de flujo a partir de la matriz procesos / clases de datos, es necesario determinar la matriz procesos clases de datos agrupados.

Sea $P = \{P_1, \dots, P_m\}$ el conjunto de procesos y $D = \{D_1, \dots, D_n\}$ el conjunto de datos que se encuentran listados en la matriz procesos – clases de datos.

En [1] se sugiere el siguiente procedimiento:

- Agrupar los procesos que tienen patrones similares para el uso de los datos mediante intercambio de filas. Esto es semejante a hacer un particionamiento del conjunto P , según ecuación (6), por procesos afines, es decir, matemáticamente se formaría un conjunto $G = \{G_1, \dots, G_k\}$ de modo que:

- Cada G_i representa un conjunto de procesos afines entre sí.

- Para todo $i = 1, \dots, k$ se tiene que $G_i \subset P$. Los elementos de cada G_i se representarán así: $G_i = \{P_{i,1}, \dots, P_{i,m_i}\}$.

- Para todo $i = 1, \dots, k; j = 1, \dots, k; i \neq j$ se cumple que $G_i \cap G_j = \emptyset$

$$\cup_{i=1}^k G_i = P \quad (6)$$

i. Para cada grupo se identifican las clases de datos creadas por los procesos que lo componen mediante una reordenación de columnas, de modo que las letras “C” se ubiquen tan cerca como sea posible de la diagonal de la matriz.

El resultado será entonces la matriz procesos – clases de datos agrupados y las submatrices con las intersecciones entre los procesos afines y las clases de datos. Dichas submatrices representarían los módulos del sistema de información (ver Tabla II).

TABLA II
REORDENAMIENTO DE LA MATRIZ PROCESOS/CLASES DE DATOS

		Datos									
		E_1	E_2	E_3	...	E_{n-k}	...	E_{n-3}	E_{n-2}	E_{n-1}	E_n
G_1	Procesos										
	$P_{1,1}$	C	U	U				
	\vdots							
	P_{1,m_1}			C	...	U	...	U			
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
G_k	$P_{ k,1}$	U	U	U	...	C	...	U			
	\vdots							C	C	C	
	P_{k,m_k}	U	U	U	...	U	...	U	C	C	C

AFIRMACIÓN 2

Si \hat{G} y G son particionamientos diferentes del conjunto P entonces las matrices procesos – clases de datos agrupados resultantes, son diferentes.

La afirmación anterior conduce a preguntar si, a partir de la información consignada en la matriz Procesos – Clases de Datos, existe alguna técnica que permita obtener un particionamiento del conjunto P .

Una alternativa para establecer un particionamiento del conjunto P , consiste en establecer una relación de equivalencia \underline{R} sobre el producto cartesiano $P \times P$. El objetivo de la construcción de \underline{R} será relacionar P_i con P_j , si y solo si, P_i y P_j son procesos afines entre sí. ¿Pero qué significa que dos procesos sean afines entre sí?

Dos procesos pueden ser afines de acuerdo a diferentes puntos de vista, como por ejemplo:

- *Por su relación administrativa* pues son procesos que o se ejecutan o están bajo la responsabilidad de una misma dependencia de la organización

- *Por su relación operativa o funcional* pues son procesos que interactúan entre sí, para la consecución de un fin específico.

- *Por su relación informática.* Si dos procesos usan la misma clase de datos o, un proceso crea una clase de datos y el otro la usa, entonces esto significa que o la toma de decisiones de estos procesos pueden estar relacionadas, o existe una relación operativa implícita entre ellos y por ende, los procesos son afines.

Manualmente, el especialista establece relaciones de afinidad entre los procesos, a partir del conocimiento que tiene de la empresa y las relaciones administrativas y operativas que identifica a partir del desarrollo de la matriz Procesos – Organización, en el marco de la aplicación de la metodología BSP durante la Planeación Estratégica del Sistema de Información. Estas relaciones de afinidad regularmente son subjetivas y dependen de la experiencia del especialista, por lo tanto, se pueden clasificar como difusas. Lo anterior permite concluir que la relación de equivalencia \underline{R} sobre el producto cartesiano $P \times P$, no puede establecerse en un modo clásico, sino que debe ser una relación de

equivalencia difusa. Pero entonces ¿cómo debe construirse la relación difusa de equivalencia \underline{R} sobre el producto cartesiano $P \times P$?

C. El algoritmo de Hsuan-Shih

El profesor Lee Hsuan-Shih abordó el problema de construir la relación difusa de equivalencia sobre el producto cartesiano $P \times P$ en un artículo titulado “Automatic clustering of business process in business system planning”, el cual fue publicado en el European Journal of Operational Research en 1998.

El algoritmo de Hsuan-Shih se basa principalmente en la posibilidad de construir una relación difusa de equivalencia \underline{R} sobre el producto cartesiano $P \times P$, teniendo en cuenta la afinidad informática de los procesos. Para ello, en primera instancia se transforma las entradas de la matriz Procesos/Clases de Datos reemplazando las letras “C” por 1 y las letras “U” por 0.5 (ver Tabla III).

TABLA III
MATRIZ PROCESOS/CLASES DE DATOS TRANSFORMADA POR EL ALGORITMO

Datos \ Proceso	D_1	D_2	D_3	D_4	D_5	D_6	D_7	...	D_n
P_1	0.5	0.5	0.5	0.5	0.5	0.5	0.5		1
P_2	1		0.5	0.5	0.5	0.5			
P_3	0.5	1	0.5	0.5	0.5	0.5			
P_4									
⋮									
P_m	0.5	0.5	0.5	0.5	0.5	0.5	1		

Sobre las entradas de la nueva matriz se define la siguiente función de pertenencia especificada en la ecuación (7):

$$\mu_{\underline{R}}(i, j) = \frac{A_i A_j}{\|A_i\| \|A_j\|} \quad (7)$$

Esta función opera sobre los datos de la matriz Procesos Clases de Datos, donde A_i se refiere a la fila i – ésima. Los valores de la función $\mu_{\underline{R}}(i, j)$ pueden representarse a través de la matriz de pertenencia M_0 de tamaño $m \times m$, donde m corresponde al cardinal del conjunto P y la entrada $M_0(i, j) = \mu_{\underline{R}}(i, j)$. La matriz M_0 es una matriz simétrica con 1's en la diagonal principal, por lo

tanto, la relación difusa asociada, es una relación difusa reflexiva y simétrica es decir, es una relación difusa de tolerancia.

La afirmación 1 indica que la relación asociada a la función de membresía $\mu_R(i, j)$ puede transformarse en una relación de equivalencia. Aplicando esta propiedad, Hsuan-Shih construye la clausura transitiva de la relación, para ello utiliza la composición de funciones max-min, y establece recursivamente las siguientes matrices, especificadas por las ecuaciones (8) y (9):

- $M_1 = [m_1(i, j)]$ donde
$$m_1(i, j) = \mu_{R^2}(i, j) = \max_k \min[\mu_R(i, k), \mu_R(k, j)]$$
 (8)

- $M_m = [m_m(i, j)]$ donde $m_m(i, j) = \mu_{R^m}(i, j) = \max_k \min[\mu_{R^{m-1}}(i, k), \mu_R(k, j)]$ (9)

(Recuerde que m corresponde al cardinal del conjunto P)

- Finalmente, establece la matriz, definida por la ecuación (10):

$$M = [m(i, j)] = \mu_{R^*}(i, j) = \max_{k \geq 1} \mu_{R^k}(i, j) \quad (10)$$

La cual determina la relación de equivalencia establecida por la ecuación (11)

$$R_\lambda^* = \{(i, j) | \mu_{R^*}(i, j) \geq \lambda\} \quad (11)$$

III. IMPLEMENTACIÓN DEL ALGORITMO DE HSUAN-SHIH

Se desarrolló un aplicativo en Java denominado UN-BSP-HS que implementa el algoritmo de Hsuan-Shih. La interfaz principal permite cargar un archivo de texto con los datos de la matriz procesos clases de datos transformados (ver Fig. 2 y Tabla III). El usuario, tiene entonces la posibilidad de aplicar el algoritmo de dos maneras: Ingresando directamente el valor de λ (Agrupamiento personalizado) o tomando el agrupamiento estándar definido por la ecuación (12)

$$\lambda = \frac{[\sum_{i=1}^m \sum_{j=1}^m m(i, j)]}{m^2} \quad (12)$$

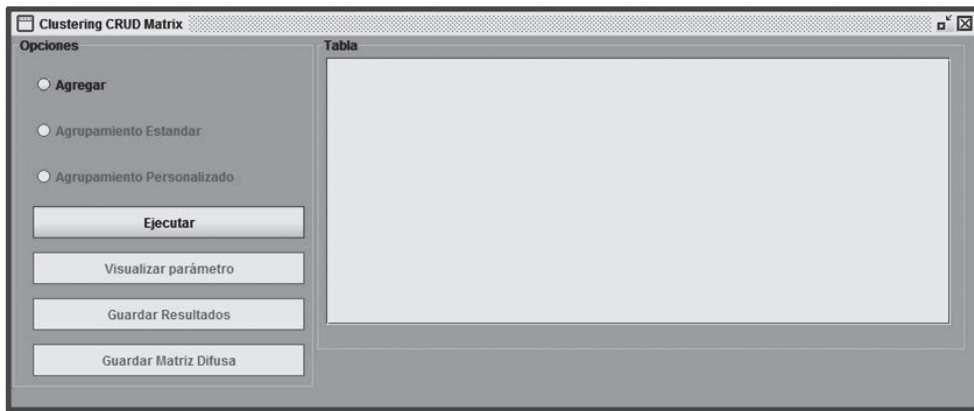


Fig. 2 Interfaz principal

A. Agrupamiento estándar

Al seleccionar la opción de Agrupamiento Estándar, el sistema mostrará la matriz Procesos –

Clases de Datos reorganizada de acuerdo a las clases de equivalencia obtenidas por el valor $\lambda = \frac{[\sum_{i=1}^m \sum_{j=1}^m m(i, j)]}{m^2}$ (Fig. 3).

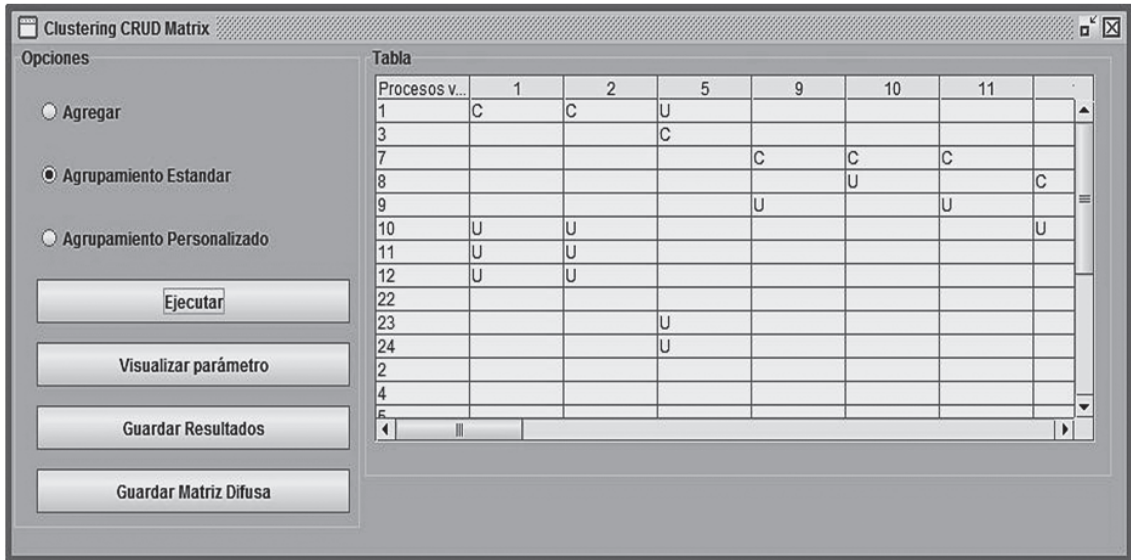


Fig. 3 Ejemplo de agrupamiento estándar

B. Agrupamiento Personalizado

Al escoger la opción de Agrupamiento Estándar, el sistema solicitará el valor de λ y mostrará la matriz procesos – clases de datos reorganizada de acuerdo a las clases de equivalencia obtenidas por el valor λ (Fig. 4 y 5).

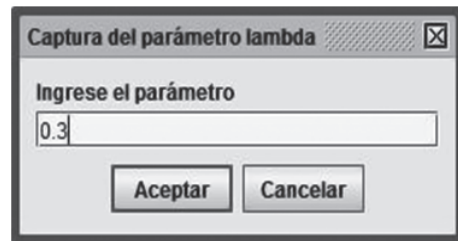


Fig. 4 Ejemplo personalización del parámetro lambda.

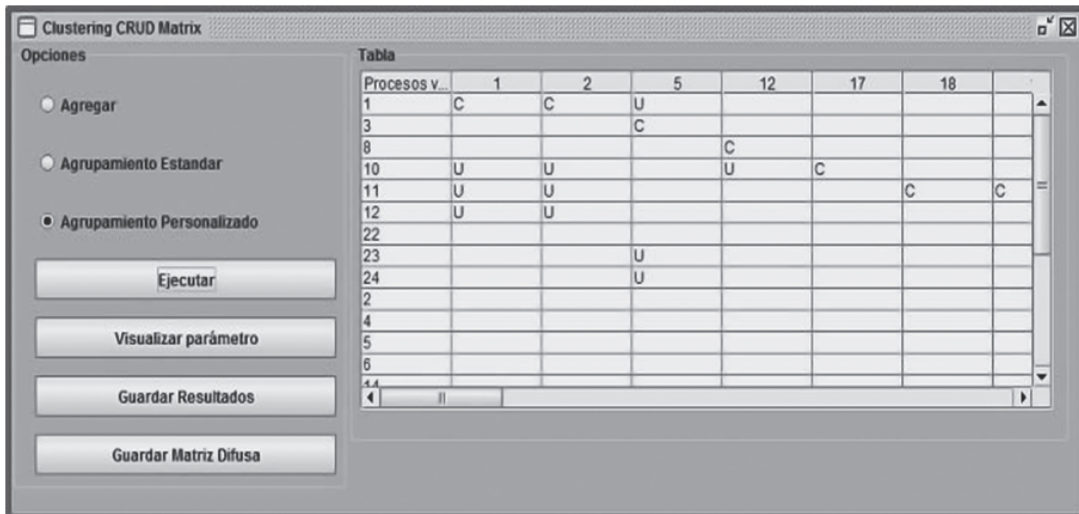


Fig. 5 Ejemplo de agrupamiento personalizado

C. Funcionalidades adicionales

Si el usuario quiere establecer el valor de la expresión $\frac{\sum_{i=1}^m \sum_{j=1}^m m(i,j)}{m^2}$, entonces debe dar click en el botón visualizar parámetro (Fig. 6)

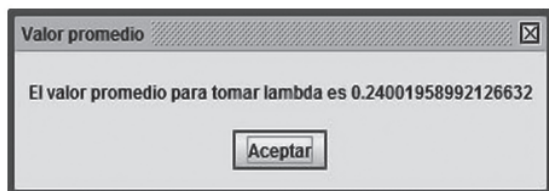


Fig. 6. Ejemplo visualización del parámetro lambda promedio.

El usuario tiene además las opciones de guardar los resultados de las matrices Procesos Clases de datos Agrupados resultantes y almacenar el valor de la matriz difusa $M = [m(i,j)]$.

IV. RESULTADOS

El aplicativo se probó en matrices Procesos/Clases de Datos obtenidas en trabajos de Planeación de Sistemas de Información anteriores. En términos generales, los hallazgos encontrados fueron los siguientes:

- El tiempo de ejecución del algoritmo es eficiente en matrices de tamaño alrededor de 100×100 .
- El empleo del algoritmo de Hsuan-Shih tomando el parámetro $\lambda = \frac{\sum_{i=1}^m \sum_{j=1}^m m(i,j)}{m^2}$, conduce a un resultado aproximado a la matriz Procesos Clases de Datos Agrupados que el especialista desarrolla manualmente. Sin embargo, a veces es posible obtener mejores aproximaciones cuando se toma un valor cercano pero mayor a dicho λ .

V. CONCLUSIONES

En los últimos años, el profesor Félix Cortés ha estado trabajando con los estudiantes de la asignatura Sistemas de Información y con los integrantes de su grupo de investigación ALGOS en el análisis, diseño e implementación de herramientas de software para facilitar el proceso de planeación de los sistemas de información en las organizaciones. Para esto se han desarrollado varios estudios de casos de Planeación Estratégica de Sistemas de Información utilizando la metodología BSP/SA de IBM.

A partir del trabajo de investigación desarrollado por el equipo de trabajo del profesor Cortés ha sido

posible automatizar dos tareas claves de la metodología BSP/SA: la matriz Procesos vs. Organización y la matriz Procesos vs. Clases de Datos, sin embargo, ¿es posible automatizar la matriz Procesos vs. Clases de Datos Agrupadas con el fin de identificar los subsistemas de información de una organización/dependencia?

La metodología BSP/SA brinda una técnica manual para la identificación de los subsistemas de información, pero este trabajo está supeditado al tipo de relaciones de afinidad entre procesos que los especialistas establezcan. El profesor Hsuan-Shih aplicó técnicas de lógica y teoría de conjuntos difusa para intentar representar la subjetividad inmersa en el proceso de identificación de subsistemas de información en la metodología BSP/SA. Para esto, definió un algoritmo que establece una relación difusa de equivalencia sobre el conjunto de procesos, mediante las conexiones informáticas que se pueden identificar en la matriz Procesos/Clases de Datos.

Este trabajo validó la eficiencia del algoritmo propuesto por el profesor Hsuan-Shih. Los resultados obtenidos permiten concluir que gracias a la información consignada en la matriz Procesos/Clases de Datos, el algoritmo de Hsuan-Shih resulta sumamente útil para generar aproximaciones en el proceso de identificación de la arquitectura de información dentro de la aplicación de la metodología BSP/SA, sin embargo, el estado actual del algoritmo aún no permite automatizar el proceso de identificación de los subsistemas de información de una organización/dependencia. Una posible explicación para esta situación obedece al hecho de que el análisis de afinidad entre procesos basado en las conexiones informáticas de éstos, asignado los valores de 1 y 0.5 respectivamente, no permiten establecer una relación difusa de equivalencia lo suficientemente precisa para efectuar el particionamiento del conjunto de procesos (P). Por lo tanto, es necesario evaluar si deben establecerse nuevas escalas numéricas en la matriz Procesos – Clases de Datos transformada (Ver Tabla III), teniendo en cuenta, por ejemplo, tres ítems: Crear, Usar y Modificar, donde los pesos mayores en las entradas de la Tabla III, reflejarán qué tan cercanos son los procesos con los datos. Un ejemplo para la determinación de estos valores puede ser a partir de la aplicación del método de eigenpesos. En la Tabla IV, se presenta una alternativa para dicha asignación de pesos.

Finalmente, pese a las semejanzas encontradas entre las salidas del algoritmo de Hsuan-Shih y los resultados de la matriz Procesos Clases de Datos Agrupados que los especialistas desarrollaron manualmente, estableciendo afinidades basadas en las relaciones operativas y administrativas entre los procesos, aún no es posible afirmar si para la definición de la arquitectura de información de una empresa/entidad, el análisis de afinidad basado en las relaciones informáticas entre procesos, resulta igual o más conveniente que los análisis de afinidad basados en las relaciones administrativas y operativas. En consecuencia, se seguirá trabajando en el tema, teniendo en cuenta que el algoritmo ofrece aproximaciones significativas a las técnicas manuales.

REFERENCIAS

- [1] International Business Corporation (IBM). *Business Systems Planning*. Georgia: IBM, 1984, pp. 1- 142
- [2] A. Mariño et al. "Herramienta de software para la enseñanza y entrenamiento en la construcción de la matriz DOFA" In *Revista Ingeniería e Investigación*, vol. 38, No. 3, 2010, pp. 159–164.
- [3] J. A. Rondón et al, "Enterprise – process: aplicación basada en computador para obtener la matriz proceso – organización durante la planeación estratégica de sistemas de información" In *Revista Ingeniería e Investigación*, vol. 2, No. 3, 2007, pp. 203–209.
- [4] J.A. Arias et al. *Software para la enseñanza y entrenamiento en la construcción de matrices para planeación estratégica de sistemas de información*. In *Revista Ingeniería e Investigación*. Vol 30 No.3, 2010, pp. 131-139.
- [5] P. Joglekar et al. *A clustering algorithm to identify information subsystems*. In *Journal of International Information Management*, 1994, pp. 129 – 141.
- [6] L. Hsuan-Shih. *Automatic clustering of business process in business system planning*. In *European Journal of Operational Research*, 1998. pp. 354-362
- [7] S. N. Sivanandam et al, *Introduction to Fuzzy Logic using Matlab*. Alemania: Springer Verlag, 2007, ch. 1–3.
- [8] S. A. Romero and J.M. Moreno. *Desarrollo de un Plan Estratégico de Informática para la Facultad de Odontología*. Bogotá: Universidad Nacional de Colombia, 2003.

- [9] G. Klir and T. Folger. *Fuzzy, Sets, Uncertainty and Information*. Estados Unidos: Prentice Hall International Editions, 1988.

BIOGRAFÍA



Dennis Elizabeth Jara Díaz es Profesora de matemáticas de la Fundación Universitaria de Compensar – Unipañamericana, sede Bogotá y docente ocasional de la Universidad Nacional de Colombia para el Departamento de Ingeniería de Sistemas y Computación de la sede Bogotá. Actualmente también trabaja en el Proyecto Interadministrativo de Articulación en la línea de ingeniería que Compensar – Unipañamericana tiene suscrito con la Secretaría Distrital de Educación, dirige y dicta los talleres titulados: Lúdica en la Matemática y Técnicas para el diseño de actividades didácticas que integren la lógica matemática con la programación. Es matemática de la Universidad Nacional de Colombia y actualmente es estudiante de la maestría en Ingeniería de Sistemas de Computación de la misma universidad.



Félix Antonio Cortés Aldana, es Profesor Asociado de la Universidad Nacional de Colombia, Sede Bogotá. Trabaja en el departamento de Ingeniería de Sistemas e Industrial. Es Doctor en Proyectos de Ingeniería e Innovación de la Universidad Politécnica de Valencia. Posee el título de Maestría en Ciencias Económicas. Especialización en Administración de la Universidad Santo Tomás. Sede Bogotá. Ingeniero de Sistemas de la Universidad

Nacional de Colombia. Sede Bogotá. Es profesor del curso Sistemas de Información de la maestría en Ingeniería de Sistemas y Computación de la Universidad Nacional de Colombia. Página web <http://www.docentes.unal.edu.co/facortesa/> Miembro del grupo de Investigación ALGOS. Página web: <http://dis.unal.edu.co/grupos/algos/>

ANEXO 1.

TABLA IV
PESOS SUGERIDOS PARA LA MATRIZ PROCESOS CLASES DE DATOS

Pesos Matriz Procesos/ Clases de Datos	MATRIZ			Eigenvalor λ_{max}	Eigenvector	Pesos (Eigenvector Normalizado)	IC	RC (Saaty 1980)	RC (Saaty 2001)
	Usar	Modificar	Crear						
Usar	1	1/2	1/7	3.00264	-0.1419	0.102513	0.00132	0.002275	0.02514
Modificar	2	1	1/3		-0.2988	0.215836			
Crear	7	3	1		-0.9437	0.68165			