Etiquetado de puntos alineados *

M. A. Garrido†, A. Márquez†, C. Iturriaga‡, J. R. Portillo†, P. Reyes† y A. Wolff§

†Departamento de Matemática Aplicada I. Universidad de Sevilla (España).
‡Computer Science Dpt. Univ. of New Brunswick (Canadá)
§Institute of Mathematics and Computer Science, University of Greifswald (Alemania)
vizuete@cica.es, almar@cica.es, citurria@unb.ca,
josera@cica.es, preyes@cica.es, awolff@mail.uni-greifswald.de

Resumen

En cartografía es frecuente encontrarse con la necesidad de situar etiquetas asociadas a elementos distintos de los mapas. Esto da lugar a multitud de problemas, entre otros el de situar etiquetas asociadas a puntos. La mayoría de los casos a que da lugar este problema son de naturaleza NP. En este trabajo presentamos algunas variantes del mismo, cuando los puntos están situados sobre una línea, tanto horizontal (o vertical) como oblicua.

1 Introducción

Muchas veces en cartografía es preciso situar etiquetas en determinados puntos de un mapa para, por ejemplo, indicar nombres de ciudades, resaltar accidentes geográficos, etc. La trascendencia de este problema ha llevado incluso a que la ACM Computational Geometry Impact Task Force [1] lo considere uno de los problemas clave. En [2] se puede ver abundante bibliografía sobre la materia.

En nuestro estudio intentamos etiquetar un conjunto de puntos en el plano, colocando etiquetas rectangulares en dichos puntos, sin que se produzcan intersecciones (solapamientos) entre ellas.

El siguiente problema general es NP–completo [3]:
ENTRADA: Un conjunto de n puntos y rectángulos con lados paralelos a los ejes, asociados a cada punto.
PREGUNTA: ¿Pueden colocarse cada uno de los rectángulos en cada uno de los puntos, haciéndolo coincidir con alguno de los vértices del rectángulo, sin que se produzcan intersecciones entre ellos?

En este trabajo presentamos algunas variantes de dicho problema, atendiendo a los siguientes aspectos:

*Parcialmente financiado por el grupo de investigación FQM—0164, adscrito al proyecto *PAI*

1. La distribución de los puntos en el plano, según una línea horizontal (H) o según una línea oblicua (O).
2. El modelo de etiquetas a utilizar para cada punto, cuadrados (C) o rectángulos (R).
3. El lugar de colocación de la etiqueta en cada punto, haciendo que el punto sea uno de sus vértices (V) o esté sobre cualquiera de sus lados (S).
4. El tamaño variable o no de las etiquetas, dependiendo de que todas las etiquetas tengan el mismo tamaño (F) o este sea variable (T).

Además de los problemas de decisión correspondientes a estos aspectos también estudiamos los problemas de optimización derivados de ellos.

2 Etiquetado sobre una línea horizontal

Hemos establecido diferentes resultados, en el caso de que los puntos se encuentren sobre una línea horizontal. Estos los tenemos en el siguiente cuadro, clasificados según la posición que tiene el punto sobre la etiqueta.

esquinas	lados vert.	lados hor.	borde
$O(n)$	$O(n)$	NP	NP
$O(n)$	$O(n)$	NP	NP

Estos tiempos son válidos si los puntos están previamente ordenados. En caso contrario habría que incrementar el coste $O(n \log n)$ correspondiente a la ordenación.

Si no se permite deslizamiento horizontal todas las variantes se pueden resolver en tiempo lineal. Para describir el procedimiento seguido en tales casos, veamos por ejemplo el siguiente problema:

HRVT($\mathcal{P}$,$\mathcal{L}$)
ENTRADA: Un conjunto $\mathcal{P}$=$\{P_1, \ldots, P_n\}$ de n puntos sobre una línea horizontal y un conjunto $\mathcal{L}$=$\{R_1, \ldots, R_n\}$ de n rectángulos con lados paralelos a los ejes.
PREGUNTA: ¿Puede colocarse en cada punto P_i el correspondiente rectángulo R_i, haciendo coincidir el punto con alguno de sus vértices, sin que se produzcan intersecciones entre ellos?

Si el conjunto $\mathcal{P}$ está ordenado utilizaremos un procedimiento incremental, de manera que añadiremos una etiqueta (rectángulo) cada vez, comenzando con P_1. Para cada realización $\mathcal{R}_\mathcal{K}$ de los k primeros rectángulos, definimos su *sombra* como el par (t_k, b_k) donde t_k (respectivamente, b_k) es la coordenada más a la derecha alcanzada por algún rectángulo sobre (respectivamente, bajo) la línea horizontal de la configuración $\mathcal{R}_\mathcal{K}$ (ver Figura 1).

Dados dos pares (t_k, b_k) y (t'_k, b'_k) correspondientes a dos configuraciones, establecemos la siguiente relación de orden parcial:

$$(t_k, b_k) \leq (t'_k, b'_k) \Longleftrightarrow \begin{cases} t_k \leq t'_k \quad y \quad b_k \leq b'_k \\ \qquad\quad \text{ó} \\ t_k \leq b'_k \quad y \quad b_k \leq t'_k \end{cases}$$

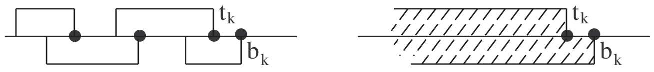

Figura 1: Sombra de una realización.

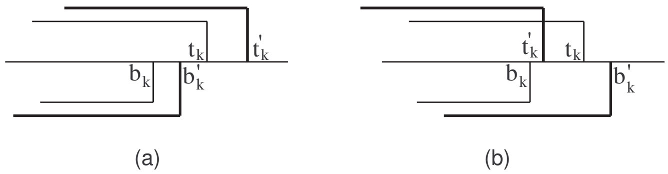

Figura 2: Dos casos en la relación de orden parcial entre sombras.

La Figura 2 ilustra gráficamente este orden entre realizaciones.

Esta relación de orden parcial nos permite hablar de realizaciones mínimas. Diremos que una realización es *mínima* si su sombra es minimal con respecto a la relación de orden definida anteriormente. Se puede probar:

Lema 1 *Si existe una realización de los $k+1$ primeros rectángulos, entonces existe una realización donde los k primeros rectángulos forman una realización mínima.*

Según este lema nos basta con llevar el control de las realizaciones mínimas en nuestro proceso incremental. Aunque al añadir el rectángulo $k+1$ a una realización mínima $\mathcal{R}_{\mathcal{K}}$, este admite en principio cuatro posiciones (pareciendo en principio que dos de ellas son mínimas), probamos que el número de sombras diferentes correspondientes a realizaciones mínimas permanece constante e igual a dos. Por lo tanto el mantenimiento de dichas realizaciones se puede hacer en tiempo lineal, obteniendo el siguiente resultado:

Teorema 1 *Los problemas de etiquetado de puntos en una línea horizontal, con etiquetas no deslizables horizontalmente se resuelven en tiempo $O(n)$.*

Si en nuestro problema permitimos desplazamientos horizontales, en cada paso se añadirán nuevas realizaciones mínimas, no quedando estas reducidas a dos, como en el caso anterior. Por lo tanto la heurística utilizada anteriormente no es válida. En realidad, no sólo dicha heurística no es válida sino que cualquier otra está condenada al fracaso, puesto que es posible hacer una reducción polinomial del problema NP–completo de la partición , enunciado en [4], a nuestro problema y por tanto se tiene:

Teorema 2 *Los problemas de etiquetado de puntos en una línea horizontal, con etiquetas deslizables horizontalmente son NP–completos.*

Estudiamos los problemas de optimización asociados a los problemas anteriores que tienen solución polinomial. Por ejemplo veamos el siguiente problema:

Max-HRVT($\mathcal{P}$,$\mathcal{L}$)

ENTRADA: Un conjunto $\mathcal{P}$=$\{P_1,\ldots,P_n\}$ de n puntos sobre una línea horizontal y un conjunto $\mathcal{L}$=$\{R_1,\ldots,R_n,\ R_i = a_i \times b_i\}$ de rectángulos con lados paralelos a los ejes.

PREGUNTA: ¿Cuál es el mayor valor de λ de manera que puedan colocarse rectángulos de dimensiones $\lambda R_i = \lambda a_i \times \lambda b_i$ en cada uno de los puntos, haciéndolo coincidir con alguno de los vértices del rectángulo, sin que se produzcan intersecciones entre ellos?

Seguiremos el siguiente proceso: 1) Encontrar el conjunto L de todos los posibles candidatos a soluciones del problema, 2) ordenar dicho conjunto y 3) mediante búsqueda binaria en este conjunto localizar la solución óptima, haciendo en cada paso una llamada al algoritmo de decisión lineal descrito en el problema anterior.

Para obtener el conjunto de soluciones posibles hemos de tener en cuenta que en la solución óptima al menos dos etiquetas deberán estar en contacto, que cada etiqueta puede contactar con cada una de las demás y que existen tres formas diferentes de contacto entre dos etiquetas (ver Figura 3).

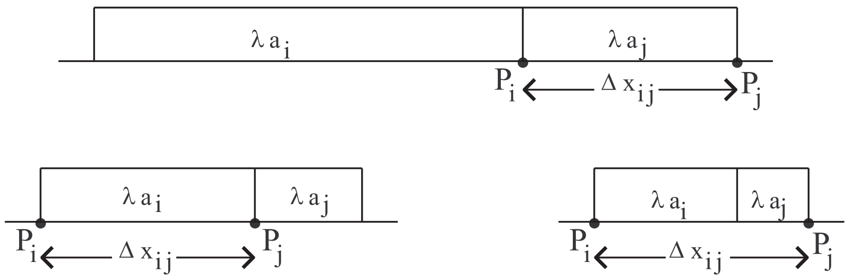

Figura 3: Contactos entre etiquetas en una realización óptima.

Por lo tanto el conjunto de soluciones posibles será:

$$L = \left\{ \frac{\Delta x_{ij}}{a_i}, \frac{\Delta x_{ij}}{a_j}, \frac{\Delta x_{ij}}{a_i + a_j}, i = 1, \ldots, n-1, j = i+1, \ldots, n \right\}, \qquad |L| = \frac{3n^2 - 3n}{2}$$

Entonces podemos probar el siguiente resultado:

Teorema 3 *El problema Max-HRVT se resuelve en tiempo $O(n^2 \log n)$.*

3 Etiquetado sobre una línea oblicua

Nos ocupamos en esta sección de estudiar el etiquetado de puntos a lo largo de una línea oblicua. Para ello veamos diferentes variantes:

OCVF($\mathcal{P}$)

ENTRADA: Un conjunto $\mathcal{P}$ de n puntos sobre una línea oblicua.

PREGUNTA: ¿Puede colocarse un cuadrado de lado unidad en cada uno de los puntos, haciéndolo coincidir con alguno de los vértices del cuadrado, sin que se produzcan intersecciones entre ellos?

Vamos a seguir de nuevo una técnica incremental. Supongamos ordenado el conjunto $\mathcal{P}=\{P_1, P_2, \ldots, P_n\}$. Definimos la *sombra* (s_{e_i}) de una etiqueta e_i como la porción del plano obtenida al prolongar su lado más alto hacia la izquierda y su lado mas oriental hacia abajo. Si tenemos una realización $\mathcal{R}_{\mathcal{K}}$ de los k primeros puntos, formada por las etiquetas $e_1, e_2, \ldots, e_k$, definimos la *sombra* de dicha realización S_{R_k} como la unión de todas las sombras de cada una de sus etiquetas (ver Figura 4).

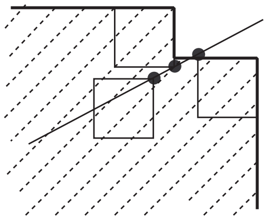

Figura 4: Sombra de una realización.

La relación de inclusión conjuntiva entre las sombras correspondientes nos permite establecer una relación de orden parcial entre diferentes realizaciones de los k primeros puntos. Trataremos de nuevo de controlar el número de sombras diferentes, correspondientes a realizaciones mínimas, que aparecen al añadir la etiqueta del punto $k+1$.

Numeremos, según se aprecia en la Figura 5, las cuatro posibles posiciones de cada etiqueta, según que el punto esté en el vértice superior izquierdo (**4**), superior derecho (**3**), inferior izquierdo (**1**) e inferior derecho (**2**).

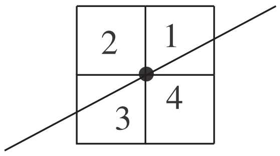

Figura 5: Numeración en las posiciones de cada etiqueta.

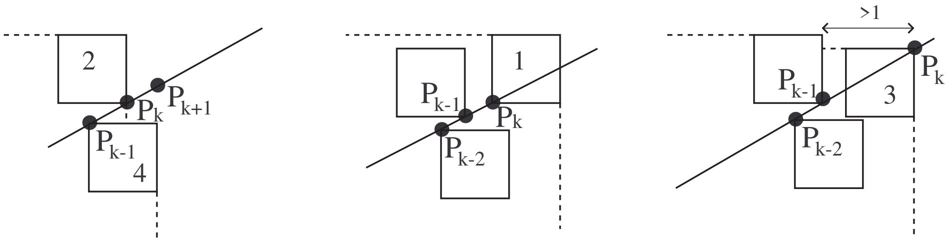

Figura 6: Los dos últimos rectángulos añadidos determinan la sombra.

Si tenemos una realización mínima $\mathcal{R}_{\mathcal{K}}$ de los k primeros cuadrados, al colocar un nuevo cuadrado sólo hemos de tener en cuenta la sombra producida por los dos últimos cuadrados, ya que los anteriores son irrelevantes pues no pueden ser intersectados por el nuevo cuadrado. Además si el último cuadrado de $\mathcal{R}_{\mathcal{K}}$ está colocado en la posición (**1**) o en la posición (**3**) sólo es relevante este último cuadrado (Figura 6). Por lo tanto, mediante el siguiente cuadro, podemos ver que en cada paso nunca hay más de dos sombras posibles correspondientes a realizaciones mínimas:

		no 3_{k+1}		
			no 2_{k+1}	
$k \setminus k+1$	3_{k+1}	2_{k+1}	4_{k+1}	no 4_{k+1}
3_k	3_{k+1}	$3_k 2_{k+1}$,$3_k 4_{k+1}$	ϕ	ϕ
2_k, 4_k	3_{k+1}	$2_k 4_{k+1}$,$4_k 2_{k+1}$	ϕ	$2_k 1_{k+1}$,$4_k 1_{k+1}$
1_k	3_{k+1}	$1_k 2_{k+1}$,$1_k 4_{k+1}$	$1_k 4_{k+1}$	Sin solución

Por lo tanto el control del número de sombras correspondientes a realizaciones mínimas se hace en tiempo lineal y queda probado el siguiente resultado:

Teorema 4 *El problema* OCVF *se resuelve en tiempo lineal (excepto ordenación).*

ORVF($\mathcal{P}$,R)

ENTRADA: Un conjunto $\mathcal{P}$ de n puntos sobre una línea oblicua y un rectángulo R de lados paralelos a los ejes.

PREGUNTA: ¿Puede colocarse un rectángulo de las dimensiones mencionadas en cada uno de los puntos, haciéndolo coincidir con alguno de los vértices del rectángulo, sin que se produzcan intersecciones entre ellos?

Este problema es idéntico al anterior, sin más que realizar una transformación afín del plano.

ORVT($\mathcal{P}$,$\mathcal{L}$)

ENTRADA: Un conjunto $\mathcal{P}=\{P_1,\ldots,P_n\}$ de n puntos sobre una línea oblicua y un conjunto $\mathcal{L}=\{R_1,\ldots,R_n\}$ de n rectángulos, de k tipos diferentes.

PREGUNTA: ¿Pueden colocarse los rectángulos indicados en cada uno de los puntos, haciéndolo coincidir con alguno de los vértices del rectángulo, sin que se produzcan intersecciones entre ellos?

En este caso, y siguiendo la heurística anterior, en las sombras producidas por las soluciones mínimas a lo sumo las dos últimas etiquetas de cada tipo tienen relevancia, por lo que el número máximo de sombras correspondientes a realizaciones mínimas será 2^k.

Tenemos entonces:

Teorema 5 *El problema* ORVT *se resuelve en tiempo $O(2^k n)$ (más ordenación).*

Max-OCVF($\mathcal{P}$)

ENTRADA: Un conjunto $\mathcal{P}$ de n puntos sobre una línea oblicua.

PREGUNTA: ¿Cuál es el tamaño máximo de los cuadrados que pueden colocarse en cada uno de los puntos, haciéndolo coincidir con alguno de los vértices, sin que se produzcan intersecciones entre ellos?

Sean $\{P_1(x_1,y_1), P_2(x_2,y_2),\ldots,P_n(x_n,y_n)\}$ los puntos sobre la recta. Es claro que en la solución óptima (máximo lado) deberá haber al menos dos etiquetas en contacto y por las características geométricas del problema (ver Figura 7) podemos garantizar el siguiente resultado:

Lema 2 *En cualquier realización del problema una etiqueta sólo podrá estar en contacto con, a lo sumo, las dos anteriores.*

El contacto entre dos etiquetas puede ser de tal manera que se toquen ambos lados o que el lado de una etiqueta contenga al otro vértice, entonces el conjunto de posibles soluciones será:

$$L = \left\{ \frac{\Delta x_{i,j}}{h}, \frac{\Delta y_{i,j}}{h}, i = 1, \ldots, n-1; j = i+1, i+2; h = 1, 2 \right\}$$

cuyo cardinal es lineal.

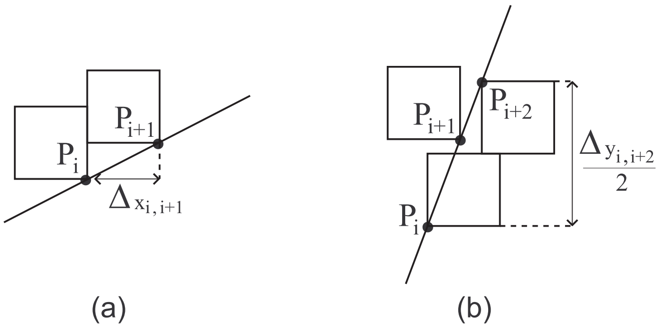

Figura 7: Algunos casos en el contacto entre etiquetas para el problema MAX-OCVF. El lado de las etiquetas en una solución óptima viene dado por (a) $\Delta x_{i,i+1}$, (b)$\frac{\Delta y_{i,i+2}}{2}$.

Y siguiendo el algoritmo descrito en el problema MAX-HRVT(ordenar los puntos, computar el conjunto L y mediante búsqueda binaria encontrar el valor óptimo, con llamadas al algoritmo descrito en el problema de decisión) queda probado el siguiente resultado:

Teorema 6 *El problema* MAX-OCVF *se resuelve en tiempo* $O(n \log n)$.

OCSF($\mathcal{P}$)

ENTRADA: Un conjunto $\mathcal{P}$ de n puntos sobre una línea oblicua.

PREGUNTA: ¿Puede colocarse un cuadrado de lado unidad en cada uno de los puntos, de manera que el punto se encuentre en cualquiera de sus cuatro lados, sin que se produzcan intersecciones entre ellos?

Volveremos a utilizar la relación de orden parcial definida en el problema OCVF (inclusión de sombras). En este sentido, supongamos que tenemos una realización óptima $\mathcal{R}_\mathcal{K}$ de las k primeras etiquetas. Si nos disponemos a añadir la etiqueta e_{k+1} correspondiente al punto $k+1$-ésimo procederemos de la siguiente manera: Colocaremos la etiqueta e_{k+1} de forma que el punto P_{k+1} se encuentre en la esquina inferior izquierda del cuadrado (posición menos ventajosa) si fuera posible, en caso contrario el problema no tendría solución. A continuación deslizaremos la etiqueta hasta que llegue a tocar a alguna de las etiquetas de los puntos $1, 2, \ldots, k$.

Según la geometría del problema (ver Figura 8) podemos enunciar el lema:

Lema 3 *Al añadir a una realización $\mathcal{R}_{\mathcal{K}}$ de las k primeras etiquetas una nueva etiqueta e_{k+1}, ésta tocará a lo sumo a las dos etiquetas anteriores.*

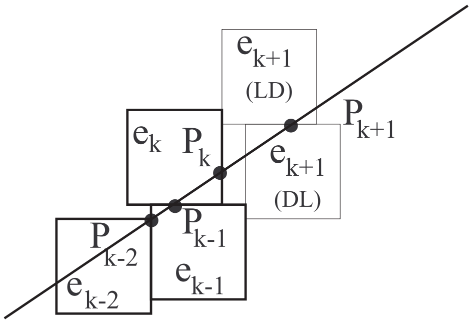

Figura 8: En el problema OCSF cada etiqueta toca, a lo más, a las dos anteriores.

Este deslizamiento lo podremos efectuar según los siguientes caminos: (**LD**) moviendo primero hacia la izquierda hasta encontrar un tope, y después hacia abajo hasta encontrar otro tope, en cuyo caso ya estaría en posición óptima, y (**DL**) si desplazamos primero hacia abajo y después hacia la izquierda.

Teniendo en cuenta esta notación podemos establecer el siguiente resultado:

Lema 4 *Si tenemos dos sombras diferentes, $S_1 \equiv \ldots-(LD)-(LD)$ y $S_2 \equiv \ldots-(LD)-(LD)$, correspondientes a realizaciones mínimas, de manera que el camino elegido para las dos últimas etiquetas es (**LD**), entonces son comparables (una de ellas tiene que estar contenida en la otra). Igualmente ocurre con dos sombras diferentes cuyas dos últimas etiquetas sean (**DL**).*

Entonces si tenemos en total s_k sombras correspondientes a realizaciones mínimas de los k primeros puntos, en algunas de ellas (s_{k_1}) la última etiqueta habrá sido colocada siguiendo el camino (**LD**) y en otras (s_{k_2}, $s_{k_1} + s_{k_2} = s_k$) siguiendo el camino (**DL**). Al añadir la etiqueta correspondiente al punto $k+1$-ésimo, ésta la podemos situar siguiendo cualquiera de los dos caminos, entonces tendremos $2 * s_k$ soluciones, pero de ellas s_{k_1} son de la forma $(\mathbf{LD})_k$–$(\mathbf{LD})_{k+1}$ en sus dos últimas etiquetas, y nos quedaríamos con la mejor de todas ellas (pudiendo computarlo en tiempo lineal, pues se trata de un conjunto totalmente ordenado). Igualmente nos quedaríamos con la mejor de las s_{k_2} que son de la forma $(\mathbf{DL})_k$–$(\mathbf{DL})_{k+1}$. Por lo tanto el número de soluciones posibles en el paso $k+1$ es a lo sumo $s_k + 2$. Entonces en cada paso añadimos a lo sumo un número finito (≤ 2) de sombras correspondientes a soluciones mínimas. Por lo tanto cada paso podemos realizarlo en tiempo lineal, quedando por tanto probado el resultado:

Teorema 7 *El problema* OCSF *se resuelve en tiempo $O(n^2)$.*

Max-OCSF($\mathcal{P}$)

ENTRADA: Un conjunto $\mathcal{P}$ de n puntos sobre una línea oblicua.

PREGUNTA: ¿Cuál es el mayor valor de s de manera que puedan colocarse cuadrados de lado s en cada uno de los puntos, estando el punto en cualquiera de sus cuatro lados, sin que se produzcan intersecciones entre ellos?

Seguiremos el mismo procedimiento descrito en el problema MAX-OCVF. Es decir encontraremos el conjunto L de posibles *candidatos* a solución del problema, y después realizando una búsqueda binaria en él y haciendo llamadas al algoritmo descrito en el problema de decisión anterior, localizaremos la mayor de todas estas soluciones.

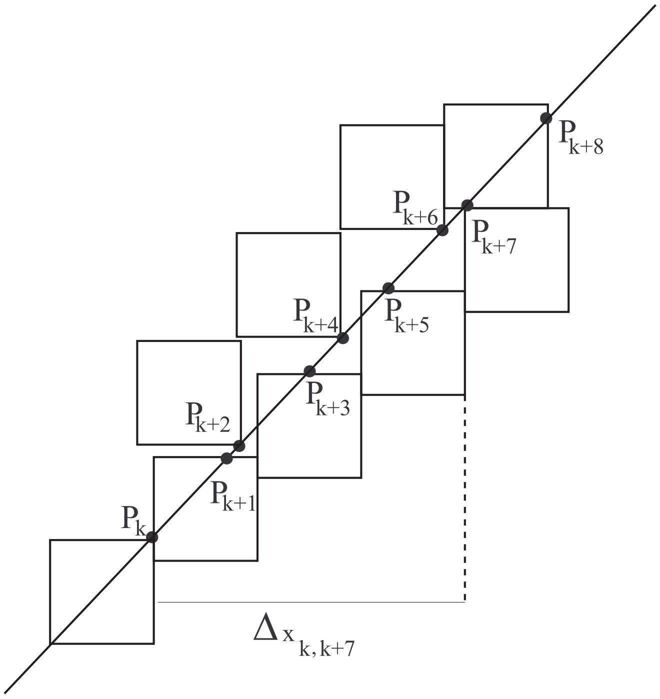

Figura 9: Ejemplo para el problema MAX-OCSF donde $s = \frac{\Delta x_{k,k+7}}{3}$

Según se aprecia en el gráfico (ver Figura 9) hemos de considerar el siguiente conjunto de posibles soluciones:

$$L = \left\{ \frac{\delta x_i x_j}{h}, \frac{\delta y_i y_j}{h}, i = 1, \ldots, n-1; j = i+1, \ldots, n; h = 1, \ldots, i-j \right\}$$

cuyo cardinal es $|L| = \frac{n^3 - n}{6}$. Por lo tanto:

Teorema 8 *El problema* MAX-OCSF *se resuelve en tiempo* $O(n^3 \log n)$.

ORVT($\mathcal{P}$,$\mathcal{W}$)

ENTRADA: Un conjunto $\mathcal{P}=\{P_1, \ldots, P_n\}$ de n puntos sobre una línea oblicua y un conjunto $\mathcal{W}=\{w_1, \ldots, w_n\}$ de n números reales.

PREGUNTA: ¿Puede colocarse en cada uno de los puntos P_i un rectángulo de altura 1 y anchura w_i, de manera que el punto se encuentre en cualquiera de sus cuatro esquinas, sin que se produzcan intersecciones entre ellos?

Sea $\mathcal{R}_{\mathcal{K}}$ una realización mínima de las k primeras etiquetas. Consideremos aquella parte s_k de la sombra producida por dicha realización mínima que tenga relevancia respecto del nuevo punto $k+1$ (es decir formada por las sombras de cada una de las etiquetas susceptibles de ser intersectadas por la nueva etiqueta e_{k+1}). Podemos entonces observar: 1) La última etiqueta e_k siempre está en la sombra. 2) No puede haber más de dos *escalones* diferentes en dicha sombra. 3) Existen a lo sumo cuatro tipos de sombra con un único escalón, dependiendo de las cuatro posiciones del último rectángulo. 4) Existen a lo sumo 16 tipos diferentes de sombra con dos escalones, atendiendo a las posiciones de los mismos.

En este caso podemos probar que al añadir la etiqueta correspondiente al punto $k+1$ el número de sombras obtenidas por realizaciones mínimas es constante (no superior a 20).

Por lo tanto:

Teorema 9 *El problema* ORVT *se resuelve en tiempo* $O(n)$.

Max-ORVT($\mathcal{P}$,$\mathcal{W}$)
ENTRADA: Un conjunto $\mathcal{P}=\{P_1,\ldots,P_n\}$ de n puntos sobre una línea oblicua y un conjunto $\mathcal{W}=\{w_1,\ldots,w_n\}$ de n números reales.
PREGUNTA: ¿Cuál es el mayor valor de λ de manera que puedan colocarse rectángulos de dimensiones $\lambda \times (\lambda w_i)$ en cada uno de los puntos P_i, de manera que el punto se encuentre en cualquiera de sus cuatro esquinas, sin que se produzcan intersecciones entre ellos?

En este caso el número posible de candidatos es cuadrático, por lo que:

Teorema 10 *El problema* MAX-ORVT *se resuelve en tiempo* $O(n^2 \, log\, n)$.

Referencias

[1] B. Chazelle, "Application challenges to computational geometry: CG impact task force report", *Technical Report*, TR–521–96, Princeton Univ., Abril, 1996.

[2] A. Wolff y T. Strijk, "A map labeling bibliography", 1996. http://www.inf.fu-berlin.de/map-labeling/papers.html.

[3] M. Formann y F. Wagner, "A packing problem with applications in lettering of maps". En Proc. 7th. ACM Symp. on Comp.Geom. (1991), pag.281-288.

[4] M. R. Garey y D. S. Johnson., "Computers and Intractability: a guide to the theory of NP—completeness", Freeman, 1979.