

DETECCIÓN DE FUENTES DE EMISIÓN A PARTIR DE SENSORES PERIMETRALES

J. E. Blazek, C. Lonbardi y E. Vetere

Facultad de Ciencias Exactas y Naturales (F. C. E. N.) – Universidad de Buenos Aires – U. B. A.
Ciudad Universitaria – Pabellón I – Av. Int. Cantillo s/n y Av. Costanera – (1428) – Capital Federal – Argentina.
jebblazek@dm.uba.ar cl5i@dc.uba.ar ev4e@dc.uba.ar

RESUMEN

El presente trabajo aborda la posible resolución de un problema de interés para varias disciplinas, entre las cuales se encuentra la física, dentro del marco de las redes neuronales.

El problema a resolver, trata de la obtención de fuentes emisoras encerradas dentro de una región acotada circular del plano; sabiendo que el efecto de las fuentes sobre los sensores colocados en la curva frontera de la región es escalar, aditivo, y que se obtiene respecto de la magnitud de la fuente con una ley que va proporcional a la intensidad de esta e inversamente proporcional al cuadrado de la distancia.

Para dicha resolución implementamos tres soluciones a partir del concepto de redes neuronales. La primera solución está basada en el aprendizaje competitivo. Con esta implementación logramos resolver el problema para dos fuentes, ~~calculando tanto la posición como la intensidad de las mismas.~~

La segunda solución consiste en un *perceptrón bicapa* entrenado mediante *back propagation*. La red implementa el modelo directo del problema, de forma tal que sus pesos representan las posiciones e intensidades de las fuentes. El entrenamiento obtiene como resultado estas incógnitas.

En la tercera solución implementamos un dispositivo que combina el ajuste competitivo de posiciones, con el ajuste de intensidades mediante un perceptrón. El perceptrón se ubica en capa de salida con función de activación lineal. La combinación de ambas redes (la competitiva y el perceptron) fue realizada según lo implementado en la primera solución, y ejecutando alternativamente una época de cada red, en donde las entradas del perceptrón serían los pesos de las unidades, ajustados por el inverso de su distancia a cada sensor.

El problema así descrito, queda resuelto si logramos encontrar la magnitud y la posición de cada una de las fuentes encerradas, esto es conocer tres números para cada fuente. Vale señalar que la cantidad de fuentes, en principio, también es desconocida. Por esta razón en cada solución propuesta se realiza un sondeo con distintas cantidades de fuentes para conocer el alcance de la solución.

DESCRIPCIÓN DE LA PRIMERA SOLUCIÓN

Implementamos una red competitiva simple, en la cual no se incluye una organización geométrica de las unidades. Cada unidad representa una fuente, y sus pesos representan su posición.

Los criterios de selección de la unidad ganadora y del vector de movimiento (que es el vector de variación de pesos) usan información del problema.

Para seleccionar la ganadora, nos basamos en el siguiente vector que se calcula para cada fuente f :

$$\sum_s \frac{\bar{x}_f - \bar{x}_s}{\|\bar{x}_f - \bar{x}_s\|} (s'_j - s_j) \frac{1}{\|\bar{x}_f - \bar{x}_s\|^2} \quad (\text{ec.1})$$

El término j -ésimo de esta sumatoria es el vector unitario que va de f al sensor j -ésimo, ponderado por un factor proporcional al error del sensor (considerando el signo de este error) e inversamente proporcional al cuadrado de la distancia entre fuente y sensor.

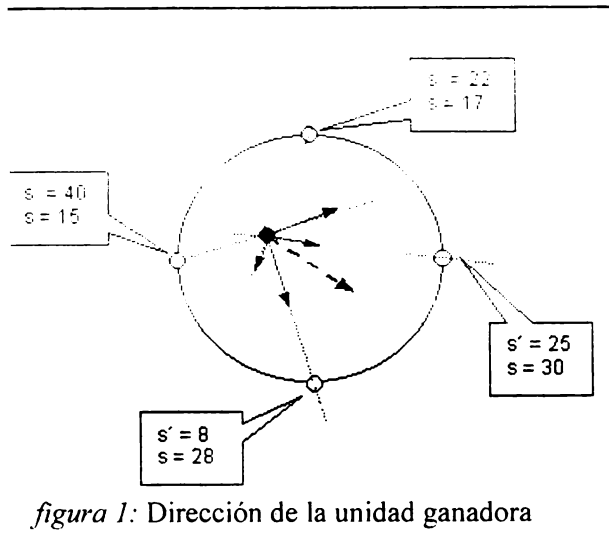


figura 1: Dirección de la unidad ganadora

La sumatoria de estos términos para todos los isores es el vector resultante para f .

Puede pensarse el cálculo del vector resultante para f imaginando que cada sensor s ejerce una "fuerza" de atracción o repulsión sobre f , en la dirección del segmento que los une. El sensor "atrae" a la fuente si el error en el sensor es negativo, y la "repele" si el error en el sensor es positivo.

Esta "atracción" o "repulsión" es proporcional al error detectado en el sensor. También la definimos matemáticamente proporcional al cuadrado de la distancia, al igual que la magnitud que miden los isores.

El vector resultante para f es la suma de esta "fuerza" que imaginamos ejerce cada sensor sobre f . La fuente cuyo vector resultante tiene mayor

módulo es declarada ganadora. La dirección en que se mueve la unidad ganadora (o sea, la dirección de su vector de variación de pesos) es la de su vector resultante (ver figura 1).

Para determinar la magnitud del vector de variación de pesos, implementamos una **búsqueda en línea**, según la idea descrita en el libro de Hertz. Lo que se implementó fue un proceso de bisección que calcula la magnitud del vector de variación que minimiza la energía.

Resultados de la primera solución

Observamos que el ajuste de posición le lleva muchas menos épocas que el ajuste de intensidades. Para tres fuentes, el método de cálculo de intensidades ideales implementado no pudo hallar los resultados correctos.

La red pudo ajustar correctamente la posición de cinco fuentes, que es la mayor cantidad de unidades con que probamos.

Un resumen de los resultados obtenidos se encuentra en la TABLA 1.

De la observación de los resultados obtenidos sin ajuste de la intensidad para cada fuente, desarrollamos una variante con intensidades fijas para las fuentes que se acumulara para generar la intensidad real.

TABLA 1: Resumen de los resultados obtenidos con la primera solución

Cantidad de Fuentes	Cantidad de Sensores	Incluye intensidad	Épocas	Observaciones
2	8	No	10	Fuentes de igual intensidad
2	8	No	--	Fuentes de distinta intensidad, entorpecen la evolución
2	8	Sí	50	Tanto para fuentes de igual intensidad como para fuentes de distintas intensidades
3	16	Sí	--	Presenta dificultades en el ajuste de intensidades
3	16	No	15	Fuentes de igual intensidad
4	32	No	65	Fuentes de igual intensidad
5	32	No	162	Fuentes de igual intensidad

Variante de la primera solución

Si consideramos la suma de las intensidades (I) de todas fuentes, y proponemos una cantidad suficientemente grande de fuentes (n), entonces a cada fuente se le asigna una intensidad igual al cociente entre la intensidad total y la cantidad (I/n).

También, en un problema real, se puede considerar la intensidad total y en vez de proponer una cantidad de fuentes grande, se utiliza la intensidad mínima que puede detectar un sensor para fijar a cada fuente y así calcular la cantidad de fuentes (o unidades), como $n=I/I_m$.

Además agregamos una heurística para los valores iniciales para que no iniciaran al azar sino con cierta evaluación del problema.

Probamos el ajuste competitivo de posiciones por separado, definiendo unidades de intensidades iguales, asignando a cada unidad inicialmente la intensidad correcta, y desactivando el cálculo de intensidades ideales.

Heurística para los valores iniciales

Lo que sigue es un estudio que permitiría acotar la cantidad de fuentes, y asignar posiciones e intensidades iniciales para las mismas, que serían cercanas a las reales.

No llegamos a implementar el resultado de este estudio. En nuestra implementación, los pesos iniciales son aleatorios.

Estudiamos en forma teórica el comportamiento que un conjunto conocido de fuentes provoca sobre el borde del círculo que las contiene. Usando la ecuación para el calculo del valor en cada sensor debido a las distintas fuentes y escribiendo la relación explícita que poseen las variables x e y con el borde del círculo se obtiene:

$$x = r \cos \theta, y = r \sin \theta, r = cte.$$

$$s(x, y) = s(\theta) = \sum_{j=1}^n \frac{I_j}{(r \cos \theta - x_j)^2 + (r \sin \theta - y_j)^2} = \frac{1}{r} \sum_{j=1}^n \frac{I_j}{(\cos \theta - u_j)^2 + (\sin \theta - v_j)^2} \quad (ec. 2)$$

$$\text{con } u_j = \frac{x_j}{r}, v_j = \frac{y_j}{r}.$$

Entonces podemos ver que la derivada del valor de los sensores respecto de la variable que recorre el borde del círculo es:

$$\frac{ds(\theta)}{d\theta} = \frac{1}{r} \sum_{j=1}^n \frac{(\cos \theta - u_j)^2 + (\sin \theta - v_j)^2}{(r \cos \theta - u_j) \cdot (r \sin \theta - v_j)} + \frac{1}{r} \sum_{j=1}^n \frac{2I_j [(\cos \theta - u_j)(-\sin \theta) + (\sin \theta - v_j) \cos \theta]}{[(r \cos \theta - u_j)^2 + (r \sin \theta - v_j)^2]}$$

de aquí podemos calcular los puntos críticos de esta función de la siguiente forma:

$$\frac{ds(\theta)}{d\theta} = 0 \Rightarrow$$

si llamamos:

$$p(\theta) = \prod_{k=1}^n [(r \cos \theta - u_k)^2 + (r \sin \theta - v_k)^2]^2$$

$$p_j(\theta) = \prod_{k=1, k \neq j}^n [(r \cos \theta - u_k)^2 + (r \sin \theta - v_k)^2]^2$$

entonces multiplicando a ambos lados de la ecuación obtenemos:

$$\sum_{j=1}^n p_j(\theta) [(\cos \theta - u_j)^2 + (\sin \theta - v_j)^2 - 2I_j] + -2 \sum_{j=1}^n I_j p_j(\theta) ((\cos \theta - u_j)(-\sin \theta) + (\sin \theta - v_j) \cos \theta) = 0$$

si llamamos $q(t)$ al polinomio de funciones trigonométricas de la ecuación vemos que el grado de este polinomio es $4(n-1)+2=4n-2$, o sea $2n-1$ posibles máximos y mínimos, entonces si tomamos las rectas posibles que pasan por dos máximos tenemos:

TABLA 2: Relación entre máximos sensores y fuentes

Cantidad de fuentes	Cantidad de sensores	Cantidad de Máximos	Observaciones
2	8	2	Fuentes alineadas
3	16	3	Ninguna alineada
3	16	2	Las tres alineadas
3	16	1	Al menos dos cerca del centro
4	32	4	Bien espaciadas y lejos del centro $r > 0,6$
4	32	3	Con $0,4 < r < 0,6$ o alineamiento de tres
4	32	2	Con $0,2 < r < 0,4$ o alineamiento de todas
4	32	1	Al menos tres cerca del centro

$$\binom{2n-1}{2} = \frac{(2n-1)!}{(2n-3)!2!} = \frac{(2n-1)(2n-2)(2n-3)!}{(2n-3)!2}$$

$$(2n-1)(n-1) = 2n-3n+1$$

es decir $(2n-1)(n-1)$ casos posibles.

En la practica hemos visto que en todos los casos el número es mucho menor y del orden de n . A continuación vemos en la TABLA 2 la relación entre cantidad de máximos, sensores y fuentes obtenida.

No hace falta decir que si todas las cargas están demasiado cerca del centro, de tal forma que entren el error de los cálculos numéricos, no encontraremos ningún máximo.

Nos parece natural encontrar estas diferencias entre el caso teórico y las observaciones, ya que la cantidad de sensores con que contamos es discreta y el segundo caso y continua en el primero. Es lógico pensar que aumentando la cantidad de sensores este problema irá desapareciendo gradualmente.

Hasta el momento podemos decir que la cantidad de máximos nos da sólo una idea de la cota inferior a la cantidad de fuentes.

En los casos donde es posible, se pueden calcular posiciones iniciales trazando rectas entre cada par de máximos y ubicando una fuente sobre cada recta de modo proporcional al valor de los máximos que la forman.

$$(x, y) = \lambda(x_i - x_j, y_i - y_j) + (x_j, y_j)$$

Para obtener el valor de λ de la ecuación obtenemos el siguiente sistema de ecuaciones:

$$\begin{cases} (x - x_i)^2 + (y - y_i)^2 = \frac{I}{s(x_i, y_i)} = \frac{I}{s_i} \\ (x - x_j)^2 + (y - y_j)^2 = \frac{I}{s(x_j, y_j)} = \frac{I}{s_j} \end{cases}$$

Si reemplazamos la ecuación en el sistema obtenemos:

$$\begin{cases} (\lambda - 1)^2 [(x - x_i)^2 + (y - y_i)^2] = \frac{I}{s_i} \\ \lambda^2 [(x - x_j)^2 + (y - y_j)^2] = \frac{I}{s_j} \end{cases}$$

del cual se obtienen las siguientes soluciones para el valor de λ .

$$\lambda_1 = \frac{s_i + \sqrt{s_i s_j}}{s_i - s_j}, \lambda_2 = \frac{s_i - \sqrt{s_i s_j}}{s_i - s_j}, s_i \neq s_j$$

$$\lambda = \frac{1}{2}, s_i = s_j$$

El valor inicial de la intensidad de la fuente se calcula una vez obtenida la posición:

$$I = \lambda^2 s_j [(x_i - x_j)^2 + (y_i - y_j)^2]$$

Resultados de la variante

Realizamos una serie de pruebas iniciales con una secuencia de fuentes y notamos que con más de 5, la energía desciende hasta un punto relativamente bajo (del orden 10^{-3}), luego comienza a subir y no se estabiliza.

Para solucionar esto modificamos el criterio de elección de la ganadora por:

$$\sum_s \frac{\bar{x}_f - \bar{x}_s (s'_j - s_j)}{\|\bar{x}_f - \bar{x}_s\|} \frac{1}{\|\bar{x}_f - \bar{x}_s\|^3}$$

La heurística se implemento como se explico antes en lo que respecta a los ángulos y luego la posición se colocó al azar entre un 0.25 a un 0.75 del radio. Como la cantidad de máximos puede ser menor a la cantidad de fuentes colocadas las restantes las ubicamos en el centro.

Además comprobamos experimentalmente la relación entre la cantidad de fuentes y la mínima cantidad de sensores que se necesitan para poder determinar dichas fuentes. Esta relación es de tres sensores para cada fuente.

La explicación nos parece razonable ya que el problema puede quedar descrito como la cantidad de circunferencias en el plano con distintos centros y radios, necesarias para que la intersección de ellas sea un solo punto. Si bien esto parece inmediato, para muchas fuentes representa un sistema de ecuaciones no lineales.

Los resultados obtenidos (ver TABLA 3) fueron con y sin usar la heurística para los valores iniciales, se repitieron los mismos valores de sensores para comprobar la eficacia de la heurística, las épocas que aparecen en los resultados son promedios de diez pruebas cada una.

TABLA 3: Resumen de los resultados obtenidos con la variante

TABLA 3: Resumen de los resultados obtenidos con la variante (continuación)

Cantidad Fuentes Reales	Cantidad Fuentes Virtuales	Utiliza la Heurística	Épocas	Energía	Observaciones
2	3	No	40000	$<10^{-6}$	Todas las fuentes de igual intensidad
3	3	No/Sí	≈ 10000	$<10^{-6}$	Intensidades distintas menores de 1
3	3	No/Sí	≈ 20000	$<10^{-6}$	Todas las fuentes de igual intensidad
3	4	No	≈ 20000	$<10^{-6}$	Intensidades distintas menores de 1
4	4	Sí	≈ 16000	$<10^{-5}$	Todas las fuentes de igual intensidad
4	6	No	≈ 17000	$<10^{-3}$	Intensidades distintas menores de 1
4	6	Sí	≈ 6000	$<10^{-3}$	Intensidades distintas menores de 1
5	5	No/Sí	≈ 5000	$<10^{-4}$	Todas las fuentes de igual intensidad
5	8	No/Sí	≈ 10000	$<10^{-3}$	Intensidades distintas menores de 1
10	10	No	≈ 8000	$<10^{-3}$	Todas las fuentes de igual intensidad
10	10	Sí	≈ 5000	$<10^{-3}$	Intensidades distintas menores de 1
10	14	No	≈ 20000	$<10^{-3}$	Todas las fuentes de igual intensidad
10	14	Sí	≈ 15000	$<10^{-3}$	Intensidades distintas menores de 1

DESCRIPCIÓN DE LA SEGUNDA SOLUCIÓN

La segunda solución propuesta es construir una red bicapa, entrenándola con back propagation con gradiente descendente.

La red construida tiene la arquitectura 2-n-1, es decir: tiene dos entradas (que representan un punto (x,y) del borde del círculo), n unidades en la capa oculta (una por cada fuente a ajustar) y 1 unidad en la capa de salida (que será la intensidad medida en el punto (x,y)).

La red implementa el modelo directo del problema: dada una posición (x,y) indicar la intensidad medida en dicho punto. Además, la red es tal que el vector de pesos de la i-ésima unidad de la capa oculta representa la posición de la i-ésima fuente; y el i-ésimo peso de la unidad de salida representa la intensidad de la i-ésima fuente.

Como implementación del modelo directo, la red resulta en una capa oculta cuya función de entrada neta está modificada, la función de activación es *Inversa Radial* y sus umbrales son cero. La capa de salida, tiene unidades con función de entrada neta standard, función de activación lineal y umbrales fijos en cero.

$$(x - w_1)^2 + (y - w_2)^2$$

Con la red así construida, tomamos el conjunto de entrenamiento formado por la información disponible de los sensores (conocidos). Así, cada elemento de entrenamiento tiene la posición (x,y)

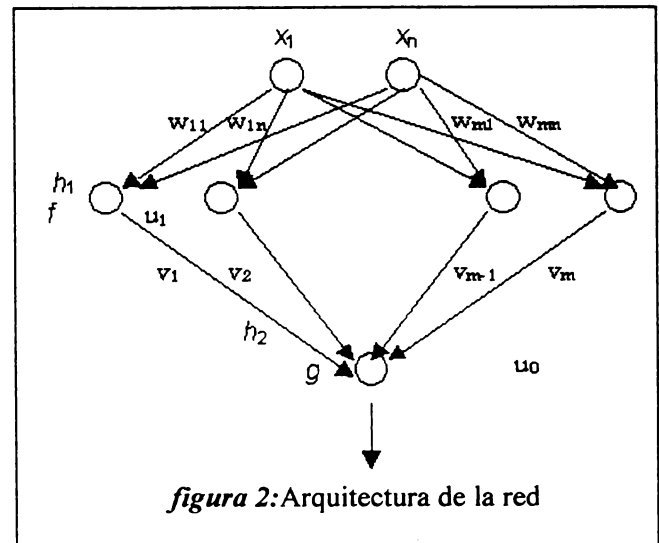


figura 2: Arquitectura de la red

del sensor como entrada, y la intensidad medida en dicho sensor como salida. El tamaño del conjunto de entrenamiento será, entonces, la cantidad de sensores.

Realizamos un aprendizaje por back propagation, mezclando el conjunto de entrenamiento para cada época. Al final del aprendizaje, deberían quedar determinadas las posiciones e intensidades de las fuentes, como se explicó antes.

Resultados de la segunda solución

Efectuamos numerosas pruebas para ajustar los parámetros, y algunas mejoras típicas al back propagation. Luego de incorporar el momentum y eta diferente para cada capa, la red ajustó

incorrectamente 3 fuentes (la red: 2-3-1) de distintas intensidades.

Después de 600 épocas, la energía cayó por debajo de 10^{-3} , siendo el error en las posiciones del orden de 10^{-3} y en las intensidades $1/4$. Luego de ~ 1200 épocas, el error en la posición había caído a $5E-4$ y las intensidades a $1E-2$.

El aprendizaje funcionó tomando momentum 0.9 con *etha* diferenciado en ambas capas. El *etha* utilizado en la capa oculta fue $5E-5$ y en la unidad de salida $1E-2$.

Las primeras pruebas fueron realizadas con una sola fuente, con *etha* uniforme y sin momentum. En esta prueba, equivalente a un ajuste por gradiente descendente, no logró ajustar en 10000 épocas.

Manteniendo el *etha* sin diferenciar, introdujimos momentum de 0.9. Esto realizó una mejora en la

TABLA 4: Resultados Generales

Cantidad de Fuentes	Cantidad de Sensores	Arquitectura de la Red (E-O-S) ^(a)	Cantidad de pesos y umbrales (P-U→T) ^(b)	Épocas	Observaciones
1	4	2-1-1	2-1→3	?	Solución alcanzada
2	8	2-2-1	4-2→6	200	Solución alcanzada
3	16	2-3-1	6-3→8	600	Solución alcanzada
4	32	2-4-1	8-4→12	1300	Solución alcanzada
5	32	2-5-1	10-5→15	3200	Mínimo local

^(a) La arquitectura de la red está notada E-O-S, donde (E) es la cantidad de entradas, (O) es la cantidad de unidades en la capa oculta y (S) es la cantidad de unidades en la capa de salida.

^(b) La notación P-U→T, significa (P) pesos, U umbrales y T total de variables involucradas en la red.

EXPLORACIÓN CON

UNA SOLUCIÓN HÍBRIDA

La tercera propuesta, como ya dijimos, consiste en una solución híbrida de la primera y de la segunda solución. Lo implementado es un dispositivo que combina una época de ajuste de posiciones según lo hecho en la primera solución, con una época de aprendizaje de un perceptrón simple que coincide con la capa de salida de la segunda solución.

El perceptrón simple usa como conjunto de entrenamiento el conjunto de los sensores, tal cual lo hicimos en la segunda solución.

Para cada elemento del conjunto de entrenamiento, que corresponde a un sensor, se muestra como entrada al perceptrón el inverso del

valor de la dirección, ubicándola correctamente, pero no ayudó a la convergencia, manteniéndose aún muy lenta.

Notando que las posiciones se ajustaban mucho más rápido que la intensidad, por esto decidimos utilizar *etha* diferenciado, un *etha* para la capa oculta y otro para la capa de salida.

En la TABLA 4 se observan los principales resultados.

Con 5 unidades de igual intensidad, no fue posible ajustar correctamente. Para intentar que funcione con 5 unidades, bajamos los *ethas* a $5E-6$ y $1E-3$. Sin embargo, la red cayó en un mínimo local muy pequeño ($E=0.067$), pero aún muy lejos de la solución.

cuadrado de la distancia entre posición estimada de la fuente y sensor.

Resultados de la tercera solución

Los resultados obtenidos en las pruebas son alentadores, y se resumen en la TABLA 5. Para cada fila se consigna el resultado obtenido con el valor de *etha* que mejor funcionó en cada caso.

En todos los casos se probó con sensores ubicados en el círculo de radio 1, e intensidades = 10 para todas las fuentes. Los pesos iniciales del perceptrón se fijaron en 1, y las posiciones iniciales al azar en el círculo de radio 0.2.

TABLA 5: Resultados obtenidos en la solución híbrida.

Unidades	Eta	Epocas	Energía	Error máximo en posiciones	Error máximo en intensidades
2	0.032	50	0.0181	0.003	0.1
3	0.03	60	0.0616	0.002	0.12
5	0.01	1200	2.9469	0.1	1.9

La prueba con cinco fuentes, fue interrumpida por falta de tiempo, estaba evolucionando razonablemente bien. Por la rapidez con que se implementó la prueba, la eficiencia deja bastante que desear, y por lo tanto no se pudo hacer una prueba con más épocas.

Con respecto a estos resultados observamos que:

- Para 2 y 3 fuentes, los resultados son superiores a los observados con la segunda solución (para tres fuentes, obtener un resultado similar con la segunda solución requiere una cantidad de épocas en el orden de las 500).
- Para 5 fuentes, el dispositivo implementado soluciona el problema. Este resultado no pudo obtenerse con ninguna de las otras dos soluciones implementadas y descritas en este trabajo.

CONCLUSIONES Y TRABAJO FUTURO

Habiendo implementado varias configuraciones basadas en enfoques distintos entre sí, pudimos observar algunos elementos para compararlos.

Estos están sugeridos por los resultados observados en las pruebas efectuadas.

Según lo observado:

- La versión de red competitiva con aprendizaje supervisado que implementamos parecer resolver mejor la cuestión de las posiciones que el enfoque feed-forward.

Asociamos esto, intuitivamente, al hecho de que en cada época la red competitiva sólo ajusta la posición de *una unidad*, que es la declarada ganadora. Mientras que en el aprendizaje con back-propagation se ajustan las posiciones de todas las unidades en todas las épocas.

Pareciera que al ajustar todas las posiciones a la vez, resulta difícil llegar a un estado estable del vector de pesos de la capa oculta,

que es el que representa las posiciones de las fuentes.

- El enfoque feed-forward parece resolver la cuestión de las intensidades razonablemente bien. Esto es debido a que la capa de salida tiene función de activación lineal y entrada neta estándar, y su incidencia en la regla de aprendizaje es lineal. Luego, si los pesos y umbrales de la capa oculta están cerca de la posición óptima y los correspondientes de la capa de salida aún no, entonces la regla de aprendizaje se va a comportar casi como si fuera lineal.

Detectamos que la combinación de la función de activación $1/x$, más la función de entrada neta $\|w-x\|^2$ probablemente complique la performance de la capa oculta del perceptrón con backpropagation.

- La variante de la red competitiva, que utiliza el efecto de aglomeración de fuentes de igual intensidad, pero algunos ordenes de magnitud más menor que la intensidad total, tiene buena performance. Además notamos que los efectos de la heurística para los valores utilizados son notables a partir de cinco fuentes reales y aumenta con la cantidad de fuentes virtuales que se utilicen en el ajuste.
- Por último la solución híbrida nos da otra variante para resolver el problema, que parece viable.

Dado lo anterior, observamos que las direcciones interesantes para seguir explorando son:

1. la solución competitiva con efecto de aglomeración.
2. la solución híbrida.

Consideramos a estas dos las propuestas más tibles para resolver un problemas con muchas fuentes. Propuestas que seguiremos estudiando.

Sin embargo nos parece interesante resaltar unas variantes posibles que no estudiaremos por nomento:

Modificar el algoritmo de backpropagation para que en cada época ajuste sólo los pesos de una unidad de la capa oculta, por ejemplo de aquella unidad para la que el módulo de su vector de variación de pesos sea máximo.

Esto le daría una característica competitiva al aprendizaje backpropagation.

Redefinir el modelo directo, o usar distintas funciones de energía, para que el perceptrón con backpropagation resultante tenga funciones de activación y entrada neta menos problemáticas. La idea básica es cambiar la formar el problema para que en vez de depender de la inversa de la distancia al cuadrado, dependa directamente de la distancia al cuadrado. Esto se podría lograr transformando el espacio de entrada de forma tal que la función de activación de la capa oculta, por ejemplo, sea lineal.

BLOGRAFÍA

John Hertz y otros, *Introduction to the Theory Neural Computation*, Addison-Wesley, EE.UU, 1.

J.F.Marar, *Polinômios Potências de Sigmóide S): Uma Nova Técnica para Aproximação de ções, Construção de Wavenets e Suas cações em Processamento de Imagens e Simais*, is Doctoral U.F.P., UNESP, Brasil, 1997.

B. Hilderbrand, *Métodos de la Matemática cada*, Manuales E.U.D.E.B.A., 1973.