

CLASIFICACION AUTOMATICA DE EMISIONES RADAR MEDIANTE COMBINACION DE ARTMAP Y FILTROS KALMAN

Trobiani, A., Rancán, C., Britos, P., García-Martínez, R.

Centro de Ingeniería del Software e Ingeniería del Conocimiento. Escuela de Postgrado. ITBA
Laboratorio de Sistemas Inteligentes. Facultad de Ingeniería. Universidad de Buenos Aires

{cranca, pbritos, rgm}@itba.edu.ar

Resumen

Las Medidas de Apoyo Electrónico (Electronic Support Measures ESM) tienen que ver con la búsqueda, interceptación, localización, análisis e identificación de energía electromagnética irradiada con propósitos militares.

Una función crítica de un sistema ESM es la identificación en tiempo real del tipo de radar asociado con cada tren de pulsos interceptado. Esta tarea constituye un desafío debido a la creciente densidad electromagnética, típica de una zona de conflicto, donde puede haber cientos de miles de pulsos por segundo, además de gran dispersión en los modos de trabajo de los radares militares.

En este trabajo se examina un sistema de reconocimiento de emisiones de radar que combina diversas fuentes de información para predecir el tipo de radar más probable. Los parámetros del pulso que caracterizan el tipo de radar se utilizan para la tarea de clasificación, mientras que los parámetros de carácter espacial (dirección de arribo y amplitud) se utilizan para separar los trenes de pulsos correspondientes a los distintos emisores activos.

El componente principal del sistema de reconocimiento es un clasificador basado en una red neuronal con capacidad de aprendizaje incremental, que se entrena para determinar el tipo de emisor radar presente en el ambiente. El sistema aprende en forma autónoma a identificar cada tipo específico de radar, directamente a partir de la información recolectada de campo.

Palabras Claves: Clasificación de emisores, red ARTMAP difusa, clustering, ESM, Radares

1. INTRODUCCIÓN

Las Medidas de Apoyo Electrónico (Electronic Support Measures ESM) involucran la búsqueda, interceptación, localización, análisis e identificación de energía electromagnética (principalmente de radar) radiada con propósitos militares. Las ESM proveen información valiosa para conocimiento de la situación en tiempo real, para detección y evasión de amenazas y para el despliegue a tiempo de contra medidas electrónicas (Browne, J. P. R, & Thurbon, M. T., 1998), (Wiley, 1993).

1.1. Descripción del problema

Una función crítica de un sistema ESM es la identificación en tiempo real del tipo de radar asociado con cada tren de pulsos interceptado. Las aproximaciones clásicas ordenan los pulsos de radar entrantes en trenes individuales, y luego comparan las características de ese tren con una librería de

descripciones paramétricas, que contiene una lista de los tipos de radares conocidos. Esta tarea constituye todo un reto debido al constante aumento de la densidad electromagnética (un radar Doppler puede transmitir cientos de miles de pulsos por segundo); así como la multiplicación y dispersión de los modos de trabajo de los radares militares, la agilidad en parámetros como el intervalo de repetición de pulsos, la frecuencia de emisión y la velocidad de barrido; modos desconocidos o reservados para los cuales no existen entradas en las librerías ESM, solapamiento entre los parámetros de diferentes tipos de radar; y ruidos o efectos de propagación que conllevan a una caracterización incompleta o errónea. Estos aspectos del problema imponen una gran demanda a los sistemas ESM.

1.2. Solución propuesta

En este trabajo examinaremos una aproximación alternativa. Un sistema de reconocimiento que combina diversas fuentes de información, para predecir el tipo de radar más probable por cada pulso interceptado. Los parámetros específicos de clase se utilizan para clasificar los pulsos de acuerdo al tipo de radar, mientras que los parámetros específicos de la posición espacial se usan para separar los pulsos provenientes de los distintos emisores activos.

El componente central del nuevo sistema de reconocimiento es un clasificador basado en una red neuronal, que se entrena para determinar los tipos de radares presentes en el medio ambiente.

El uso de una red neuronal con capacidad de aprendizaje incremental ofrece un marco de trabajo apto para refinar los tipos de radares conocidos y agregar los desconocidos, en pleno funcionamiento.

En esta aplicación se utiliza una red neuronal fuzzy ARTMAP (Carpenter, G.A., Grossberg, S., Markuzon, N., Reynolds, J. H., & Rosen, D. B., 1992) para clasificar los pulsos en base a los parámetros específicos de clase, y clustering por el método del vecino más próximo, basado en un banco de filtros Kalman (Bar-Shalom, Y., & Li, R. A., 1993); (Blackman, 1986), para separar los pulsos a partir de los parámetros específicos de la posición espacial.

2. DESCRIPCIÓN GENERAL DEL MODELO

Las señales radar son interceptadas en forma pasiva por la etapa receptora del sistema ESM. En un teatro de operaciones típico, las señales interceptadas son una mezcla de pulsos electromagnéticos transmitidos desde varias fuentes. La iluminación simultánea de estas fuentes produce solapamiento y entremezcla de los pulsos recibidos. Tras la detección del pulso, la mayoría de los sistemas ESM miden la amplitud del pulso (Pulse Amplitud PA), el ancho de pulso (Pulse Width PW), frecuencia de emisión de la onda portadora (RF), tiempo de llegada (Time Of Arrival TOA) y la dirección de arribo (Direction Of Arrival DOA). Una vez que los parámetros del pulso fueron medidos, se digitalizan y ensamblan en una estructura denominada firma del pulso (Pulse Signature PS). En la figura 1 se ilustra el diagrama en bloques de la solución propuesta.

En primer lugar un separador trata de descubrir periodicidades en el TOA de las PS entrantes. Si la agrupación de pulsos resulta posible por este método, se forman clusters y se les asigna un número de track y PRI a cada pulso del grupo (se entiende por track a un grupo de pulsos provenientes del mismo emisor). La agrupación se realiza mediante la técnica clásica de histogramas de diferencia de TOA.

Posteriormente las PS se dividen en dos cadenas. Una cadena constituye los parámetros que caracterizan los aspectos funcionales del radar. Estos parámetros incluyen la RF, el PW y PRI, y

son los que habitualmente se compilan en las librerías ESM porque son directamente útiles para reconocer el tipo de radar. La otra cadena consiste en los parámetros específicos del contexto, tales como la DOA y la amplitud, que reflejan la posición del emisor en el espacio. Estos parámetros son menos útiles que los anteriores para el reconocimiento del tipo de radar, pero son importantes para agrupar los pulsos en tracks.

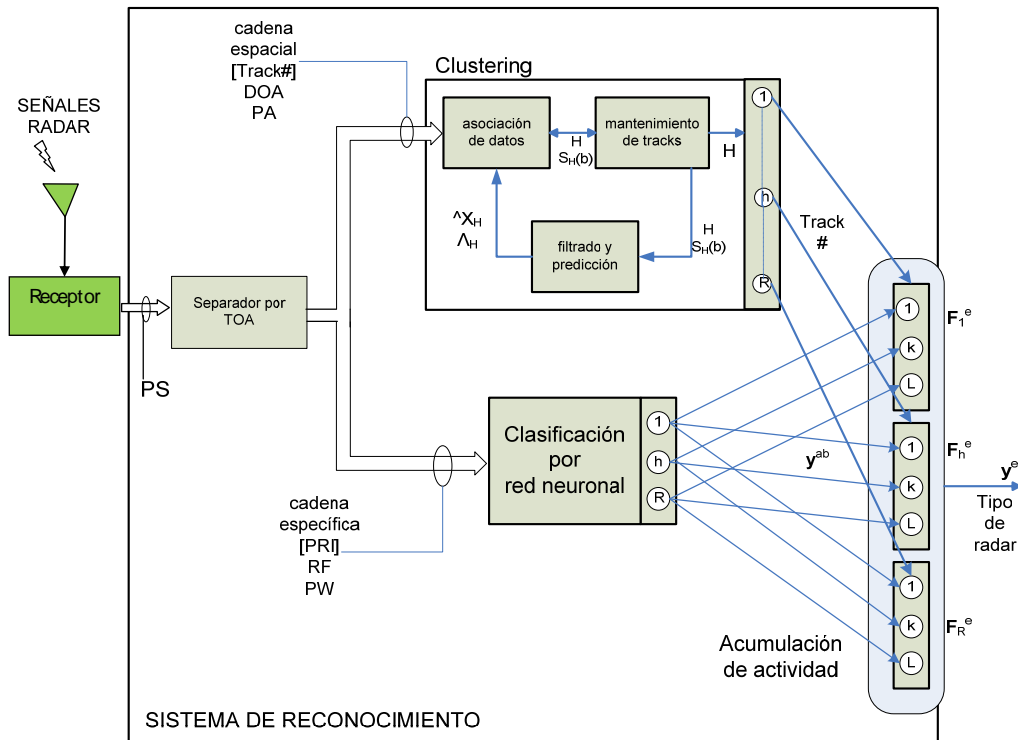


Fig. 1. Diagrama en bloques de un sistema ESM
(los corchetes indican que el campo puede estar vacío para algunos pulsos)

Antes de la operación en línea, la red neuronal se entrena mediante aprendizaje supervisado, usando datos de radares colectados en campo y etiquetados con su respectivo tipo de radar. Sólo los parámetros específicos de tipo se utilizan en el entrenamiento, incluyendo el PRI si está disponible. Durante la operación online, el sistema de reconocimiento procesa la secuencia de PS correspondiente a los pulsos interceptados, así como el número de track y PRI de cada pulso agrupado por el separador por TOA. Los parámetros específicos de tipo son enviados a la red neuronal y los espaciales al subsistema de clustering. Para cada pulso entrante el subsistema clasificador emite una predicción del tipo de radar. Esta predicción tiene la forma de un patrón denotado por y^{ab} . A su vez, el subsistema de clustering intenta agrupar los pulsos en tracks basándose en los parámetros espaciales. Si el número de track ya fue provisto por el separador por TOA, la salida de este subsistema es descartada.

La fusión de las respuestas de los subsistemas de clasificación y de clustering se realiza en el módulo de acumulación de actividad. Los patrones de respuesta y^{ab} obtenidos por el subsistema de clasificación son acumulados a lo largo del tiempo según los tracks, es decir según los agrupamientos surgidos de la información espacial. Los números de track obtenidos en el módulo de clustering indican los emisores a los que se asocian las PS. La fusión se implementa mediante un conjunto de campos de acumulación de actividad, donde cada campo F_h^e corresponde a un track $h = 1, 2, \dots, R$. La asignación de un track $h = H$ a una PS habilita un campo de acumulación de actividad F_H^e para que acumule los patrones de salida y^{ab} del módulo de clasificación. La

La arquitectura ARTMAP simplificada se obtiene combinando una red ART no supervisada con un campo de mapeo. La variante fuzzy ARTMAP puede procesar información tanto analógica como digital, empleando fuzzy ART como red. Una red neuronal Fuzzy ART consiste en dos capas de nodos totalmente conectados: una capa de entrada F_1 con M nodos, y una capa competitiva F_2 con N nodos. F_1 y F_2 están conectadas mediante un conjunto de pesos de valores reales $\mathbf{W} = \{w_{ij} \in [0, 1]: i = 1, 2, \dots, M; j = 1, 2, \dots, N\}$. Cada nodo j de F_2 representa una categoría de reconocimiento que aprende un vector prototipo $\mathbf{w}_j = (w_{1j}, w_{2j}, \dots, w_{Mj})$. La capa F_2 está conectada, a través de vínculos asociativos, a un campo de mapeo F^{ab} de L nodos, donde L es el número de clases en el espacio de salida. Un conjunto de pesos binarios $\mathbf{W}^{ab} = \{w_{jk}^{ab} \in [0, 1]: j = 1, 2, \dots, N; k = 1, 2, \dots, L\}$ se asocia con las conexiones de F_2 a F^{ab} . El vector $w_j^{ab} = (w_{j1}^{ab}, w_{j2}^{ab}, \dots, w_{jL}^{ab})$ vincula el nodo j de F_2 con una de las L clases de salida. Durante el entrenamiento la red hace aprendizaje supervisado del mapeo entre los vectores de entrenamiento $\mathbf{a} = (a_1, a_2, \dots, a_m)$ y las salidas $\mathbf{t} = (t_1, t_2, \dots, t_L)$, donde $t_k = 1$ si K es la clase para \mathbf{a} , y cero en caso contrario.

La siguiente secuencia describe el algoritmo de aprendizaje fuzzy ARTMAP:

1. **Inicialización.** Inicialmente todos los nodos F_2 están libres, todos los pesos w_{ij} son iguales a 1 y todos los pesos w_{jk}^{ab} son iguales a 0. Un nodo de F_2 se dice comprometido cuando es seleccionado para codificar a un vector de entrada \mathbf{a} , y vinculado con un nodo de F^{ab} . Los parámetros de la tasa de aprendizaje $\beta \in [0, 1]$, la selección $\alpha > 0$ y la vigilancia $\rho \in [0, 1]$ son establecidos a su valor inicial.
2. **Codificación del patrón de entrada.** Cuando un par de vectores de entrenamiento (\mathbf{a}, \mathbf{t}) es presentado a la red, \mathbf{a} sufre una transformación denominada codificación complementaria, que duplica el número de sus componentes. El patrón de entrada así codificado tiene $M = 2m$ dimensiones y está definido por $\mathbf{A} = (\mathbf{a}, \mathbf{a}^c) = (a_1, a_2, \dots, a_m; a_1^c, a_2^c, \dots, a_m^c)$, donde $a_i^c = (1-a_i)$, y $a_i \in [0, 1]$. El parámetro de vigilancia ρ es reseteado a su valor base.
3. **Selección del prototipo.** El patrón \mathbf{A} activa la capa F_1 y se propaga a través de los pesos de conexión \mathbf{W} a la capa F_2 . La activación de cada nodo j en la capa F_2 está determinada por la función de selección de Weber:

$$T_j(\mathbf{A}) = \frac{|\mathbf{A} \wedge \mathbf{w}_j|}{\alpha + |\mathbf{w}_j|} \quad (1)$$

Donde $|\cdot|$ es el operador norma, $|\mathbf{w}_j| \equiv \sum_{i=1}^M |w_{ij}|$

\wedge es el operador difuso AND, $(\mathbf{A} \wedge \mathbf{w}_j)_i \equiv \min(A_i, w_{ij})$

α es el parámetro de selección definido por el usuario.

La capa F_2 produce, por medio de una competencia winner-take-all, un patrón binario de actividad $\mathbf{y} = (y_1, y_2, \dots, y_N)$ tal que solamente el nodo $j = J$ con el mayor valor de activación $J = \operatorname{argmax}\{T_j : j = 1, 2, \dots, N\}$ permanece activo; entonces $y_J = 1$ e $y_j = 0 \forall j \neq J$. Si más de un T_j es máximo, es seleccionado el nodo j con el menor índice. El nodo J propaga su expectativa descendente, o vector prototipo \mathbf{w}_J , de vuelta hacia F_1 y se realiza la prueba de vigilancia. Esta prueba compara el grado de semejanza entre \mathbf{w}_J y \mathbf{A} contra el parámetro de vigilancia adimensional ρ :

$$\frac{|\mathbf{A} \wedge \mathbf{w}_J|}{M} \geq \rho \quad (2)$$

Si se supera la prueba el nodo J permanece activo, y se dice que se alcanzó la resonancia. De lo contrario, la red inhibe el nodo activo de F_2 (o sea que T_j es puesto a 0 hasta que sea presentado a la red el próximo par de entrenamiento (\mathbf{a} , \mathbf{t})) y busca otro nodo J que pase la prueba de vigilancia. Si ese nodo no existe, se activa un nodo libre. La profundidad de la búsqueda, antes de que se seleccione un nodo libre está determinada por el parámetro de selección α .

4. **Predicción de clase.** El patrón de clase \mathbf{t} es aplicado directamente al campo de mapeo F^{ab} , mientras la categoría \mathbf{y} es enviada de F_2 al campo de mapeo a través de los pesos asociativos \mathbf{W}^{ab} . La capa F^{ab} genera un patrón binario de actividad $\mathbf{y}^{ab} = (y_1^{ab}, y_2^{ab}, \dots, y_L^{ab})$, en el cual el nodo K más activo produce la predicción de clase $K = k(J)$. Si el nodo K constituye una predicción de clase incorrecta, una señal de seguimiento de la semejanza aumenta el parámetro de vigilancia ρ , justo lo suficiente para inducir otra búsqueda entre los nodos de F_2 , en el paso 3. Esta búsqueda continúa hasta que es activado un nodo libre de F_2 (y el aprendizaje sigue con el paso 5), o hasta que se active un nodo J que haya aprendido previamente la predicción correcta.
5. **Aprendizaje.** El aprendizaje de la entrada \mathbf{a} involucra la actualización del vector de pesos \mathbf{w}_j , y si J corresponde a un nodo recientemente comprometido, la creación de un vínculo asociativo con F^{ab} . El vector de pesos del nodo J de F_2 es actualizado de acuerdo a

$$\mathbf{w}'_J = \beta(\mathbf{A} \wedge \mathbf{w}_j) + (1 - \beta)\mathbf{w}_j \quad (3)$$

donde β es la tasa de aprendizaje. El algoritmo puede configurarse para aprendizaje lento con $0 < \beta < 1$, o para aprendizaje rápido con $\beta = 1$. Con codificación complementaria y aprendizaje rápido, fuzzy ART representa a una categoría j como un hiper rectángulo R_j m-dimensional, que tiene el tamaño justo para contener el cluster del patrón de entrenamiento \mathbf{a} , al que ha sido asignado. Para aprender una nueva asociación entre el nodo J de F_2 y el nodo K de F^{ab} ($k(J) = K$), se hace $w_{jk}^{ab} = 1$ para $k = K$, donde K es la clase deseada para \mathbf{a} , o cero para $k \neq K$. Una vez que los pesos \mathbf{W} han convergido para el conjunto de patrones de entrenamiento, ARTMAP puede predecir la clase para un patrón de entrada cualquiera, ejecutando los pasos 2, 3, y 4 sin ninguna vigilancia, o prueba de semejanza.

2.2. Clustering

El objetivo del módulo de clustering en la arquitectura de reconocimiento ilustrada en la figura 1, es agrupar en tracks los patrones de la cadena espacial. Las señales entrantes contienen información acerca del estado del emisor, que puede cambiar con el tiempo. Es deseable que este clustering online tenga la capacidad de inicializar nuevos tracks cada vez que es detectado un emisor nuevo, de ajustar los tracks en respuesta a los cambios de los emisores y de eliminar tracks cuando los emisores desaparecen o dejan de transmitir.

Varias técnicas pueden realizar clustering secuencial online. Por ejemplo los algoritmos de cuantificación adaptativa de vectores (Gray, 1984) pueden funcionar bien cuando el valor de los parámetros varía lentamente. Si los parámetros exhiben variaciones rápidas lineales o no lineales, se requieren algoritmos más sofisticados.

En este caso el clustering online de los parámetros DOA (Direction Of Arrival) y PA (Pulse Amplitud), es implementado combinando el método del vecino más próximo con filtrado Kalman lineal. Un filtro Kalman es un algoritmo de optimización recursivo. En la figura 1 se muestra una descomposición del proceso recursivo requerido para el clustering online. Las funciones básicas son asociación de datos, mantenimiento de tracks y filtrado y predicción.

2.2.1. Asociación de datos

Cuando llega un patrón entrante \mathbf{b} ($\mathbf{b} = \{\text{DOA}, \text{PA}\}$) de la cadena espacial se intenta asociar inicialmente con algún track existente. La asociación implica computar una semejanza $s_h(\mathbf{b})$ entre la entrada \mathbf{b} y la próxima posición predicha de cada track ($h = 1, 2, \dots, R$) en el entorno espacial. Supongamos que las posiciones de los tracks son establecidas en forma independiente e idéntica, donde cada track lleva asociada una distribución gaussiana. Entonces la semejanza puede tomarse como una probabilidad:

$$s_h(\mathbf{b}) = \frac{1}{(2\pi)^{\frac{M}{2}} |\Lambda_h|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\mathbf{b} - \hat{\mathbf{x}}_h)^T \Lambda_h^{-1} (\mathbf{b} - \hat{\mathbf{x}}_h)\right\} \quad (4)$$

Donde M es el número de dimensiones en el entorno espacial, y $\hat{\mathbf{x}}_h$ y Λ_h son respectivamente la posición predicha y la matriz de covarianza para el track h . Suponiendo que todos los tracks tienen igual probabilidad a priori, el track $h = H$ que maximiza la ecuación (4) es asociado con el patrón \mathbf{b}

$$\mathbf{H} = \text{argmax}_h \{s_h(\mathbf{b}) : h = 1, 2, \dots, R\} \quad (5)$$

El filtrado Kalman es utilizado para predecir la próxima posición $\hat{\mathbf{x}}_h$ y la matriz de covarianza Λ_h de cada track h . Recordemos que el clustering es obviado cuando \mathbf{b} corresponde a una PS que ha sido asignada a un track previamente establecido por el separador por TOA. En ese caso \mathbf{b} retiene su track y no se realizan la asociación de datos ni el mantenimiento de track. El filtrado y predicción Kalman sin embargo se sigue realizando para sustentar una descripción coherente de todos los emisores activos del medio.

2.2.2. Mantenimiento de tracks

Una vez asociado con el patrón \mathbf{b} , el track H atraviesa dos pruebas. En la primera, la semejanza $s_h(\mathbf{b})$ se compara contra un umbral δ_c que regula la creación de nuevos tracks, $\delta_c \in [0, 1]$. Si $s_h(\mathbf{b}) \geq \delta_c$ la prueba se considera superada. En la segunda prueba se computa el promedio acumulativo de los valores de semejanza S_H :

$$S_H = \frac{\sum s_H(\mathbf{b})}{Q_H} \quad (6)$$

Donde Q_H es el número de patrones a los cuales se les asignó el track H . S_H es comparado contra otro umbral δ_d que regula la calidad global de los tracks existentes, $\delta_d \in [\delta_c, 1]$. Si $S_H \geq \delta_d$ la prueba se considera superada. Si el track H pasa ambas pruebas, es asignado al patrón \mathbf{b} .

Si cualquiera de las dos pruebas falla, un nuevo track es iniciado para \mathbf{b} . Cuando un nuevo track H es iniciado, $s_h(\mathbf{b}) = 1$, $\hat{\mathbf{x}}_h$ es igualado a \mathbf{b} y Λ_h es igualado a $\sigma^2 I_M$, donde I_M es la matriz identidad, y

$\sigma = (\sigma_1, \sigma_2, \dots, \sigma_M)$ representa la resolución de la medición de los parámetros espaciales. Además, si la segunda prueba falla (es decir $S_H < \delta_d$), el track H previamente establecido es eliminado.

Después de la asignación de \mathbf{b} a un track existente, o recientemente iniciado, cualquier otro track que no haya sido asignado a un patrón de entrada, por un período de tiempo mayor a un parámetro $\tau > 0$, es eliminado. Es decir que un track es eliminado si:

$$\text{TOA}(\mathbf{b}) - \text{TOA}_h > \tau \quad (7)$$

Donde TOA_h es el tiempo en el cual el track h fue asignado por última vez a una entrada. Borrar tracks libera recursos y reduce las probabilidades de futuras asignaciones erróneas. El número de track H es la salida del proceso de mantenimiento de tracks.

Un track de alta calidad es aquel al cual se le asignaron un gran número de pulsos transmitidos por el mismo emisor. Para asegurar tracks de alta calidad, nuevos tracks se inician rápidamente estableciendo δ_c cerca de 1, a la vez que los tracks de baja calidad son eliminados más lentamente (a medida que S_H declina progresivamente). Esto reduce la ambigüedad durante la asignación de tracks.

2.2.3. Filtrado y predicción Kalman

El filtrado y predicción se implementa con un banco de filtros Kalman estándar, uno por track. Cada track h es asociado con un filtro Kalman y representado por una distribución Gaussiana unimodal. Después de la asignación del track H a \mathbf{b} , el filtro de H es empleado para predecir la próxima posición y la matriz de covarianza.

2.3. Fusión de la información específica y espacial

La fusión de la información específica y espacial es implementada por medio de campos idénticos de acumulación de actividad $F_1^e, F_2^e, \dots, F_R^e$, donde cada campo F_h^e está conectado a un track h , y replica el campo de salida de la red neuronal, es decir que contiene L nodos, uno por cada tipo de radar. Los nodos de salida de la red neuronal están conectados a sus respectivos nodos en todos los campos F_h^e , $h = 1, 2, \dots, R$. Cada campo F_h^e incorpora una memoria de corto plazo capaz de acumular sus patrones de entrada. La memoria de F_h^e es caracterizada por un patrón de acumulación $\mathbf{T}_h^e = (T_{h1}^e, T_{h2}^e, \dots, T_{hL}^e)$.

Tras la inicialización de un track h , \mathbf{T}_h^e se iguala a $\mathbf{0}$. Cuando el track $h = H$ se asigna a un patrón \mathbf{b} , F_H^e se activa. El patrón de actividad \mathbf{y}^{ab} de salida del clasificador se acumula en F_H^e de acuerdo a:

$$(\mathbf{T}_H^e)' = \mathbf{T}_H^e + \mathbf{y}^{ab} \quad (8)$$

La acumulación de patrones de actividad en F_h^e continúa hasta que el track h es eliminado. Para una PS dada de entrada, el patrón de actividad \mathbf{y}^e de salida del módulo de acumulación es igual a \mathbf{T}_H^e . El tipo de radar se predice como:

$$K^e = \text{argmax}_k \{T_{Hk}^e; k^e = 1, 2, \dots, L\} \quad (9)$$

Además de la predicción discreta, los campos de acumulación de actividad pueden ser usados para alimentar una tabla de emisores. Los campos también pueden describir múltiples tipos de radar que comparten las mismas características espaciales (misma posición geográfica), lo cual ayuda a vincular los emisores con las plataformas que los transportan.

La acumulación secuencial de actividad mejora la tasa de clasificación global del sistema de reconocimiento, debido a que la predicción acumulada K^e es tolerante a los errores cometidos por la red neuronal (predicción K). El concepto de predecir clases desde múltiples "vistas" de un emisor, que son acumuladas en el tiempo, ha sido exitosamente desarrollado en varias arquitecturas neuronales (Granger, E., Rubin, M. A., Grossberg, S., & Lavoie, P. 2000), (Baloch, A. A., & Waxman, A. M., 1991). La efectividad de la acumulación de actividad depende de la calidad de los tracks computados.

2.4. Simulaciones comparativas

La arquitectura propuesta fue comparada mediante simulaciones por computadora con los clasificadores kNN y RBF para referencia paramétrica y no paramétrica respectivamente. Previa a cada simulación, los datos de los pulsos radar fueron particionados en dos subconjuntos de entrenamiento y prueba. El 50% de la información de cada tipo de radar fue seleccionado al azar para formar el subconjunto de entrenamiento. Luego, los vectores de entrenamiento \mathbf{a} , junto con la etiqueta del tipo de radar \mathbf{t} , fueron presentados repetidamente hasta la convergencia de la red. El mismo orden aleatorio fue usado en todas las presentaciones. La convergencia se alcanzaba cuando la suma cuadrática del cambio fraccional del vector de pesos \mathbf{W} era menor que 0.001, para dos ciclos sucesivos. Un ciclo se define como la presentación de todo el subconjunto de entrenamiento al clasificador, por orden de TOA. El clasificador RBF usado en esta comparación selecciona subconjuntos de patrones de entrenamiento, uno por uno para codificar capas ocultas (Chen, S., Cowan, C. F. N., & Grant, P. M., 1991). En este caso la convergencia se alcanzaba cuando la suma cuadrática de los errores entre la salida real y la deseada caía por debajo de 0.01. Después de la convergencia, el subconjunto de prueba era presentado al clasificador ya entrenado, para probar la predicción.

La performance de cada clasificador fue evaluada en términos de cantidad de recursos consumidos y precisión predictiva. La cantidad de recursos asignados durante el entrenamiento se mide de 3 maneras. La *compresión* se refiere a la relación promedio de patrones de entrenamiento a nodos comprometidos en la capa F_2 . La *memoria* es el número de registros normalizados necesarios para almacenar el conjunto de vectores de pesos, siendo un registro normalizado un conjunto de bits suficientes para almacenar los valores reales del clasificador, como a_i , w_{ji} , ρ , etc. *Tiempo de convergencia* es el número de ciclos requeridos para que el clasificador converja. La precisión predictiva sobre el subconjunto de prueba se midió usando la *tasa de clasificación*, o sea la relación de patrones correctamente clasificados sobre el total de patrones de prueba.

La tabla 1 contiene el promedio de 20 simulaciones. Los parámetros de la red fueron elegidos por prueba y error para lograr la mejor tasa de clasificación, con la menor memoria y tiempo de convergencia del entrenamiento. Los resultados indican que la red fuzzy ARTMAP logró consistentemente una buena tasa de clasificación promedio, comparable con las obtenidas por los clasificadores kNN y RBF. Las arquitecturas ARTMAP alcanzan la misma tasa de clasificación con mayor compresión (menos memoria física para almacenar los vectores) y menos ciclos de entrenamiento para la convergencia.

Clasificador	Criterio de evaluación			
	Precisión	Recursos		
		Tasa de clasificación	Compresión	Memoria
k-Nearest-Neighbor ($k=1$, $d_{\text{Euclídea}}$)	99,53%	1	80340	N/A
RBF (Spread del kernel = 0.05)	99,57%	4.2	6370	2125
Fuzzy ARTMAP ($\varepsilon=10^{-4}$, $\alpha=.001$, $\beta=1$, $\bar{p}=0$)	99,45%	220	733	3.6

Tabla 1. Resultados de la comparación.

También se realizaron ensayos con datos de pulsos faltantes y etiquetas de clase faltantes durante el entrenamiento y durante la prueba, todos con resultados muy aceptables.

Existen varias extensiones del algoritmo básico fuzzy ARTMAP que aportan sustanciales mejoras al rendimiento. Entre estas cabe mencionar: ARTMAP-IVS *estrategia del vector indicador*

(Granger, E., Rubin, M. A., Grossberg, S., & Lavoie, P., 2000) para el manejo de datos faltantes, ARTMAP-FD *discriminación de familiaridad* (Granger, E., Grossberg, S., Rubin, M. A., & Streilein, W. W., 1999) para la detección de pulsos de clases desconocidas, ARTMAP-MT-*seguimiento negativo de la semejanza* para mejorar la convergencia y evitar la proliferación de categorías y ARTMAP-LUC *aprendizaje de clases desconocidas* para continuar ajustando los pesos durante la fase de operación.

3. CONCLUSIONES

Se ha evaluado una arquitectura de ESM alternativa para el reconocimiento e identificación de emisiones de radar. Esta arquitectura combina un clasificador basado en una red neuronal, un algoritmo de clustering online y un módulo de acumulación de actividad. Una vez entrenada con muestras tomadas de campo, la red neuronal clasifica y predice el tipo de radar de los pulsos interceptados, basándose en los parámetros específicos de tipo. A su vez, el algoritmo de clustering separa los pulsos de cada emisor, basándose en los parámetros espaciales. El módulo de acumulación de actividad permite fusionar las respuestas del clasificador con las estimaciones del algoritmo de clustering, y de esta manera predecir la clase del emisor más probable. La implementación particular del sistema consiste en una variante de la red fuzzy ARTMAP para clasificación, y el método del vecino más próximo junto con un banco de filtros Kalman para clustering online.

Las simulaciones demostraron que la red fuzzy ARTMAP obtuvo consistentemente un alto nivel de precisión y compresión, aún con una cantidad de datos de entrenamiento muy limitada. Comparada con los clasificadores de referencia RBF y kNN, esta implementación presenta una tasa de clasificación semejante, requiriendo menor cantidad de recursos (menor tiempo de convergencia y menos memoria para almacenar los prototipos).

Los resultados obtenidos son promisorios para abordar aplicaciones de sistemas ESM autónomos, que deban enfrentarse a datos de emisores complejos, incompletos y superpuestos.

4. BIBLIOGRAFÍA

- Anderberg, M. R. (1973). *Cluster Analysis for Applications*. New York, NY: Academic Press.
- Baloch, A. A., & Waxman, A. M. (1991). Visual Learning, Adaptive Expectation, and Behavioral Conditioning of the Mobile Robot MAVIN. *Neural Networks*, 4, 271-302.
- Bar-Shalom, Y., & Li, R. A. (1993). *Estimation and Tracking: Principles, Techniques and Software*. Boston, MA: ARtech House.
- Blackman, S. S. (1986). *Multiple-Target Tracking with Radar Applications*. Norwood, MA: Artech House.
- Browne, J. P. R., & Thurbon, M. T. (1998). *Electronic Warfare*. London, UK: Brassey's.
- Carpenter, G. A., and Grossberg S. (1987). A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine. *Computer Vision, Graphics, and Image Processing*, 37, 54-115.

- Carpenter, G. A., Grossberg, S., & Reynolds, J. H. (1991). ARTMAP: Supervised Real Time Learning and Classification of Nonstationary Data by a Self-Organizing Neural Network. *Neural Networks*, 4, 565-588.
- Carpenter, G. A., Rubin, M. A., & Streilein, W. W. (1997a). ARTMAP-FD: Familiarity Discrimination Applied to Radar Target Recognition. *Pro. Int'l Conference on Neural Networks, III*, pp. 1459-1464.
- Carpenter, G.A., Grossberg, S., Markuzon, N., Reynolds, J. H., & Rosen, D. B. (1992). Fuzzy ARTMAP: A Neural Network Architecture for Incremental Supervised Learning of Analog Multidimensional Maps. *IEEE Transactions on Neural Networks*, 3 (5), 698-713.
- Chen, S., Cowan, C. F. N., & Grant, P. M. (1991). Orthogonal Least Square Learning Algorithm for Radial Basis Function Networks. *IEEE Transactions on Neural Networks*, 2 (2), 302-309.
- Davies, C. L., & Hollands, P. (1982). Automatic Processing for ESM. *Proceedings of IEE*, 129 (3), 164-171.
- Granger, E., Grossberg, S., Rubin, M. A., & Streilein, W. W. (1999). Familiarity Discrimination of Radar Pulses. (M. Press, Ed.) *Advances in Neural Information Processing Systems*, 11, 875-881.
- Granger, E., Rubin, M. A., Grossberg, S., & Lavoie, P. (2000). Classification of Incomplete Data Using the Fuzzy ARTMAP Neural Network. *Proc. Int'l Joint Conference on Neural Networks, IV*, pp. 35-40.
- Granger, E., Rubin, M. A., Grossberg, S., & Lavoie, P. (2000). *A What and Where Fusion Neural Network for Recognition and Tracking of Multiple Radar Emitters*. Technical Report CAS/CNS-TR-2000-029, Boston University, Boston, MA.
- Gray, R. M. (1984). Vector Quantization. *IEEE ASSP Magazine*, 1 (2), 4-29.
- Grossberg, S. (2000). *Adaptive Resonance Theory*. Technical Report CAS/CNS-2000-024, Boston University, Center for Adaptive Systems and Department of Cognitive and Neural Systems, Boston, MA.
- Wiley, R. G. (1993). *Electronic Intelligence: The Analysis of Radar Signals* (2nd ed.). Norwood, MA: Artech House.