

# Clustering Dinámico para Tiempo de Encendido Mínimo en Redes Inalámbricas de Sensores (CLUDITEM) Análisis de la Definición de Clusters

Rosa Corti<sup>1</sup>, Estela D'Agostino<sup>1</sup>, Enrique Giandomenico<sup>1</sup>,  
Javier Belmonte<sup>1</sup>, Roberto Martinez<sup>1</sup>

<sup>1</sup> Facultad de Cs. Exactas, Ingeniería y Agrimensura, Universidad Nacional de Rosario,  
Avenida Pellegrini 250, 2000-Rosario. Argentina  
{rcorti, estelad, giandome, belmonte, romamar}@fceia.unr.edu.ar

**Resumen.** El gran avance en comunicaciones y nanotecnologías ha permitido el desarrollo e implementación de Redes Inalámbricas de Sensores Inteligentes (RISI). En este trabajo se analiza la fase de definición de clusters de un algoritmo jerárquico para adquisición periódica de datos en aplicaciones de supervisión ambiental. Este algoritmo se gestó teniendo como objetivo disminuir y distribuir adecuadamente el consumo de energía, para lo cual trabaja con clustering dinámico y mantiene apagados los transeptores de los nodos el mayor tiempo posible. Se describen las simulaciones de la etapa de definición de clusters y se reportan las conclusiones obtenidas respecto a la configuración de la red y a la modalidad de postulación de las cabeceras. Finalmente se proponen líneas de trabajo futuro teniendo en cuenta las conclusiones alcanzadas.

**Palabras clave:** redes de sensores inalámbricas, algoritmos de encaminamiento, clustering dinámico.

## 1 Introducción

El gran avance en comunicaciones y nanotecnologías ha permitido el desarrollo y aplicación de Redes Inalámbricas de Sensores Inteligentes (RISI). Estas redes se utilizan para adquirir información en ambientes de características muy diversas. En particular se las incorpora en aplicaciones industriales, médicas, agrícolas, de preservación del medio natural o creación de ambientes inteligentes, entre otras. Las RISI están constituidas por nodos que además de recolectar datos del ambiente, son capaces de procesarlos y colaborar con sus vecinos para transmitirlos hacia la/las estaciones base. Estas redes se auto-organizan para adaptarse a topologías cambiantes, y deben trabajar bajo fuertes restricciones de energía, tratando de maximizar su tiempo de vida útil [1] [2].

Una RISI consta de un gran número de nodos constituidos por cuatro módulos: adquisición de datos, procesamiento, transceptor y potencia [3]. Cada nodo, de acuerdo a las tareas que realice, puede estar en uno de los siguientes estados: transmisión, recepción, en escucha y apagado. En transmisión, el nodo ha ganado acceso al medio y envía información a la red. En recepción, detecta la llegada de un mensaje enviado por un vecino, en estado de escucha está analizando el medio y cuando se apaga, no realiza actividad alguna consumiendo el mínimo de energía posible. Por lo tanto, cuanto más tiempo permanezca un nodo en este último estado, consumirá menos y su tiempo de vida será más prolongado [4].

En este trabajo se analiza la formación de clusters de un algoritmo de encaminamiento, desarrollado por el Grupo de Redes Inalámbricas de Sensores Inteligentes (GRISI) [5]. El entorno a supervisar se divide en una cuadrícula uniforme para lograr una distribución favorable de las cabeceras de clusters (CH). Además un objetivo de CLUDITEM es maximizar el tiempo que los nodos están apagados, como se explica en la sección cuatro.

Este artículo se organiza de la siguiente forma: en la sección 2 se mencionan otros trabajos relacionados con el aquí presentado, en la sección 3 se describe brevemente el algoritmo desarrollado, en la sección 4 se discute el consumo de energía con el fin de determinar el momento de rotación de cabeceras, y en la sección 5 se describen las simulaciones realizadas. Finalmente en las secciones 6 y 7 se enumeran las conclusiones alcanzadas y se proponen líneas de trabajo futuro.

## 2 Trabajos Relacionados

Los algoritmos de encaminamiento jerárquicos basados en clusters, proporcionan importantes ventajas para las RISI, respecto a su escalabilidad y eficiencia en la comunicación entre los nodos [6]. La definición de clusters permite realizar agregación de datos, que prolonga la vida útil de la red al reducir la cantidad de transmisiones y disminuye la congestión acotando el número de colisiones [7].

En particular, CLUDITEM utiliza la idea de rotación de cabeceras de cluster propuesta por Heinzelman en su algoritmo LEACH [8]. La rotación de CH permite distribuir el consumo de energía entre los nodos, ya que el rol de cabecera es exigente respecto al uso de recursos. Sin embargo, un inconveniente importante de LEACH es que al seleccionar las cabeceras en forma aleatoria, las mismas pueden quedar concentradas en algún sector de la red. Esto podría ocasionar la desconexión de un número indeterminado de nodos, como se puntualiza en [6]. Para prevenir esta situación, CLUDITEM divide el área bajo estudio en una cuadrícula virtual cuyo uso ya fue propuesto por Al-Karaki en [9]. En su algoritmo, los clusters coinciden con las divisiones de la cuadrícula, y el rol de cabecera se rota entre los nodos miembros de la misma teniendo en cuenta su energía remanente. En CLUDITEM, cuando se rotan las CH, se vuelven a definir los clusters, pudiendo incorporarse nodos de otras divisiones de la cuadrícula. Esto involucra un mayor número de mensajes de control, pero dichos mensajes son muy breves. La importante ventaja que se obtiene, es lograr clusters de menor profundidad, reduciendo el intercambio de mensajes de datos, de mayor longitud, en la comunicación multisalto hacia los CH.

En resumen, los objetivos de CLUDITEM al combinar una grilla virtual con la definición dinámica de clusters son lograr una distribución uniforme de CH en el área bajo estudio, y un consumo de energía menor y mejor distribuido entre los nodos de la red.

### **3 Algoritmo de Clustering Dinámico con Tiempo de Encendido Mínimo (CLUDITEM)**

El algoritmo desarrollado se orienta hacia aplicaciones de monitoreo ambiental. En las mismas la adquisición de datos se realiza en forma periódica, lo que implica que en cada período de sensado ( $T$ ), todos los nodos envían información útil a la estación base. Los nodos de la red son idénticos, respecto a recursos y energía inicial. Se distribuyen manualmente, en forma uniforme, una única vez y son fijos. La estación base (sink) es única y se encuentra fuera del área bajo estudio.

Una de las métricas más importantes para evaluar la performance de las RISI es el tiempo de vida de la red. Para utilizar esta métrica, es necesario fijar cuando se considera que la red ya no es útil, para lo cual se pueden establecer diversos criterios. [8]. Entre otros se puede citar: cuando muere el primer nodo, cuando la pérdida de mensajes es superior a un porcentaje establecido antes de poner en funcionamiento la red, o cuando una de las cuadrículas en que se dividió el área a supervisar queda desconectada. En CLUDITEM se adopta un porcentaje de mensajes perdidos, que se determina en base a los requerimientos de la aplicación abordada.

La comunicación entre los nodos de la red es por radiofrecuencia. En este algoritmo se propone que los nodos permanezcan apagados el mayor tiempo posible, para ahorrar energía.

#### **3.1 Generalidades**

CLUDITEM se desarrolló para dominios donde los valores de las variables de interés poseen una fuerte correlación espacio/temporal entre nodos vecinos. Aprovechando esta característica se propone utilizar clusters aplicando alguna función de agregación, que dependerá de la aplicación particular.

El área a supervisar se divide en una grilla virtual donde cada nodo está identificado. La identificación consiste en un byte del cual los cuatro bits más significativos corresponden a la cuadrícula (IDHi), y los cuatro menos significativos representan al nodo (IDLi). Este identificador se utiliza para definir algunos parámetros del algoritmo.

Por último, el funcionamiento de CLUDITEM se divide en dos rondas bien diferenciadas. La primera se ocupa del establecimiento del árbol de encaminamiento y la segunda del envío de datos al sink. Estas rondas se repiten periódicamente durante el tiempo de vida de la red [5].

### 3.2 Definición de Clusters

El árbol de encaminamiento jerárquico se define en dos niveles. El primer nivel establece la estructura de cada cluster. El segundo nivel se ocupa de la definición del árbol de CH, que se encarga de enviar los datos agregados hasta la estación base. Esta fase se considera crítica, ya que un fallo en la misma puede producir subredes desconectadas que afectan seriamente el funcionamiento de la red.

Cada cluster se define seleccionando en primer lugar su CH. Las cabeceras se postulan respetando un esquema de tipo TDMA. Los nodos comunes adhieren a una cabecera y eligen su nodo de enlace (NDE). Dentro del cluster la comunicación es multisalto y por lo tanto cada nodo común que escucha al CH reenvía el mensaje recibido con el objetivo de definir la estructura del cluster [5]. Luego los nodos que no cumplen el rol de cabecera duermen hasta la fase de envío de datos.

El comportamiento de un nodo durante la definición de los cluster se detalla en la figura 1. En la misma se muestra que cada nodo de una cuadrícula se postula como CH, respetando el TDMA definido, sólo si hasta ese momento no ha conseguido un NDE. Esta regla de postulación de CH sería efectiva si los relojes de los nodos estuvieran perfectamente sincronizados, y las postulaciones en las distintas celdas resultaran simultáneas. Sin embargo, no es realista suponer la existencia de un tiempo global para la red, y cada nodo posee un tiempo local desfasado respecto de sus vecinos. Esta situación ocasiona que al postularse el primer CH, un gran número de nodos se unen a su cluster, resultando una red con muy pocos nodos cabecera, como se muestra en la figura 2 obtenida por simulación. Esta situación no es deseable ya que no cumple con el objetivo de CLUDITEM de lograr una distribución uniforme de CH. Por lo tanto, el proceso de definición de clusters se modificó estableciendo una nueva regla para postulación de los nodos. De acuerdo al TDMA definido, el nodo se propone como cabecera si carece de NDE, o posee un nivel mayor a 1 y su NDE corresponde a una celda distinta a la suya. Esta nueva condición para postulación de los nodos permitió lograr una mejor distribución de CH en la red, situación que se verificó por simulación y se presenta en la sección cinco.

### 3.3 Rotación de Cabeceras

Las postulaciones se realizan en forma secuencial de acuerdo al IDLi de cada nodo.

El rol de cabecera de cluster es exigente respecto del consumo de energía, por lo cual se hace necesario rotarlo periódicamente con el fin de distribuir el consumo en forma más uniforme. Cuando un nodo asume este rol lo cumple durante una cierta cantidad de rondas de envío de datos (X) y luego se realiza la rotación de CH. Si las cabeceras mantienen su rol demasiado tiempo, se agotan y se producen desconexiones en la red. Si X es demasiado pequeño, se consume mucha energía en el rearmado del árbol de encaminamiento. Por lo tanto debe existir un valor de X que equilibre el gasto de energía en la red [8].

Cada nodo de la celda se postulará al inicio de la fase de armado del cluster, calculando el momento de envío de su postulación (TPosi) de acuerdo a la fórmula (1)

$$TPosi = (TP * IDLi) * (1 + n * NChi). \quad (1)$$

En (1),  $n$  es el número estimado de nodos en la cuadrícula, y  $NCH_i$  la cantidad de veces que el nodo  $i$  ha sido CH. El tiempo para transmitir un mensaje a la máxima cantidad de saltos es  $TP$ , y puede calcularse como (2)

$$TP = SM * (TM + TMAC). \tag{2}$$

En (2),  $SM$  es la cantidad de saltos estimada,  $TM$  es el tiempo de transmisión de una trama y  $TMAC$  contempla el tiempo de acceso al medio.

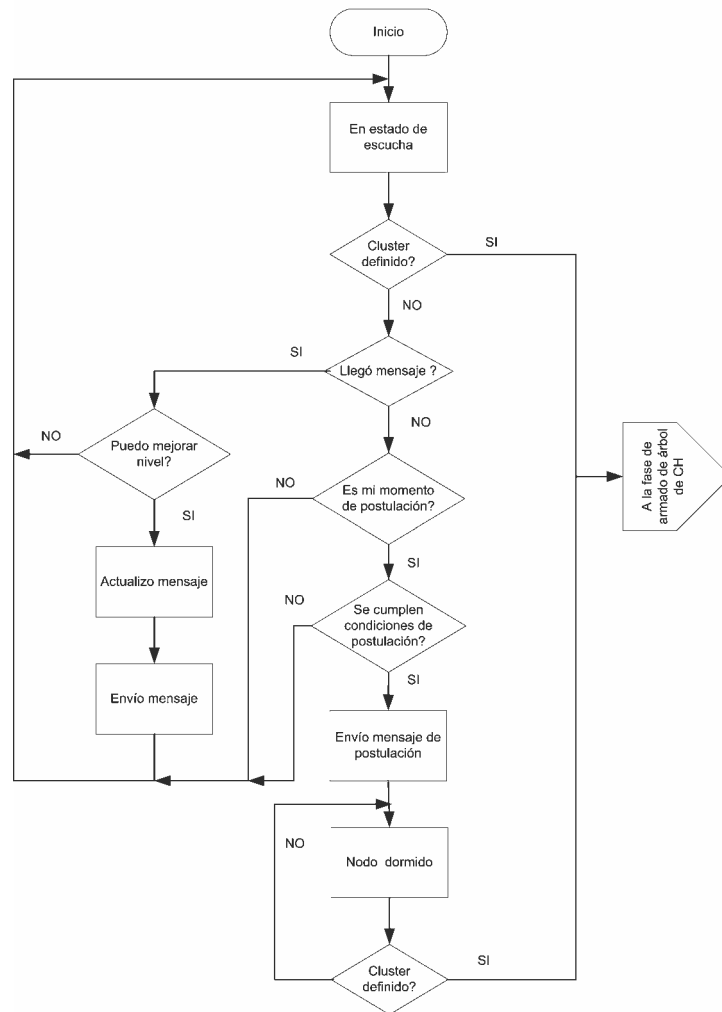


Fig. 1. Comportamiento de un nodo durante la fase de armado de clusters.

La primera vez que un nodo se postula  $NCH_i$  es cero, y luego agrega retardo al momento de postulación del nodo  $i$  con el fin de que el rol lo asuma otro nodo que

aún no fue CH. Cada nodo vuelve a cero el valor de su  $NCH_i$  luego que todos han cumplido el rol de cabecera. Se debe destacar, que si un nodo no consigue un NDE, se postulará como CH, por lo cual no quedan nodos desconectados en las celdas.

### 3.4 Definición del Árbol de Cabeceras y Envío de Datos

El segundo nivel del árbol de encaminamiento define la ruta de comunicación al sink de todas las cabeceras de cluster, los nodos comunes permanecen apagados. Este árbol se establece con el objetivo de que la información agregada alcance al nodo sink. Cada nodo cabecera adopta como NDE, al CH que escucha con menor nivel en el árbol.

Al finalizar la definición del árbol de encaminamiento, se procede al envío de datos hacia la estación base [5]. CLUDITEM, tiene una modalidad de envío de datos, periódica por salto, de acuerdo a la clasificación propuesta por Solis en [10].

## 4 Modelo de Energía

En este trabajo se propone transmitir con dos niveles de potencia distintos, uno más bajo para la comunicación intracluster y otro mayor para la comunicación intercluster.

### 4.1 Potencia de Transmisión

Se fija un radio de transmisión  $r$  en la transmisión intracluster, y las divisiones de la cuadrícula tienen una longitud  $2r$ . La mayor distancia entre un CH y un nodo común en una celda ( $d$ ), se presenta cuando el CH está un vértice y el nodo común se ubica en el vértice en diagonal. De aquí se concluye (3)

$$d = \sqrt{(2r)^2 + (2r)^2}. \quad (3)$$

$$d = 2r * \sqrt{2} < 3r. \quad (4)$$

Entonces, según (4) la distancia  $d$  resulta menor a 3 saltos. Como en los clusters pueden intervenir nodos de celdas adyacentes, no se puede asegurar que todos los nodos alcancen su CH con una cantidad de saltos menor o igual a tres, pero según los resultados de simulación, la máxima distancia entre un nodo común y su CH no excede a cuatro saltos.

### 4.2 Energía durante el Periodo de Rotación de CH

Un nodo de la red consume su energía, de acuerdo al estado en que se encuentra. En los estados de transmisión, recepción y escucha, el consumo resulta comparable, por el rango de potencia de transmisión que se maneja. La energía consumida en

procesamiento y en período de apagado se considera despreciable para el análisis. Por estos motivos CLUDITEM busca apagar los nodos el mayor tiempo posible.

Además, la energía inicial del nodo debe ser mayor o igual que la necesaria para un número entero de rotaciones completas de CH en una celda ( $Z$ ), para lograr una distribución uniforme del consumo (5). Esto garantiza que todos los nodos han sido CH la misma cantidad de veces ( $Z$ ).

$$EIN \geq Z * (E_{AECH} + E_{CH} * X + E_{AENC} * (n - 1) + E_{ENC} * (n - 1) * X). \quad (5)$$

$EIN$ : Energía inicial de un nodo.

$E_{CH}$ : Energía gastada por un CH para transmitir datos en un período de envío ( $T$ ).

$E_{NC}$ : Energía gastada por un NC para transmitir datos en  $T$ .

$E_{AECH}$ : Energía gastada por un CH para armar el árbol de encaminamiento.

$E_{AENC}$ : Energía gastada por un NC para armar el árbol de encaminamiento.

El armado del árbol de encaminamiento implica un gasto de energía considerable, y es deseable reducir la cantidad de veces que se lleva a cabo. Por lo tanto, para lograr que todos los nodos sean CH la misma cantidad de veces y minimizar el consumo, se asume que cada nodo es una única vez cabecera en su celda. Esto significa que  $Z$  es igual a uno. En consecuencia, a partir de la expresión anterior, se obtiene el valor más conveniente para  $X$  según (6):

$$X \leq \frac{EIN - E_{AECH} - E_{AENC} * (n - 1)}{E_{CH} + E_{ENC} * (n - 1)}. \quad (6)$$

## 5 Resultados de Simulación

Se simuló el armado de los clusters para una red constituida por nueve celdas con 10 nodos en cada una ( $n$ ). Se trabajó con el simulador desarrollado por el grupo GRISI, disponible en [11].

Para cada simulación se recogieron los siguientes datos: cantidad de CH por celda, cantidad de celdas sin CH, cantidad de nodos que pertenecen a una celda distinta a la de su CH y tiempo necesario para que los clusters estén definidos.

El objetivo de analizar los puntos 1 y 2, fue verificar la condición de que los CH resultaran adecuadamente distribuidos en el área bajo estudio. Esto es interesante ya que la adopción de una cuadrícula virtual tiene como meta una distribución uniforme de CH en el entorno supervisado. El punto 3 evidencia el hecho que si un nodo se encuentra a menor cantidad de saltos de un CH que no pertenece a su celda, respecto del que si pertenece, adopta al primero pues mejora su nivel. El punto 4 se plantea con el objetivo de investigar el número adecuado de postulantes para lograr un balance entre consumo y cantidad de nodos desconectados.

Las decisiones inicialmente plasmadas en el algoritmo, llevaron a una distribución de CH como la que se muestra en la figura 2. En la misma se destacan con un círculo, los dos únicos nodos que asumieron el rol de cabeceras. Por lo tanto, la red resultó conformada por pocos clusters y no se cumplió el objetivo de lograr una distribución

uniforme de CH, como se puntualizó en la sección 3.2. En la misma sección se propuso la modificación de las condiciones de postulación, y la nueva configuración de la red se muestra en la figura 3. Las figuras 2 y 3 se obtuvieron como resultado de dos simulaciones, entre todas las realizadas.

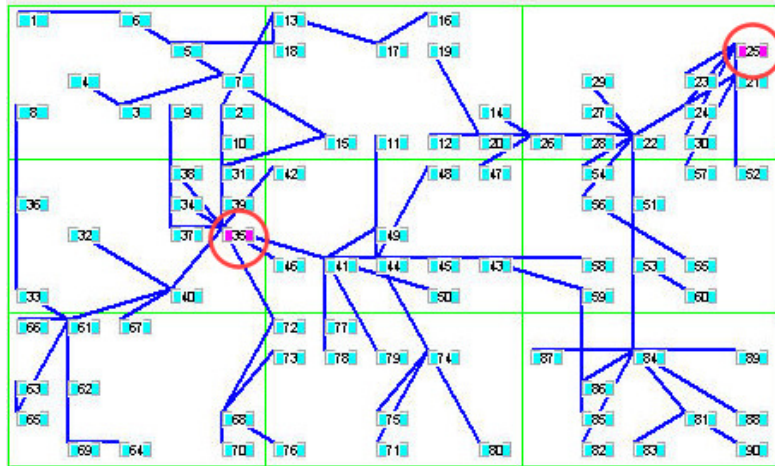


Fig. 2. Estructura de la red a partir de la primera versión de definición de clusters

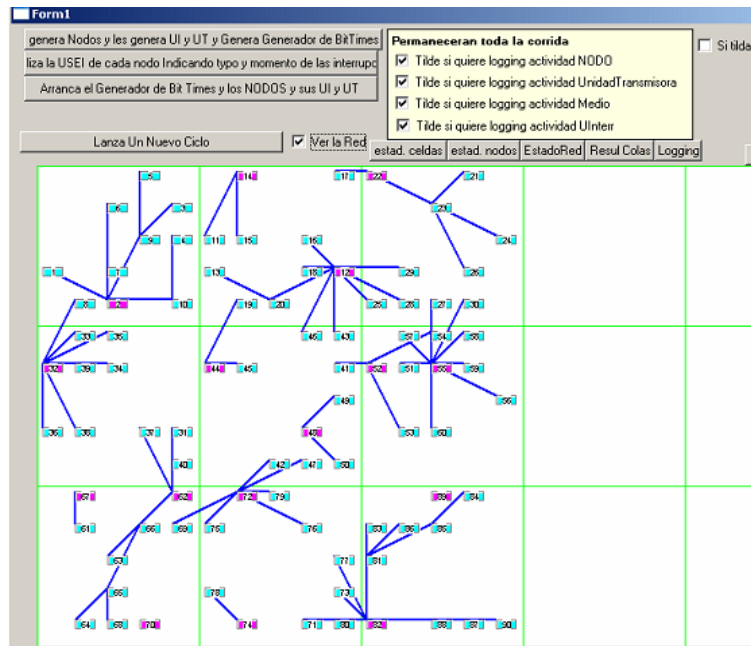


Fig. 3. Estructura de la red a partir de la segunda versión de definición de clusters

## 6 Conclusiones

La tabla 1 resume los resultados de todas las simulaciones realizadas y permite obtener conclusiones sobre varios aspectos del comportamiento de CLUDITEM.

En la segunda columna se observa que la cantidad promedio de CH por celda resulta siempre superior a uno. Sería deseable que el promedio de CH resultara lo más cercano a uno que fuera posible. En las simulaciones realizadas no quedan celdas sin CH, como muestra la columna cuatro. Sin embargo de acuerdo a la columna tres, lo que se observa en la configuración de la red es la presencia de celdas con más de un cabecera, hasta un máximo de cuatro. Esta situación lleva a un gasto de energía innecesario y a un agotamiento prematuro de algunos nodos de la red.

**Tabla 1.** Resultados de simulación obtenidos para la segunda modalidad de postulación de CH.

<i>Celda</i>	<i>Prom CH</i>	<i>Máx CH</i>	<i>Mín CH</i>	<i>Prom CH en celda</i>	<i>Máx tiempo</i>
1	1,83	4	1	10,86	5507
2	1,80	3	1	9,09	6221
3	1,66	3	1	9,71	6289
4	1,94	3	1	10,00	4561
5	1,60	3	1	9,34	5502
6	1,34	2	1	10,34	6187
7	1,57	3	1	10,14	5509
8	1,83	3	1	10,63	6244
9	2,00	3	1	9,89	5444

El hecho de que se defina más de un CH por celda, se debe a que los nodos que tienen un NDE de otra celda y se encuentran a más de un salto del CH que han elegido, se postulan a pesar de estar ya conectados. En la columna cinco se muestra el promedio de nodos que reportan a CH en la celda, estos nodos adoptan como CH a alguno de los existentes en su división. Sería deseable lograr un único CH por celda y que los valores de la columna cinco se aproximaran a 10, con el fin de obtener clusters más homogéneos y favorecer la distribución uniforme del consumo de energía.

Para mejorar la situación descrita, se trabajó sobre las condiciones de postulación de CH y la definición del contexto en el cual un nodo decide incorporarse a un cluster. A partir de este análisis se hicieron algunas modificaciones preliminares para la definición de los clusters y las simulaciones asociadas presentaron resultados más favorables. En este sentido se evaluó que es importante seguir avanzando con esta línea de trabajo y se espera realizar en breve un estudio más exhaustivo, probando por simulación las modificaciones que se propongan. Respecto al tiempo necesario para que los clusters queden definidos se observó que en todos los casos era muy inferior al tiempo total de postulaciones de todos los nodos, como muestra la columna seis. Esto lleva a pensar que no es necesario que todos los nodos deban potencialmente postularse. Restringiendo la cantidad de postulaciones es posible disminuir el tiempo de encendido de los nodos y de esta forma maximizar el tiempo de vida de la red.

## 7 Trabajo Futuro

En primer lugar se realizará el estudio detallado de las modificaciones propuestas para la definición de los clusters y se las ajustará teniendo en cuenta los resultados de simulación. Estas modificaciones tienden a lograr una única cabecera por celda y una estructura de clusters con una cantidad similar de nodos miembros. Esto es deseable para obtener un consumo uniforme de energía mejor distribuido entre los nodos de la red.

La cantidad de postulantes a CH por celda, puede disminuirse sin afectar seriamente la conectividad de la red de acuerdo a los resultados de las simulaciones realizadas. Por lo tanto se propone trabajar en este sentido para lograr un equilibrio entre cantidad de postulantes y número máximo de nodos aislados.

Finalmente se avanzará en la simulación y depuración de las etapas de definición del árbol de cabeceras y de envío de datos hacia la estación base.

## Bibliografía

1. Aboelaze M., Aloul F.: Current and Future Trends in Sensor Networks: A Survey. In: Proceedings of Wireless and Optical Communications Networks. WOCN05, Dubai (2005)
2. Cook D., Das S.: Smart Environments: Technologies, Protocols, and Applications. Wiley-Interscience, ISBN 0-471-54448-5, (2004)
3. D'Agostino E., Giandomenico E., Corti R., Martinez R., Belmonte J.: Algoritmo de encaminamiento para redes inalámbricas de sensores inteligentes. In: 36° Jornadas Argentinas de Informática 36 JAIIO. Mar del Plata, Argentina (2007)
4. Boukerche A., Cheng X., Linus J.: A Performance of a Novel Energy-Aware Data-Centric Routing Algorithm in Wireless Sensor Networks. Wireless Networks, vol. 11, pp 619--635. Springer (2005)
5. Corti R., D'Agostino E., Giandomenico E., Martinez R.: Algoritmo de encaminamiento para redes de sensores inalámbricas utilizando técnicas de agregación. In: XIII Congreso Argentino de Ciencias de la Computación XIII CACIC. Corrientes y Resistencia, Argentina (2007)
6. Al-Karaki J. N., Kamal A. E.: Routing techniques in Wireless Sensor Networks: a Survey. IEEE Wireless Communications, vol. 11, pp. 6--28. IEEE (2004)
7. Shrivastava N., Buragohain C., Agrawal D.: Medians and Beyond: New aggregation Techniques for Sensor Networks. University of California. In: Proceedings of the Second ACM Conference on Embedded Networked Sensor Systems SenSys 2004. ACM (2004)
8. Heinzelman W. Application-specific Protocol Architectures for Wireless Networks. In PhD Thesis, Massachusetts Institute of Technology (2000)
9. Al-Karaki J. N., UI-Mustafa R., Kamal A. E.: Data Aggregation in Wireless Sensor Networks - Exact and Approximate Algorithms. In: Workshop on High Performance Switching and Routing, pp. 241--245. IEEE (2004)
10. Solis I., Obraczka K.: In-network Aggregation Trade-offs for Data Collection in Wireless Sensor . International Journal of Sensor Networks, vol. 1, No 3/4 pp. 200--212. (2006)
11. Depto. de Sistemas e Informática FCEIA. Universidad Nacional de Rosario. Argentina <http://www.dsi.fceia.unr.edu.ar/downloads/sensores/Simulador.zip>