R. Rios, J. Cuellar, and J. Lopez, "Ocultación de la estación base en redes inalámbricas de sensores", *XI Jornadas de Ingeniería Telemática (JITEL 2013)*, pp. 481-486, 2013.
NICS Lab. Publications: https://www.nics.uma.es/publications

Ocultación de la estación base en redes inalámbricas de sensores

Ruben Rios
Lenguajes y Ciencias de la Computación
Universidad de Málaga
Email: ruben@lcc.uma.es

Jorge Cuellar
Siemens Corporate Technology
Múnich, Alemania
Email: jorge.cuellar@siemens.com

Javier Lopez
Lenguajes y Ciencias de la Computación
Universidad de Málaga
Email: jlm@lcc.uma.es

***Resumen*—La estación base es el elemento más importante en un red de sensores y, por tanto, es necesario evitar que un atacante pueda hacerse con el control de este valioso dispositivo. Para ello, el atacante puede valerse tanto de técnicas de análisis de tráfico como de la captura de nodos. En este trabajo presentamos un esquema que consta de dos fases, la primera está dedicada a homogeneizar los patrones de tráfico y la segunda encargada de perturbar las tablas de rutas de los nodos. Ambas fases permiten mantener a la estación base fuera del alcance del atacante con un coste computacional insignificante y un consumo energético moderado. La validez de nuestro esquema ha sido validada analíticamente y a través de numerosas simulaciones.**

***Palabras Clave*—Redes de sensores, análisis de tráfico, captura de nodos, seguridad, privacidad de localización.**

I. Introducción

Las redes inalámbricas de sensores (WSNs) [1] son redes ad-hoc compuestas por cientos de pequeños dispositivos inalámbricos (nodos sensores o sensores) alimentados con baterías y capaces de medir ciertas propiedades físicas en su entorno como temperatura, humedad o radiación. Estas mediciones son luego enviadas a un dispositivo, llamado estación base, que se encarga de procesar y analizar los datos recolectados.

Dado que las WSNs se encuentran limitadas en términos de consumo energético, los sensores se valen de sus vecinos para hacer llegar sus mediciones a la estación base. Asimismo, para prolongar aún más el tiempo de vida de la red, los paquetes suelen enviarse utilizando el menor número de intermediarios posible, lo que da lugar a marcados patrones de tráfico (véase Fig. 1). Esto hace que un atacante que se limite a observar el número de paquetes enviados y recibidos en su entorno puede determinar información sensible sobre la red a pesar de que el contenido de los paquetes se encuentre debidamente protegido mediante técnicas criptográficas. En particular, el atacante puede distinguir entre los nodos que generan tráfico, los nodos a los que está destinado y los nodos que sirven de meros intermediarios.

El origen del problema es inherente a las redes de sensores y radica en su particular modelo de comunicación. Supongamos, por ejemplo, que un grupo de biólogos decide desplegar una red de sensores acuáticos para monitorizar el comportamiento y paso de cetáceos por las aguas de un determinado país. La información recopilada por la red es enviada a la estación base que se encuentra abordo de un barco donde los investigadores estudian los datos. En este escenario, existen dos tipos de atacantes que podrían estar interesados en identificar bien el origen o bien el destino de los mensajes. Los atacantes interesados en localizar los nodos origen podrían ser

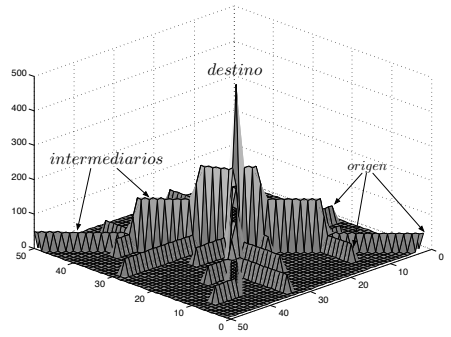


Fig. 1. Tasa de envío en una WSN con 15 emisores

pescadores furtivos porque este tipo de sensor les conduce directamente al cetáceo. Por otra parte, para un grupo de piratas sería atractivo determinar la localización de la estación base ya que esto supondría encontrar el barco. El primer tipo de atacante utilizaría el ángulo de llegada de los paquetes para avanzar hacia el nodo que envía un paquete. Repitiendo este proceso a cada salto, el atacante sería capaz de alcanzar el origen de la comunicación. El segundo tipo de atacante puede valerse de técnicas de monitorización de la tasa de envío y tiempo de envío de paquetes entre vecinos ya que esta información le indica la dirección en la que se encuentra la estación base.

Hasta la fecha los esfuerzos de investigación se han centrado en contrarrestar al primer tipo de atacante [2], [3], [4] mientras que el segundo tipo de atacante ha recibido menos atención [5], [6]. Además, la mayoría de soluciones propuestas para proteger a la estación base son demasiado costosas desde un punto de vista energético o son incapaces de proporcionar un nivel de protección suficiente. Por último es interesante resaltar que un atacante capaz de capturar nodos y obtener sus tablas de rutas daría al traste con los mecanismos de protección propuestos por estas soluciones ya que las tablas de rutas contienen información acerca de la dirección hacia la estación base.

En este trabajo proponemos un protocolo capaz de proteger la localización a la estación base frente a atacantes que realizan análisis de tráfico así como de atacantes capaces de capturar nodos. El protocolo consta de dos esquemas complementarios, uno dedicado a la homogeneización de los patrones de tráfico y el otro a la perturbación de las tablas de rutas de los nodos de manera que estas sean lo suficientemente correctas como para que los paquetes lleguen

al destino y al mismo tiempo oculten la dirección hacia la estación base. A nuestro entender este protocolo es el primero en proporcionar una solución unificada capaz de hacer frente a ambas amenazas.

La organización del trabajo es la siguiente. En la Sec. II hacemos un repaso de las soluciones existentes en la literatura que abordan el problema de la ocultación de la estación base. La Sec. III está dedicada a presentar de manera detallada las características de la red de sensores así como las habilidades de los distintos modelos de atacante considerados en este trabajo. Seguidamente, en la Sec. IV, presentamos una visión general de la solución y los pormenores de los esquemas que la integran. La Sec. V está dedicada a la evaluación de la solución tanto desde el punto de vista de la sobrecarga que introduce en la red como desde el punto de vista del nivel de protección que proporciona. Finalmente, en la Sec. VI se presentan las conclusiones y se esbozan posibles lineas de trabajo futuro.

II. Trabajos relacionados

En [7], [5] se proponen varias técnicas de balanceo del tráfico para solventar el problema de localización de la estación base. En concreto, se presenta un protocolo de encaminamiento en el que a cada salto los nodos envían los paquetes a un nodo arbitrario de entre los más cercanos a la estación base. De esta forma se evita que siempre sean los mismos nodos los que reciben los paquetes, pero el nivel de protección proporcionado es insuficiente. Para tratar de mejorar la solución se propone el envío de paquetes falsos[1] en rutas aleatorias de manera ocasional.

La creación de zonas que reciben un alto volumen de tráfico falso [5], [6] es otra de las técnicas utilizadas para distraer a posibles atacantes. Sin embargo, este tipo de soluciones no sólo requiere una elevada tasa de mensajes sino que además es sólo una medida temporal ya que una vez que el atacante alcanza la zona, la puede descartar. También en [8] se utiliza una gran cantidad de paquetes falsos para hacer que todos los nodos de la red envíen siempre el mismo número de paquetes independientemente de su distancia a la estación base. Esta estrategia es muy costosa ya que implica que todos los nodos de la red estén constantemente generando tráfico falso. Otros trabajos [9] utilizan un enfoque distinto y se basan en que la estación base se comporte como un nodo ordinario o la mueven a otra posición aparentemente más segura. Sin embargo, no siempre es posible mover a la estación base ni sencillo determinar si la nueva posición será realmente segura.

Jian et al. [10] proponen un esquema parecido al nuestro en cuanto a las técnicas de prevención de análisis de tráfico. Su solución se basa, al igual que la nuestra, en enviar los paquetes de datos usando un *biased random walk* (camino aleatorio sesgado) que tratan de ocultar con el envío de paquetes falsos en el sentido contrario con cierta probabilidad. Sin embargo, en ocasiones el atacante puede determinar cuando un paquete es falso y, por tanto, es capaz de obtener la dirección hacia la estación base.

En una versión preliminar de este trabajo [11] conseguimos solucionar algunos de los problemas presentes en trabajos anteriores. Esta nueva versión introduce además un mecanismo de protección capaz de soportar ataques de captura de nodos. Ninguno de los trabajos anteriores había considerado esta amenaza como un problema para la ocultación de la estación base.

[1]Distinguimos entre paquetes (reales) de datos y paquetes (falsos) que contienen basura y cuyo único cometido es despistar al atacante.

III. Descripción del problema

En esta sección presentamos de manera detallada las características de la red así como los modelos de atacante considerados en el resto del artículo.

A. Modelo de red

En este trabajo consideramos WSNs compuestas por un gran número de sensores y una única estación base. Se trata de una red dedicada a la monitorización de eventos y, por tanto, tan pronto como se detecta un fenómeno de interés se envía un mensaje a la estación base.

Asumimos que la conectividad de la red es elevada y que cada nodo conoce a todos sus vecinos gracias a un protocolo de descubrimiento de rutas. Esto permite a los nodos construir sus tablas de rutas de forma que los vecinos que se encuentran más arriba en la tabla son los nodos más próximos a la estación base. En concreto cada nodo puede tener tres tipos de vecinos según su distancia a la estación base: más cercanos, a la misma distancia, o más alejados. Nos referiremos a cada uno de estos grupos como $L^{\mathcal{C}}$, $L^{\mathcal{E}}$ y $L^{\mathcal{F}}$ respectivamente.

Además, supondremos que los nodos comparten claves criptográficas con sus vecinos que les permiten ocultar el contenido de los paquetes. Por tanto, los mensajes con datos reales serán indistinguibles de mensajes falsos.

Finalmente, supondremos que la distancia entre nodos es lo suficientemente amplia como para evitar que el atacante puede observar todas las comunicaciones de manera simultánea. A continuación se dan más detalles sobre las capacidades y estrategias del atacante.

B. Modelo de atacante

El modelo de atacante considerado es capaz de realizar tanto ataques pasivos (análisis de tráfico) como activos (captura de nodos). En ambos casos se trata de un atacante con un ámbito de actuación local y capaz de desplazarse de una lugar a otro de la red.

El rango de acción del atacante *pasivo* viene determinado por el número de nodos que puede observar de manera simultánea. De esta forma, podemos definir $\mathcal{ADV}_n$ como aquel capaz de observar las transmisiones de todos los nodos a distancia menor o igual que n. En general, en la literatura se considera un atacante $\mathcal{ADV}_1$, que tiene un alcance similar al de un nodo ordinario. Tras observar las comunicaciones en su entorno el atacante decide moverse hacia otro nodo que le permita reducir su distancia hasta el destino. Esta decisión depende de si el atacante opta por un ataque por correlación de tiempos o un ataque por volumen de tráfico.

En el ataque por correlación de tiempos (time-correlation) se observa el tiempo de envío de paquetes de un nodo y sus vecinos. Dado que un nodo reenvía un paquete inmediatamente después de recibirlo, el atacante puede deducir la dirección hacia la estación base. El ataque por volumen de tráfico (rate-monitoring) se basa en que la tasa de envío de los nodos más cercanos a la estación base es mayor. El atacante se mueve hacia aquellos nodos con mayor tasa de envío. Es un atacante

menos eficiente ya que requiere hacer varias observaciones antes de tomar la decisión de moverse.

El modelo de atacante *activo* considerado está interesado únicamente en capturar nodos con el fin de obtener sus tablas de rutas ya que con ellas puede determinar qué vecinos del nodo se encuentran más cercanos a la estación base. Tras realizar varias capturas el atacante obtiene información fiel sobre la dirección a seguir para encontrar su objetivo. En la literatura no existe una estrategia de captura claramente definida para la protección de la estación base. Sin embargo, es posible encontrar varios trabajos [12], [13] dedicados al modelado y mitigación de estos ataques durante la distribución de claves. Algunos autores consideran la captura aleatoria de nodos mientras que otros optan por la captura de (algunos o todos) los nodos en una región. En este trabajo consideramos que el atacante es más exitoso si centra su esfuerzo en una región y avanza según la información obtenida. Nótese, que dado el esfuerzo que supone un ataque de este tipo, el atacante sólo podrá comprometer un número reducido de nodos.

IV. Esquema de ocultación

Nuestro esquema de ocultación consta básicamente de dos elementos complementarios que tienen como objetivo alterar los patrones de tráfico y las tablas de rutas de los nodos.

A. Visión general

El protocolo de transmisión consiste básicamente en un *biased random walk* que es ocultado con cantidades controladas de tráfico falso. Ante la recepción de un paquete de datos, el nodo reenvía este paquete hacia la estación base con cierta probabilidad sesgada. Por cada paquete de datos se genera un paquete falso que oculta la dirección del flujo de paquetes reales y la tasa de paquetes reales enviados por cada vecino. De esta forma se homogeneiza localmente el tráfico sin introducir un retraso excesivo en la llegada de paquetes a la estación base.

El algoritmo de perturbación consiste en reordenar la tabla de rutas de cada nodo para que si un atacante tiene acceso a ésta no sea capaz de alcanzar fácilmente la estación base al tener la certeza de que los nodos más próximos se encuentran más altos en la tabla. El nivel de perturbación de la tabla introduce incertidumbre en el atacante pero al mismo tiempo repercute negativamente en el tiempo de llegada de los paquetes.

B. Protocolo de transmisión

El protocolo de transmisión debe cumplir una una serie de propiedades para garantizar la seguridad y usabilidad del sistema. Debemos asegurar que los paquetes de datos alcanzan su destino (Prop. 1) al mismo tiempo que la tasa de envío de paquetes se distribuye uniformemente entre los vecinos (Prop. 2). Finalmente, dado que nuestro protocolo envía parejas de mensajes, la Prop. 3 garantiza que cada uno de estos se envía a un nodo diferente.

Propiedad 1 (Convergencia)**.** *Sea x un nodo arbitrario y BS la estación base. Sea también $neigh(n)$ el conjunto de vecinos de un nodo n. Entonces se dice que el camino es convergente si x elige al siguiente nodo $x' \in neigh(x)$ tal que:*

$$E(dist(x', BS)) < E(dist(x, BS))$$

Algorithm 1 Protocolo de transmisión

Input: $packet \leftarrow receive()$
Input: $combs \leftarrow combinations(sort(neighs), 2)$
Input: $FAKE\_TTL$
1: $\{neigh1, neigh2\} \leftarrow select\_random(combs)$
2: **if** $isreal(packet)$ **then**
3: $\quad send\_random(neigh1, packet, neigh2, fake(FAKE\_TTL))$
4: **else**
5: $\quad TTL \leftarrow get\_time\_to\_live(packet) - 1$
6: $\quad$ **if** $TTL > 0$ **then**
7: $\quad\quad send\_random(neigh1, fake(TTL), neigh2, fake(TTL))$
8: $\quad$ **end if**
9: **end if**

donde E representa el valor esperado y $dist$ es una función de la distancia entre dos nodos.

Propiedad 2 (Homogeneidad)**.** *Sea x un nodo arbitrario y $neigh(n)$ el conjunto de vecinos de un nodo n. Se dice que una transmisión del nodo x mantiene la propiedad de homogeneidad si:*

$$\forall y, z \in neigh(x) \quad Frec_m(x, y) \simeq Frec_m(x, z)$$

donde $Frec_m(x, y)$ representa el total de mensajes enviados por x a y.

Propiedad 3 (Exclusión)**.** *Sean m y m' un par de mensajes y t un tiempo de transmisión determinado. Denotemos $send(m, x, y, t)$ al hecho de transmitir el mensaje m de x a y en el instante t. La propiedad de exclusión establece:*

$$\forall m, m', x, y, t \; send(m, x, y, t) \; \wedge \; m \neq m' \Rightarrow \neg send(m', x, y, t)$$

Dado que cada transmisión consta de dos paquetes, las combinaciones sin repetición de dos elementos de la tabla de rutas es un mecanismo ligero capaz de conseguir una pareja de destinatarios de manera consistente con lo establecido por la Prop. 3. Además, si las tablas están ordenadas (i.e., $[L^{\mathcal{C}}, L^{\mathcal{E}}, L^{\mathcal{F}}]$) se consigue que, con alta probabilidad, el primer elemento de la combinación sea un nodo más próximo a la estación base. Por tanto, si el paquete real lo mandamos al primer vecino y el falso al segundo, estaremos satisfaciendo la propiedad Prop. 1. Finalmente, la Prop. 2 se mantiene si, de entre todas las combinaciones generadas, cada vez se elige una de manera aleatoria.

En Alg. 1 se muestra de manera programática el comportamiento de nuestro protocolo de transmisión. Los argumentos de entrada al algoritmo son el paquete a reenviar, las combinaciones sin repetición de la tabla de rutas ordenada y el parámetro $FAKE\_TTL$, que controla el tiempo de vida de los mensajes falsos en la red y que depende del rango de escucha del adversario. Cuando un nodo recibe un paquete real, se elige una combinación aleatoria de dos vecinos que recibirán el mensaje real y uno falso (líneas 1 a 3). El mensaje falso se reenviará durante $FAKE\_TTL$ saltos. Si el paquete recibido es un paquete falso aún vigente, se reduce su tiempo de vida y se envían dos mensajes falsos (líneas 5 a 7). Además, las parejas de paquetes se envían en un orden aleatorio para evitar que el atacante puede determinar de forma trivial cuál de los paquetes es el real.

C. Perturbación de tablas

Mantener el orden de las tablas de rutas es fundamental para el correcto funcionamiento de nuestro protocolo de transmisión. Sin embargo, esto puede permitir a un atacante determinar qué vecinos se encuentran más próximos a la estación base con solo capturar el nodo y obtener su tabla de rutas. Por ello, es fundamental crear cierta incertidumbre aunque esto conlleve un aumento en el tiempo de entrega de los paquetes.

Definición 1 (Tabla de rutas)**.** *Sea $L^* = L^{\mathcal{C}} \cup L^{\mathcal{E}} \cup L^{\mathcal{F}}$ la lista de todos los vecinos de un nodo n, donde*
$L^{\mathcal{C}} = \{c_1, c_2, c_3, \ldots\}$ son los vecinos más cercanos,
$L^{\mathcal{E}} = \{e_1, e_2, e_3, \ldots\}$ son los vecinos a igual distancia, y
$L^{\mathcal{F}} = \{f_1, f_2, f_3, \ldots\}$ son los vecinos más alejados.

Una tabla de rutas es una biyección $r : \{N-1, \ldots, 1, 0\} \rightarrow L^*$, donde N es el número total de vecinos.

Es decir, una tabla de rutas es una ordenación concreta de los vecinos de un nodo. De manera similar, podemos definir $pos : L^* \rightarrow \{N-1, \ldots, 1, 0\}$ como la inversa de r, de manera que, dado un vecino particular, pos devuelve la posición que éste ocupa en la tabla.

En este punto podemos estudiar bajo qué circunstancias una tabla de rutas está correctamente sesgada (*biased*), esto es, qué ordenaciones permiten la llegada de los paquetes de datos a la estación base.

Teorema 1. *Una tabla de rutas está correctamente sesgada sii* $\sum_{n \in L^{\mathcal{C}}} pos(n) > \sum_{n \in L^{\mathcal{F}}} pos(n)$

En otras palabras, la tabla cumple la propiedad si y sólo si $\mathbb{P}(n_1 \in L^{\mathcal{C}}) > \mathbb{P}(n_1 \in L^{\mathcal{F}})$. Es decir, si al elegir una combinación, la probabilidad de mandar el paquete real a un nodo más próximo a la estación base es mayor que la probabilidad de mandarlo a un nodo más alejado estamos ante una tabla correctamente sesgada.

Demostración:

Asumamos que elegimos aleatoriamente una dupla (n_1, n_2) de vecinos tal que $pos(n_1) > pos(n_2)$. La probabilidad de que n_1 pertenezca al subconjunto $L \subseteq L^*$ viene dada por:

$$\mathbb{P}(n_1 \in L) = \frac{1}{C} \sum_{n \in L} pos(n) \tag{1}$$

donde $C = N * (N-1)/2$ es el total de combinaciones sin repetición de dos elementos de L^*.

Ahora, si escribimos como una lista de duplas todas las combinaciones como una lista de duplas ordenada lexicográficamente, tenemos:

(r(N-1), r(N-2)),	*(r(N-1), r(N-3)),*	*(r(N-1), r(N-4)),*	*…,*	*(r(N-1), r(0))*
	(r(N-2), r(N-3)),	*(r(N-2), r(N-4)),*	*…,*	*(r(N-2), r(0))*
		(r(N-3), r(N-4)),	*…,*	*(r(N-3), r(0))*
				…
				(r(1), r(0))

Observamos que el primer vecino, $r(N-1)$, aparece como primer elemento en $N-1$ duplas, el segundo, $r(N-2)$, en $N-2$, y así sucesivamente. En concreto, el número de veces que un nodo aparece como primer elemento es exactamente la posición que ocupa en la tabla. Con lo que tenemos $(N-1) + (N-2) + (N-3) + \ldots + 1 = C$ duplas.

Algorithm 2 Algoritmo de perturbación

Input: $br \leftarrow \{L^{\mathcal{C}}, L^{\mathcal{E}}, L^{\mathcal{F}}\}$
Input: $bias$, MAX_ITER
1: $E \leftarrow energy(bias, br)$
2: $i \leftarrow 0$
3: **while** $(i < MAX\_ITER) \wedge (E \neq 0)$ **do**
4: $\quad br' \leftarrow swap(br)$
5: $\quad E' \leftarrow energy(bias, br')$
6: $\quad$ **if** $(E' < E)$ **then**
7: $\quad\quad br \leftarrow br'$
8: $\quad\quad E \leftarrow E'$
9: $\quad$ **end if**
10: $\quad i \leftarrow i + 1$
11: **end while**
12: **return** br

Si elegimos cualquier dupla (n_1, n_2) tal que $pos(n_1) > pos(n_2)$, esto equivale a elegir cualquier dupla de la lista anterior. Por tanto, la probabilidad de que un nodo n_1 aparezca como primer elemento de la dupla equivale al número total de elementos en r que se encuentran por debajo de n_1 dividido por el total de combinaciones. Esto es exactamente $pos(n_1)/C$, de donde la Ec. 1 se deduce directamente. ■

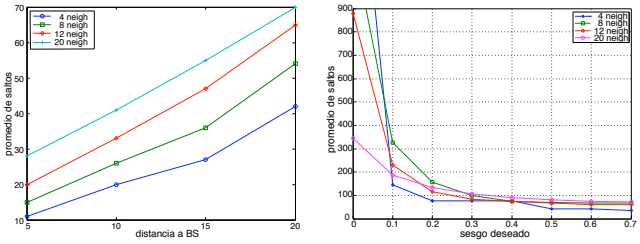
Finalmente, es necesario cuantificar el sesgo de una tabla de rutas, $bias(r) \in [-1, 1]$, ya que es un indicador del tiempo de llegada de los paquetes a la estación base. Cuanto más próximo a 1 más probable es que el siguiente nodo de la ruta se encuentre más próximo a la estación base, mientras que valores próximos a -1 indican que el siguiente nodo se encontrará más alejado. Formalmente puede calcularse como:

$$bias(r) = \frac{1}{C}\left(\sum_{n \in L^{\mathcal{C}}} pos(n) - \sum_{n \in L^{\mathcal{F}}} pos(n)\right) \tag{2}$$

Es sencillo comprobar que si $L^* \equiv L^{\mathcal{F}}$, entonces $bias(r) = -1$ ya que $\sum_{n \in L^{\mathcal{F}}} pos(n) = C$. Del mismo modo, si $L^* \equiv L^{\mathcal{C}}$, entonces $bias(r) = 1$.

Nuestro algoritmo de perturbación recibirá como parámetros un valor de sesgo deseado y una tabla de rutas, y devolverá la tabla reordenada conforme al sesgo dado. En Alg. 2 puede observarse que hemos modelado este algoritmo como un problema de optimización donde la función objetivo (línea 1) depende del valor de sesgo deseado y la ordenación actual de la tabla. En concreto, el algoritmo se inspira en estrategias evolutivas donde intercambiamos dos elementos de la tabla de rutas (línea 4) y comprobamos si así se reduce la distancia al sesgo deseado (línea 6). El proceso se repite por un número máximo de iteraciones o bien hasta que se genere una ordenación acorde al sesgo.

La principal ventaja de utilizar este tipo de estrategia frente a un algoritmo de búsqueda determinista se encuentra en el tiempo necesario para encontrar una solución (seudo-) óptima al problema, que dependiendo del tamaño del espacio de búsqueda puede diferir varios órdenes de magnitud. Sin embargo, su principal desventaja es que, al contrario de los algoritmos deterministas, este tipo de algoritmos puede no encontrar la solución óptima al problema, aunque converge a ella. Nótese, que la perturbación introducida es difícilmente reversible si el valor de sesgo no es conocido, más aún cuando el algoritmo es no determinista.



(a) Numero esperado de saltos (b) Impacto del sesgo

Fig. 2. Tiempo de entrega de paquetes

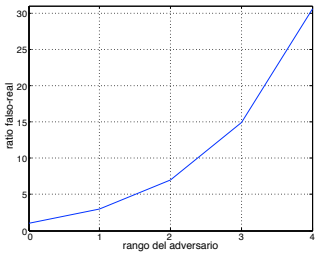
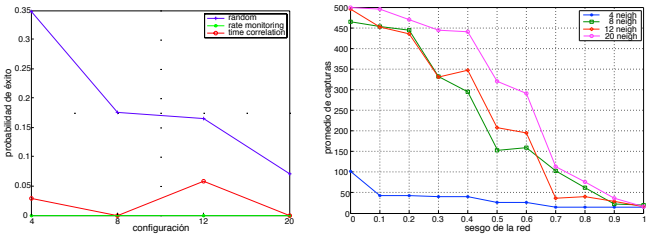


Fig. 3. Ratio de tráfico falso



(a) Ataques de análisis de tráfico (b) Ataques de captura de nodos

Fig. 4. Tasa de éxito de diferentes adversarios

V. Evaluación

En esta sección se evalúa la viabilidad de nuestra solución en relación a la sobrecarga que introduce y al nivel de protección que proporciona frente a distintos modelos de atacante. Las simulaciones se han realizado con MatLab sobre cuatro configuraciones de red en la que variamos el radio de transmisión para conseguir un número promedio de vecinos (4, 8, 12 y 20) diferente por cada nodo.

A. Impacto sobre el tiempo de llegada

La naturaleza probabilística de nuestro protocolo de transmisión influye sobre el tiempo de llegada de los datos a la estación base. En particular, nuestro protocolo puede modelarse como un *biased random walk* donde las probabilidades de enviar hacia la estación base depende del número de vecinos de cada tipo que tenga el nodo.

En la Fig. 2 mostramos el número esperado de saltos para las cuatro configuraciones. En concreto, la Fig. 2a presenta los resultados para nodos origen situados a diferentes distancias (5, 10, 15 y 20 saltos) de la estación base. Como era de esperar, a mayor distancia y mayor conectividad de los nodos, mayor es el número esperado de saltos. Sin embargo, es interesante observar que la velocidad de entrega de los paquetes disminuye cuando los paquetes se acercan a su destino. Esto se debe a que en las proximidades de la estación base hay un mayor número de vecinos $L^{\mathcal{F}}$.

En la Fig. 2b se muestra el impacto que tiene el algoritmo de perturbación sobre el tiempo de entrega. En este experimento todos los nodos están situados a distancia 20. Observamos que a medida que el sesgo se aproxima a cero el tiempo de entrega aumenta siendo este aumento considerablemente mayor para configuraciones con un menor número de vecinos. Esto se debe a que las configuraciones con menos vecinos tienen menos formas de modificar las tablas de rutas. En concreto, cuando el sesgo deseado es cero, el sesgo promedio de la red para la configuración de cuatro vecinos es ligeramente inferior a cero, mientras que para la configuración de veinte vecinos el sesgo promedio está próximo a 0.1. En general, para un sesgo superior a 0.2 la longitud media de los caminos es inferior a 100 saltos.

B. Sobrecarga de tráfico falso

Nuestro protocolo de transmisión se basa en el envío de paquetes falsos para ocultar el flujo de tráfico real. Sin embargo, es necesario controlar la propagación de estos paquetes para evitar el consumo excesivo de energía. Para ello definimos un parámetro, $FAKE\_TTL$, cuyo valor depende del rango de escucha del atacante y que limita el número de paquetes falsos generados.

En la Fig. 3 mostramos el ratio de mensaje falsos frente al tráfico real dependiendo del rango de escucha del atacante. Cuando el adversario sólo escucha los paquetes en su entorno inmediato, $\mathcal{ADV}_0$, el ratio es 1 porque cada mensaje real va acompañado de un mensaje falso que no vuelve a propagarse. A medida que el rango de escucha del adversario (n) aumenta, el ratio lo hace en el orden de $\mathcal{O}(2^{n+1})$.

Nótese que el modelo de atacante más usual en la literatura es $\mathcal{ADV}_1$, es decir, aquel con un rango de escucha similar al de un nodo ordinario.

C. Proteccion frente atacantes

Con el fin de validar la robustez de nuestra solución hemos lanzado simulaciones con atacantes que realizan análisis de tráfico o captura de nodos. En la Fig. 4a se muestra como un modelo de atacante que se mueve de manera aleatoria, sin tener en cuenta las comunicaciones, tiene más probabilidades de llegar a la estación base que aquellos que recurren a técnicas de monitorización del tiempo y tasa de envío de paquetes. Además, como era de esperar, su tasa de éxito es mayor en configuraciones con un promedio de vecinos más bajo. Obsérvese que del total de simulaciones lanzadas, el atacante que realiza monitorización de la tasa de envío nunca llega a la estación base mientras que el que realiza correlación de tiempos lo consigue en limitadas ocasiones. Las ocasiones en las que éste localiza a la estación base se debe a que inicialmente se encuentra a distancia 5 y a la naturaleza de nuestro simulador, que es incapaz de determinar exactamente qué paquete es enviado antes. Por tanto, este atacante elige el siguiente salto de forma aleatoria entre los vecinos que envían mensajes.

En la Fig. 4b, el adversario comienza en un punto del extremo de la red y puede capturar hasta 500 nodos para llegar a la estación base. Además, asumimos que el atacante puede moverse al siguiente vecino tras obtener su identificador

aunque en un escenario real puede necesitar capturar a los vecinos del nodo para saber a cuál de ellos corresponde el identificador encontrado. La estrategia del atacante es moverse al primer nodo de la tabla de rutas que ha visitado un menor número de veces para evitar quedar atrapado en bucles. Los resultados muestran que, a medida que el sesgo de la red se acerca a cero, el adversario necesita capturar un mayor número de nodos para llegar a su destino. Sin embargo, un sesgo bajo influye negativamente en el tiempo de llegada de los paquetes a la estación base (ver Sec. V-A). En general, si consideramos que un atacante podría capturar hasta una décima parte de los nodos de la red, sería seguro utilizar un valor de sesgo menor o igual a 0.5. Nótese que el número de nodos de la red es respectivamente de 400, 1600, 1600 y 3600 para las configuraciones de 4, 8, 12 y 20 vecinos, respectivamente.

VI. Conclusiones

En este trabajo hemos presentado un esquema que permite ocultar la localización de la estación base para protegerla así de posibles ataques. Nuestra solución consta de dos esquemas complementarios capaces de hacer frente a atacantes que realizan tanto análisis de tráfico como capturas de nodos de la red. El primer esquema es un protocolo de transmisión que utiliza cantidades moderadas de mensajes falsos para ocultar el flujo de datos. En concreto, el protocolo preserva tres propiedades (convergencia, homogeneidad y exclusión) lo cual garantiza la llegada de paquetes a la estación base al mismo tiempo que interfiere con los ataques de correlación de tiempos y de volumen de tráfico. El segundo esquema se trata de un algoritmo evolutivo que tiene como objetivo perturbar las tablas de rutas de los nodos para evitar que si un atacante es capaz de obtener estas tablas pueda determinar con facilidad en qué sentido avanzar para encontrar la estación base. Este algoritmo de perturbación está regido por un valor de sesgo, que determina la cantidad de perturbación introducida en las tablas. Este valor introduce un compromiso entre el nivel de protección obtenido y el tiempo medio de espera para recibir los paquetes en la estación base.

La viabilidad de nuestra solución has sido validada analíticamente y a través de simulaciones. En concreto, hemos estudiado el impacto que tiene la conectividad de la red sobre la convergencia paquetes y el nivel de protección de la estación base. Además, hemos analizado el tiempo medio de llegada de los paquetes de datos a su destino y la sobrecarga que supone la inyección de mensajes falsos. Finalmente, hemos evaluado el nivel de protección obtenido frente a atacantes capaces de realizar ataques activos y pasivos.

Como trabajo futuro tenemos como objetivo investigar mecanismos para reducir el número de mensajes falsos requerido para proteger a la estación base de atacantes con un amplio rango de escucha. Además, queremos explorar la robustez de nuestro esquema frente a atacantes más inteligentes. Para ello, en primer lugar será necesario definir una serie de estrategias basadas en el conocimiento del adversario acerca de la red y el esquema de protección utilizado. Este tipo de adversario podría modificar su estrategia de ataque dependiendo del contexto. Para hacer frente a este tipo de atacantes puede ser necesario desarrollar nuevos mecanismos de protección más sofisticados que los considerados hasta la fecha. Finalmente, entre nuestros objetivos se encuentra el desarrollar un sistema de protección integral, que al mismo tiempo sea capaz de hacer frente a atacantes interesados en determinar la localización de la estación base y atacantes cuyo objetivo sea obtener la localización de los nodos origen de eventos.

Agradecimientos

Este trabajo ha sido parcialmente financiado por la Comisión Europea a través del proyecto NESSoS (FP7 256890) y el Ministerio de Innovación y Ciencia a través de los proyectos SPRINT (TIN2009-09237) e IOT-SEC (ACI2009-0949). SPRINT está co-financiado con fondos FEDER. El primer autor es becario FPU del Ministerio de Educación.

Referencias

[1] C. Gómez, J. Paradells, and J. E. Caballero, *Sensors Everywhere: Wireless Network Technologies and Solutions*, Fundación Vodafone España, Ed. Fundación Vodafone España, 2010, iSBN 978-84-934740-5-8. [Online]. Available: http://fundacion.vodafone.es/static/fichero/pre_ucm_mgmt_002618.pdf

[2] C. Ozturk, Y. Zhang, and W. Trappe, "Source-Location Privacy in Energy-Constrained Sensor Network Routing," in *2nd ACM workshop on Security of ad hoc and sensor networks (SASN '04)*, 2004, pp. 88–93.

[3] R. Rios and J. Lopez, "Exploiting Context-Awareness to Enhance Source-Location Privacy in Wireless Sensor Networks," *The Computer Journal*, vol. 54, no. 10, pp. 1603–1615, 2011.

[4] B. Alomair, A. Clark, J. Cuellar, and R. Poovendran, "Statistical Framework for Source Anonymity in Sensor Networks," in *IEEE Global Telecommunications Conference (GLOBECOM 2010)*, 2010, pp. 1 –6.

[5] J. Deng, R. Han, and S. Mishra, "Decorrelating wireless sensor network traffic to inhibit traffic analysis attacks," *Pervasive and Mobile Computing*, vol. 2, no. 2, pp. 159–186, 2006.

[6] S. Chang, Y. Qi, H. Zhu, M. Dong, and K. Ota, "Maelstrom: Receiver-Location Preserving in Wireless Sensor Networks," in *Wireless Algorithms, Systems, and Applications*, ser. LNCS. Springer, 2011, vol. 6843, pp. 190–201.

[7] J. Deng, R. Han, and S. Mishra, "Countermeasures Against Traffic Analysis Attacks in Wireless Sensor Networks," in *1st International Conference on Security and Privacy for Emerging Areas in Communications Networks (SECURECOMM '05)*, 2005, pp. 113–126.

[8] B. Ying, J. R. Gallardo, D. Makrakis, and H. T. Mouftah, "Concealing of the Sink Location in WSNs by Artificially Homogenizing Traffic Intensity," in *1st International Workshop on Security in Computers, Networking and Communications*, 2011, pp. 1005–1010.

[9] U. Acharya and M. Younis, "Increasing base-station anonymity in wireless sensor networks," *Ad Hoc Networks*, vol. 8, no. 8, pp. 791–809, 2010.

[10] Y. Jian, S. Chen, Z. Zhang, and L. Zhang, "Protecting receiver-location privacy in wireless sensor networks," in *26th IEEE International Conference on Computer Communications (INFOCOM 2007)*, 2007, pp. 1955–1963.

[11] R. Rios, J. Cuellar, and J. Lopez, "Robust Probabilistic Fake Packet Injection for Receiver-Location Privacy in WSN," in *17th European Symposium on Research in Computer Security (ESORICS 2012)*, ser. LNCS, M. Y. S. Foresti and F. Martinelli, Eds., vol. 7459, Springer. Pisa, Italy: Springer, Sept. 2012, pp. 163–180.

[12] X. Chen, K. Makki, K. Yen, and N. Pissinou, "Node Compromise Modeling and its Applications in Sensor Networks," in *12th IEEE Symposium on Computers and Communications (ISCC 2007).*, july 2007, pp. 575 –582.

[13] T. M. Vu, R. Safavi-Naini, and C. Williamson, "Securing wireless sensor networks against large-scale node capture attacks," in *Proceedings of the 5th ACM Symposium on Information, Computer and Communications Security*, ser. ASIACCS '10. New York, NY, USA: ACM, 2010, pp. 112–123.